

**Implementation and applications of gas chromatography/
atmospheric pressure chemical ionization time-of-flight
mass spectrometry in metabolomics**

Dissertation

zur Erlangung des Doktorgrades der Naturwissenschaften (Dr. rer. nat.)

an der Fakultät für

Chemie und Pharmazie

der Universität Regensburg



vorgelegt von

Christian J. Wachsmuth

aus Neustadt a. d. Donau

im Jahr 2015

Diese Doktorarbeit entstand in der Zeit von April 2011 bis Januar 2015 am Institut für Funktionelle Genomik der Universität Regensburg.

Die Arbeit wurde angeleitet von Prof. Dr. Peter J. Oefner.

Promotionsgesuch eingereicht am 29. Juni 2015.

Für meine Eltern

“Nothing worth having comes easy.” Dr. Bob Kelso (Scrubs)

Danksagung

Zu guter Letzt freue ich mich, die letzte Seite noch voll zu machen und dabei zum Abschluss einer Vielzahl an Menschen zu danken. Denjenigen, die mich schon vor der Dissertation unterschützt haben und genauso den lieben Menschen, die während der Zeit des Promovierens hinzugekommen sind und mich begleitet haben.

An erster Stelle danke ich **Prof. Dr. Peter J. Oefner**, dass ich an diesem interessanten Thema arbeiten konnte und für die fachliche Begleitung und Unterstützung dabei. Ich habe den gewissen zugestanden Freiraum für eigene Ideen sehr zu schätzen gewusst genauso wie die Erfahrungen, die ich auf den zahlreichen Konferenzen sammeln durfte.

Des Weiteren gilt mein Dank **Prof. Dr. Frank-Michael Matysik** für die freundliche Übernahme des Zweitgutachtens sowie die stets äußerst kooperative und unkomplizierte Art. **Prof. Dr. Joachim Ruther** und **Prof. Dr. Joachim Wegener** bin ich dankbar für die Übernahme der Drittprüferschaft bzw. des Prüfungsvorsitzes.

Ein herzliches „Vergelt's Gott“ an **Dr. Katja Dettmer-Wilde** für die direkte Betreuung und Begleitung meiner wissenschaftlichen Pubertätsphase und den dabei weitergegebenen Chromatographie-Erfahrungsschatz. Darüber hinaus für die Vorbereitung auf das spätere „harte Arbeitsleben“, was mir sicherlich gut weiterhelfen wird.

Ich bin überaus froh gewesen, **Magda Waldhier** und **Franzi Vogl** am IFG als Wegbegleiterinnen gehabt zu haben. Vielen Dank euch beiden für das offene Ohr, die aufbauenden Worte und Komplimente und den Glauben an mich. Magda danke ich insbesondere noch für die überaus bereitwillige Weitergabe ihres Wissens bei meinen ersten GC-MS Schritten, die abendlichen Wiener-Breaks und die Geduld mit mir. Franzi danke ich dafür, dass sie nach unzähligen Bachelorettes und Bachelors zur Konstante an meiner rechten Seite wurde und für ihr heiteres und liebenswertes Wesen.

Ebenso werde ich meinen Notduschen- und Tanz-Buddy **Nadine Nürnberger** und Märchentante **Claudia Samol** im Alltag vermissen. Nadine, du warst mir darüber hinaus eine hervorragende Ausbildungsleiterin für soziale Belange, stets an meiner Seite für gewünscht mal mehr und mal weniger tiefsinnige Gespräche und hast mir viele

Glücksmomente als Beifahrer beschert. Claudi, du fehlst mir wegen deiner treuen und liebenswerten Art und von niemandem sonst lass ich mir so gern Horrorgeschichten erzählen, die niemals passiert sind.

Inzwischen schon ein Weilchen her, aber nicht weniger herzlichen Dank an **Dr. Martin Almstetter** für die große Hilfe am Anfang meiner IFG-Zeit, die Aufnahme in die erlebte Jean-Pierre Runde und ein aufrichtiges „Sorry“ für die Sabotage deiner Laborarbeiten durch das Verstecken der GC Crimp-Vials.

Sehr bedanken möchte ich mich auch bei den Mitgliedern der Proteomics-Gruppe. Allen voran bei **Dr. Nadine Assmann** für Gurkenbrote zum richtigen Zeitpunkt, Gesellschaft zu „ungewöhnlichen“ Bürozeiten und der anschließenden Heimfahrt, um mich vor Räubern und Verbrechern zu schützen. Insgesamt lässt mich das darüber hinweg sehen, dass ich das versprochene Borschtsch-Rezept niemals erhalten habe, werde es aber beim nächsten Wiedersehen erneut einfordern. Des Weiteren danke an die übrigen „Proteomiker and friends“ **Elke Perthen** und **Hans Simbürger** für die gute Laune, **Dr. Jörg Reinders** und **Dr. Alex Riechers** für viele Tipps und wissenschaftliche Zerstreuung, **Dr. Yvonne Reinders** für die Alkoholzufuhr in Kuchenform (der Rumkuchen bleibt legendär) und **Niklas „Ramon“ Reinders** fürs Teilen von Kinderbüchern auf langen Autofahrten.

Gerne erinnere ich mich auch an die NMR-ler **Dr. Matthias Klein** (danke für die angenehme gemeinsame Bürozeit), **Helena Zacharias**, meinen TuMet-Buddy **Jochen Hochrein** sowie **Philipp Schwarzfischer**. Ihnen bin ich insbesondere dankbar für Hilfestellungen im Zusammenhang mit Statistik und **Prof. Dr. Wolfram Gronwald** dazu noch für viele freundliche Ratschläge, seine große Hilfsbereitschaft und die gemeinsam geteilte Vorliebe für Kuchen aller Art.

Ferner besonders bedanken möchte ich mich bei allen Kooperationspartnern, vor allem von Bruker Daltonik (Bremen) und der Uni-Klinik Regensburg, den weiteren ehemaligen IFG-lern **Carry Louis** (für die Ananas-Vitamin/Zuckerschocks), **Dr. Wenni Zhu**, **Dr. Axel Stevens**, **Sophie Schirmer**, **Anja Thomas**, **Conny Feuchtinger**, **Sabine Botzler** und **Michaela Schröpfer**, den noch recht „Neuen“ **Lisa Ellmann** und **Trixi von Schlippenbach**, mit denen es zwar ein kurzes Vergnügen war aber sicherlich gut geworden wäre, sowie **Dr. Inka Appel**, **Franziska Taruttis** und **Christian Kohler** aus der Bioin-

formatik-Abteilung für ihre Hilfe. Dazu noch **Jutta Schipka** und **Susanne Schwab** für das eine oder andere nette Gespräch beim Essen und **Eva Engl** für Organisatorisches und den Schmetterling an meiner Tür. Allen weiteren Mitarbeitern und Ehemaligen des IFGs danke ich für jegliche Hilfe und das angenehme Klima in der Gruppe.

Ich hatte das Vergnügen, im Laufe meiner Zeit am IFG neben der im Zusammenhang mit einer Promotion immer wieder genannten vielen Arbeit und benötigten Ausdauer auch viele schöne Momente zu erleben. Deshalb werde ich mich auch mit Freude zurück erinnern und gerne die Verbindung zu den liebgewonnenen Menschen aufrechterhalten.

Danke an meine Freunde von außerhalb und meine Familie für die aufgebrachte Geduld, das offene Ohr und den ständigen Zuspruch. Der größte Dank überhaupt gebührt zuletzt meinen Eltern **Karin und Herbert Wachsmuth** für die große Unterstützung, den Zusammenhalt in schwierigen Phasen und das Gefühl, angenommen und geliebt zu werden, ganz egal, ob ich diese letzten Zeilen und somit die Dissertation jemals zu Ende gebracht hätte oder nicht.

1 Table of contents

1	TABLE OF CONTENTS	I
2	ABBREVIATIONS AND ACRONYMS.....	VI
3	MOTIVATION.....	1
4	BACKGROUND.....	6
4.1	METABOLOMICS.....	6
4.1.1	CHALLENGES	7
4.1.2	STRATEGIES	8
4.2	ESTABLISHED GC-MS APPROACHES IN METABOLOMICS	9
4.2.1	DERIVATIZATION STRATEGIES	10
4.2.1.1	MeOx-TMS derivatization.....	10
4.2.1.2	Derivatization with alkyl chloroformates	12
4.2.2	GC STATIONARY PHASES	13
4.2.3	GC-MS WITH EI AND CI SOURCES.....	14
4.2.4	COMPREHENSIVE TWO-DIMENSIONAL GAS CHROMATOGRAPHY – TIME-OF-FLIGHT MASS SPECTROMETRY	16
4.3	ATMOSPHERIC PRESSURE CHEMICAL IONIZATION FOR GC-MS.....	18
4.3.1	FUNDAMENTALS	18
4.3.2	REVIEW OF GC(×GC)-APCI-MS APPLICATIONS.....	29
4.4	EXPERIMENTAL DESIGN AND DATA ANALYSIS IN METABOLOMICS	33
4.4.1	SETTING UP AN ANALYSIS.....	34
4.4.2	DATA ANALYSIS WORKFLOW	37

4.4.3	RECALIBRATION OF HIGH-RESOLUTION GC-APCI-TOFMS DATA AND CALCULATION OF ELEMENTAL FORMULAS.....	39
4.5	IDENTIFICATION OF METABOLITES	41
4.5.1	GAS CHROMATOGRAPHY.....	41
4.5.2	MASS SPECTROMETRY.....	44
5	EXPERIMENTAL SECTION	47
5.1	MATERIALS.....	47
5.2	SAMPLE PREPARATION	48
5.2.1	PREPARATION OF CALIBRATION AND CELL CULTURE SAMPLES.....	48
5.2.2	METHOXIMATION–TRIMETHYLSILYLATION.....	49
5.2.3	METHYL CHLOROFORMATE/METHANOL DERIVATIZATION	50
5.3	INSTRUMENTATION	50
5.3.1	GC-APCI-TOFMS	50
5.3.2	MISCELLANEOUS.....	52
5.4	DATA ANALYSIS	52
5.4.1	SOFTWARE	52
5.4.2	CALIBRATION CURVES.....	54
5.4.3	RECALIBRATION OF MASS SPECTRA.....	54
5.4.4	CALCULATION OF ELEMENTAL FORMULAS	56
5.5	VALIDATION METHODS	59
5.5.1	ROC CURVE	59
5.5.2	BLAND-ALTMAN PLOT	61
6	INVESTIGATIONS ON THE EFFECTS OF CONTINUOUS WATER INFUSION ON APCI OF DERIVATIZED METABOLITES	62
6.1	INTRODUCTION.....	62

6.2 MATERIALS AND METHODS.....	63
6.2.1 MATERIALS.....	63
6.2.2 SAMPLE PREPARATION.....	63
6.2.3 GC-APCI-TOFMS ANALYSIS.....	64
6.2.4 CROSS-VALIDATION OF AMINO ACIDS.....	64
6.2.5 FEATURE EXTRACTION AND ALIGNMENT.....	65
6.3 RESULTS AND DISCUSSION.....	66
6.3.1 EFFECTS OF CONTINUOUS WATER INFUSION ON APCI MASS SPECTRA OF MCF AND MeOX-TMS DERIVATIVES.....	66
6.3.2 OPTIMIZATION OF WATER INFUSION RATE.....	68
6.3.3 METHOD EVALUATION VIA SPIKE-IN EXPERIMENT.....	70
6.3.4 EVALUATION OF FOLD CHANGES AND MASS AND ISOTOPE ACCURACY.....	74
6.3.5 ANALYSIS OF METABOLIC FINGERPRINTS OF CANCER CELL EXTRACTS.....	75
6.3.6 COMPARATIVE ANALYSIS OF CONTROL AND 17-DMAG -TREATED SAMPLE GROUPS.....	79
6.3.7 CROSS-VALIDATION OF DISCRIMINATING AMINO ACIDS.....	82
6.4 CONCLUSIONS.....	83
7 ASSESSMENT OF MATRIX EFFECTS IN GC-APCI-MS.....	85
7.1 INTRODUCTION.....	85
7.2 MATERIALS AND METHODS.....	86
7.2.1 SPIKE-IN EXPERIMENT.....	86
7.2.2 CO-ELUTING ANALYTE PAIRS.....	88
7.3 RESULTS AND DISCUSSION.....	89
7.3.1 SELECTION OF INTERNAL STANDARDS FOR QUANTIFICATION OF SPIKE-IN METABOLITES.....	89

7.3.2	RECOVERY RATES OF 15 SPIKED METABOLITES IN THREE DIFFERENT MATRICES.....	91
7.3.3	SLOPES OF CALIBRATION CURVES OF SIX ANALYTES IN THE PRESENCE OF CO-ELUTING COMPOUNDS	95
7.4	CONCLUSIONS	98
8	PERFORMANCE EVALUATION OF A REDESIGNED APCI SOURCE	100
8.1	INTRODUCTION.....	100
8.2	MATERIALS AND METHODS.....	101
8.2.1	MATERIALS	101
8.2.2	INSTRUMENTATION	101
8.2.3	DATA ANALYSIS.....	104
8.2.3.1	FAMES.....	104
8.2.3.2	Cell culture supernatant samples	104
8.3	RESULTS AND DISCUSSION.....	106
8.3.1	GC-APCI-HRTOFMS ANALYSIS OF FAMES WITH THE APCI I AND II SOURCES	106
8.3.2	COMPARISON OF GC-APCI-HRTOFMS CALIBRATION CURVES	110
8.3.3	ANALYSIS OF CELL CULTURE SUPERNATANTS.....	114
8.4	CONCLUSIONS	123
9	DEVELOPMENT OF AN ENANTIOSELECTIVE QUANTITATIVE PROFILING METHOD FOR THE ONCOMETABOLITE D-2-HYDROXYGLUTARATE	124
9.1	INTRODUCTION.....	124
9.2	METHODS.....	127
9.2.1	SAMPLE PREPARATION.....	127
9.2.2	ANALYTICAL APPROACHES	128
9.3	RESULTS AND DISCUSSION.....	131

9.3.1 SELECTION OF A SPECIFIC QUANTIFIER ION TRACE FROM APCI-TOFMS VERSUS EI-QMS MASS SPECTRA.....	131
9.3.2 IONIZATION BEHAVIOR OF D/L-2-HG IN GC-APCI-TOFMS INFUSING WATER AS MODIFIER.....	135
9.3.3 METHOD VALIDATION	137
9.3.4 LACTONIZATION OF D/L-2-HG IN DERIVATIZED STANDARDS AND BIOLOGICAL SAMPLES.....	140
9.4 CONCLUSIONS	143
10 CONCLUSIONS AND PERSPECTIVES	145
11 REFERENCES	148
12 APPENDIX	156
13 CURRICULUM VITAE	186
14 PUBLICATIONS AND PRESENTATIONS	187
14.1 PUBLICATIONS	187
14.2 ORAL PRESENTATIONS.....	189
14.3 POSTER PRESENTATIONS	190
15 SUMMARY	191
16 ZUSAMMENFASSUNG.....	194

2 Abbreviations and Acronyms

(v/v)	Volume/volume percent
[M+H] ⁺	Protonated molecule of metabolite M
17-DMAG	17-Dimethylaminoethylamino-17-demethoxygeldanamycin
2-HG	2-Hydroxyglutarate
3-P-glycerate	3-Phosphoglycerate
5-HIAA	5-Hydroxyindoleacetate
Ala	Alanine
AML	Acute myeloid leukemia
ANOVA	Analysis of variance
AP	Atmospheric pressure
APCI	Atmospheric pressure chemical ionization
APCI I	First-generation APCI source
APCI II	Second generation APCI source
APCI(+)	Atmospheric pressure chemical ionization positive mode
APCI/+H ₂ O	APCI with water infusion
APCI/-H ₂ O	APCI without water infusion
API	Atmospheric pressure ionization
APLI	Atmospheric pressure laser ionization
APPI	Atmospheric pressure photo ionization

Asn	Asparagine
Asp	Aspartate
AUC	Area under the curve
CD	Cyclodextrin
CI	Chemical ionization
CID	Collision-induced dissociation
Cys	Cysteine
D/L-2-HG	D and L enantiomers of 2-hydroxyglutarate
Da	Dalton
DA-APLI	Dopant-assisted atmospheric pressure laser ionization
DA-APPI	Dopant-assisted atmospheric pressure photo ionization
DART	Direct Analysis in Real Time
DB	Database
DIP	Direct Inlet Probe
<i>E. coli</i>	<i>Escherichia coli</i>
EI	Electron ionization
EIC	Extracted ion chromatogram
ESI	Electrospray ionization
FAME	Fatty acid methyl ester
FC	Fold change
FDA	Food and Drug Administration
FDR	False discovery rate

FID	Flame ionization detector
FMF	Find Molecular Features
FN	False negative
FP	False positive
FPR	False positive rate
FTICR	Fourier transform ion cyclotron resonance
G6P	Glucose-6-phosphate
GABA	Gamma amino butyrate
GC	Gas chromatography
GC×GC	Comprehensive two-dimensional gas chromatography
GLC	Gas liquid chromatography
Gln	Glutamine
Glu	Glutamate
Gly	Glycine
Glycerol-1-P	Glycerol-1-phosphate
GSC	Gas solid chromatography
HDA	Heptadecanoate
His	Histidine
HMDB	Human Metabolome Database
HPLC	High performance liquid chromatography
HRTOF	High-resolution time-of-flight
Hsp90	Heat-shock protein 90

IAS	Ion activation stage
IDH	Isocitrate dehydrogenase
Ile	Isoleucine
IPT	Isopentenyl phosphotransferase encoding gene
IS	Internal standard
KEGG	Kyoto Encyclopedia of Genes and Genomes
LC	Liquid chromatography
Leu	Leucine
LLOQ	Lower limit of quantification
LOD	Limit of detection
$\log P_{\text{octanol/water}}$	Logarithm of partition coefficient
LOQ	Limit of quantification
LR	Linear range
Lys	Lysine
<i>m/z</i>	mass-to-charge ratio
MCF	Methyl chloroformate
Me	Methyl
MeOx	Methoximation
Met	Methionine
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry

MS ^E	Simultaneous acquisition mode of high and low collision energy spectra
MS ⁿ	Tandem mass spectrometry
MSTFA	N-methyl-N-(trimethylsilyl) trifluoroacetamide
NA	Nonanoate
NAA	N-acetylaspartate
NADP ⁺ /NADPH	Nicotinamide adenine dinucleotide phosphate (oxidized and reduced form)
NDA	Nonadecanoate
netCDF	Network Common Data Format
NIST	National Institute of Standards and Technology
NMR	Nuclear magnetic resonance
NPAH	Nitrated polycyclic aromatic hydrocarbon
Nval	Norvaline
OH	Hydroxy
Orn	Ornithine
PAH	Polycyclic aromatic hydrocarbon
PBM	Probability-based matching
PBS	Phosphate-buffered saline
PCA	Principal component analysis
PCF	Propyl chloroformate
PDA	Pentadecanoate
PFPP	Pentafluorophenyl propyl

Phe	Phenylalanine
PLS-DA	Partial least squares discriminant analysis
ppm	parts per million
ppmV	parts per million by volume
Pro	Proline
PTV	Programmed-temperature vaporization
PU	<i>E. coli</i> double-mutant (PntAB-UdhA) strain
qMS	Quadrupole mass spectrometry
qTOF	Quadrupole- Time-of-flight hybrid mass spectrometry
R ²	Square of the linear regression coefficient R
RI	Linear retention index
ROC	Receiver operator characteristic
RSD	Relative standard deviation
Rt- γ DEXsa	2,3-di-acetoxy-6-O-tert-butyl-dimethylsilyl gamma CD doped into 14% cyanopropylphenyl/86% dimethyl polysiloxane
S/N	Signal-to-noise ratio
SD	Standard deviation
Ser	Serine
SIL-IS	Stable isotope-labeled internal standard
SIM	Selected ion monitoring
SP	Stationary phase
SVM	Support vector machines
TCA	Tricarboxylic acid

TDA	Tridecanoate
Thr	Threonine
TIC	Total ion current
TMS	Trimethylsilyl
TN	True negative
TOCSY	Total correlated spectroscopy
TOF	Time-of-flight
TP	True positive
TPR	True positive rate
Trp	Tryptophan
TSP	Trimethylsilyl propanoic acid
Tyr	Tyrosine
U- ¹³ C	Uniform ¹³ C-labeled
U- ² H	Uniform ² H-labeled
UDA	Undecanoate
ULOQ	Upper limit of quantification
UTI	Urinary tract infection
Val	Valine
VOC	Volatile organic compound
w_h	Peak width measured at half the peak height
wt	wild type
Δm	mass error

3 Motivation

Metabolomics aims at the measurement of the largest possible number of metabolites in a single analysis. The ultimate goal is the identification and quantification of those that differ significantly between groups. Knowing the identity and quantity of discriminating metabolites allows the interpretation of metabolic changes in the context of metabolic pathways, which is essential for understanding the underlying biology.

Gas chromatography – electron ionization – mass spectrometry (GC-EI-MS) in combination with a derivatization is a valuable tool in metabolomics due to high analytical reproducibility, low detection limits, and the availability of large mass spectral libraries that also include trimethylsilyl derivatives that are commonly used. Nevertheless, many derivatives of metabolites are not found in standard commercial libraries rendering the annotation of corresponding signals in the GC-MS chromatogram cumbersome. Elemental formulas can be calculated from the accurate mass of molecular ions as a first step towards the *de-novo* identification of unknown metabolites, but electron ionization causes strong fragmentation and does not always yield the molecular ion. Hence, soft ionization techniques, such as chemical ionization (CI), must be applied. In the 1970s, Horning et al. introduced atmospheric pressure chemical ionization (APCI) as a soft ionization technique for coupling GC to MS [1,2], but it was not widely used at that time because of the required expensive instrumentation and it did not reach maturity for commercialization. Another 30 years later, McEwen and McKay as well as Schiewek et al. continued the studies on APCI and developed independently ion sources suited for coupling mass spectrometry with both gas and liquid chromatography [3,4]. Their work led in 2008 to the introduction of commercially available APCI sources for GC-MS and their hyphenation with high-resolution mass analyzers to exploit such powerful features as accurate mass measurement and tandem mass spectrometry for the identification of unknowns. The great potential of high-resolution time-of-flight mass spectrometry (HRTOF-MS) in metabolomics is not only based on accurate mass measurement for

structural assignment of unknown species but also on the option to perform targeted quantification simultaneously. The latter benefitted from the introduction of the analog-to-digital converter, which led to a distinct increase in the dynamic range of TOFMS instruments [5]. This was highly needed as metabolites vary tremendously in their concentration levels. Since GC-APCI-TOFMS has shown high detection sensitivity in initial studies [6], it appeared to be promising for quantitative metabolic profiling.

Prior to my master thesis in 2010 and this doctoral research work, there had been few reports on the use of GC-APCI-HRTOFMS. Reported applications included the identification of unknown compounds in pharmaceutical research [7], analysis of foodstuff [8], and metabolomics [9-11]. Its performance had not been compared to established GC-MS approaches. The technique had only been applied to study metabolite composition in human cerebrospinal fluid [9], but comparative metabolic fingerprinting of several sample classes had not been performed to that date. Furthermore, the outstanding quantitative capabilities of GC-APCI-HRTOFMS had only been demonstrated using metabolite standards [9,10] and not been used for metabolic profiling, *i.e.*, the quantitative analysis of known metabolites, in biological specimens.

My master thesis and the research done by Carrasco et al. in 2009 [9] were the first GC-APCI-HRTOFMS studies that investigated the ionization behavior of a broad range of metabolites. In my master thesis, 43 metabolites from different compound classes, such as amino acids, organic acids, sugars, and alcohols, were analyzed. APCI source parameters were optimized and the quantitative performance of GC-APCI-HRTOFMS was comprehensively compared to that of GC×GC-EI-TOFMS, GC-EI-TOFMS, GC-CI-quadrupole mass spectrometry (qMS) and GC-EI-qMS, respectively. In addition, comparative metabolic fingerprinting was performed in wild type and mutant *Escherichia coli* (*E. coli*) strains and a workflow for the identification of unknown metabolites was established based on the accurate mass and isotopic pattern of their protonated molecules ($[M+H]^+$). For validation, results were compared to a previous study that had employed GC×GC-EI-TOFMS [12]. Although the general applicability of GC-APCI-HRTOFMS in metabolomics could be shown, the quantitative performance of GC-APCI-HRTOFMS was less than satisfying, mainly because of inconsistent ionization. Factors influencing

ionization efficiency and reproducibility needed to be studied in more detail, such as the infusion of water into the ionization source and potential matrix effects due to co-eluting compounds. Hence, the initial results of my master thesis formed the basis for this doctoral thesis.

Aim #1: Investigation and optimization of factors influencing APCI

1.1 Water infusion

Already in the course of my master thesis it became obvious that humidity in the laboratory critically affects reproducibility of APCI. Therefore, the first aim of this thesis was a systematic assessment of the effect of water infusion into the APCI source on efficiency and repeatability of APCI for the two most commonly used derivatization protocols, namely alkylation using methyl chloroformate (MCF) and methoximation-silylation, in GC-MS based metabolome analysis. To that end, 20 metabolite standards and five different water infusion rates between 0.1-0.5 mL/h were tested. As a proof of principle, MCF derivatization and GC-APCI-TOFMS w/o water infusion were applied to the detection of changes in metabolite abundance in pancreatic cancer cells upon treatment with the heat-shock protein 90 (Hsp90) inhibitor 17-dimethylaminoethylamino-17-demethoxygeldanamycin (17-DMAG).

1.2 Matrix effects

In liquid chromatography (LC)-MS, APCI is known to be prone to matrix effects hampering accurate quantification. Hence, the aim of this project was to study whether biological matrices affect GC-APCI-HRTOFMS. To investigate potential interferences from matrix compounds, an *E. coli* extract as well as human serum and urine specimens were spiked with 15 metabolite standards and recovery rates determined. In addition, for three pairs of co-eluting compounds, possible interferences between the analyte of interest and defined amounts of its co-eluent were studied. The purpose of this experiment was to reveal whether APCI might be an actual source of ion suppression or enhancement in the case of co-eluting analytes in the absence of a biological matrix.

1.3 APCI source type

During my thesis, Bruker Daltonics (Bremen, Germany) introduced a redesigned “GC-APCI II” source that supposedly addressed shortcomings of the original GC-APCI I source such as gas turbulences, cold spots, and insufficient heating, which had exerted a negative impact on ionization and ion transfer into the MS. The new source was therefore subjected to comprehensive testing to evaluate detection sensitivity, repeatability and suitability for metabolic fingerprinting. Initially, a commercial 37-component FAME mixture was used to investigate the repeatability of retention time, peak width and ionization efficiency for structurally similar compounds over a wide boiling point range. To assess the quantitative potential of APCI II in comparison to its predecessor, figures of merit for metabolic profiling were determined from calibration curves of 20 MeOx-TMS derivatized metabolite standards and nine ISs. Both source types were also employed in the analysis of three biological replicates of cell culture supernatants with the objective of comparing their performance with regard to technical variability and identification of metabolites.

Aim #2: Application of GC-APCI-HRTOFMS to the enantioselective quantitative profiling of D/L-2-HG

The goal was the development of an enantioselective, sensitive, and reliable quantification approach for 2-hydroxyglutarate (2-HG). Analysis of 2-HG has become increasingly important in cancer biology due to the observed accumulation of D-2-HG in several types of cancer mostly as a result of neomorphic mutations in the isocitrate dehydrogenase (IDH) 1 and 2 genes. To separate D and L enantiomers of 2-HG, MCF derivatization and a chiral cyclodextrin (CD)-based Rt- γ DEXsa capillary column were tested based on previous experiences in our laboratory with the enantioselective GC analysis of amino acids. To prove its reliability, the method was validated by means of spike-in experiment in serum and cross-validation of total 2-HG levels in human urine specimens was carried out using HPLC-ESI(-)-MS/MS.

Parts of this thesis have been already published [13-17]. The funding by the Bavarian Genome Network BayGene of the Bavarian State Ministry of Sciences, Research and the Arts (Munich, Germany) and the DFG (KFO 262, DE 835/2-1) is gratefully acknowledged. All persons mentioned in this thesis are current or former members of the Institute of Functional Genomics of the University of Regensburg, Regensburg, Germany, unless stated otherwise.

4 Background

Parts of this chapter were published in book chapters [15] and [16]. The focus is primarily on GC-MS based metabolomics.

4.1 Metabolomics

Metabolomics aims at providing an in-depth view of chemical changes in cells, tissues, organs or organisms evoked by cellular processes in response to genetic and environmental causes [18]. It is an integral part of systems biology and provides a direct link between an external stimulus and the non-structural phenotype or physiology of a biological system [19,20]. Since the metabolome represents the endpoint of the “omics” cascade (**Figure 4.1**), it yields a more predictive phenotype than the genome, transcriptome and proteome, respectively. It also captures the dynamic changes that occur in response to genetic and environmental constraints much better, with the composition of the metabolome having been observed to change within seconds, whereas turnover in the proteome and transcriptome (minutes to hours) is considerably slower [21].

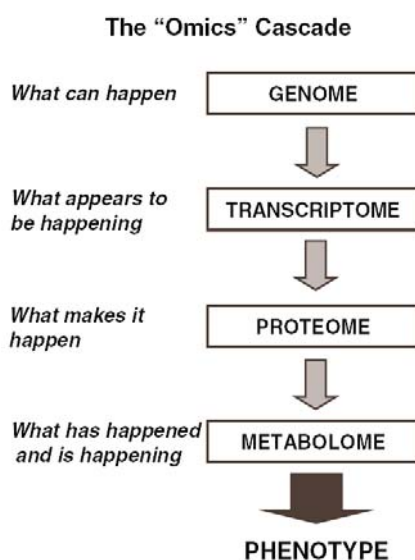


Figure 4.1 The metabolome represents the endpoint of the "omics" cascade. Reprinted with permission from [22].

4.1.1 Challenges

Metabolome analysis comprises both the qualitative and the quantitative assessment of low-molecular mass compounds (<1,000 Da), which show tremendous diversity in their chemical and physical properties. Moreover, metabolite concentrations range over up to ten orders of magnitude. The physicochemical diversity and wide range of abundance of metabolites in biological specimens constitute major challenges in the comprehensive determination of the metabolic state of biological systems, requiring the combination of several analytical platforms such as nuclear magnetic resonance (NMR) spectroscopy and direct-infusion or hyphenated mass spectrometry (MS). Proton NMR presents the closest to a universal nondestructive detector that provides in combination with carbon-13 NMR an indispensable tool for the structural elucidation of unknown metabolites. However, it lacks the detection sensitivity of MS required for the determination of metabolites in the submicromolar concentration range. For that reason, MS in combination with chromatographic separation techniques such as LC and GC has become a widely used, albeit still to be perfected tool in targeted and non-targeted metabolomics. Existing shortcomings include instrumental drift, ion suppression and the lack of comprehensive mass spectral libraries for LC-MS and in case of GC-MS the frequent need for derivatization, which carries pitfalls such as the degradation of metabolites and the lack

of metabolite mass spectral libraries for derivatives other than silylated compounds. Moreover, reference compounds are only available for a subset of metabolites. One way of overcoming the lack of comprehensive mass spectral libraries is the use of soft ionization techniques that generate gas-phase ions without extensive fragmentation. In combination with high-resolution MS, measurement of the accurate mass of non-fragmented protonated molecules allows calculation of elemental formulas that can then be searched against metabolite databases for putative metabolite identification. In addition to cross-platform approaches to cope with analyte diversity, various metabolite databases can be searched (e.g. HMDB [23], LIPID MAPS [24], and Metlin [25]). Once identified, metabolites may be visualized within their metabolic pathways (e.g. provided by KEGG database) to facilitate biological interpretation. Finally, metabolomics data can be incorporated with results obtained by the other –omics methods to obtain a global picture of the biological entity under study.

Apart from the high diversity of metabolites, undesired variation in metabolite levels might be introduced at each stage of the metabolomics workflow and pose an additional challenge to the analytical chemist. In order to minimize variability introduced during sample preparation, standardized protocols for quenching, extraction and derivatization steps are essential and should be followed as proposed by Sumner et al. in 2007 [26]. Still, it is almost impossible to avoid analyte losses during sample preparation. Therefore, stable isotopically labeled internal standards (SIL-ISs) are commonly used to account for analyte losses, and apart from that for injection variability, instrumental drift and ion suppression.

4.1.2 Strategies

Two different strategies are mostly used in metabolomics [22,27]. These are metabolite profiling or targeted analysis of a small number of analytes on the one hand, and metabolic fingerprinting on the other hand. For metabolite profiling, i.e., the quantitative analysis of pre-selected metabolites, analytical methods are tailored for the separation and sensitive detection of the target analytes. Reference substances in known concentrations and an appropriate set of internal standards are used to construct calibration

curves to quantify the target analytes in biological specimens. Alternatively, isotope dilution analysis can be employed for quantification. Stable isotopically labeled analogs of the target analytes are spiked into the samples in known concentration and quantification is performed by comparing the peak area of the target analyte to that of the labeled standard. For metabolic fingerprinting, the complete GC-MS or LC-MS chromatogram is exported and analyzed. One can either search for signals that change significantly between sample groups or use the chromatograms to classify samples. This does not require knowledge of the identity of the signals, often referred to as features or markers. Metabolite identification is then restricted to those features that yielded significant differences in signal intensity between sample groups. Aiming at the simultaneous detection of the highest possible number of metabolites, sample preparation is typically kept to a minimum in metabolite fingerprinting to prevent analyte losses. In targeted studies, on the other hand, sample preparation is tailored to the specific analytes and, thus, typically more extensive to reduce chemical complexity of the samples prior to analysis.

Apart from targeted and untargeted analyses, the analysis of metabolic flux using stable isotopically labeled substrates has been a valuable tool in the study of dynamic processes in a cell. Instead of measuring only metabolite abundances, flux experiments can reveal the flux of carbon (or nitrogen) atoms derived from substrates such as glucose, glutamine or other amino acids, through different metabolic pathways and detect changes in rates of metabolic reaction due to certain perturbations. The number of publications reporting flux experiments has increased steadily in recent years [28,29].

4.2 Established GC-MS approaches in metabolomics

Gas chromatography coupled to mass spectrometry (GC-MS) is an indispensable tool in metabolomics. It is perfectly suited for the analysis of small volatile metabolites. Capillary GC is capable of resolving very complex mixtures and MS provides highly selective and sensitive detection of metabolites along with their identification. Capillary GC can be easily interfaced with an MS instrument as carrier gas flow rates of about 1 mL/min

are compatible with modern MS vacuum systems. Moreover, analytes have already been transferred into the gas phase before reaching the ion source of the MS.

4.2.1 Derivatization strategies

Major prerequisites for GC-MS analysis include a sufficient vapor pressure and thermal stability of metabolites. Most metabolites, such as amino acids, organic acids, sugars and alcohols, possess polar functional groups, which have to be derivatized prior to GC analysis. Apart from increasing volatility and thermal stability, derivatization of the analytes can improve chromatographic properties. It can provide additional structural information and improve both, lower limits of quantification and selectivity for quantitative measurements. However, derivatization prolongs overall analysis time and adds variance to the analysis. It can also hamper identification of metabolites, derivatives of which are not contained in mass spectral libraries. Overall, derivatization should be simple, fast, automatable, efficient, reproducible, and spectral libraries of the derivatives are needed. Halket and Zaikin extensively reported on derivatives for mass spectrometry in a series of reviews, which in each case deals with a single derivatization reaction, e.g., silylation [30]. Among others, silylation, acylation, alkylation/esterification, the formation of cyclic derivatives and generation of mixed derivatives of polyfunctional compounds has been performed for GC-MS.

4.2.1.1 MeOx-TMS derivatization

In 2011, Koek et al. [31] reviewed a large number of metabolomics studies with respect to the pursued derivatization strategy. The majority of GC(-MS) methods employed silylation with or without prior oximation. **Figure 4.2** provides the example of a typical two-step derivatization strategy including methoximation followed by silylation using N-methyl-N-(trimethylsilyl)-trifluoroacetamide (MSTFA).

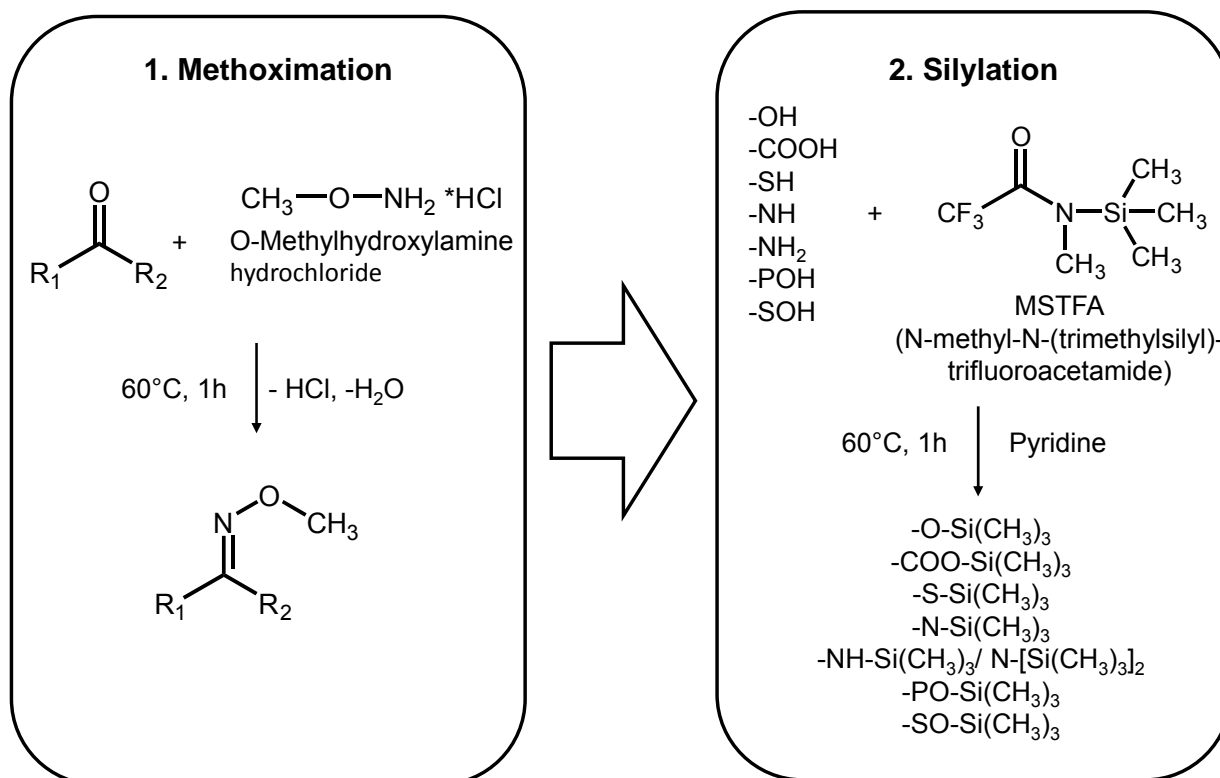


Figure 4.2 Commonly used two-step derivatization procedure for GC-MS based metabolomics.

Oximation reagents, such as hydroxylamines or alkoxyamines, react with aldehyde and keto groups. They react with open-chain reducing sugars thereby preventing ring formation of reducing sugars. The derivatization results in the formation of two stereoisomers (syn and anti form) that can be separated chromatographically. Further, decarboxylation of α -ketocarboxylic acids and keto-enol tautomerism is hindered. MSTFA or N,O-bis-(trimethylsilyl)-trifluoroacetamide (BSTFA) are typically used reagents for silylation. They replace active hydrogen atoms in functional groups, such as -COOH, -OH, -NH and -SH, with a trimethylsilyl (TMS) group. In case of silylation reagents that contain chlorine atoms, small amounts of a catalyst, e.g. 1% of trimethylchlorosilane, are added. Silyl derivatives are prone to hydrolysis in the presence of traces of water. Consequently, anhydrous conditions have to be employed. Bulkier silylation reagents, such as N-methyl-N-(tert-butyldimethylsilyl)-trifluoroacetamide (MTBSTFA), have been shown to be less susceptible to hydrolysis [32]. Moreover, tert-butyldimethylsilyl derivatives show characteristic $[M-57]^+$ ions in EI-MS, which may facilitate the identification of unknown compounds [33], but requires harsher reaction conditions.

Oximation followed by silylation is often employed in non-targeted metabolomics [12,17] because of the ready derivatization of many different functional groups and the availability of large spectral libraries [31]. However, degradation or rearrangement reactions have been reported, e.g. arginine is decomposed to form ornithine [34], and glutamate rearranges to pyroglutamate. Partial silylation yielding multiple peaks for amino acids is a well-known problem, which can be addressed by either weighting correctly the multiple peaks [35] or by summing up the responses of all derivatives [36].

Modern GC instruments equipped with advanced sample robots (e.g., from CTC Analytics AG or Gerstel) enable the sequential automated derivatization of sample analytes immediately prior to their analysis instead of off-line en bloc derivatization, thereby minimizing hydrolysis and variability in the ratio of partially derivatized metabolites.

Large numbers of derivatizable groups may lead to indefinite retention of an analyte on the GC column or the molecular mass of the derivative might exceed the mass range of the detector [30]. Native lactose, for example, has a molecular mass of 342 Da, which increases to 947 Da after methoximation and the introduction of 8 trimethylsilyl groups. Furthermore, silylation is neither simple nor fast, as it cannot be performed in aqueous solution or in the presence of protic solvents and usually requires heating for some time, especially in case of amines.

4.2.1.2 Derivatization with alkyl chloroformates

In 2011, Villas-Bôas et al. [37] compared silylation to derivatization with methyl chloroformate. The authors concluded, that a combination of both derivatization approaches leads to a more comprehensive coverage of the metabolome. Alkyl chloroformates are used as reagents to esterify carboxylic groups to form the corresponding alkyl ester, while carbamates are derived from amino groups (**Figure 4.3**). This derivatization has emerged as an attractive alternative to silylation, especially for amino acids [38,39] and non-amino organic acids [40]. The reaction is fast and feasible in aqueous media at room temperature, which makes it easy to handle. Kaspar et al. [38] demonstrated its potential for full automation. Nevertheless, alkyl chloroformate derivatization is less versatile compared to silylation. Moreover, comprehensive mass spectral libraries are

still lacking and need to be developed, which renders methoximation/silylation overall the method of choice in GC-MS based metabolomics.

Interestingly, in contrast to the study by Koek et al. in 2006 [41], Villas-Bôas et al. [37] described a poor analytical performance for TMS derivatives. This underscores the importance of strictly enforced standard operation procedures such as the regular replacement of GC-inlet liners and trimming of the inlet end of the capillary column, as well as the addition of several SIL-ISs to monitor performance of extraction and derivatization [41]. More details on experimental strategies including quality control are provided in subchapter 4.4.

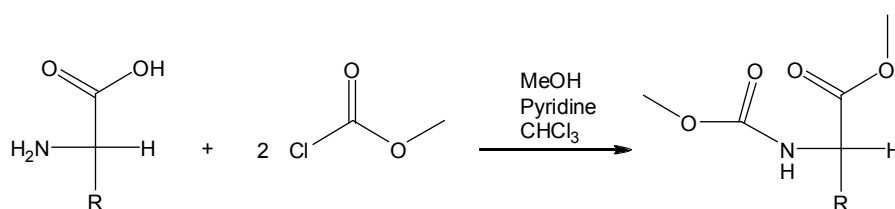


Figure 4.3 Reaction scheme for the derivatization of amino acids with methyl chloroformate.

4.2.2 GC stationary phases

For GC separation, analytes have to be transferred into the gas phase, which is typically achieved by a hot split/splitless injector or by a programmed-temperature vaporizer (PTV). Nowadays, GC separation is commonly performed using capillary columns. These are open tubular columns with an inner diameter of less than 1 mm and a length of 1 – 100 m. The stationary phase is coated to the inner wall of the column. The stationary phase can either be a highly viscous liquid (WCOT – wall coated open tubular column), with analyte separation attained by partitioning of the analytes into the liquid (GLC – gas liquid chromatography), or a solid adsorbent (PLOT – porous layer open tubular column) for adsorption-based separation (GSC – gas solid chromatography). GSC is mainly used for highly volatile analytes such as permanent gases. In metabolomics, GLC prevails. A variety of stationary phases (SPs) are available. Non-polar SPs provide separation of analytes as a function of mere differences in vapor pressure, while polar SPs allow the exploitation of additional polar interactions. Common SPs are polysiloxanes, which are modified with both non-polar methyl groups (PDMS – polydime-

thylsiloxy) and polar phenyl- or cyanopropyl groups to tune the polarity of the phase. The recently introduced ionic liquid SPs are an attractive alternative to polysiloxane SPs. Furthermore, direct stereoselective analyses can be carried out using chiral SPs, such as the widely used Chirasil-L-Val SP or CD-based SPs, which separate enantiomers by formation of inclusion complexes. However, due to the huge chemical diversity of metabolites there is no single ideal column chemistry capable of covering entire metabolomes. Rather, multiple SP chemistries tailored to different classes of metabolites or their derivatives have to be employed sequentially to achieve comprehensive coverage.

4.2.3 GC-MS with EI and CI sources

Capillary GC lends itself to interfacing with MS, as analytes have been already transferred into the gas phase and carrier gas flow rates are compatible with modern vacuum systems. Hyphenation of GC and MS is typically accomplished by means of electron ionization (EI), which is a hard ionization technique under high vacuum. Electrons are emitted from a heated wire filament usually at electron energies of 70eV. In an initial step, radical cations are generated from gas phase molecules. Subsequent fragmentation reactions, losses of neutral groups and rearrangement of secondary ions occur because electron energies are by far exceeding ionization energies of organic molecules. This yields predictable ion patterns, which can be interpreted by the experienced scientist. Moreover, EI is highly reproducible and, therefore, huge commercial spectral libraries containing currently up to 242,000 unique compounds (NIST 14 Mass Spectral Library) have been built for routine identification of analytes. However, structural isomers often show similar EI spectra, thus precluding unambiguous compound identification in the absence of additional information such as differences in chromatographic retention. Further, both the still incomplete coverage of natural metabolites in spectral libraries and the frequent absence of molecular ions in EI mass spectra continue to constitute major bottlenecks in the identification of group discriminating features. Therefore, soft ionization techniques that preserve intact protonated molecules have become prominent in recent years, such as chemical ionization (CI) and APCI (subchapter 4.3). Operation in CI mode takes place under low vacuum in the presence of a reagent gas,

e.g. methane, isobutane or ammonia. Electrons emitted from a heated wire filament initially ionize the reagent gas, which is present in great excess. Subsequently, analytes are ionized by ion-molecule reactions (**Figure 4.4**).

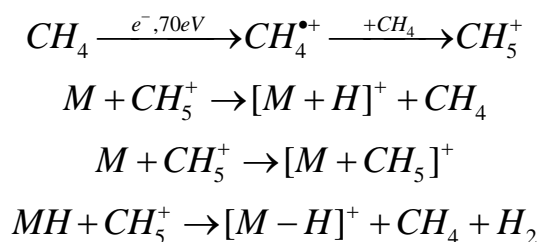


Figure 4.4 Primary reaction gas (methane) ion formation and secondary analyte (M) ion formation in positive chemical ionization mode.

However, extensive soft ionization mass spectral libraries have not been built yet. In case of CI this is owed to the strong dependence of mass spectra on the nature and pressure of the reactant gas. Hence, a lot of manual effort by the user is still required during data analysis.

The appropriate choice of mass analyzer depends on the aim of the study. Mainly quadrupole (qMS) and time-of-flight mass spectrometers (TOFMS) are applied to metabolomics. Quadrupole mass analyzers enjoy widespread use. They are inexpensive, robust and easy to handle. But they are limited in mass range and scan speed and, in most cases, restricted to nominal mass resolution. Therefore, quadrupoles are commonly used in routine analysis either in full scan mode or selected ion monitoring (SIM) mode, with the latter achieving lower limits of quantification due to a longer dwell time on pre-selected masses. Even more powerful is the sequential combination of three quadrupoles, which offers highly specific and sensitive quantitative analysis. In addition, structural information on the analytes is provided as an asset by tandem mass spectrometry (MS/MS). TOFMS instruments provide not only low limits of quantification but also wide dynamic ranges over up to four orders of magnitude. Fast-acquisition TOFMS instruments, which acquire up to 500 spectra per second, are a prerequisite for the successful resolution and integration of the very narrow chromatographic peaks realized by fast GC or GC×GC. Typically, however, they provide only nominal mass resolution. On the other hand, TOFMS instruments that are operated at acquisition rates of 20 – 50

spectra/s provide enhanced mass spectral resolution and accuracy of ion masses, which facilitates the identification of unknown compounds. However, TOFMS instruments are sensitive to voltage or temperature fluctuations requiring external as well as internal (re-)calibration of the mass scale on a regular basis. Even better suited for the identification of unknowns are hybrid qTOFMS instruments that combine one or more quadrupoles with a TOF mass analyzer. This MS type enables the determination of parent and fragment ions with high mass accuracy.

4.2.4 Comprehensive two-dimensional gas chromatography – time-of-flight mass spectrometry

Despite high chromatographic resolution, overlapping peaks and coelutions are still observed in the analysis of complex samples by one-dimensional GC. Comprehensive two-dimensional GC (GC×GC) is a powerful analytical tool to tackle this problem and its application to metabolomics has been recently reviewed [42]. It uses two columns with orthogonal separation characteristics that are connected via a modulator. Typically, a nonpolar/polar column combination is employed. Hence, separation on the first column is essentially driven by differences in vapor pressure. Subsequent resolution of isovolatile sample components in the second dimension arises from the different strengths and types of polar interactions with the stationary phase. A number of different modulators are available that mostly employ thermal modulation. For example, the dual-stage four-jet modulator from LECO Corporation (St. Joseph, MI, USA) uses alternately nitrogen-cooled cold jets for the periodic trapping of effluent from the first column in small segments that are then transferred to the very short second column upon hot jet remobilization. Commonly, because of their high data acquisition rate, TOF mass analyzers are used to record the narrow peaks (peak widths of 50 – 200 ms) created by the focusing step and the fast separation in the second dimension. Electron ionization is used to generate fragment mass spectra that are then matched against available spectral libraries. GC×GC-MS chromatograms are displayed as so-called 2D contour plots. The first and second column retention times are represented on the x-axis and y-axis, respectively, and a color code depicts the detector response. An exemplary GC×GC-MS chromatogram of a methanolic avocado pulp extract is given in **Figure 4.5**. Peak anno-

tation was based on spectral matching to the NIST 05 library using a cut-off score of 700.

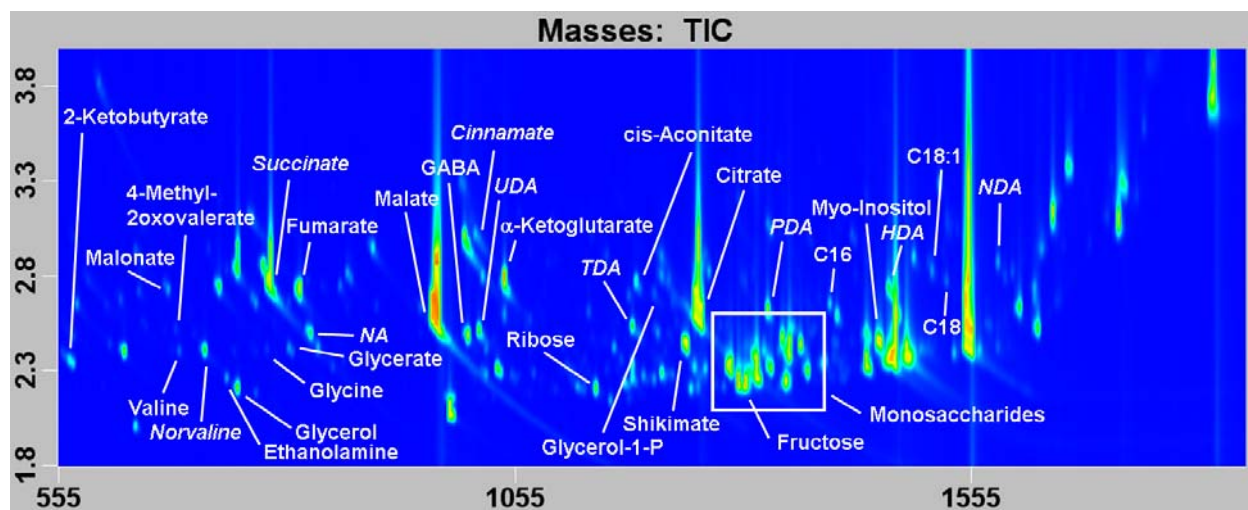


Figure 4.5 GC×GC-TOFMS chromatogram of an avocado sample. Avocado pulp (40 mg) was homogenized and extracted using cold methanol, extract was dried, derivatized (methoximation/silylation), and analyzed by GC×GC-TOFMS [12]. For peak annotation spectral matching to the NIST 05 library was performed using a cut-off score of 700. Signals marked in italics refer to internal standards. C16, palmitate; C18, stearate; C18:1, oleate. Reprinted with permission from K. Dettmer et al, in W. Weckwerth, G. Kahl (Editors), *The Handbook of Plant Metabolomics*, Wiley-VCH, Weinheim, 2013, p. 77.

Comprehensive two-dimensional GC separations not only offer enhanced resolution and a multiplicative increase in peak capacity over conventional GC methods, but also narrower peaks due to modulation and, hence, improved detection sensitivity. The latter was demonstrated for a standard mixture of 43 metabolites that was analyzed by both, one-dimensional and comprehensive two-dimensional GC-TOFMS [17]. For most metabolites, lower limits of quantification were at least one order of magnitude lower using GC×GC-TOFMS. Still in need of improvement are convenient and practicable ways of data analysis to make full use of GC×GC-TOFMS for metabolic fingerprinting. In a first step, the raw data is pre-processed, for instance, in LECO's ChromaTOF software. This includes background subtraction, deconvolution, picking of peaks and their integration, as well as combining several modulations belonging to one compound. Further, alignment tools must be capable of correctly aligning chromatograms within and across batches. This requires correction for shifts in retention time to recognize the same analyte across different samples. In addition to shifts in retention time, incorrect peak recognition can be caused by poor mass spectral deconvolution. These effects can

result in missing values in the alignment table despite the presence of the peak in the chromatogram, which impedes discriminate analysis. Quantitative analysis using GC×GC-TOFMS data can also be time-consuming. Employing peak based quantification tools, the peaks from all modulations must be integrated to determine the response of an analyte. Since the software may fail in the proper integration of minor subpeaks, manual inspection and correction of integrated peaks may be necessary [43].

4.3 Atmospheric pressure chemical ionization for GC-MS

In this subchapter, the fundamentals of GC-APCI-MS and recent applications are described, with an emphasis on the role of clusters in atmospheric pressure ionization.

4.3.1 Fundamentals

Historical outline (taken from [44])

APCI was first introduced by Horning et al. in 1973 [1] using ^{63}Ni as a beta-emitter for ionization that was later replaced by a corona discharge ion source [2,45]. A range of applications were described by then, such as GC-APCI-MS in negative ion mode for environmental studies [46] as well as APCI-MS for gas analysis [47]. However, the technique was not widely used, in part because of the expensive instrumentation and due to already commercially available “plug and play” EI and CI GC-MS systems.

APCI again attracted attention during the 1980s, when the introduction of electrospray ionization for LC-MS [48] caused a breakthrough in the development of LC-MS ion sources for generating ions at atmospheric pressure. Unlike previous LC-MS interfaces, such as particle beam or thermospray, where ions were generated under vacuum conditions, molecules are ionized at atmospheric pressure in ESI and APCI. However, APCI was initially only used as an ionization interface for LC-MS.

The renaissance of atmospheric pressure ionization techniques for GC-MS started just recently, when both McEwen and McKay [4] and Schiewek et al. [3] modified an API source on a commercial LC-MS instrument to enable an easy switch between various

ionization modes and to use both LC and GC with one certain mass spectrometer. The exciting potential of the latter investigations has led to the development of commercially available APCI sources that allow the hyphenation of GC to high-resolution TOF or QTOF mass spectrometers and tandem mass spectrometers that were previously only coupled with LC.

Hence, distinctly lower quantification limits than previously obtained with quadrupole instruments can be achieved and features such as accurate mass measurements for chemical formula determination of unknown analytes can be exploited. A further advantage of GC-MS is the ability to observe compounds of low polarity that are not amenable to LC-MS. These benefits are especially important in the study of complex mixtures as in the field of metabolomics. A major drawback remains the lack of convenient libraries [44].

GC-APCI-TOFMS instrumental setup

One of the first successful attempts to couple GC to TOFMS was made by Erickson et al. in 1990 [49] by applying time array detection, which yielded data acquisition rates of up to 20 spectra/s and, hence, the optimization of chromatography for speed of analysis with no loss in chromatographic resolution or analyte detectability. Furthermore, the more widespread use of GC-TOFMS was fostered by the introduction of soft ionization techniques [4,50-53] for the determination of the elemental composition of unknown analytes and analog-to-digital converters distinctly increasing the dynamic range of TOFMS instruments [5].

In the following, the GC-APCI I-TOFMS instrumental setup employed in this thesis is described (adapted from [44]).

The first APCI source (APCI I) commercialized by Bruker Daltonics consists of a rectangular aluminium chamber with five different ports A-E (compare **Figure 4.6**, [3]). Port A is mounted on the ion source mounting flange of the Bruker MS. The GC is connected to the ion source via a heated transfer line assembly at port B opposite to port A. The transfer line is fixed and heated externally by a controller to prevent compound conden-

sation. Port E, located at the top of the ion source, contains the LC vaporizer stage, which is used for external hot nitrogen supply (illustrated in yellow in **Figure 4.7**) in case of GC to compensate for the relatively large pumping speed of the MS and to generate the heat required for the APCI process. Ports C and D are both arranged perpendicular to the longitudinal axis of the GC transfer line. Port C is sealed with an optical window and the corona needle assembly is mounted on Port D. This places the needle tip in close vicinity to the GC effluent and sheath gas flow. **Figure 4.7** shows a general scheme of the GC-APCI-TOFMS setup. A sheath gas flow (nitrogen, illustrated in orange in **Figure 4.7**) is introduced through coaxial holes around the outlet of the GC capillary and envelopes eluting analytes on their way through the source. Moreover, it directs the generated ions towards the MS. The source is purged with a clean dry gas (N_2) to prevent reduction of ions by contaminants [3]. However, part of the dry gas is not directed into the APCI source but soaked into the transfer capillary (illustrated in red in **Figure 4.7**), which causes gas turbulences in front of the end plate of the capillary. Furthermore, the vaporizer gas is a major source of the overall extensive gas turbulences in the APCI I source, so that it was omitted for the second-generation APCI II source (see section 8.2.2). With regard to the ionization several APCI source parameters have to be optimized for a specific application namely drying gas (nitrogen) temperature and flow rate, vaporizer temperature and gas (nitrogen) pressure, current of corona needle discharge and capillary end plate voltage (compare **Figure 4.7**).

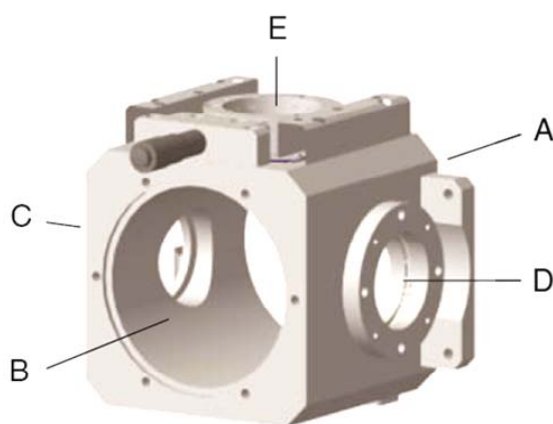


Figure 4.6 Scheme of the APCI I source body. Reprinted with permission from [3].

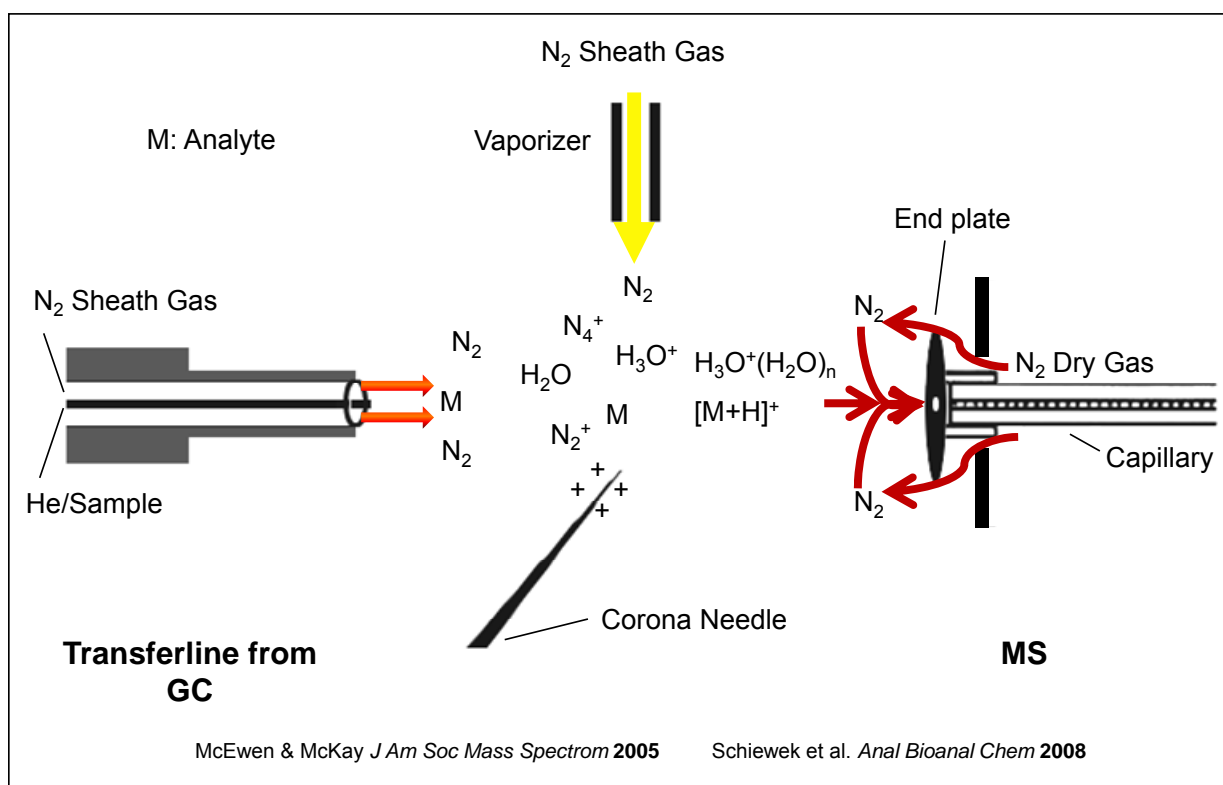


Figure 4.7 GC-APCI-TOFMS setup and ion species present during the ionization procedure [3,4,45]. The flow direction of sheath gas, vaporizer gas, and dry gas is indicated by orange, yellow, and red arrows, respectively. Modified and reprinted from [15] with permission from Elsevier.

In addition to the ionization in the APCI source, different stages of the ion transfer are crucial for the quality of the obtained APCI mass spectrum and have to be pointed out. Since analytes are ionized at atmospheric pressure, but ion separation in the mass analyzer is achieved under high vacuum conditions, pressure has to be reduced stepwise in consecutive stages within the mass spectrometer. As discussed by Klee et al. [54,55] up to seven different domains can be distinguished within the overall ion generation and detection process. For the microTOF (Bruker Daltonics) employed here, this includes regions of primary and secondary ion generation, ion transfer to the MS, three different vacuum pumping stages and ion optics. The domains are characterized by varying chemical composition, temperature and pressure, which lead to rapid switching of the dynamic entity of analyte ions and surrounding matrix.

Apart from the GC-MS approach followed in this thesis, APCI has been used for analyzing compounds from solid or liquid surfaces. A detailed overview of AP surface sam-

pling/ ionization techniques for mass spectrometry was compiled by van Berkel et al. [56], including Atmospheric-pressure Solids Analysis Probe (ASAP) and Direct Analysis in Real Time (DART), which make use of a hot nitrogen gas stream and excited helium atoms for vaporization and desorption of the sample, respectively, and subsequent ionization by APCI. Krieger et al. [57] developed a Direct Inlet Probe (DIP)-APCI ion source with temperature-programmed heating of the sample to separate analytes based on boiling point differences and, hence, reduce ion suppression. While DART and DIP-APCI sources are used for analyzing solid samples and small amounts of liquid samples, the APCI source for LC-MS, in comparison to that of GC-MS, has an extra heated APCI nebulizer unit used for efficient nebulization and desolvation of the LC effluent.

APCI ionization mechanism (partly taken from [44])

As stated above, APCI gained in importance first in LC-MS as an alternative to ESI, which is well suited for polar compounds and, especially, for large biomolecules. APCI, in comparison, is mostly used whenever small and medium-polar molecules are investigated. Both ionization modes generate mainly protonated molecules with very little fragmentation. In case of ESI, evaporation of the eluent and ion formation from micro droplets through Coulomb explosions can be considered as top down processes, whereas APCI in the gas phase is mediated through proton-bound solvent clusters that are build bottom up from primary ions. "In positive-ion mode, APCI is initiated by corona discharge through application of a high voltage (2-6 kV) to a sharp metal electrode (corona-discharge electrode). The high electric field at the tip of the electrode attracts electrons and the surrounding nitrogen gas is ionized by electron ionization. This leads finally to plasma formation and together with traces of water vapour from air or nitrogen the following reactions (**Figure 4.8**) [45] occur in the plasma region." [44]

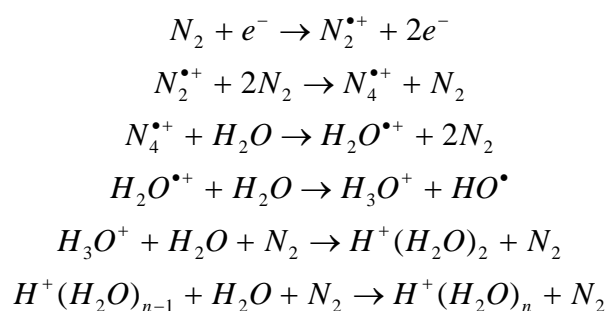


Figure 4.8 Ion molecule reactions in positive atmospheric pressure chemical ionization mode [45].

This series of reactions leads to the formation of clusters, which are typically proton-bound water clusters in GC-MS, whereas proton-bound LC solvent or mixed water/solvent clusters are present in LC-MS. These clusters finally lead to the dominantly observed protonated molecules as described later in this section.

It should be noted that gas-tight API interfaces have been designed in the context of ionization mechanism studies that favor charge exchange rather than protonation [58]. Furthermore, formation of molecular radical cations depends on the type of analyte. For instance, it was observed in the analysis of pesticides by means of GC-APCI-QTOFMS [59]. However, N_2 transfer conditions resulting in the formation of M^+ species [59] did not apply to the instrumental setup and analytes investigated in this thesis. The protonated molecule, however, was always formed (termed as “proton transfer conditions” [59]).

It should be mentioned that APCI in negative ion mode is also possible [6], but this was of no practical relevance in the context of the current thesis.

Furthermore, among the different API techniques, both direct and dopant-assisted atmospheric pressure photo ionization (APPI/DA-APPI) and atmospheric pressure laser ionization (APLI/DA-APLI) are applied to medium- to nonpolar and aromatic nonpolar analytes, respectively. Both LC-MS [54] and GC-MS [3,60] applications have been reported. Unlike in APCI, primary reagent ions are formed from photons instead of electrons, but, importantly, ionization of analytes is also based on chemical ionization by proton-bound cluster species [54,61]. Hence, it appears appropriate to derive universal

conclusions from the role of clusters in API and their fate in the ion transfer process from individual APCI, APPI or APLI studies.

APCI conditions

Atmospheric pressure chemical ionization takes place under completely different physical and chemical conditions than CI and especially EI. Hereinafter, given that the described ionization conditions and processes in analyte ion generation are common to APCI, APPI and APLI, the expression API was used as a general term.

Most relevant to the analytical chemist is the reduced degree in fragmentation in API and a concomitant increase in protonated molecules $[M+H]^+$. The benefit of APCI for metabolite identification is shown in **Figure 4.9** that depicts mass spectra of fructose-1MeOx-5TMS acquired with GC-EI-qMS and GC-APCI-TOFMS, respectively. Information on the entire molecule is only preserved in the latter case, and very few fragments are observed, e.g. resulting from the loss of a neutral $\text{Si}(\text{CH}_3)_3\text{OH}$ (m/z 90) group. Protonated molecules preferentially transfer excess energy to neutral molecules through collision processes lowering fragmentation (“collisional deactivation”) [62], whereas unimolecular decay reactions occur under vacuum conditions due to much lower densities of neutral molecules. Furthermore, source residence times of ions in API may reach up to several hundred milliseconds [55,63] and their transfer to the MS is dominated by gas flow instead of electric fields [64]. As opposed to that, ions formed under vacuum conditions are transported out of the source region within less than one micro-second by the applied electric fields [55]. Due to the prolonged residence time of ions in the APCI source, the formation of a silylation adduct was observed in my initial GC-APCI-TOFMS study [17], which was formed only in case of compounds with a carboxylic group and, thus, helped to strengthen a putative identification, e.g., nonanoate-1TMS (**Figure 4.9C**) showed adduct formation as opposed to glycerol-3TMS (**Figure 4.9D**), which lacks a carboxylic group.

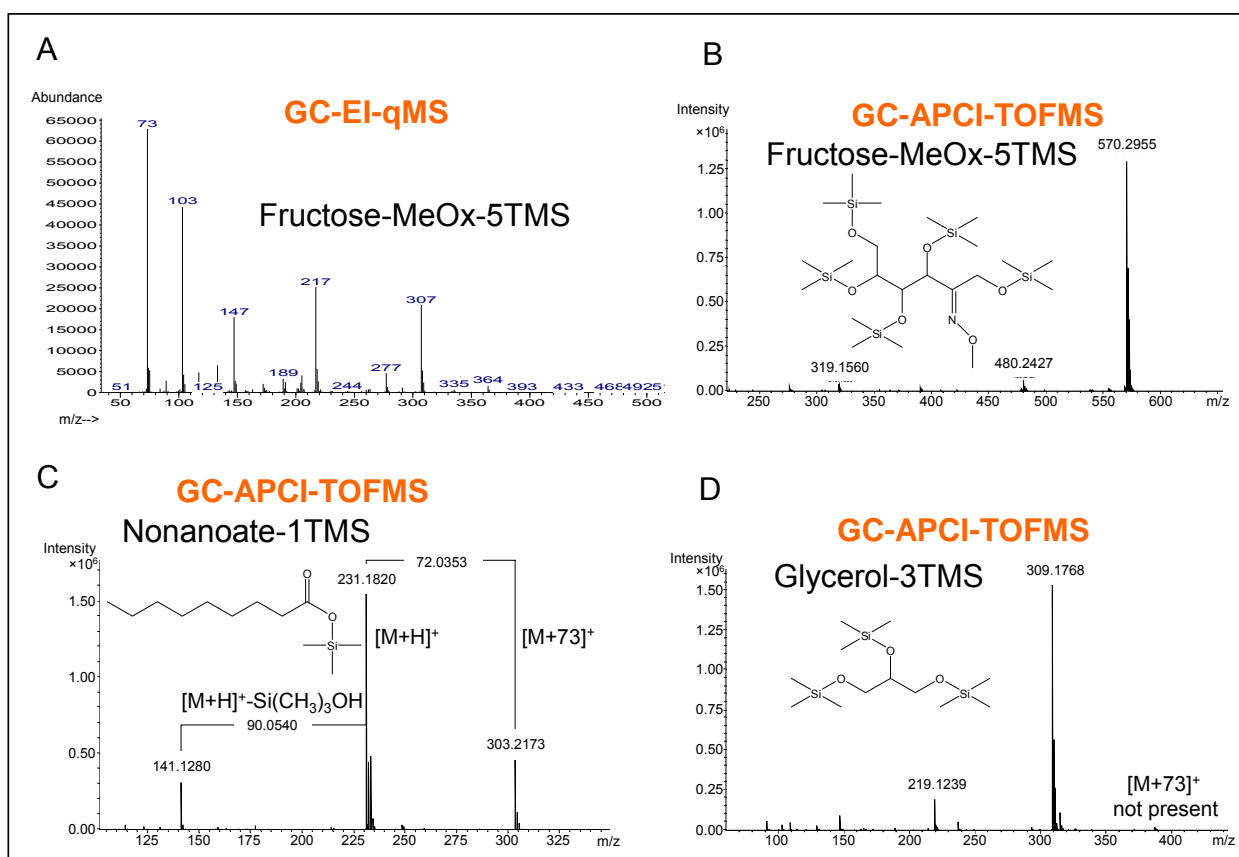


Figure 4.9 GC-MS mass spectra of standards (62.5 μ M) using hard and soft ionization techniques as described elsewhere [17]. Identification of fructose-1MeOx-5TMS requires spectral libraries for GC-EI-qMS (A) or can be achieved starting from the protonated molecule in case of GC-APCI-TOFMS (B), where structure-dependent adduct formation may serve as plausibility check as exemplified for (C) nonanoate-1TMS and (D) glycerol-3TMS with the latter showing no adduct formation as it lacks a carboxylic group. MeOx, group introduced by methoximation; TMS, trimethylsilyl group; $[M+73]^+$, silylation adduct of analyte M. [16] – Reproduced by permission of The Royal Society of Chemistry.

Considering the long residence times of ions in the source compared to typical ion-molecule reaction rates and the great numbers of neutral molecules as potential reactants (typically nitrogen, oxygen and water molecules in GC-MS), chemical (bi- and termolecular) reactions inevitably occur in API sources. Under such conditions, as recently pointed out by several authors [55,61], chemical processes are typically thermodynamically controlled, i.e. the minimum of the Gibbs function G of the reaction system is reached within the dwell time of ions in the source. Hence, ionization at atmospheric pressure leads to thermally equilibrated ion populations, whereas ionization under vacuum conditions is kinetically driven [63]. However, equilibrated ion populations may drastically change en route to the detection system, not only because of interactions

with neutral molecules that reach the first ion transfer stages but also due to association and dissociation reactions of the ion-bound cluster systems.

In order to establish optimum API conditions, the ion-bound clusters increasingly became the focus in API mechanistic studies [54,55], because they were found to play a more complex role in the generation of protonated molecules than originally assumed. For instance, mainly $[M+H]^+$ ions were detected in DA-APPI- and DA-APLI-MS even when aprotic dopants were used [54,65]. Previously, the difficulty in elucidating API ionization mechanisms lay in the fact that clusters are typically not observed using commercial API mass spectrometers, because cluster dissociation occurs due to the electrical voltages applied in the ion transfer.

Water plays a major role in cluster ion chemistry. Due to its polar nature, cluster formation with other reactants is energetically efficient and, furthermore, water is omnipresent in API sources above levels of 1 ppmV even if the latter sources are purged with ultrapure nitrogen for several days. Numerous working groups have provided kinetic and thermodynamic data on the proton-bound water cluster system [66-69]. It has to be pointed out that LC solvents such as methanol or acetonitrile participate in clusters as well, but this was irrelevant in the context of the current thesis. The water cluster system consists of different $H^+(H_2O)_n$ species, with a varying number of n water molecules bound to a core cluster, which reach a thermal equilibrium in the API source. The equilibrium cluster size distribution of the system is typically between $n = 3-9$ with a maximum between $n = 4-6$, depending on the temperature and water content in the source [70]. In fact, with an increase in the background water concentration the maximum of the latter distribution shifts towards a higher number of bound water molecules [54]. The water cluster system is highly dynamic and clusters undergo numerous collisions in the source and also in the different stages of the ion transfer to the MS, where changes in pressure and ion temperature also affect cluster size and, hence, reactivity. This applies to the entire high or intermediate-pressure regions down to between 10^{-4} and 1 mbar [54,55].

Analyte ion generation

Proton-bound water clusters interact with analyte molecules to yield the dominantly observed protonated molecules. Previously, it was assumed that these clusters protonate molecules with gas-phase basicities higher than that of water in bi-molecular reactions in the API source [69]. Recently, an initial analyte capture by water ion cluster ligand exchange / association reactions followed by dissociation of water molecules further downstream yielding bare $[M+H]^+$ ions has been proposed [54,71]. The unlikelihood of direct acid-base-type protonation reactions is supported by the low gas-phase acidity of typically occurring water ion clusters with at least four water molecules. In fact, clusters of that size have been considered as essentially unreactive toward analytes with moderate gas-phase basicity [71]. For instance, mixed analyte-water ion clusters instead of the $[M+H]^+$ ions of N,N-diisopropylethylamine were formed under thermal equilibrium in DA-APLI [54], demonstrating that even an amine with a high gas-phase basicity rather participates in ligand exchange/ association reactions than being directly protonated.

Analyte molecules with suitable chemical and physical properties such as polarity/dipole moment can be captured as ligands in cluster equilibria [54]. All matrix components act as potential competitors in the cluster/ligand association reaction and the impact increases with increasing matrix concentrations. Higher concentrated matrix species, even if less suitable in terms of polarity, may prevent analyte molecules from entering the cluster equilibria. Vice versa, omnipresent dopants in DA-API-MS that show a low tendency towards ligand exchange / association reactions (e.g., toluene) may lead to more efficient analyte protonation than those with a high tendency (e.g., anisole). Using toluene and anisole as dopants in DA-APPI-MS analysis, testosterone was ionized more efficiently through proton-bound water clusters in the case of toluene [72].

As ligand exchange between clusters and background matrix is highly dynamic, analyte molecules may be replaced by other matrix molecules before cluster activation, i.e. the reduction of the number of cluster ligands to finally obtain protonated molecules. Analyte elimination from clusters may occur throughout ion generation and passage along the detection pathway from the ionization source to the intermediate-pressure region

down to between 10^{-4} and 1 mbar, where the number of collisions is still high enough to induce changes in the existing ion population. Klee et al. [54,55] distinguished up to seven different domains in API MS instruments, namely corona discharge region, thermalized region within the API source, ion sampling port and transfer to the MS, three consecutive pressure reduction stages and the vacuum analyzer region. The turbulent flow within the API source and ion transfer to the MS causes strong mixture of all system components, followed by gas expansion and concomitant cluster growth and re-equilibration in the first vacuum pumping stage with a sustained background pressure of a few mbar. Ions are further guided to the analyzer region by electrical fields, whereas neutral molecules are continuously removed. The applied electrical fields also lead to increasing kinetic energies of the ions, which in turn results in enhanced collisions and, hence, ion activation. Furthermore, fragmentation may occur in the case of fragile analyte ions as another consequence of the increased number of collisions. Apart from electrical field mediated collision-induced dissociation (CID) of analyte-water ion clusters, stepwise CID of equilibrated water ion clusters may be a minor pathway leading to protonated molecules. In the process, smaller, more acidic water ion clusters are formed, which then protonate neutral analytes that reach the first pressure reducing stages [54,71]. The two above described mechanisms of ion activation are a favorable side effect of ion guiding to the mass analyzer, but the aim of ion guiding is a high ion transmission rather than controlled ion activation. Keeping that in mind, an ion activation stage (IAS) for controlled ion activation has been recently introduced [61]. In fact, the prevailing water-ion cluster distribution can be shifted towards reactive species for subsequent analyte protonation by the IAS, which still awaits a systematic evaluation of its practical relevance.

In summary, according to Klee et al. [54], critical prerequisites for efficient generation of $[M+H]^+$ ions include analyte properties that favour participation in ligand exchange/association reactions, e.g. polarity. This must be accompanied by the ability of the analyte molecule to be retained as a ligand in the cluster system until the very final stages of cluster activation to end up as a bare protonated molecule. As direct protonation of analytes by thermally equilibrated water ion clusters appears negligible in comparison to cluster mediated protonation, thermodynamic properties, such as gas-phase acidity and

basicity of the analyte are probably less relevant for predicting the formation of $[M+H]^+$ ions than its polarity. However, several further analyte properties, e.g., steric hindrance, certainly play a role in the overall ionization efficiency, requiring more research and evaluation in the future [54].

4.3.2 Review of GC(xGC)-APCI-MS applications

This section focuses on applications in positive APCI mode over the last decade. Hereinafter, metabolomics studies are discussed in detail (partly taken from [44]) and a brief overview on GC(xGC)-APCI-MS studies of environmental samples, foodstuffs, pharmaceutical impurities, pesticides, steroids and VOCs is provided in **Table 12.1** in the appendix. In summary, recent publications have demonstrated the advantages of GC-APCI-TOFMS with regards to the range of amenable compounds, detection sensitivity, compound identification and the provision of complementary information compared to conventional GC-MS approaches.

Only few metabolomics studies employing GC-APCI-MS have been published so far. Ishimaru et al. in 2008 [11] used APCI for headspace GC-MS to reduce complexity of mass spectra of volatile metabolites in bacterial samples. Three different bacterial species (*Escherichia coli*, *Streptococcus pneumoniae* and *Staphylococcus aureus*) were included in the study, which revealed 156 out of 163 peaks detected in total (95.7%) to be species specific. However, the potential of GC-APCI-MS for the rapid identification of pathogens might be compromised by a low repeatability of the presented approach, or the identified peaks might not have originated from the bacterial species. In fact, only 54.7% to 75.3% (on average 68.0%) of detected peaks overlapped in the analyses of the three strains upon repeated cultivation, which was not further addressed in the paper. In addition to that, the peaks that were thought to be specific to one of the bacterial species were selected according to visual judgement rather than statistical analysis of an adequate number of biological replicate samples [11].

“In 2009, Carrasco et al. [9] tested GC-APCI coupled to a Bruker microTOF mass spectrometer using a standard mixture with 31 compounds from diverse chemical classes (amino acids, organic acids, alcohols, etc.). The complete analytical procedure, includ-

ing source and instrument parameters, was optimized and evaluated with respect to characteristic analytical parameters. LODs were within 11.8 to 72.5 nM, which is considerably lower than previously published for the determination of these compounds by GC-EI-qMS. LOQs were in the sub-micromolar range (between 0.05 and 0.24 μM) and the linear range extended over three orders of magnitude. Furthermore, good interday repeatabilities below 7% (RSD) were achieved. The developed method was applied to the analysis of human cerebrospinal fluid samples. Adequate values of mass error and sigma value (below 2 mDa and 10, respectively) made it possible to identify more than 300 distinct features in cerebrospinal fluid (CSF) samples [9].

Pacchiarotta et al. (2010) applied the well-known concept of parallel detection for metabolomics studies using APCI-TOFMS and FID [10]. Since APCI works at ambient pressure, parallel detection is fairly straightforward while parallel detection with conventional EI-MS detectors requires restriction capillaries in order to maintain proper column flow to both detectors. In this study, MS combined with APCI ionization provided high detection sensitivity with LOQs of silylated amino acids of about 0.5 μM and additional structural information. In contrast to that, FID appeared to be more robust, and provided a wider dynamic range that is especially helpful for quantification of compounds present at concentrations above saturation of the MS detector. Furthermore, FID covers all organic peaks, some of which might be missed by GC-APCI-MS, but it ideally requires baseline-resolved analytes, which is often not achieved by one-dimensional GC for complex biological samples. Nevertheless, the authors claim that MS and FID provided complementary information that might be helpful to expand the set of analyzable compounds in metabolic profiling studies [10].” [44]

The results of my master thesis were compiled in the beginning of my doctoral thesis and published in 2011 [17]. In this work, GC-APCI-TOFMS was compared to GC \times GC-EI-TOFMS, GC-EI-TOFMS, GC-CI-qMS and GC-EI-qMS in terms of repeatability, dynamic range, limit of detection and quantification using a mix of 43 metabolites and 12 internal standards. Lower limits of quantification for GC-APCI-TOFMS ranged between 0.06 and 7.81 μM , and relative standard deviations for calibration replicates were between 0.4 and 8.7%. For all compounds and techniques, except in four cases, R square

values were above 0.99. With regards to limits of quantification, GC-APCI-TOFMS was only inferior to GC×GC-EI-TOFMS, but outperformed all other techniques tested. GC-APCI-TOFMS was further applied to metabolic fingerprinting of *E. coli* wt strain MG1655 and the *double*-mutant PntAB-UdhA. The latter mutant lacks two genes encoding for two transhydrogenase isoforms that are involved in the electron transfer from NADH to NADP⁺ and vice versa. It was reported that this leads to a substantial re-routing of the metabolic flux from glycolysis to the pentose-phosphate pathway to meet the cellular demand for NADPH [73]. Out of 45 features that differed significantly (false discovery rate <0.05) between the two *E. coli* strains, 25 metabolites, mainly from the Krebs Cycle and related pathways, were identified due to highly accurate and reproducible mass measurement ($\Delta m \pm SD$ below 5mDa over m/z 190-722). Starting from the protonated molecule, six additional metabolites were identified that had not been found in a previous study using GC×GC-EI-TOFMS and the NIST 05 mass spectral library for identification purposes. In fact, identification based on their EI spectra was not possible because the derivatives of dihydroorotate, N-acetylneuraminate and N-acetylputrescine were contained in neither the used NIST 05 nor the commercial Fiehn metabolite library. Silylation adducts formed in the APCI source helped the identification of unknown compounds as its formation proved to be structure-dependent as it was not observed for compounds lacking a carboxylic group. It was concluded that GC-APCI-TOFMS is a potentially useful tool for metabolomics. However, ionization at atmospheric pressure was influenced by multiple factors. For example, moisture content in the source was highly variable and correlated to the humidity in the laboratory. This requires future investigations to evaluate and optimize factors influencing reproducibility during GC-APCI-TOFMS analysis. This might benefit lower limits of quantification, as well as more consistent results in comparative studies in metabolic fingerprinting.

Part of the reason why APCI has not been more widely adopted in GC-MS based metabolomics is the lack of commercial mass spectral libraries. Pacchiarotta et al. (2013) addressed this issue by building a web-based 150-component GC-APCI-QqTOF spectral library [74], which also proved to be useful in the current thesis (see chapter 8). Compounds most commonly found in biofluids and/or food specimens were initially included in the library, such as amino acids, fatty acids, organic acids and sugars. Chal-

lenges in the GC analysis of above mentioned metabolite classes with their common fragmentation patterns are described in the article, which also provides a protocol for *de novo* identification of unknowns on a routine basis comprising the steps of (1) spectra calibration, (2) calculation of elemental formulas considering the isotopic distribution, (3) MSⁿ experiments to reduce the number of putative candidates, and (4) verification with an analytical standard if available [74]. The spectral library and *de novo* identification protocol were employed in follow-up studies on avocado fruit extracts [75,76] and human urine specimens [77] by the same authors. GC-APCI-TOFMS proved as complementary tool to GC-EI-qMS in the characterization of the avocado fruit metabolome, with a wider range of amenable compounds in terms of polarities and molecular weights that could be analyzed. However, this might be partly explained by the different intrinsic detection capabilities of the employed mass analyzers. In fact, GC-APCI was coupled to an orthogonal-accelerated Qq-TOF MS, whereas a single quadrupole was used for GC-EI-MS analysis [75]. Untargeted metabolomics by GC-APCI-TOFMS was further performed in two studies to either investigate the effects of ripening degree and fruit varieties on the metabolic profile of avocado fruits [76] or urinary tract infection (UTI)-induced changes of the human urinary metabolome [77]. Two-class PLS-DA models were built first using the information on the degree of avocado ripening and the infectious status of the *E. coli* UTI patients, either diseased on the day of enrolment or recovered after 30 days, as sample class characteristic. This was followed by the allocation of peaks that accounted for the observed group separation in each case. Several of the identified metabolites could be corroborated by previous studies and information on new possibly relevant metabolites was gained, demonstrating, according to the authors, the usefulness of the applied GC-APCI-MS platform in exploratory studies for these applications. However, a major limitation was the incomplete annotation of important group discriminating peaks, which might improve with the future expansion of the employed GC-APCI-MS spectral library [76,77].

GC-APCI-MS has also been exploited as a complementary tool to GC-EI-MS for unknown metabolites that had escaped previously identification using EI libraries. Strehmel et al. (2014) derived a strategy for the pair-wise assignment of GC-EI-qMS and GC-APCI-QTOFMS mass spectra based on 102 known components in MeOx-TMS

derivatized extracts of *Arabidopsis thaliana* [78]. GC separation was achieved under almost identical conditions, which facilitated establishing a prediction model for GC-APCI-QTOFMS retention times based on GC-EI-qMS data. Furthermore, in order to identify the corresponding APCI spectrum of an unidentified EI spectrum in the GC-APCI-QTOFMS data, the mass of the protonated molecule could be hypothesized for 80% of the reference compounds. This was based on the EI molecular and/or $[M-CH_3]^+$ ion, or a pair of consistent EI and APCI fragment ions. From the protonated molecule elemental formulas were then calculated and additional information from experiments with deuterated derivatization reagents and CID mass spectra was incorporated, ultimately leading to 14 structural hypotheses for a total of 25 unknown compounds [78]. A similar strategy for the pair-wise assignment of GC-APCI-TOFMS and GC×GC-EI-TOFMS mass spectra was applied in chapter 8 in the current thesis. Moreover, Jaeger et al. employed GC-APCI-TOFMS and the above-mentioned concept of retention time prediction / elemental formulae calculation to confirm putatively annotated metabolites by GC-EI-MS in the study of global metabolomic changes in *Caenorhabditis elegans* lifespan mutants [79]. The number of metabolites identified by APCI was found to be comparable to EI, with a total of 95% (108 of 114) primary GC-EI-MS annotations confirmed by the elemental sum formulas obtained from high-resolution APCI-TOFMS data. Importantly, this included several metabolites with a low EI spectral match factor or high RI deviation relative to the library reference compound. In the further course of the study, PCA, hierarchical clustering and ANOVA followed by post-hoc testing were performed on the GC-EI-MS data, revealing a critical role of peroxisomes in aging [79].

4.4 Experimental design and data analysis in metabolomics

A metabolomics study comprises several experimental steps, subsequent data analysis, and, finally, biochemical interpretation of the results. Sampling, quenching, extraction of metabolites and derivatization have to be performed. Samples are further analyzed using optimized and validated analytical methods. Common data analysis steps include data pre-processing, (optionally) manual peak integration, statistical analysis, identifica-

tion of metabolites, and validation of the results. Finally, the study outcome has to be interpreted in the particular biological context. In such a study, the observed variability of an acquired data set results from biological, pre-analytical and analytical sources. The latter two are preferably reduced to a minimum. Established protocols for quenching, extraction of metabolites and derivatization were used to cope with pre-analytical variability as described in chapter 5 and the experimental sections of the following chapters. In this subchapter, strategies are described that were applied in this thesis to eliminate analytical variability. Furthermore, an overview on data analysis is given, with an emphasis on the evaluation of high-resolution GC-APCI-TOFMS data.

4.4.1 Setting up an analysis

Proper priming of the analytical system is a key factor in obtaining data of high quality. Whenever parts of the GC system (e.g., ferrules, connectors, pre-column) are exchanged, the system has to be checked for leaks and the oven temperature should be ramped stepwise to the maximum temperature tolerated by the stationary phase to clean and equilibrate the system. It is also advisable to start a sequence with a blank sample containing, e.g., the derivatization reagents to equilibrate the system. A test mixture can be used to check proper and expected performance of the system. GC column performance should be monitored on a regular basis because contamination of the column or column bleeding over time may deteriorate chromatographic separation. Some degradation of the stationary phase, typically more pronounced for thick film stationary phases, is unavoidable. It is mainly visible as an elevated baseline at higher column temperatures, which is caused by the formation of linear and cyclic polysiloxanes in case of siloxane-based stationary phases. In particular oxygen in the carrier gas in combination with higher column temperatures will increase column bleeding. This background noise negatively affects the quality of MS spectra and the limit of detection. Nowadays, most manufacturers offer low bleed column that are suitable for GC-MS.

In case of GC-MS instruments with nominal mass resolution, tuning of the mass scale and ion intensities is performed with perfluorotributylamine (PFTBA) on a regular basis using automated routines in the MS software. For high-resolution MS, however, tuning

is more elaborate. Initial external calibration with a vendor standard mixture that contains masses in the range of the planned application is advised. Ions monitored for calibration should not go into saturation, and experimental masses should not deviate more than 1 ppm from their theoretical values. In addition, mass spectra should be recalibrated after every analysis, because mass accuracy of high-resolution TOFMS and qTOFMS instruments suffers from slight drifts in temperatures and/or voltages between runs. This can be accomplished by the addition of standards to each sample (for details see section 5.4.3). Moreover, a lock mass that is present in every spectrum can be used to recalibrate the mass scale afterwards.

Samples should be analyzed in random order. Blanks, quality control samples such as pools of individual samples, and calibration controls should be interspersed. The addition of blanks helps to identify contamination peaks that originate from solvents, extraction solutions, impurities from internal standards or the column. The blanks should undergo the complete analytical procedure. Quality control samples are helpful especially in case of large sample sets that cannot be analyzed in a single batch. Between-block effects, also called batch effects, are most critical in metabolic fingerprinting. Quality control samples that are repeatedly measured may be used to correct for variance introduced by instrumental drifts. Calibration controls should also be included to ensure proper quantitation across a batch. They may contain two to three calibration standards well in the linear range of the calibration curves of the target compounds and are used to calculate recoveries.

An important aspect of metabolomics is the generation of quantitative data. In order to account for matrix effects and to achieve required levels of accuracy and precision, SIL-ISs are commonly included in quantitative analysis. Ion suppression or enhancement in case of the stable isotopically labeled analogue is expected to be equal to that of the sample analyte given the almost identical chromatographic and ionization behavior. One should note that deuterated internal standards, depending on the number of deuterium atoms incorporated, might show a partial separation from the unlabeled analyte. Using unique quantification traces, the mass spectrometer is capable of distinguishing analyte and labeled analogue. For quantification, the response [peak area normalized by the

corresponding internal standard (IS) peak area and internal standard concentration: $\text{peak area} \times \text{IS concentration} / \text{IS peak area}$] is plotted instead of the peak area value. At the initial stage of method development, it is important to evaluate matrix effects, which have been reported mostly for LC-MS [80,81], but also for GC-APCI-MS [82]. Spike-in experiments should be performed to reveal ion suppression [83]. Definite amounts of the studied analyte are added to the extracted samples. Plotting the response values of calibrators and spiked sample extract against each other should yield a linear curve with a slope of one and an intercept of zero. However, possible over-quantification may go unnoticed in spike-in experiments [39]. For that reason, absolute amounts obtained by a novel method should always be validated by an established technique or the analysis of certified reference materials.

Furthermore, the addition of standards before extraction and injection, respectively, allows the detection of extraction losses and/or problems with sample injection. The addition of a homologous series of standards further enables the calculation of retention indices. In their plant metabolomics studies, Gullberg et al. [84] included 11 SIL-ISs representing different compound classes and optimized their extraction and derivatization protocols using a design-of-experiments approach. Koek et al. [41] distinguished three performance classes of silylated metabolites and recommended the inclusion of representative compounds from each class, preferably labeled with stable isotopes, for method optimization and validation. This classification was based on derivatization efficiency, analytical repeatability and reproducibility, linear range and quantification limits. Sugars and organic acids, which showed the best performance, were classified as group I compounds, followed by metabolites containing amine or phosphoric functional groups (group II compounds), whereas those with amide, thiol or sulfonic functional groups (group III compounds) were regarded as the most critical to analyze [41].

However, SIL-ISs are expensive and their availability is limited. In addition, they can complicate several data processing steps, such as deconvolution, peak picking and integration, especially in the case of GC \times GC-MS [43]. Lien et al. [85] employed deuterated MSTFA as a derivatizing reagent (d9-MSTFA) for a mixture of standard compounds to overcome the limited number of available SIL-ISs. This deuterated

standard mix was then spiked into MSTFA derivatized samples prior to GC-MS/MS analysis. However, this strategy does not correct for metabolite losses during sample preparation and derivatization. Alternative approaches to overcoming the lack of commercially available SIL-ISs include the use of d3-MCF/d3-MeOH to generate stable isotopically labeled methyl chloroformate derivatives [86] and the *in vivo* biosynthesis by feeding microorganisms stable isotopically labeled substrates [87]. However, such approaches will cover only parts of more complex metabolomes observed for example in plants and mammals.

4.4.2 Data analysis workflow

Analysis of the acquired mass spectral data depends on the intended strategy. For a quantitative method, mass traces (quantifier and qualifier ions) of the target compounds have to be specified, along with their retention time, and an internal standard has to be assigned for each analyte that is ideally the labeled analogue. Concentration levels of the standards have to be defined, further parameters may be defined (e.g. 1/x weighing for linear regression analysis or smoothing), and finally calibration curves can be calculated and integrated together with the samples. If SIL-ISs are used, isotope dilution analysis can be performed for quantification.

For comparative analysis of multiple sample groups in an untargeted approach, data files are pre-processed first, e.g., baseline correction, recalibration of the mass scale, deconvolution and peak picking are performed, and then peaks across chromatograms are aligned in one data matrix that contains their m/z values, retention times and area integrals. In the process of data processing and alignment, one has to be aware of several pitfalls, such as batch effects or missing values in the data matrix. The latter can be imputed by dedicated software algorithms provided that they are missing completely at random [88]. Prior to multivariate analysis, it is important to minimize contributions from unwanted biases and experimental variance by using normalization methods. The simplest of these would be normalization to an appropriate internal standard or the sum of all integrals in the sample, whereas more advanced methods include quantile normalization or variance stabilization approaches in the absence of significant between-group

differences in total signal intensity. Depending on the nature of specimens analyzed, further normalization may be conducted to correct for undesired biological variance, such as normalization to creatinine concentration in case of urine specimens as a consequence of differences in fluid intake or to the number of cells or protein content in the analysis of cell extracts. Since the number of peaks, features, or “dimensions” that a given dataset represents may number in the hundreds, dimension reduction techniques such as principal component analysis (PCA) are commonly used to reduce data complexity and, thus, facilitate visualization of differences between samples. Selection of differential features for subsequent assignment to metabolites using different strategies for identification (see subchapter 4.5) could be initially based on two-sided t-tests or analysis of variance (ANOVA) followed by post-hoc tests. In case of large feature numbers correction for multiple testing is needed, for example, by controlling the false discovery rate [89]. The ultimate goal is the reliable detection of features (biomarkers) and corresponding pathways that distinguish a group. In recent years, a number of different classification algorithms, such as nearest neighbour, linear discriminant analysis and classification trees, have been developed and applied to the classification of samples based on transcriptome data [90]. Their applicability and performance for LC/MS- or GC/MS-based metabolite fingerprinting data still awaits systematic evaluation. Recent work performed in our laboratory on ¹H NMR-based metabolite fingerprints of different physiological fluids from different species compared Elastic Net, Nearest Shrunken Centroids, Partial Least Squares-Discriminant Analysis (PLS-DA), Random Forests, Top scoring pairs, and Support Vector Machines (SVM) as classification tools [91]. Random Forests and SVM combined with *t*-score based feature filtering were found to be overall the best suited techniques for classification of samples [91].

To facilitate the biological interpretation of results, multiple databases are available. Major metabolite databases are the Human Metabolome Database (HMDB) [23], the Madison Metabolomics Consortium Database (MMCD) [92], Metlin [25], and LIPID MAPS [24]. They contain huge numbers of fully annotated metabolite entries tailored to studies on specific compound classes/pathways, e.g. lipid metabolism, or a particular organism. These databases incorporate many sources of information. The HMDB, for example, includes chemical, clinical, and molecular biology/biochemistry data for over

7900 fully annotated metabolites. It also contains spectra information for more than 800 compounds. Search entries are linked to other databases, e.g. KEGG, and still more dedicated software and search algorithms have been added to the latest version (V 2.5). However, each of the mentioned databases only covers a fraction of the metabolome. Therefore, one should search multiple databases. Zhou et al. [93] recently implemented a web-based software tool, which searches the four major metabolite databases listed above, and finally combines the results for a more comprehensive coverage of possible candidates for mass-based metabolite identification. In addition to searching databases, mapping of identified metabolites to corresponding metabolic pathways is helpful. In a recent work by Leader et al. [94], the authors described a freely available tool called “Pathos”, which uses as input raw m/z values or KEGG IDs, if the metabolites have been already identified, to search the KEGG database. Apart from returning a pathway map with the identified metabolites highlighted, it can also visualize changes in their experimental abundance if relative quantification data under different experimental conditions is provided. This allows the scientist to obtain a quick survey over regions of metabolism of potential interest. Finally, metabolomics data can be combined with results from other –omics fields to yield a global picture of the biological entity under study.

4.4.3 Recalibration of high-resolution GC-APCI-TOFMS data and calculation of elemental formulas

Especially in the case of GC-high-resolution MS (GC-HRMS), mass recalibration can be difficult. For external calibration, the APCI source can be replaced with an ESI source and calibration can be performed using an ESI tune mix. Alternatively, FAMES can be infused from the top of the APCI source at the beginning of each analytical run. However, post-acquisition recalibration of mass spectra is not as straightforward as in LC-MS. Internal standards that are added to every sample prior to analysis can be used for internal calibration, which was carried out in this thesis (compare section 5.4.3).

Accurate mass measurements are especially beneficial for compounds that are not included in spectral libraries. In that case, soft ionization methods coupled to HRMS

yield protonated molecules, from which elemental formulas can be derived and databases searched after elimination of groups that have been introduced by derivatization. This concept has been applied in the current thesis and is described in more detail in section 5.4.4. Nevertheless, high mass accuracy suffices rarely to achieve an unambiguous assignment of a single molecular formula to an unknown, especially in higher mass regions, where the number of elemental combinations for a given mass and the defined mass error is many times higher. Therefore, isotopic pattern matching is required as an orthogonal filter to remove false candidate formulas. Kind and Fiehn in 2006 [95] demonstrated via query of over 1.6 million elemental formulas that an MS with 3 ppm mass accuracy and 2% error for isotopic abundance patterns outperforms a hypothetical mass spectrometer with 0.1 ppm mass accuracy that does not take isotope information into account.

Tools for calculating elemental formulas from accurate mass measurements, e.g. the SmartFormula tool from Bruker Daltonics, consider chemical and heuristic rules as proposed by Kind and Fiehn in 2007 [96] in addition to comparing the isotopic patterns, which is assessed by the mSigma value (between 0 in case of perfect match and 1000 in case of no match). These rules include, among others, constraints on number and ratio of elements as well as LEWIS and SENIOR rules, and their application has resulted in correct metabolite annotations with 98% probability for 6,000 compounds in target databases [96]. For calculating elemental formulas, the search space can be reduced by assuming a lower experimental mass error. But, the correct formula may be missed, whereas if a larger mass error is tolerated, hit lists become more extensive and are more likely to include hits that are closer to the experimental data than the correct formula. SmartFormula3D additionally relies on the accurate mass and isotope pattern of the fragments from a parent ion in MS/MS experiments. Accordingly, parent ions that cannot be related to the fragments are excluded, thus restricting the search space drastically up to a single remaining elemental formula in some cases. In summary, in the identification process of unknowns from high-resolution GC-APCI-TOFMS data still a lot of manual intervention is required as reported in section 5.4.4.

4.5 Identification of metabolites

Recent years have seen great progress in the identification of spectral features, still considered one of the biggest bottlenecks in metabolomics (see <http://fiehnlab.ucdavis.edu/staff/kind/Metabolomics-Survey-2009/>). For targeted metabolomics, only very few peaks are of interest, so-called “known unknowns” that have been either described in the literature or included in databases, but have not been reported for the organism and/or type of sample at hand. Their identification is typically accomplished by means of commercial standards. In semi-targeted or untargeted metabolomics, the use of reference compounds for the identification of features becomes cumbersome as the number of features recorded in a single run may easily run into the hundreds. Therefore, routines for automatic feature annotation and identification of so-called “unknown unknowns” are required. The rapid progress in MS technologies and resulting improvement in detection sensitivity has revealed a vast number of such compounds that still await positive identification. This holds especially true for secondary metabolites of plants and fungi, whose metabolomes have been estimated to contain tens of thousands of individual metabolites [27]. However, these numbers are only rough estimates [97].

In 2007, Sumner et al. [26] proposed four levels of metabolite identification depending on the amount of information collected on a feature of interest. The authors distinguished between positively (level 1) and putatively (level 2) identified compounds, compounds that are putatively assigned to a class of compounds (level 3), and unknown compounds (level 4). For positive identification (level 1), matching two independent parameters, such as retention time index and mass spectrum, relative to an authentic standard is required, which emphasizes the need for chromatography as an additional dimension to MS. Several excellent reviews have appeared in recent years on advances in metabolite identification by means of chromatography and MS [98-101].

4.5.1 Gas chromatography

Separation of analytes in a time dimension constitutes added value to the identification process. GC-MS retention times are very stable and, therefore, of great benefit to identi-

fication. Typically, retention time of a compound is given relative to a reference compound within the chromatogram to exclude variation due to slight instrument changes in-between consecutive analyses and make retention times more comparable to other laboratories. In 1958, Kováts introduced the concept of calculating retention indices (RIs) [102] based on a series of alkanes as reference compounds for isothermal GC. This concept was later modified and applied to temperature-programmed GC by Van den Dool and Kratz [103]. Temperature programmed RIs are often referred to linear retention indices. Retention indices are a normalized measure of retention time. They are calculated relative to compounds from a homologous series of alkanes or fatty acids that are added to every sample. In temperature-programmed GC, members of a homologous series show a linear increase in retention time with increasing number of carbon atoms for a given stationary phase. Per definition, members of the homologous series are assigned a retention index by multiplying the respective carbon number with 100. The linear retention index I of an analyte x is calculated using the following formula:

$$I_x = 100n + 100 \frac{t_{R(x)} - t_{R(n)}}{t_{R(n+1)} - t_{R(n)}}$$

$t_{R(x)}$	Retention time of the analyte
$t_{R(n)}$	Retention time of the reference compound eluting before the analyte
$t_{R(n+1)}$	Retention time of the reference compound eluting after the analyte
n	Carbon number of the reference compound eluting before the analyte

Traditionally, n-alkanes have been used as RI markers [104,105], albeit fatty acid methyl esters (FAMES) [106] or any other homologous series can be used. Selection of the reference system depends on column polarity, e.g. n-alkanes are poorly dissolved in polar stationary phases, and the detector used, e.g. n-alkanes evade ionization using an APCI source for GC-MS. Retention indices are partly reported in MS databases, such

as NIST, HMDB [23], Golm Database [107], and Fiehn Metabolome Database [106]. Physicochemical properties of not yet annotated metabolites are collected in the Golm DB as mass spectral tags (MSTs), which were introduced recently [108,109] in analogy to expressed DNA sequence tags (ESTs). MSTs now await their structural elucidation.

For databases to be useful, they need to be updated regularly. Insufficient resources for database curation still limit the utility of HMDB and Golm DB. Hence, NIST is still the most widely used database. The latest NIST 14 version contains EI-MS spectra for over 242,000 unique compounds, and for more than 56,000 compounds RI values are provided. However, the majority of entries do not refer to metabolites.

In the absence of RI values for metabolites of interest, they may be predicted. In 2011, Kumari et al. [110] introduced a workflow for re-identifying 29 trimethylsilylated metabolites from accurate mass GC-CI-TOFMS. Elemental formulas were derived from the protonated molecules in an initial step, followed by in-silico derivatization and prediction of RIs using the NIST RI algorithm with a correction factor for TMS groups [111]. In a subsequent step, accurate mass EI spectra were predicted for remaining structures and scored against experimental spectra. As a result, the correct structure was obtained in 73% within the top-5 hits of the proposed candidates [110]. Earlier, Mihaleva et al. [112] made use of multiple linear regression (MLR) and support vector regression (SVR) models for RI prediction, which cut the hit lists obtained from matching experimental spectra of the NIST 05 MS library in half.

The concept of RIs has also been transferred to GC×GC [113], but it is not used routinely. Nevertheless, GC×GC facilitates identification of compounds belonging to different chemical classes as they occupy different regions within the 2D separation space. Hence, structure-retention relationships serve as plausibility check for the identification of unknown signals in the chromatogram. Quantitative structure-retention relationship (QSRR) models have also been developed to correlate molecular structures to their retention times [114]. In summary, retention times are essential in compound identification, especially in case of isomers, which cannot be distinguished by their mass spectrum only. To check for plausibility, one should not only rely on the mass spectrum, but also consider chromatographic retention of a putatively identified metabolite.

4.5.2 Mass spectrometry

Mass spectrometry is the method of choice in the identification of unknown metabolites. Matching EI fragment patterns to spectral libraries has been traditionally performed in MS-based metabolite identification. A second strategy involves the use of soft ionization techniques in conjunction with high-resolution MS. Elemental formulas are derived from the accurate mass and isotopic pattern of the protonated molecules in an initial step, groups introduced by derivatization are removed, and curated formulas can be fed into databases for a putative identification (compare section 5.4.4).

Successful identification of known unknowns based on mass spectral libraries requires mass spectra of high quality, powerful spectral matching algorithms to reveal significant degrees of spectral convergence and, particularly, comprehensive and reliable reference spectra libraries. Successful library identification strongly depends on the quality of the input spectrum. Spectral quality suffers often from poor deconvolution, ion source overload and/or varying acquisition parameters. Mass spectral deconvolution has to be performed in case of mixed mass spectra originating from unresolved peaks. Local peak maxima are analyzed for all ion traces and ions that share a peak maximum (retention time) are combined in one mass spectrum. This process gives a “pure” mass spectrum that can be matched against reference spectra. Of course, deconvolution is only successful if overlapping peaks display a slight difference in peak maxima and if enough data points across the peak are available to distinguish these slight differences. Undoubtedly, deconvolution software packages such as AMDIS supplied with the NIST library, ChromaTOF (LECO), or AnalyzerPro (SpectralWorks) constitute an added value for proper analysis of highly complex samples. However, deconvolution may generate outright false or impure spectra. In 2008, Lu et al. [115] evaluated the performance of AMDIS, ChromaTOF and AnalyzerPro using a mixture of 36 endogenous compounds, which exhibited varying chemical structures and concentrations. Numerous failures of AMDIS and ChromaTOF occurred, which resulted in incorrect identifications [115], emphasizing the need for further improvements or novel approaches.

For spectral matching, several algorithms such as INCOS, probability-based matching (PBM) and dot-product scoring have been described [116]. In case of INCOS and PBM,

the number of mass/intensity pairs of library spectra considered for matching is restricted, which might result in failure to retrieve as first hit the correct library entry for a searched full spectrum of an unknown. The dot-product algorithm, in contrast, is first carried out on condensed spectra before finally considering entire spectra for definite metabolite identification. Stein and Scott in 1994 [117] found dot-product scoring superior (75% accuracy) to four other algorithms (Euclidean distance, absolute value distance, PBM and Hertz et al. similarity index) based on matching test spectra against the NIST library, which contained 12,592 spectra of about 8,000 unique compounds at that time.

The library search delivers a match factor, which is a measure for the agreement between search and library spectrum. A NIST search delivers both, a match factor (direct match) and a reverse match factor. The reverse search algorithm compares the library spectra against the search spectrum and non-matching peaks in the submitted spectrum are ignored. Match factors above 900 indicate an excellent match with 999 being a perfect match. Match factors above 800 are still considered good, while match factors below 700 have to be considered with care. In general, top candidates from the hit list should be checked for plausibility by considering information on retention indices, or reference standards should be run for verification.

It is also possible to create custom libraries with commercial standards as accomplished by Smart et al. for methyl chloroformate derivatives [118]. *In silico* mass spectral fragmentation routines can be used if reference spectra are not available. However, in 2009, Schymanski et al. [119] tested three commercial programs for *in silico* generation of mass fragments and concluded that available software was not practical yet.

Alternatively, efficient and sensitive identification of metabolites may be accomplished by soft ionization techniques, accurate mass measurement and tandem MS. GC-APCI-MS principles and applications (subchapter 4.3) as well as data analysis strategies (subchapters 4.4 and 5.4) including (re-)calibration of the mass scale and calculation of elemental formulas are reported elsewhere. Furthermore, GC-CI-TOFMS has also been employed for metabolite identification [50,110]. Abate et al. [50] reported, that higher ion intensities improved accurate mass and isotopic abundance measurements and thus,

allowed the determination of elemental formulas with higher confidence. In these GC-CI-TOFMS studies, high-mass tuning [50] as well as the careful selection of the reagent gas [110] proved critical in obtaining higher ion intensities.

According to Sumner et al. [26], two independent/orthogonal parameters, e.g. retention index/mass spectrum or accurate mass in combination with MS/MS information, relative to a commercial standard are required to positively identify compounds. Whenever standards are not available, confidence in a putative identification can be increased by the use of isotope labeling to distinguish background ions from ions that originate from a true biological source. Furthermore, with the commercial availability of fraction collectors for GC, compound isolation for metabolite identification by NMR spectroscopy has become feasible. However, stability of derivatized metabolites is critical while collecting fractions by cold trapping and a huge discrepancy exists between amounts that are subjected to GC-MS analysis and those required for NMR analysis. Once a single or several metabolites have been identified, proper biological or medical interpretation is promoted by metabolic pathway databases such as the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [120].

5 Experimental section

5.1 Materials

The mixture of fatty acid methyl esters (FAMES #CRM47885) in methylene chloride was purchased from Supelco (Belafonte, PA, USA). Compounds included in the metabolite standard mixtures used in chapters 6, 7, and 8, except for glycerol, as well as citric acid, aminoadipic acid, malic acid, tryptophan, lysine, threonine, cysteine, histidine, tyrosine, arginine, asparagine, lauric acid, myristic acid, stearic acid, serine, ornithine, glycine, palmitic acid, N-acetyl-L-aspartic acid, odd-numbered, saturated straight-chain fatty acids (C9-C19), pyridine, methyl chloroformate (MCF), methoxylamine hydrochloride, 3-picoline, propyl chloroformate, fetal bovine serum, L-glutamine, sodium dodecyl sulfate, the FluoroProfile kit for determination of protein content in cell pellets, the disodium salt of D/L racemate of 2-hydroxyglutarate, disodium D- and L-2-hydroxyglutarate, (R)-5-oxotetrahydro-2-furancarboxylic acid, formic acid, 3-trimethylsilyl-2,2,3,3-tetradeuteriopropionate (TSP) in D₂O, ethanol (ACS reagent grade), urease type III (#U1500), [U-¹³C]glucose, [²H₇]trans-cinnamic acid, norvaline, [U-¹³C]β-hydroxybutyric acid, [2,4,4,4-²H₄]citric acid, and alkane standard solutions (C8-C20 and C21-C40) were from Sigma-Aldrich/Fluka (Taufkirchen, Germany). N-Methyl-N-(trimethylsilyl)-trifluoroacetamide (MSTFA) was purchased from Macherey-Nagel (Dueren, Germany), sodium dihydrogen phosphate monohydrate, tri-sodium phosphate dodecahydrate, dipotassium hydrogenphosphate trihydrate, potassium dihydrogen-phosphate, and chloroform (analytical grade) from Merck KGaA (Darmstadt, Germany), and [U-¹³C, U-¹⁵N] cell-free amino acid mix, [U-¹³C]lactic acid, sodium [U-¹³C]pyruvate, [U-¹³C]fumaric acid, and [U-²H]succinic acid from Euriso-top (Saint-Aubin Cedex, France). [4,6,7-²H₃]5-hydroxyindole-[²H₂]acetic acid, D/L-[2,3,3-²H₃]malic acid, and the disodium salt of D/L-[2,3,3-²H₃]2-hydroxyglutarate were delivered from C/D/N Isotopes Inc. (Quebec, Canada), and [U-¹³C]lactose from Omicron Biochemicals (South Bend, IN, USA). Methanol

(MeOH; LC-MS grade), ethyl acetate (analytical grade), methylene chloride (LC-MS grade) and propanol (LC-MS grade) were from Fisher Scientific GmbH (Ulm, Germany), and glycerol, isooctane and acetonitrile (HPLC grade) from BDH Prolabo (VWR International, Vienna, Austria). Purified water from a PURELAB Plus system (ELGA LabWater, Celle, Germany) was used in all experiments. MiaPaCa-2 pancreatic cancer cells (ATCC #CRL-1420) were obtained from ATCC (Manassas, VA, USA). DMEM medium, glucose, and phosphate-buffered saline (PBS) were purchased from PAA Laboratories GmbH (Coelbe, Germany), and penicillin/streptomycin from Invitrogen (Karlsruhe, Germany). The Hsp90 inhibitor 17-DMAG was obtained from Invivogen (Toulouse, France).

A surrogate solution was prepared containing the 12 internal standards [2,2,4,4-²H₄]citrate, [U-¹³C]fumarate, [U-¹³C]glucose, [U-¹³C]β-hydroxybutyrate, [U-¹³C]lactate, [2,3,3-²H₃]malate, [U-¹³C]pyruvate, [U-²H]succinate, [²H₇]trans-cinnamate, [U-¹³C]lactose, [4,6,7-²H₃]5-hydroxy-indole-[²H₂]acetate and norvaline at 1 mM each in methanol. The odd-numbered, saturated straight-chain fatty acids (C9-C19) were dissolved in isooctane, and combined, hereinafter referred to as fatty acid mixture (1 mM of each compound).

5.2 Sample preparation

5.2.1 Preparation of calibration and cell culture samples

For the acquisition of calibration curves in chapter 8, 75 or 100 μL of a metabolite standard and 10 μL of the surrogate solution of ISs were combined in a 2-mL vial with a 400-μL glass insert and evaporated, respectively, only surrogate solution or 100 μL of pure methanol was dried in case of blank samples. Subsequently, MeOx-TMS derivatization was performed. To assess derivatization repeatability, a calibration standard at a concentration of 31.25 μM was derivatized in quintuplicate.

MiaPaCa-2 pancreatic cancer cells were grown in 75 cm² cell culture flasks (PAA Laboratories GmbH) in 25 mL of DMEM medium supplemented with 15% (v/v) fetal bovine

serum and enriched with 2 mM glutamine, 25 mM glucose, and standard penicillin/streptomycin solution. Cells were cultured in a humidified incubator at 37°C with 5% CO₂. For assessing the effect of 17-DMAG on cell metabolism, 1 × 10⁶ cells were plated into 10-cm Petri dishes and treated for 72 h with either medium containing 100 nM 17-DMAG or control medium only. Cells were washed thrice with cold PBS. For quenching of cell metabolism as well as combined cell harvesting and extraction of metabolites, direct solvent scraping in MeOH/H₂O (80:20, v/v) [121] was performed on ice and with cold solutions. The aforementioned experiments were carried out by Dr. Maria Mycielska from the Department of Surgery/University Hospital Regensburg.

Pellets were re-extracted twice more with 500 µL of MeOH/H₂O (80:20, v/v). The supernatants were combined and fully evaporated. Dried extracts were redissolved in 200 µL of MeOH/H₂O (80:20, v/v), and 60 µL were taken for MCF derivatization in chapter 6 (N=5 biological replicates per group). For the determination of cellular protein content according to the instructions provided with the FluoroProfile kit from Sigma-Aldrich, protein pellets were dissolved in 1 mL of a 20 mM sodium dihydrogen phosphate buffer, pH 4.5, supplemented with 1% (v/v) sodium dodecyl sulfate and diluted 1:50 (v/v) with water.

Cell culture supernatants from the above-described experiment of MiaPaCa-2 pancreatic cancer cells grown for 72 hrs under standard conditions (N=3 biological replicates) were used in chapter 8. Hundred µL of each supernatant sample were further prepared as follows: Ten µL of 1 mM [²H₇]trans-cinnamate were added to 100 µL of supernatant. Proteins were precipitated with 600 µL cold methanol, centrifuged for 6 min at 4°C and 9500 × g and the supernatant was collected. The protein pellet was washed twice with 200 µL of methanol. The supernatants were combined and the solvent was evaporated. MeOx-TMS derivatized cell culture supernatant samples were diluted 50:50 (v/v) in pyridine prior to injection.

5.2.2 Methoximation–trimethylsilylation

A two-step MeOx-TMS derivatization protocol was applied to dried standards (chapters 6, 7, and 8), spiked serum samples (chapter 6), spiked matrix samples (chapter 7), and

cell culture supernatants (chapter 8). Fifty μL of 20 mg/mL methoxylamine hydrochloride in pyridine were transferred to the dried residues and incubated at 60°C for 60 min. Prior to silylation, 10 μL of the fatty acid mixture were added, for internal recalibration of GC-APCI-TOFMS mass spectra as well as the calculation of retention indices. Silylation was carried out by the addition of 50 μL of MSTFA, and again incubating for 60 min at 60°C . Derivatization was carried out manually.

5.2.3 Methyl chloroformate/methanol derivatization

Manual MCF derivatization was performed in chapters 6 and 9 based on a modified MeOH/MCF protocol published previously [39]. Depending on the experiment, water was added to the standard, IS, and sample or sample extract already containing the IS, to obtain a final volume of 275 μL of aqueous phase. For the analysis of sample extracts, the mixture was dried and reconstituted in 275 μL of water. Next the aqueous phase was fortified with 167 μL of MeOH, and 34 μL of pyridine, followed by adding two times 20 μL of MCF vortexing for 10 s after each addition. Derivatives were extracted with 300 μL CHCl_3 under vortexing for 10 s. After 30 s, extraction was completed and an aliquot of the lower chloroform phase was taken for GC analysis.

5.3 Instrumentation

5.3.1 GC-APCI-TOFMS

A model 450-GC (Bruker Daltonics GmbH, Bremen, Germany) with an autosampler (model PAL COMBI-xt from CTC Analytics, Zwingen, Switzerland) for sample injection with a 10- μL Hamilton syringe was coupled to a microTOF orthogonal acceleration TOF mass spectrometer (Bruker Daltonics) via an APCI source. For separation of analytes, a Phenomenex ZB-AAA column (15 m \times 0.25 mm i.d. \times 0.1 μm film thickness, Torrance, CA, USA) was used with a 2 m deactivated pre- and a 0.5 m post-column of matching inner diameter (Agilent Technologies, Palo Alto, CA, USA). The oven temperature program was started at 50°C and held for 1 min, and temperature was ramped at 8 K/min

to 300°C and held for 15 min. A sample volume of 1 µL was introduced by means of an 1177 split/splitless injector at 280°C performing hot needle splitless injection (pre-injection dwell of 1 s) with a splitless time of 1 min. Helium served as the carrier gas at a constant flow rate of 1 mL/min. The transfer line to the MS instrument was maintained at 300°C, and the protrusion of the GC capillary from the exit of the transfer line was approximately 1 mm. The APCI source was operated as follows, with deviating parameter settings in chapter 7 in brackets: ionization mode, positive; drying gas (nitrogen) temperature, 160°C; drying gas flow rate, 2.0 L/min; APCI vaporizer temperature, 300°C (450°C); nebulizer gas (nitrogen) pressure, 4.0 bar (1.7 bar); current of the corona discharge needle, +3000 nA; capillary voltage, -2900 V; end-plate offset, 0 V. Water was introduced through the top ESI inlet into the APCI source by means of a syringe pump (KD Scientific Inc., Holliston, MA, USA) in chapters 6 and 9 unless otherwise stated in the text. Mass range for acquisition of spectra with a rate of 3 spectra/s was from 50 to 1000 *m/z*. Initial external mass calibration was performed with an electrospray ionization tuning mix (Agilent) and an ESI source, and an additional post-acquisition processing routine was performed for internal recalibration of each mass spectrum (see section 5.4.3).

Diverging from the instrumental and operational settings above, an Agilent model 6890 GC instrument equipped with an Auto Liquid Injector (model 7683B) was used in chapter 7, and data were acquired with the “GC-APCI II” source from Bruker in chapters 8 (in addition to measurements with the APCI I source) and 9. Chromatographic separation was achieved using a Rxi-5MS capillary column (30 m × 0.25 mm i. d. × 0.25 µm film thickness, Restek, Bad Homburg, Germany) in chapters 7 and 8, whereas an Rt- γ DEXsa (2,3-di-acetoxy-6-O-tert-butyl-dimethylsilyl gamma CD doped into 14% cyanopropylphenyl/86% dimethyl polysiloxane) column (30 m × 0.25 mm ID, 0.25 µm film thickness; Restek) was used in chapter 9. The flow rate of helium as carrier gas was set to 2 mL/min (constant flow) and the injector was kept at 250°C in chapter 9. In the same chapter, except for initial experiments to optimize the temperature ramp, the oven program started at 70°C, held for 1 min, was then ramped at 5 K/min to 190°C, and finally held for 7 min.

The APCI II transfer line to the mass spectrometer was maintained at 290°C, and the post-column capillary was mounted flush with the outlet of the transfer line. APCI II source parameters were as follows: ionization mode, positive; drying gas (nitrogen) temperature, 150°C; drying gas flow rate, 2.0 L/min; nebulizer gas (nitrogen) pressure, 2.5 bar; current of the corona discharge needle, +4000 nA; capillary voltage, -3000V; and end-plate offset, -500V. The APCI II source was kept at 300°C. No water was infused into the APCI source in projects presented in chapters 7 and 8.

Details on the “GC-APCI II” source setup are provided in section 8.2.2. The source parameters for APCI I and II were individually optimized using metabolite standards as part of my master thesis and the bachelor thesis by Thomas Hahn, respectively.

5.3.2 Miscellaneous

Solvents were evaporated to complete dryness by means of a vacuum evaporator (CombiDancer, Hettich AG, Bäch, Switzerland). Other lab equipment used in the course of this doctoral research work included a model MP220 pH meter from Mettler-Toledo GmbH (Greifensee, Switzerland) for the preparation of buffer solutions, a vortexer (lab dancer, IKA-Werke GmbH, Staufen, Germany), a heater with two heating blocks (Haep Labor Consult, Bovenden, Germany), a model GS-15R centrifuge (Beckman Coulter GmbH, Krefeld, Germany), and a FLUOstar OPTIMA multi-mode microplate reader (BMG Labtech, Ortenberg, Germany) for the determination of cellular protein content based on fluorescence measurements.

5.4 Data analysis

5.4.1 Software

Bruker DataAnalysis versions 4.0 to 4.1 (Bruker Daltonics) were employed for processing and manual inspection of GC-APCI-MS chromatograms and mass spectra, compound extraction, internal recalibration of mass spectra, and calculation of accurate

masses, elemental formulas and mSigma values. For peak integration and generation of calibration curves, exported data files were loaded into MassLynx V4.1 (Waters Inc., Milford, MA, USA), with m/z of $[M+H]^+$ for standard compounds and internal standards (see **Table 12.2**, **Table 8.2** and **Table 9.1** referring to chapters 6, 8, and 9, respectively), or m/z of a particular significant feature used as quantifier mass. ChemStation version D.02.00.275 (Agilent) was used for manual inspection of GC-EI-qMS chromatograms (section 9.3.1). GC×GC-EI-TOFMS raw data were processed by ChromaTOF software version 4.50 from LECO Corp. (St. Joseph, MI, USA) for annotating compounds in cell culture supernatants (chapter 8). Compounds extracted in DataAnalysis along with their retention times, experimental masses, area integrals, peak widths measured at half the peak height (w_h), and S/N ratios were further exported into Excel (Microsoft Corporation, Redmond, WA, USA). Generation of bucket tables, feature extraction by means of the FMF algorithm, and alignment were performed in ProfileAnalysis V2.1 (Bruker Daltonics). Bucket tables were then exported into Excel for generation of ROC curves and fold change plots (chapter 6) as well as cumulative distributions of RSDs (chapter 8). Excel was also used to create Bland-Altman plots and to perform basic statistics, e.g. the calculation of relative standard deviations (RSDs) or paired student's t tests. The statistical computing package R [122] was used to perform principal component analysis (PCA), the Kolmogorov-Smirnov test, Shapiro-Wilk testing, the Wilcoxon signed-rank test, ANOVA followed by post hoc tests using LIMMA [123], to impute missing values by MISSMDA [124], and to calculate false discovery rates (FDRs) according to Benjamini and Hochberg [89] using MULTTEST [125]. The partition coefficients of the MCF and MeOx-TMS derivatives of the standard compounds and fatty acids analyzed in chapter 6 were estimated using ACD/Laboratories V12.01 (Advanced Chemistry Development Inc., Toronto, Canada) and expressed as the $\log P_{\text{octanol/water}}$, i.e., the logarithm of the ratio of the concentrations of the un-ionized solute in octanol and water, respectively. Retention indices were calculated in Excel on the basis of alkanes or fatty acid derivatives according to equation (8) from a previous work by van den Dool and Kratz [103].

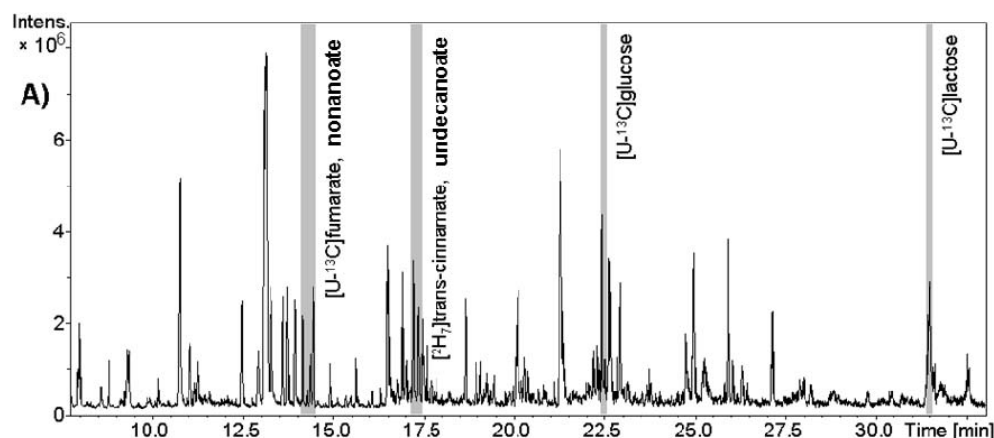
5.4.2 Calibration curves

Peak areas of standards were normalized to those of their corresponding SIL-ISs or a closely eluting or structural similar IS (see **Table 8.2** and **Table 9.1**). Hence, the response (area*IS concentration/IS area) was plotted as a function of the absolute amount of standard to yield the calibration curve for each standard. 1/x weighting was applied. In case that multiple chromatographic peaks were obtained for a metabolite, due to a varied extent of silylation in case of amino acids or cis-trans isomers of methoximated compounds, the most intense signal was exclusively evaluated. All data points that spanned the linear range of the calibration curve, determined by LLOQ and ULOQ, exceeded a peak S/N ratio of 8:1 and did not deviate more than 20% from the curve, according to the FDA Guide for Bioanalytical Method Validation [126]. The correlation coefficient R was obtained by linear regression analysis. To assess derivatization repeatability, response values from derivatization replicates were exported and RSDs calculated in Excel, which was also used to determine the order of magnitude of the linear concentration range based on a log-10 scale. For quantification, calibration curves were exported in MassLynx and subsequently used for sample processing.

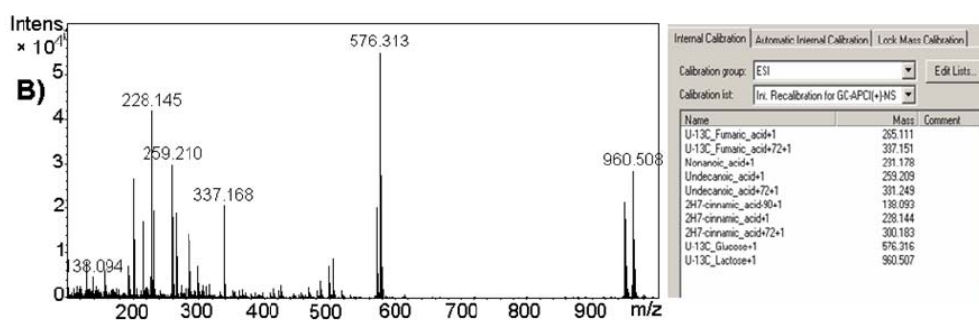
5.4.3 Recalibration of mass spectra

Post-acquisition recalibration of APCI mass spectra was based on SIL-ISs and fatty acids that had been added to every sample prior to analysis. For recalibration, a previously in-house developed Visual Basic script was employed. Briefly, an average mass spectrum was created from regions within each chromatogram that corresponded to the internal standards nonanoate, [U-¹³C]fumarate, [²H₇]-trans-cinnamate, undecanoate, [U-¹³C]glucose, and [U-¹³C]lactose (**Figure 5.1A**). The mass scale was then recalibrated according to the theoretical masses of fragment ions, protonated molecules and adducts that originated from the standards (**Figure 5.1B**). In total, ten masses over a mass range from *m/z* 138.093 to *m/z* 960.507 were used in chapter 8. The entire procedure is illustrated in **Figure 5.1**. Internal recalibration improved mass accuracy for cis-aconitate from 22.9 mDa before recalibration to 4.0 mDa after recalibration (**Figure 5.1C**). In chapter 6, mass spectra were recalibrated using solely the *m/z* values of the [M+H]⁺

ions of the odd-numbered, saturated fatty acid derivatives (C9-19), which are listed along with the corresponding retention times in **Table 12.2**.



An average mass spectrum is created from defined chromatogram areas



Recalibration is performed with a predefined mass list using ions of internal standards

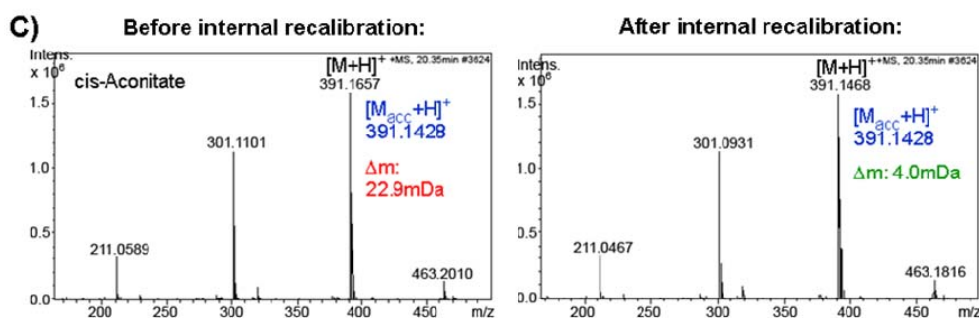


Figure 5.1 Internal mass recalibration of GC-APCI-TOFMS data. An average mass spectrum is generated from defined areas of the chromatogram containing internal standards (A). The average spectrum is used to recalibrate the complete run (B). The improvement in mass accuracy is shown for cis-aconitate in (C). M_{acc} , (calculated) accurate mass of analyte M. Reprinted from [15] with permission from Elsevier.

5.4.4 Calculation of elemental formulas

For the identification of features that were significantly differently regulated between sample groups (chapter 6) or to annotate features extracted from cell culture supernatants (chapter 8), the respective protonated molecule of a recalibrated data file was queried using the SmartFormula tool included in DataAnalysis [9,127]. Parameters for calculating elemental formulas were as follows for MeOx-TMS and MCF derivatives, respectively: positive adducts, $M+H$; $\Delta m \leq \pm 5$ mDa; charge, 1; even electron number; filter H/C element ratio ≤ 3 ; and check ring plus double bonds from -0.5 to 40. Apart from C, H, N, O atoms, which are considered automatically by the algorithm, P (0-3), S (0-3) and Si (at least 1) atoms were included for calculation in the case of MeOx-TMS derivatives, whereas the following restrictions were applied for MCF derivatives: $0 \leq n_N \leq 5$; $0 \leq n_P \leq 1$; $0 \leq n_S \leq 2$. Calculated formulas with an mSigma value below 50 were further considered. From the calculated elemental formula, the proton and possible groups introduced by derivatization had to be eliminated, such as trimethylsilyl groups as well as groups introduced by methoximation in the case of MeOx-TMS derivatization. The sum formula of the native metabolite was then searched in the HMDB [23]. The overall workflow is exemplified in **Figure 5.2** for N-acetylaspartate [17]. Candidate metabolites that were obtained by database search were only kept if they possessed the functional groups for forming the derivatives that had been the basis for the identification process.

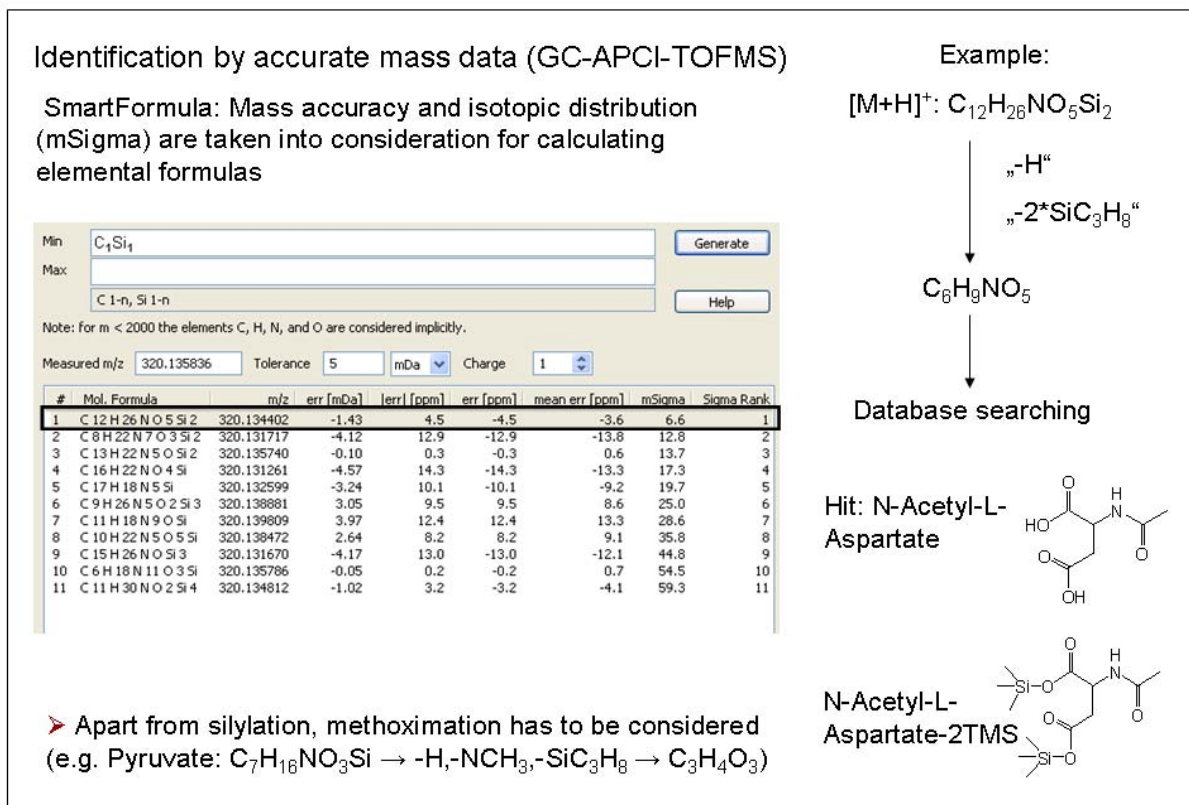


Figure 5.2 Scheme for initial steps towards identification of unknowns by GC-APCI-TOFMS. Reprinted from [15] with permission from Elsevier.

However, identification of unknowns required manual intervention. The corresponding spectrum had to be reviewed for proper selection of the protonated molecule used to calculate an elemental formula. Accidentally subjecting masses of fragments or adducts to the elemental formula calculator would have resulted in significant errors or an inconclusive database search. As exemplified for N-acetyl-neuraminate-1MeOx-6TMS in **Figure 5.3**, putative annotation via calculation of elemental formulas and subsequent database search initially failed (step 3 on the left-hand side), because feature extraction routine and statistical analysis (step 2) performed in comparative metabolic fingerprinting of wild type and double-mutant (PntAB-UdhA) *E. coli* strains [17] pointed to a fragment mass instead of the mass of the protonated molecule. Manual review of the corresponding mass spectra finally led to positive identification (step 3 on the right-hand side).

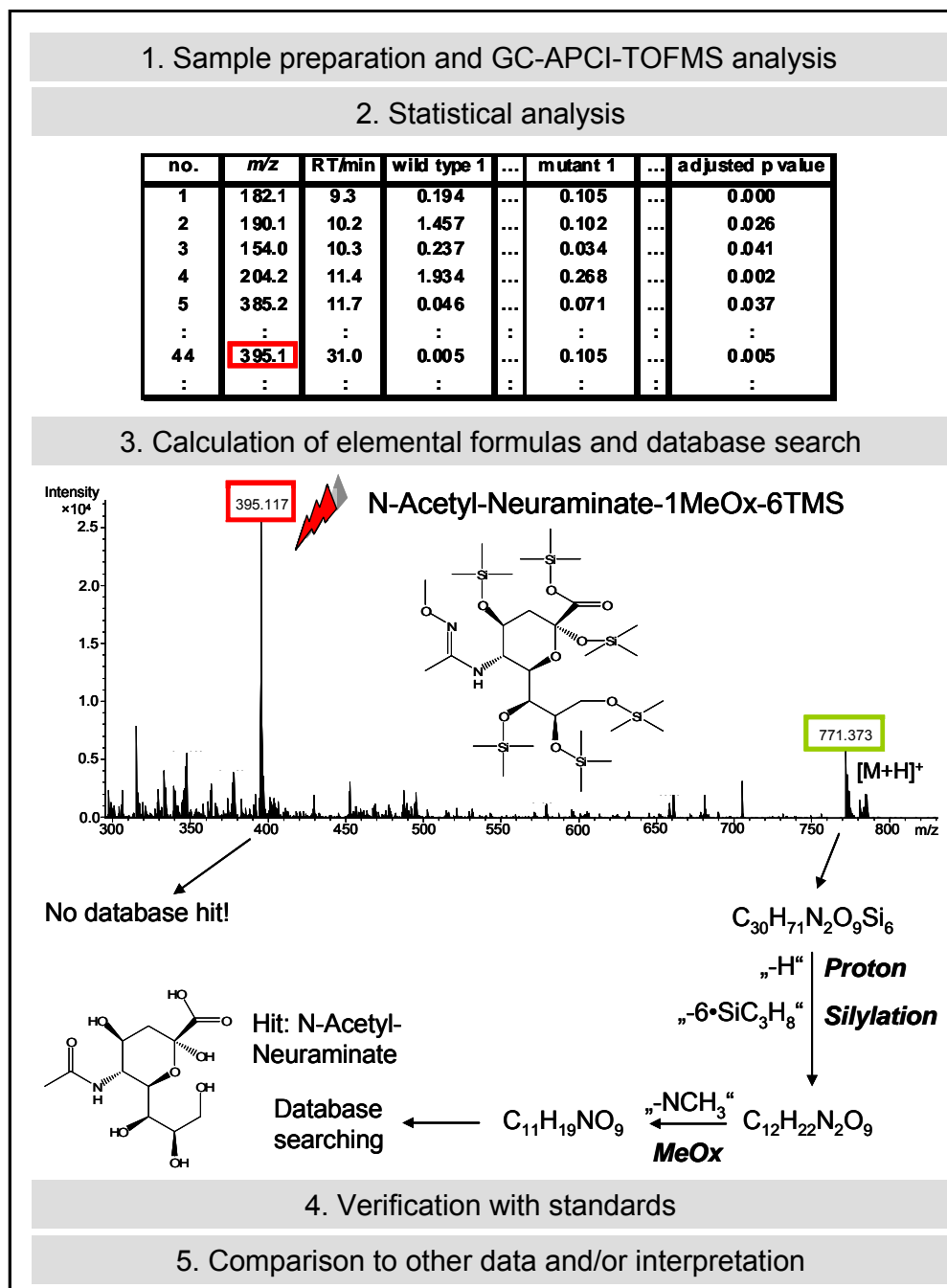


Figure 5.3 Manual review of GC-APCI-TOFMS mass spectra in the identification procedure of unknown features distinguishing wt and mutant (PntAB-UdhA) *E. coli* strains [17]. The example of N-acetyl-neuramate-1MeOx-6TMS is given, as it emphasizes the need for manual intervention. Annotation based on the accurate mass of a fragment ion that had been extracted by the feature extraction routine yielded initially no database match; manual selection of the protonated molecule finally led to positive identification. [16] – Reproduced by permission of The Royal Society of Chemistry.

Commercial standards were run on GC-APCI-TOFMS For tentatively identified metabolites, and mass spectra and linear retention indices were compared for final confirma-

tion, which required that a relative difference of 1.0% established by Strehmel et al. [128] was not exceeded.

5.5 Validation methods

5.5.1 ROC curve

Receiver operator characteristic (ROC) curves were plotted in chapter 6 to assess the effect of water infusion on the ability of GC-APCI-TOFMS to detect concentration fold changes of metabolite standards that were spiked into extracts of human serum and derivatized afterwards. Initially, features were extracted and aligned into a single data matrix for each spike-in data set, e.g. all MCF-derivatized spike-in samples that were acquired by GC-APCI-TOFMS in the absence of water infusion. Features were assigned as true positives (TPs) or true negatives (TNs). Then the different spike-in levels were compared by t tests between all features, and the latter were listed in ascending order of P values. Features corresponding to spike-in compounds should appear at the top of the list. After each feature, starting from top to bottom, true positive rate (TPR or sensitivity) and, similarly, false positive rate (FPR or 1-specificity) were determined and plotted against each other in receiver operating characteristic (ROC) curves. The true positive rate is also referred to as sensitivity, i.e., the number of detected TPs up to a respective row compared to total number of TPs. In the ideal case, i.e., a perfect discrimination of two different abundance levels, all features corresponding to spike-in compounds are ranked higher than signals from matrix compounds, and an area under the curve (AUC) value of 1 is obtained. **Figure 5.4** depicts three ROC curves as an example, which were obtained from six spike-in compounds and six matrix compounds that were compared between two arbitrary abundance levels in each case. Discrimination of the two levels gets worse from (A) to (C) as indicated by a decreasing AUC value (**Figure 5.4**).

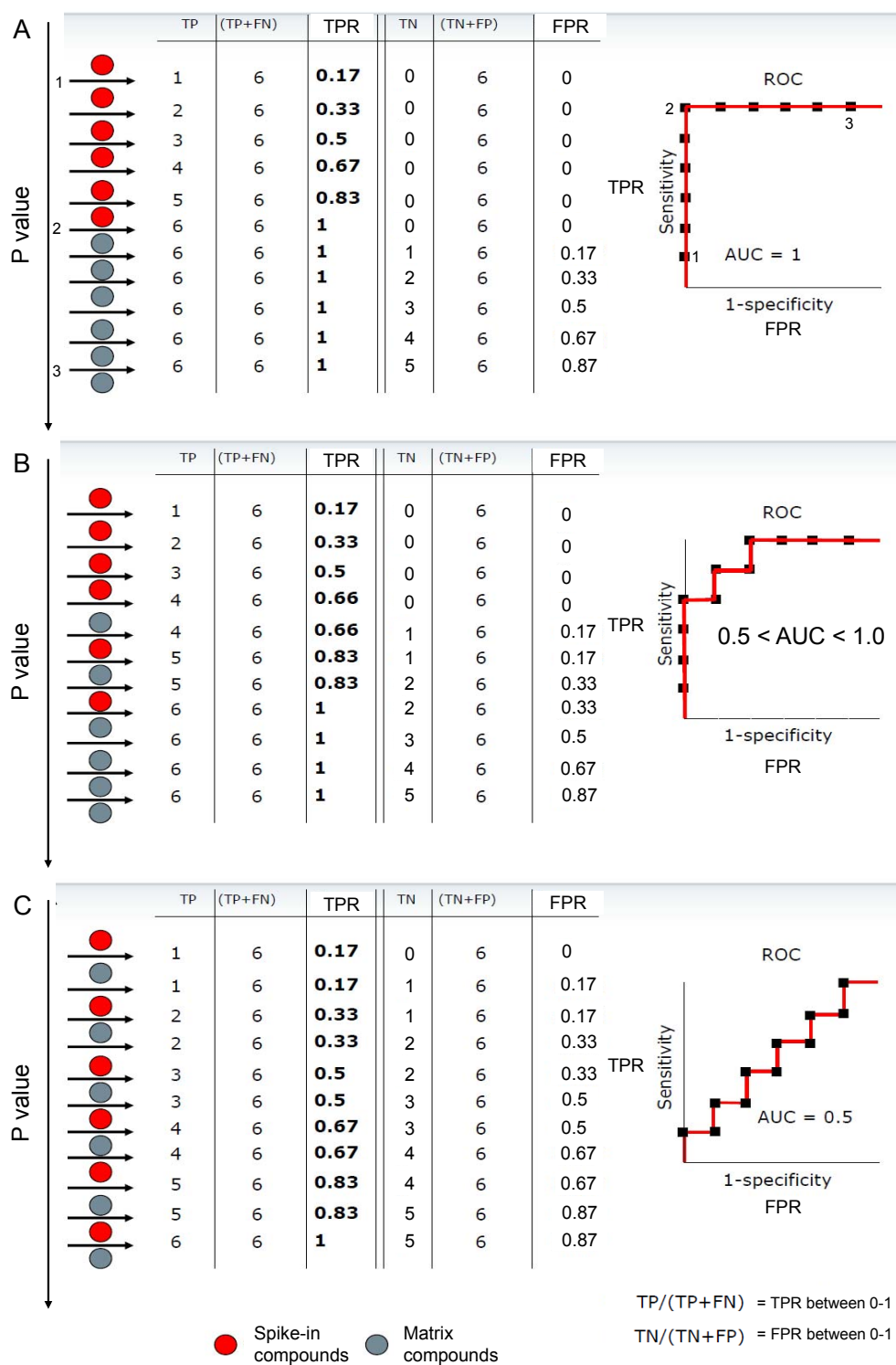


Figure 5.4 Example of three ROC curves based on six spike-in compounds and six matrix compounds that were compared between two arbitrary abundance levels in each case. AUC values decrease from (A) to (C) indicating that discrimination of the two levels gets worse. In (A) three data points of the curve are assigned to the underlying data matrix. FN, false negative; FP, false positive. Adapted and reprinted with permission from Bruker Daltonics.

5.5.2 Bland-Altman plot

A Bland-Altman plot is a method of data plotting to determine agreement between two different analytical approaches. It can be also used to assess repeatability of a single analytical method or to compare measurements by two operators. It was introduced by Bland and Altman in the 1980s [129,130] as a better alternative to correlation and displays the difference in output variable c obtained by the two approaches A and B ($y = c_A - c_B$) against the corresponding average ($x = (c_A + c_B)/2$) for each sample. To allow a quick visual inspection, mean difference (\bar{d}) as well as upper and lower limits of agreement ($\bar{d} \pm 1.96 \cdot SD$) are represented as horizontal lines in the graph. The latter limits refer to the 95% confidence interval of the differences assuming normal distribution, indicating how far measurements by the two approaches were apart for most samples. According to Kaspar et al. [126], six different types of Bland-Altman plots can be distinguished based on absolute and relative mean differences as well as whether differences scatter randomly or proportionally with x (relative mean difference corresponds to the mean difference divided by the averaged mean of all couples and multiplied by 100). Ideally, absolute and relative mean differences are almost zero and $\leq 15\%$, respectively, and individual differences scatter randomly (type A, compare [126]).

6 Investigations on the effects of continuous water infusion on APCI of derivatized metabolites

6.1 Introduction

It has been recognized as early as 1976 that water promotes the formation of protonated molecules in APCI [45]. Notable improvements in peak abundance and selectivity achieved by the infusion of water into the APCI source have been demonstrated recently for gas chromatography-tandem mass spectrometry of underivatized pyrethroid insecticides [82] and GC-APCI-TOFMS analysis of dissolved standards of methyl stearate, benzophenone, methyl palmitate, and cocaine [131]. Whether such improvements also hold true for derivatives of metabolites, and, moreover, are associated with better reproducibility of APCI were the objectives of the present study. To that end, water was continuously infused from the top into the APCI I source and the effects of water infusion on APCI-TOF mass spectra of MCF and MeOx-TMS derivatives of 20 standard compounds and the ability of GC-APCI-TOFMS to detect differences in metabolite concentration levels in a human serum extract were assessed. Subsequently, MCF derivatization and GC-APCI-TOFMS were applied to the comparative metabolic fingerprinting of extracts of pancreatic cancer cells that had been cultivated in the absence and presence, respectively, of the Hsp90 inhibitor 17-DMAG.

This chapter was published in [14].

6.2 Materials and methods

6.2.1 Materials

For each standard compound an individual stock solution was prepared in either methanol, water, or a mixture thereof, except for the even-numbered (C12-C18), saturated straight-chain fatty acids, which were dissolved in propanol. The compounds from the fatty acid mixture (C9-19, compare subchapter 5.1) were used as retention index markers and for recalibration of the mass scale according to the procedure described in section 5.4.3. The spike-in mixture (**Table 6.1**) contained 20 compounds at a concentration of 1 mM each in MeOH.

6.2.2 Sample preparation

For the spike-in experiment aliquots of 10 μL of human serum were extracted with 50 μL of MeOH, according to an established protocol in our laboratory [132]. Extraction standards, were not added. Pellets were re-extracted twice more with 50 μL of methanol. Extracts were pooled and then split into aliquots for the preparation of spike-in samples. Amounts of 960 and 160 μL each were taken from the pool for MCF and MeOx-TMS derivatization, respectively, which corresponded to 60 and 10 μL of extracted serum. Since the final volume after MCF derivatization (300 μL) is three times as high as that after MeOx-TMS derivatization (100 μL), a three times higher extract volume was needed assuming that efficiency of both derivatization strategies was comparable. In addition, the volume was doubled to obtain two times more concentrated MCF-derivatized serum compounds. A two-fold higher concentration for MCF-derivatized serum compounds, which was also used for MCF-derivatized spike-in compounds, was required to detect a sufficient number of MCF -derivatized spike-in and matrix compounds in the absence of water infusion into the APCI source.

Six different spike-in levels were prepared (N=5 derivatization replicates), ranging in final concentrations of the spike-in compounds from 133 μM to 200 μM , 267 μM , 333 μM , 400 μM , and 467 μM , respectively, for MCF derivatization, and 67 μM to 100 μM , 133 μM , 167 μM , 200 μM , and 233 μM , respectively, for MeOx-TMS derivatization.

Prior to derivatization, on the basis of the desired final standard concentration, the respective volume of standard stock solution, spike-in mixture, or an appropriate dilution was transferred either to a 2-mL glass vial together with 20 μ L of fatty acid mixture in case of MCF derivatization, or to a 2-mL glass vial with a 400- μ L glass insert for MeOx-TMS derivatization. For example, 120 μ L and 20 μ L, respectively, of a 0.5 mM standard stock solution were taken for preparation of 200 μ M MCF- and 100 μ M MeOx-TMS-derivatized standard samples in initial experiments and for optimizing the water infusion rate. Depending on the sample set, cell extract or an aliquot of the serum extract pool was added as well, and the entire sample was then dried using a vacuum evaporator followed by either MCF or MeOx-TMS derivatization according to sections 5.2.2 and 5.2.3. In the case of cell extracts and MCF derivatization, 250 μ L of the chloroform phase were carefully dried and reconstituted in 50 μ L of chloroform.

6.2.3 GC-APCI-TOFMS analysis

See subchapter 5.3. Spike-in samples along with blanks and cell extracts were measured in random order to avoid systematic error. A blank sample was always run between changes in water flow rates for equilibration of APCI source conditions.

6.2.4 Cross-validation of amino acids

Targeted quantification of amino acids was performed by high-performance liquid chromatography – electrospray ionization – tandem mass spectrometry (HPLC-ESI-MS/MS) in positive ionization mode as previously described [133]. Briefly, 10 μ L of aqueous cell extract together with 10 μ L of SIL-IS mix in water were subjected to propyl chloroformate (PCF) derivatization. Separation of PCF derivatives was carried out on a Phenomenex EZ:faast 4u AAA-MS (250 \times 3 mm i.d., 4 μ m, Torrence, CA, USA) reversed-phase column using gradient elution with 10 mM ammonium formate and 0.1% (v/v) heptafluorobutyrate in water and methanol. One selective transition each for the unlabeled and labeled amino acid was programmed for operating the MS/MS in MRM mode. For quantification, calibration curves were obtained from standards by plotting nominal concentration ratio (analyte/internal standard) against peak area ratio (analyte/internal standard), whereby the corresponding labeled analogue was included for each target

analyte as internal standard, with the exception of ornithine, for which labeled methionine was used. Acquisition and analysis of data was carried out employing Analyst version 1.5.1 from AB Sciex (Darmstadt, Germany).

6.2.5 Feature extraction and alignment

For the analysis of data from the spike-in experiment, optimized Find Molecular Features (FMF) algorithm and bucketing parameters for feature extraction and alignment within ProfileAnalysis were as follows for MCF/ MeOx-TMS derivatives, with deviating parameters for MeOx-TMS derivatives in brackets: S/N threshold, 2 (5); correlation coefficient threshold, 0.6 (0.7); minimum compound length, 8 (10); smoothing width, 2 (3); additional smoothing, enabled; proteomics, enabled; bucketing basis, M+H. Retention time range, 8-25 min (8-17 min); mass range, m/z 100-250 (m/z 100-500); advanced bucketing tolerance parameters, 0.1 min (0.4 min) and 10 mDa for retention time and mass, respectively; split buckets with multiple compounds, disabled; value count of group attribute (spike level) within bucket as bucket filter, enabled at least three; and other parameters, always none or disabled.

Each of the four different bucket tables (0.0 mL/h H₂O/MCF, 0.4 mL/h H₂O/MCF, 0.0 mL/h H₂O/ MeOx-TMS, and 0.4 mL/h H₂O/ MeOx-TMS) listed features in rows with accompanying information on retention time and m/z value as well as the feature intensities over all spike-in samples from entire spike-in levels. To reduce the number of zero values within the bucket table, bucket filtering was applied: at least three out of five area integrals had to be available for each spike-in level; otherwise the respective feature was excluded. Further data analysis proceeded in Excel as described in 6.3.

For metabolite fingerprinting in cell extracts, individual data matrices were generated from GC-APCI/+H₂O-TOFMS and GC-APCI/-H₂O-TOFMS measurements of MCF-derivatized cell extract samples of control and 17-DMAG -treated groups, applying with a few exceptions the same FMF and bucketing parameters as for the MCF spike-in experiment. Exceptions were as follows: the mass range for bucketing was extended to m/z 100-300, advanced bucketing tolerance parameter for retention time was increased to 0.15 min, and bucket filtering criteria allowed only the inclusion of features into the

bucket table, for which at least 3/5 area integrals were available in at least one group. Further data analysis was carried out in Excel as described in 6.3.

6.3 Results and discussion

6.3.1 Effects of continuous water infusion on APCI mass spectra of MCF and MeOx-TMS derivatives

Initial experiments on the effect of continuous water infusion on APCI mass spectra of MCF and MeOx-TMS derivatives were performed at an infusion rate of 0.4 mL/h. **Table 12.2** lists the most prominent ions observed in APCI/+H₂O mass spectra of MCF and MeOx-TMS derivatives of 20 standard compounds as compared to spectra acquired without water infusion (APCI/-H₂O). Compounds included in the standard mixture were selected as to cover a variety of chemical classes, including organic acids, amino acids, and analytes containing amide and thiol functional groups, according to the three performance classes introduced by Koek et al. [31] for silylated derivatives.

APCI/-H₂O mass spectra of MCF derivatives were typically dominated by fragment ions. The [M+H]⁺ ion constituted only a minor component with a relative intensity of <30% for 12 of the 20 investigated compounds. Notably, the six MCF derivatives that exhibited [M+H]⁺ as the base peak were, aside from malonate and fumarate, the four aromatic compounds benzoate, hippurate, nicotinate, and phenylacetate, indicating that MCF-derivatized aromatic compounds might be less prone to in-source fragmentation. In contrast, APCI/+H₂O mass spectra exhibited dominant [M+H]⁺ ions for all MCF derivatives. Major fragment ions of MCF derivatives could be attributed to neutral losses including CH₃OH (32.026 Da), [CH₃OH + CO] (60.021 Da), and combinations thereof. Besides, [2M+H]⁺-type adduct ions (relative intensity <5%) for several amino acids and [M+H+H₂O]⁺-type adduct ions for fumarate and the fatty acids were detected. The strongest effect of water infusion on suppression of in-source fragmentation and, consequently, the formation of abundant [M+H]⁺ ions was noticed for dicarboxylic acids, as shown exemplarily for suberate in **Figure 6.1**. The APCI/-H₂O mass spectrum of suber-

ate was dominated by the $[M+H-CH_3OH]^+$ fragment ion, while the relative intensity of the $[M+H]^+$ ion amounted only to 8%. Upon water infusion, $[M+H]^+$ became the dominant ion species. Hence, APCI/+H₂O provides softer ionization than conventional APCI, possibly due to an increased formation of $[H(H_2O)_n]^+$ ion clusters of higher order, which exhibit a higher proton affinity [134,135]. However, in analogy to the use of two different reagent gases in GC-Cl-MS metabolite profiling [136], APCI/-H₂O and APCI/+H₂O may complement each other in the identification of unknowns. While APCI/+H₂O is more likely to generate $[M+H]^+$ ions of sufficient intensity for the calculation of sum formulas, fragment rich APCI/-H₂O mass spectra will provide additional structural information and, thus, narrow the list of sum formulas generated from accurate mass measurements. Other than that, APCI/+H₂O is more suited for quantification and the $[M+H]^+$ ions may also serve as precursor ions for MS/MS analyses. In fact, protonated molecules have been preferred over EI fragment ions in terms of abundance and selectivity in previous metabolite profiling studies [82,86].

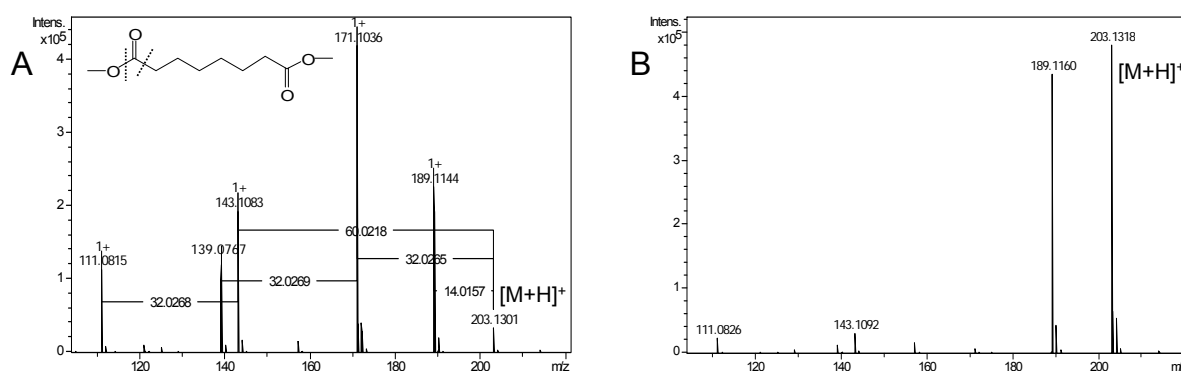


Figure 6.1 GC-APCI-TOFMS mass spectra of the MCF derivative of suberate (conc = 200 μ M) acquired without (A) and with (B) infusion of water at 0.4 mL/h. Neutral losses yield a series of fragments in the case of APCI/-H₂O as annotated in the left spectrum; $[M+H]^+$ abundance is distinctly promoted by water infusion. Reprinted from [14].

Water infusion, interestingly, exerted almost no effect on MeOx-TMS derivatives, i.e., the $[M+H]^+$ ions dominated for 19/20 compounds even in the APCI/-H₂O mass spectra. Typical fragment ions were caused by neutral loss of C₃H₈Si [TMS] (72.040), C₃H₉SiOH [TMSOH] (90.050 Da), [TMSOH + CO] (118.045 Da), and combinations thereof. Overall, MeOx-TMS mass spectra were less rich in fragment ions compared to their MCF analogues (compare **Figure 6.1**, **Figure 12.1** and **Figure 12.2** for suberate and valine,

respectively). For some compounds, including benzoate, fumarate, and phenylacetate, a $[M+H+TMS]^+$ -type adduct was observed, as recently reported [78].

6.3.2 Optimization of water infusion rate

Next the effects of different water infusion rates on the ionization behavior of both standard mixtures of 20 MCF- and MeOx-TMS -derivatized compounds were investigated, in 0.1 mL/h increments up to 0.5 mL/h. Details are presented in **Table 12.3**. The geometric means of fold changes in $[M+H]^+$ peak area relative to no water infusion differed significantly (ANOVA, $p < 0.05$) for MCF-derivatized spike-in standards (N=20) across water infusion rates of 0.1-0.5 mL/h. On average, peak areas increased the most for organic acids with increasing rates of water infusion, followed by amino acids, whereas (hetero) aromatic compounds yielded the smallest increases. In contrast, geometric means of RSD values were fairly similar across the water infusion rates tested, ranging from 16.3 % to 23.2%. Ultimately, a water infusion rate of 0.4 mL/h was chosen for all subsequent experiments, as it had yielded overall the highest fold change in $[M+H]^+$ peak area (16.6-fold) and the lowest RSD (16.3 %), and $[M+H]^+$ constituted the base peak for all 20 MCF derivatives. **Figure 12.2** shows mass spectra of MCF-derivatized suberate and valine acquired without and at five different rates of water infusion. The standard mixture was also spiked at 200 μ M into a serum extract and analyzed without and with 0.4-mL/h water infusion, followed by normalization of the $[M+H]^+$ area values against undecanoate. Relative standard deviations improved significantly at 0.4 mL/h water infusion (**Table 6.1**), e.g. from 25.3% to 11.3% for isoleucine. In addition, an average 28-fold higher S/N ratio was observed for the $[M+H]^+$ ions.

On the other hand, in the case of MeOx-TMS derivatives only 6 of the 20 standard components were significantly affected over the different water infusion rates (ANOVA, $p < 0.05$). Significant increases in fold changes in $[M+H]^+$ peak area relative to no water infusion were limited to (hetero) aromatic compounds and adipate only. Some compounds were even negatively affected as evidenced by a fold change < 1 (**Table 12.3**). Overall, geometric means for fold-changes and RSDs did not differ markedly across the

water-infusion rates tested. Compared to MCF derivatives, RSDs for MeOx-TMS derivatives were generally higher.

Table 6.1 Impact of 0.4 mL/h water infusion on APCI of MCF derivatized spiked serum extracts. Mean fold changes in peak area and S/N ratio of $[M+H]^+$ ions of the 20 standard compounds were obtained by comparing 0.4 mL/h water infusion to no water infusion. In addition, RSD values and average S/N ratio of $[M+H]^+$ ions for APCI/-H₂O and APCI/+H₂O are given. For calculation of FCs and RSDs, peak areas were normalized to undecanoate. Water infusion increased $[M+H]^+$ peak areas on average by 11-fold and significantly decreased RSD values. Standard concn = 200 μ M; N=5 derivatization replicates; A=amide group; E=methyl ester group. Reprinted from [14].

MCF derivative	FC peak area 0.4/0.0 mL/h	RSD 0.0 mL/h (%)	RSD 0.4 mL/h (%)	avg. S/N $[M+H]^+$ 0.0 mL/h	avg. S/N $[M+H]^+$ 0.4 mL/h	FC S/N $[M+H]^+$ 0.4/0.0 mL/h
Adipate-2E	166.9	34.6	12.0	20	2955	147.7
Ala-1A-1E	27.7	36.5	26.6	84	2257	26.9
4-Aminobutyrate-1A-1E	^a	^a	7.3	^a	259	^a
Asp-1A-2E	10.1	22.0	9.3	257	2408	9.4
Benzoate-1E	1.3	33.9	11.2	2145	3370	1.6
Fumarate-2E	0.6	38.8	13.3	1154	1041	0.9
Glu-1A-2E	9.6	45.6	5.1	104	1494	14.4
Hippurate-1E	5.1	23.6	15.3	713	5052	7.1
Ile-1A-1E	14.5	25.3	11.3	290	4362	15.0
Leu-1A-1E	11.0	19.8	16.5	416	5384	12.9
Malonate-2E	44.0	17.1	10.6	28	1767	63.1
Met-1A-1E	18.1	46.0	11.9	174	3892	22.4
Methyl-malonate-2E	66.0	23.0	12.6	30	2387	80.1
Nicotinate-1E	1.1	35.6	20.9	1170	1483	1.3
Nval-1A-1E	18.6	32.5	24.3	132	2985	22.7
Phe-1A-1E	9.3	45.9	9.8	367	5314	14.5
Phenylacetate-1E	5.1	17.3	12.8	467	3161	6.8
Pro-1A-1E	8.4	28.5	10.7	519	4606	8.9
Suberate-2E	62.6	25.5	4.9	97	6868	71.0
Val-1A-1E	17.3	24.0	21.4	254	3005	11.8

^a S/N of $[M+H]^+$ of 4-Aminobutyrate below 20 in case of no water infusion.

Ionization of analytes is not only due to direct protonation but also ligand switching [69], i.e., association of analytes with $[H(H_2O)_n]^+$ water ion clusters followed by dissociation of water molecules. As evident from **Table 12.2**, log P values are significantly lower for MCF derivatives compared to their MeOx-TMS analogues, indicating that MCF derivatives are more polar. This might facilitate the formation of MCF derivative-water ion

clusters and explain the observed increase in signal intensities upon infusion of water into the APCI source.

6.3.3 Method evaluation via spike-in experiment

Metabolic fingerprinting sets high requirements on quantitative precision as well as extraction and alignment of features. To test whether APCI/+H₂O improves the detection of differences in metabolite concentration of MCF and MeOx-TMS derivatives, a serum extract was spiked with the 20 standard compounds at six different concentration levels (N=5 derivatization replicates) ranging from 133-467 μ M and 67-233 μ M, respectively, thus generating a total of 15 concentration fold changes (FCs) between 1.17- and 3.5-fold. A 2-fold higher concentration for MCF spike-in compounds and volume of serum extract was chosen in comparison to the MeOx-TMS spike-in experiment to extract features for an increasing number of low-abundant MCF -derivatized spike-in compounds in the case of APCI/-H₂O. Four data matrices were obtained in total, and features were assigned as true positives (TPs) or true negatives (TNs), by comparing serum blank and individual standard samples. All multiple features of a single spike-in compound, e.g., protonated molecule, fragments, adducts, or different derivatives, were further considered as true positives, and ROC curves were obtained as described in detail in section 5.5.1. For MCF derivatives, 91 (53 TPs, 38 TNs) and 178 (97 TPs, 81 TNs) features were obtained for APCI/-H₂O and APCI/+H₂O data matrices, respectively, with a maximum of 3% zero values after exclusion of one run each due to technical problems. **Figure 6.2** shows how water infusion influences accuracy of detection of differential features by presenting ROC curves for 6 of the 15 possible FCs. AUC values were noticeably improved in the case of APCI/+H₂O (compare **Figure 6.2A,B** and **Table 12.4**). Furthermore, the AUC values of APCI/-H₂O and APCI/+H₂O for all 15 possible FCs were plotted against each other (**Figure 6.3A**). Higher AUC values for APCI/+H₂O were found in almost all cases. Interestingly, the AUC value for an FC of 1.2 (spike-in levels 4 vs 5) in the case of APCI/-H₂O distinctly exceeded the corresponding value for APCI/+H₂O. Upon pairwise leaving out one of the five replicates for spike-in levels 4 and 5 (25 combinations in total), a mean AUC value \pm SD of 0.66 \pm 0.08 and 0.47 \pm 0.05 was obtained for APCI/-H₂O and APCI/+H₂O, respectively. These greater relative standard

deviations implied that AUC values for an FC of 1.2 were less stable compared to a higher FC of 1.67 (spike-in levels 3 vs 5), for which AUC values of 0.76 ± 0.04 (APCI/-H₂O) and 0.84 ± 0.01 (APCI/+H₂O) were obtained. Because lower FCs are more prone to error due to measurement fluctuations, the FC of 1.2 was regarded as random deviation. Hence, only AUC values >0.7 were considered trustworthy. Using this cutoff, APCI/+H₂O proved more sensitive in the detection of discriminating features, as it succeeded in distinguishing an FC of 1.33 from the background, whereas AUC values exceeding constantly 0.7 were obtained for FCs of 1.67 and higher in the case of APCI/-H₂O (**Table 12.4**). In addition to that, no $[M+H]^+$ ions were included in the data matrix in the case of APCI/-H₂O for 11 spike-in compounds and no features at all were extracted for four analytes, which might complicate or entirely prevent the detection of discriminating MCF -derivatized metabolites without water infusion.

While assigning features as TPs or TNs, the number of TNs was distinctly higher than that of TPs and therefore the retention time range for generation of the bucket table was reduced to 8-17 min in the MeOx-TMS spike-in experiment to decrease the number of TNs. Otherwise, in the case of an enormous excess of TNs over TPs, small P values for many TNs would have occurred by chance and these TNs appeared at random on top of the sorted list, which would have prevented similar fold changes from being distinguished by means of their ROC curves. Comparable AUC values for APCI/-H₂O and APCI/+H₂O were obtained in the case of MeOx-TMS derivatives. This was in accordance with the observed lack of any influence of water infusion on RSD values. **Table 12.4** lists AUC values for both APCI/-H₂O and APCI/+H₂O of MeOx-TMS derivatives, and the respective AUC vs AUC plot is depicted in **Figure 6.3B**. It was concluded that water infusion held no advantage for APCI-MS of MeOx-TMS derivatives in metabolite fingerprinting.

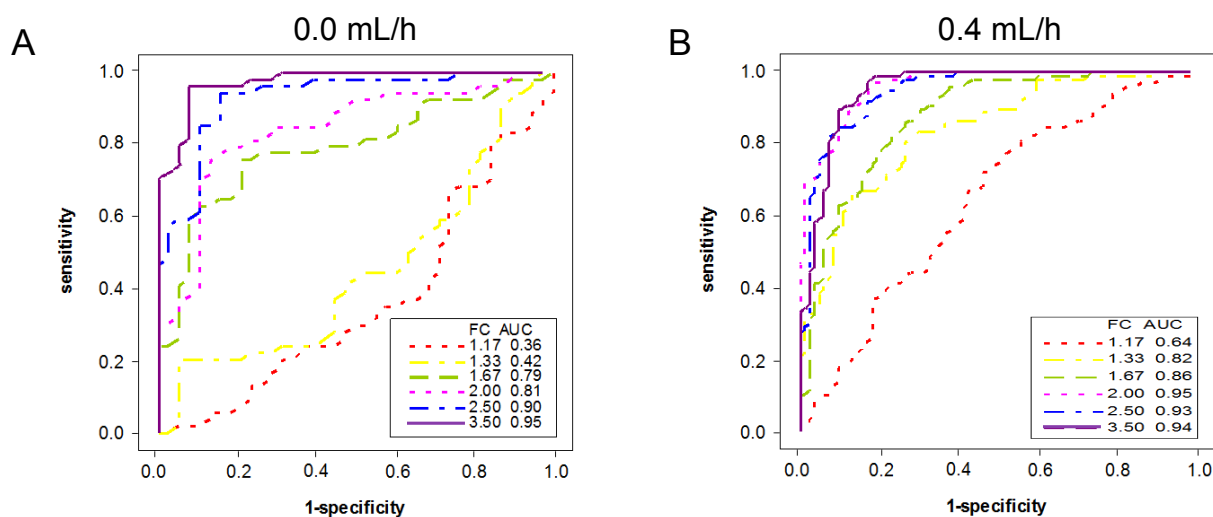


Figure 6.2 Impact of water infusion on the detection of concentration differences of MCF -derivatized metabolites illustrated by receiver operating characteristic (ROC) curves. Twenty known compounds were spiked at six different concentrations (133-467 μM) into aliquots of serum extract and analyzed by GC-APCI-TOFMS without (A) and with water infusion (B). Water infusion increased area under the curve (AUC) values. Hence, it improved the ability to distinguish genuine changes in metabolite concentration from background signals of the serum matrix. Reprinted from [14].

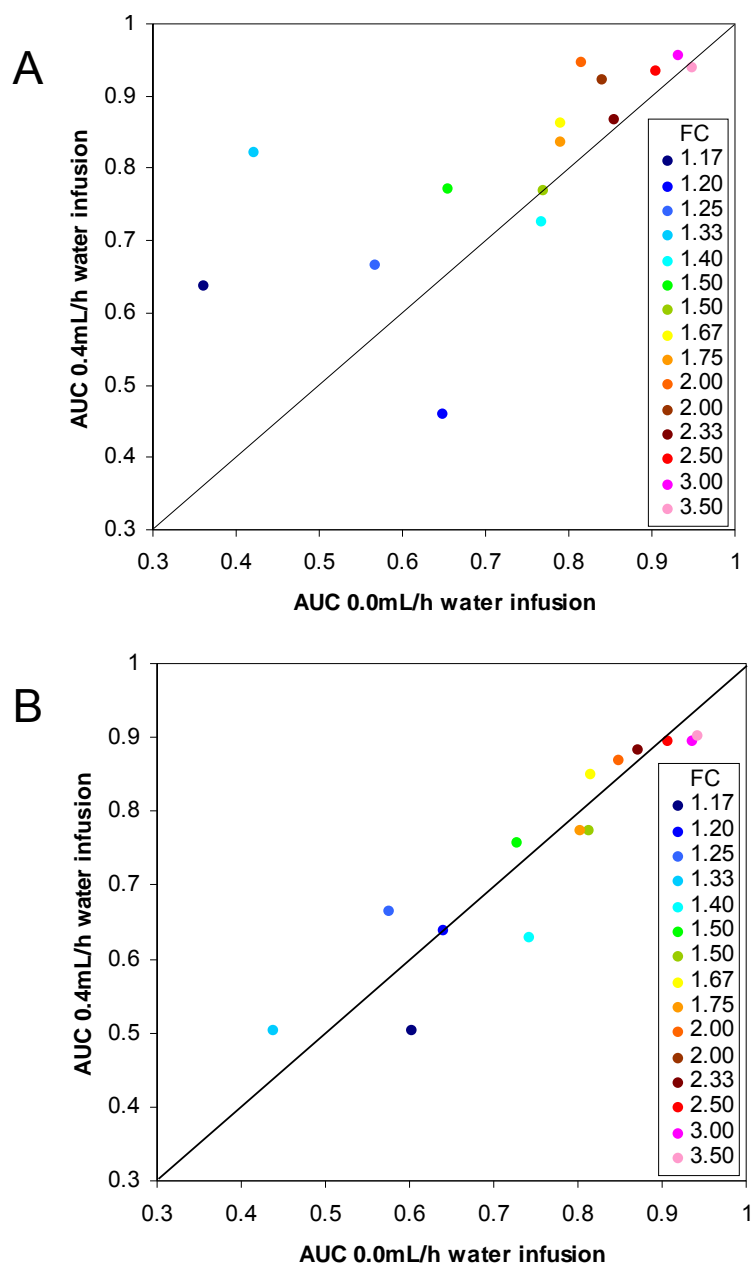


Figure 6.3 Impact of water infusion on the detection of differential MCF and MeOx-TMS derivatized features illustrated by area under the curve comparison plot. AUC values from the corresponding ROC curves for all pairwise fold changes are compared for MCF (A) and MeOx-TMS (B) derivatives. ROC curves were generated from 20 compounds spiked at six different concentrations into aliquots of a serum extract, followed by derivatization (N=5 derivatization replicates per spike level), and GC-APCI-TOFMS analysis without and with water infusion at 0.4 mL/h. AUC values were overall increased by water infusion in the case of MCF derivatives, whereas those corresponding to MeOx-TMS derivatives were not significantly affected. Reprinted from [14].

6.3.4 Evaluation of fold changes and mass and isotope accuracy

Next the accuracy and repeatability of FCs detected with MCF derivatives under water infusion was studied. To that end, following subtraction of blank feature area integrals, all possible pairwise FCs from area integrals of preferentially the $[M+H]^+$ ion of a specific compound within the spike-in data matrix were plotted against their expected FCs, as shown exemplarily for benzoate in **Figure 6.4**. Then univariate linear regression analysis was performed. The respective slopes and intercepts of the regression equations, regression coefficients, and mean RSDs over all FCs are listed in **Table 12.5**. Upon water infusion, accuracy and repeatability of FCs were distinctly enhanced, with ranges of regression coefficients and mean RSDs improving from 0.350-0.945 (median value of 0.774) to 0.709-0.944 (0.883) and 19.5-67.2% (34.3%) to 15.9-29.9% (22.5%), respectively.

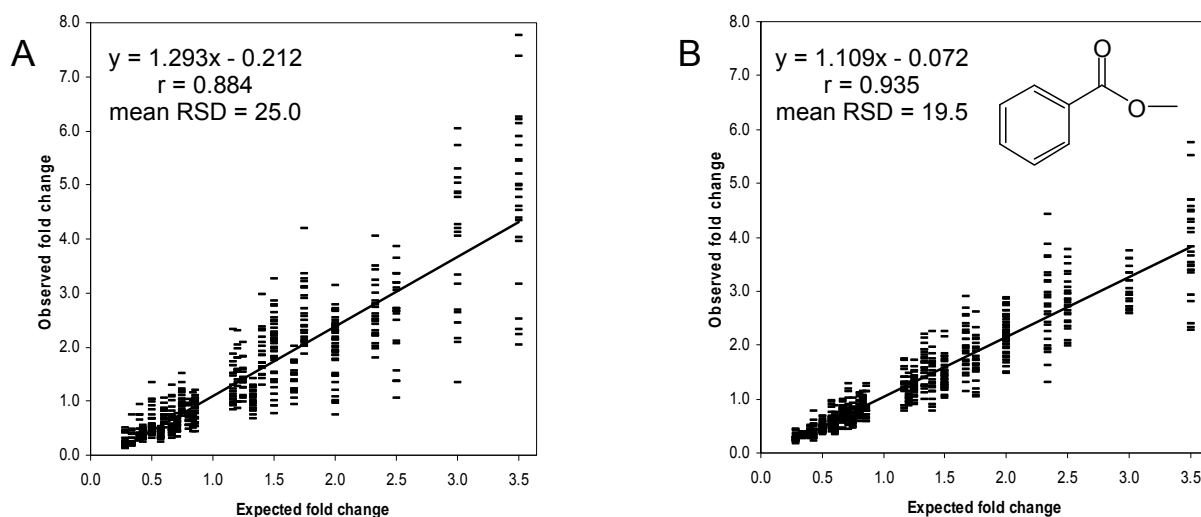


Figure 6.4 Linear dependency between expected and observed fold changes for the MCF derivative of benzoate. Blank feature area values had been subtracted as “background” from area integrals of $[M+H]^+$ of benzoate (m/z 137.0597 Da) for generation of FC plots derived from analyses of spiked serum extracts acquired at a water infusion rate of (A) 0.0 mL/h, and (B) 0.4 mL/h. All possible fold changes are plotted, and the respective linear regression equation, regression coefficient, and mean RSD value over all fold changes are given in the graphs. Accuracy and repeatability of FCs were distinctly improved by water infusion. Reprinted from [14].

Reliability of accurate mass measurements was verified by means of retrieving elemental formulas from known MCF-derivatized standards spiked into serum extracts at a concentration of 200 μ M each. Results are listed in **Table 12.6**. Upon internal recalibration excellent mass accuracy over a mass range of 133-238 m/z for all spike-in com-

pounds was obtained irrespective of water infusion, as confirmed by Wilcoxon signed-rank test for each spike-in compound except 4-aminobutyrate, with absolute mean mass errors (\pm SD) of 0.64 (\pm 0.72) mDa and 0.66 (\pm 0.68) mDa, respectively, for APCI/-H₂O and APCI/+H₂O at 0.4 mL/h. In general, the absolute mass error was below 3 mDa for all measured ions. In contrast, isotope patterns of [M+H]⁺ ions were markedly enhanced by water infusion because of the increased signal intensities (**Table 12.6**). In the case of APCI/-H₂O, mSigma values, which decrease with increasing goodness-of-fit between the theoretical and experimental isotope pattern, were inadequate, ranging from 33-80. Further, for more than half of the compounds, the first isotope of the [M+H]⁺ ion was not detectable. Upon water infusion, the mean mSigma value was 10.0 \pm 1.0. As a consequence, given the relevance of isotopic patterns for the exclusion of incorrect elemental combinations and the selection of the correct formula among the top-hits [50,95], average rank of correct formula within all calculated formulas was 2.7 and 1.9 for APCI/-H₂O and APCI/+H₂O, respectively. APCI/+H₂O yielded excellent isotope ratio accuracies for the first isotope with a mean error of 0.66 \pm 0.68%. According to our own results and those of previous studies [50,78] for MeOx-TMS derivatives that had been investigated by means of GC-APCI-QTOFMS [78] as well as GC-EI-TOFMS and GC-CI-TOFMS [50], there is a dependency of isotopic pattern accuracy on ion abundance in contrast to mass accuracy. Nevertheless, for both APCI/-H₂O and APCI/+H₂O, the correct elemental formula was always derived, except for 4-aminobutyrate whose [M+H]⁺ ion had not been detected by APCI/-H₂O. In conclusion, APCI/+H₂O yielded a number of notable improvements in the quantification and identification of MCF derivatives that might benefit metabolite fingerprinting.

6.3.5 Analysis of metabolic fingerprints of cancer cell extracts

As a proof of principle, MCF derivatization and GC-APCI-TOFMS were applied to metabolic fingerprinting of pancreatic cancer cells treated with 17-DMAG. The latter is an inhibitor for Hsp90, which acts as a molecular chaperone for many oncoproteins and thus has become a target in clinical trials. However, little is known about whether Hsp90 has a direct impact on cancer cell metabolism. As evidenced by the overlaid total ion current (TIC) chromatograms in **Figure 6.5**, ionization efficiency was distinctly en-

hanced by water infusion. Feature extraction and alignment yielded a total of 217 and 211 features for APCI/-H₂O and APCI/+H₂O, respectively. In the latter case, one would expect an increase in detectable [M+H]⁺ ions, which should benefit metabolite identifications. For instance, this was the case for malate, which eluted at 14.9 min in the chromatogram depicted in **Figure 6.5B** (peak no. 9). Twenty of the metabolites identified in extracts of untreated pancreatic cancer cells are represented in **Figure 6.6A**, with mean [M+H]⁺ areas ± SD given in the graph. With the exception of the fatty acid derivatives, which formed water adducts at the expense of [M+H]⁺ ions under water infusion, all metabolites showed an increase in mean [M+H]⁺ area. Furthermore, SDs were lower for most metabolites. Importantly, of the four metabolites that could solely be identified by water infusion (**Figure 6.6**), malate, serine, and amino adipate were found to show significant discrimination between control and 17-DMAG treated cell extracts.

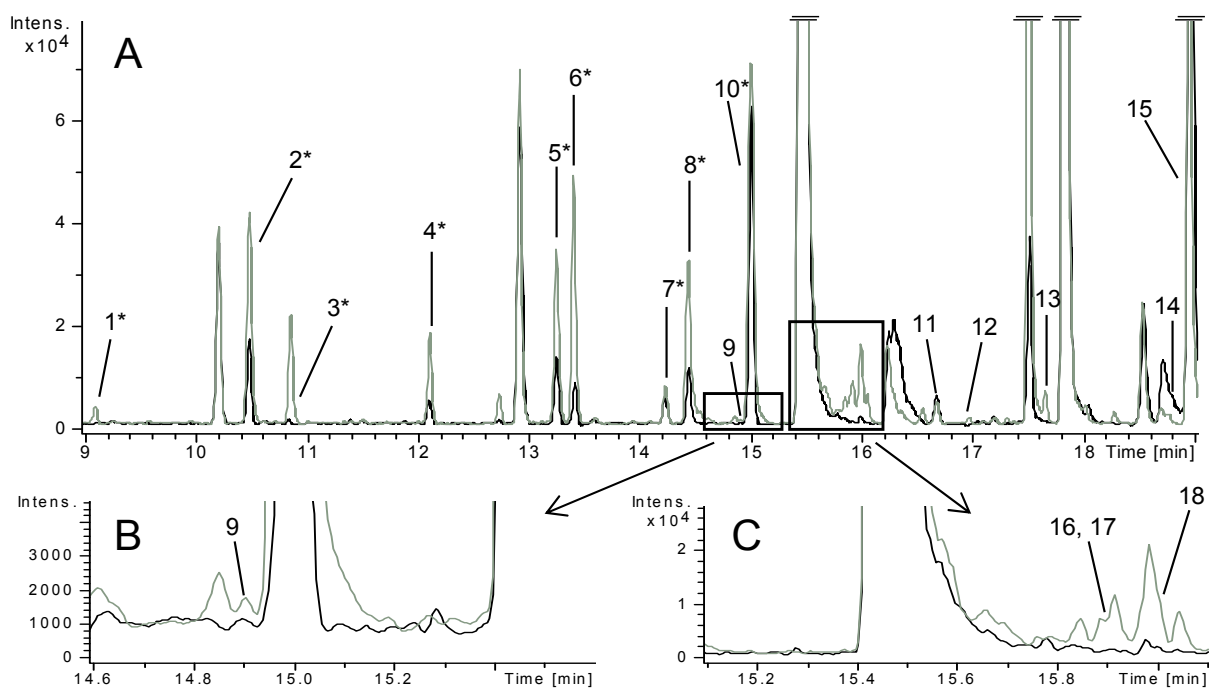


Figure 6.5 Impact of water infusion on ionization efficiency as exemplified by the overlay of 10 min sections of the TIC chromatograms (A) and enlarged sections deriving therefrom (B, C) acquired by GC-APCI-TOFMS without (chromatogram in black) and with (0.4 mL/h) water infusion (chromatogram in gray) for an MCF-derivatized metabolite extract of untreated pancreatic cancer cells. Metabolites, whose abundance was found by means of GC-APCI/+H₂O-TOFMS to be significantly different in 17-DMAG -treated cells versus controls, are annotated, with asterisks indicating metabolites found to be discriminating irrespective of water infusion. Experimental details are given in chapter 5. As a result, ionization efficiency was distinctly enhanced by water infusion. Peak identification: 1, lactate; 2, alanine; 3, glycine; 4, valine; 5, leucine; 6, isoleucine; 7, laurate; 8, threonine; 9, malate; 10, proline; 11, myristate; 12, serine; 13, methionine; 14, amino adipate; 15, palmitate; 16, aspartate; 17, citrate; 18, N-acetyl-L-aspartate. Reprinted from [14].

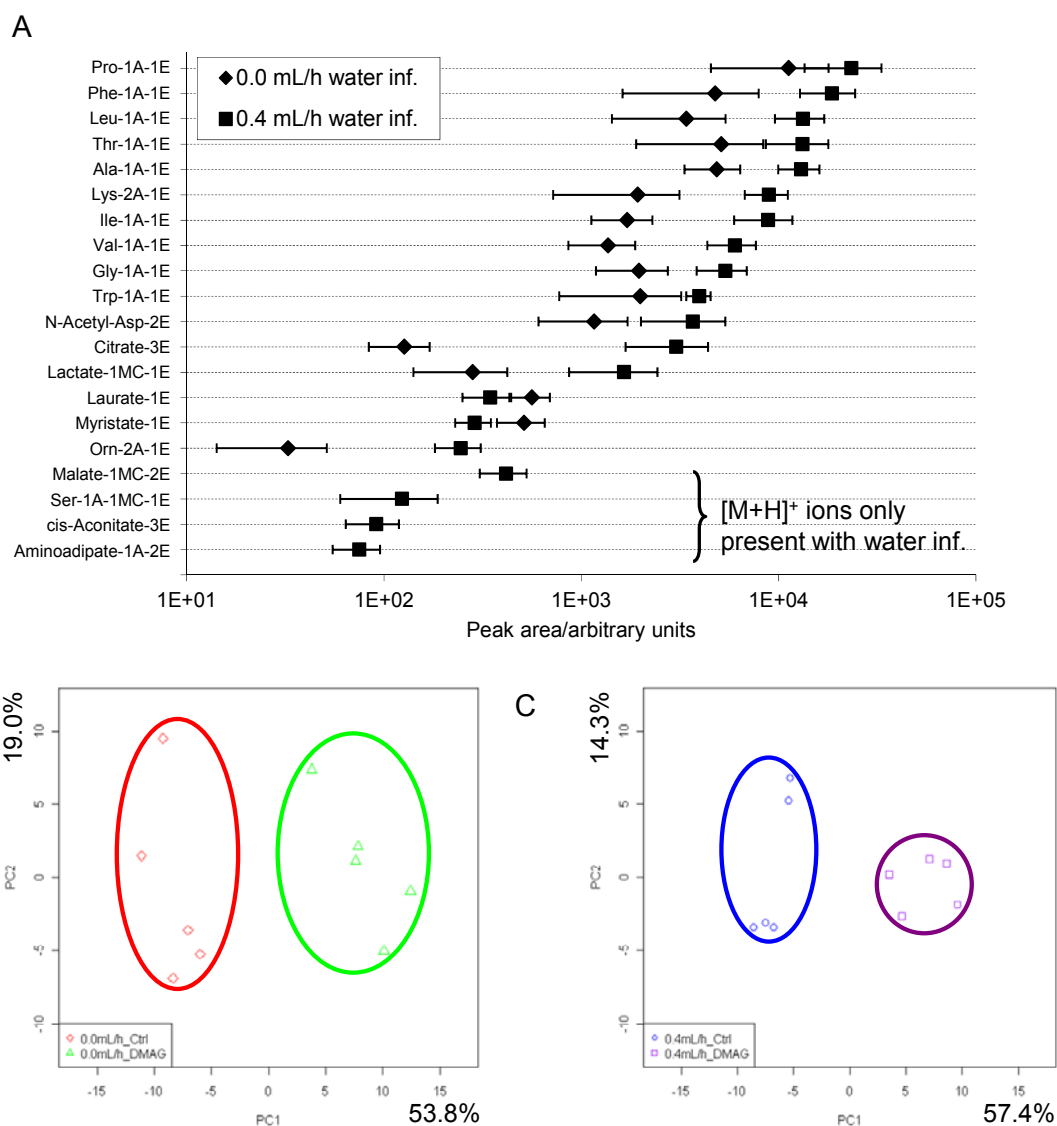


Figure 6.6 Influence of water infusion on magnitude and repeatability of signal intensities of selected MCF-derivatized metabolites and on the clustering of biological replicates in principal component analysis. (A) Arithmetic means \pm SD of peak areas of $[M+H]^+$ ions of selected metabolites detected in MCF-derivatized cell extracts (N=5) from the control group without and with water infusion. (B, C) Principal component analysis of metabolic fingerprints of MCF-derivatized cell extracts from control and 17-DMAG-treated pancreatic cancer cells (N=5 biological replicates) acquired (B) without and (C) with water infusion at 0.4 mL/h, respectively. $[M+H]^+$ abundance of identified compounds was distinctly promoted by water infusion, except for the fatty acid derivatives, and technical variability overall reduced as indicated by tighter clustering of the biological groups. A, amide group; MC, methoxycarbonyl group; E, methyl ester group. Reprinted from [14].

6.3.6 Comparative analysis of control and 17-DMAG -treated sample groups

Following extraction and alignment of features for both control and 17-DMAG -treated cells that had been acquired by APCI/-H₂O and APCI/+H₂O, respectively, peak area integrals were normalized against protein content followed by log 2 transformation, and missing values were imputed. Further, FDRs according to Benjamini and Hochberg [89] assuming equal variances in control and treated groups were calculated. A total of 18 and 77 significant features (FDR <0.05) that distinguished control from 17-DMAG -treated cells were obtained by APCI/-H₂O and APCI/+H₂O, respectively. To minimize false positive features, areas of significant features were reintegrated manually using MassLynx to exclude incorrect feature integration by the automated data analysis routine. Furthermore, GC-MS [M+H]⁺ features for amino acids that were also independently quantified by HPLC-MS/MS (**Table 6.3**) as well as [M+H]⁺ features of metabolites displayed in **Figure 6.6A** were included in the final bucket tables, in case the features had not been extracted previously by the routine. Final bucket tables contained in total 218 (APCI/-H₂O) and 195 (APCI/+H₂O) features after the exclusion of retention index markers at this stage. Data analysis was repeated, leaving 35 and 64 significant features, respectively, for APCI/-H₂O and APCI/+H₂O with an FDR <0.05 and an S/N ratio >20 for corresponding peaks across all samples of at least one group. For visualization, principal component analysis was performed based on the final bucket tables. While a group separation between control and 17-DMAG -treated cell extracts along PC1 was obtained based on both APCI/-H₂O- and APCI/+H₂O-TOFMS data, variability within the two biological groups was distinctly reduced in the latter case, as presented in **Figure 6.6B,C**. Among the top loadings of PC1, features corresponding to cysteine, threonine, alanine, glycine, and proline were always found, irrespective of water infusion.

The identities of discriminating features were determined whenever possible. One has to keep in mind that the number of detected features does not necessarily equal that of metabolites, as features corresponding to isotopes, adduct and fragment ions might have been included as well. Eventually, 27 out of 35 (APCI/-H₂O) and 41 out of 64 (APCI/+H₂O) significant features could be assigned to 13 and 23 metabolites, respectively.

Table 6.2 lists all discriminating metabolites along with figures of merit from the identification procedure. Because, according to Sumner et al. [26], at least two independent and orthogonal data derived from a standard are required for the definite (level 1) identification of an unknown, van den Dool retention indices (RIs) were compared in addition to the mass spectra. Differences in RIs (Δ RI) were between 0 and 6 index units corresponding to relative Δ RIs of 0-0.5%. This was below the established threshold of 1.0% for transferring RIs between different methods by Strehmel et al. [128], who had found Δ RI to increase with retention time and, therefore, had suggested to use relative rather than absolute RIs values. Significant features from APCI/+H₂O analysis were either not detectable with APCI/-H₂O or not extracted by the FMF routine despite a S/N threshold of 2. Further reasons for the smaller number of significant features from APCI/-H₂O analysis included missing values within group attributes of extracted features, which entailed exclusion of the respective features, as well as insufficient repeatability of peak areas. With a rapid turnover of intracellular metabolites, biological variances are increased, thus putting even higher requirements on analytical precision. Water infusion made it possible to extract features with overall better repeatability, e.g., RSDs for peak areas of lysine (m/z [M+H]⁺: 277.1394 Da) in control and treated samples improved from 32% and 53% to 21% and 24%, respectively, thus enabling its identification as a metabolite differing in abundance between untreated and treated cells. The same held true for ornithine and methionine with fold changes <2, that had been missed by APCI/-H₂O. Finally, the identity of four significant metabolites could not have been revealed in the case of APCI/-H₂O. Malate (**Figure 6.7**), serine, and aminoadipate only formed an [M+H]⁺ ion under water infusion, whereas for ornithine, mSigma values >50 obtained from APCI/-H₂O mass spectra would have entailed exclusion of the correct elemental formula during the identification procedure.

Table 6.2 Figures of merit for the metabolites identified by GC-APCI/+H₂O-TOFMS that showed significant discrimination (FDR <0.05) between control and 17-DMAG treated pancreatic cancer cells (N=5 biological replicates). Compounds marked with an asterisk (*) were revealed by GC-APCI-TOFMS irrespective of water infusion. Water infusion almost doubled the number of significantly regulated metabolites that were identified to a total of 23. A, amide group; MC, methoxycarbonyl group; E, methyl ester group. Reprinted from [14].

assigned metabolite	elemental composition [M+H] ⁺	retention time (min)	Δ RI ^a	m/z expt	Δ m ^b (mDa)	mSigma	adjusted p-value ^c	fold ^d
Gly-1A-1E*	C ₅ H ₁₀ NO ₄ ⁺	10.87	-2	148.0610	-0.6	6.4	0.0015	3.9
Trp-1A-1E*	C ₁₄ H ₁₇ N ₂ O ₄ ⁺	27.74	^e	277.1173	1	12.0	0.0015	1.9
Leu-1A-1E*	C ₉ H ₁₈ NO ₄ ⁺	13.25	-3	204.1230	0.0	7.9	0.0023	2.4
Ala-1A-1E*	C ₆ H ₁₂ NO ₄ ⁺	10.51	-5	162.0764	-0.3	7.3	0.0030	3.5
Pro-1A-1E*	C ₈ H ₁₄ NO ₄ ⁺	15.00	-3	188.0913	0.4	7.4	0.0033	3.5
Cys-1A-1MC-1E*	C ₈ H ₁₄ NO ₆ S ⁺	19.17	1	252.0532	0.5	8.9	0.0033	7.5
Phe-1A-1E*	C ₁₂ H ₁₆ NO ₄ ⁺	19.09	-4	238.1067	0.7	10.0	0.0060	2.8
Thr-1A-1E*	C ₇ H ₁₄ NO ₅ ⁺	14.45	-1	192.0863	0.4	9.0	0.0070	3.7
His-2A-1E*	C ₁₁ H ₁₆ N ₃ O ₆ ⁺	24.07	^e	286.1025	0.9	14.0	0.0077	2.7
lactate-1MC-1E*	C ₆ H ₁₁ O ₅ ⁺	9.08	^e	163.0600	0.1	7.8	0.0103	3.9
laurate-1E*	C ₁₃ H ₂₇ O ₂ ⁺	14.22	0	215.2000	0.6	15.6	0.0116	2.4
Val-1A-1E*	C ₈ H ₁₆ NO ₄ ⁺	12.13	-3	190.1067	0.7	9.8	0.0196	2.5
Ile-1A-1E*	C ₉ H ₁₈ NO ₄ ⁺	13.41	-3	204.1229	0.1	9.1	0.0214	2.0
stearate-1E	C ₁₉ H ₃₉ O ₂ ⁺	21.09	2	299.2938	0.6	9.2	0.0003	3.3
Tyr-1A-1MC-1E	C ₁₄ H ₁₈ NO ₇ ⁺	24.96	^e	312.1085	-0.7	3.3	0.0018	2.6
Lys-2A-1E	C ₁₁ H ₂₁ N ₂ O ₆ ⁺	23.08	^e	277.1390	0.5	12.0	0.0020	2.5
myristate-1E	C ₁₅ H ₃₁ O ₂ ⁺	16.67	0	243.2311	0.7	9.9	0.0057	2.2
palmitate-1E	C ₁₇ H ₃₅ O ₂ ⁺	18.94	3	271.2624	0.7	5.1	0.0061	2.7
Orn-2A-1E	C ₁₀ H ₁₉ N ₂ O ₆ ⁺	21.96	6	263.1217	2.1	15.1	0.0270	1.9
Met-1A-1E	C ₈ H ₁₆ NO ₄ S ⁺	17.65	-3	222.0777	1.7	4.1	0.0323	1.9
Ser-1A-1MC-1E	C ₈ H ₁₄ NO ₇ ⁺	16.95	-1	236.0766	-0.1	11.1	0.0128	1.3
aminoadipate-1A-2E	C ₁₀ H ₁₈ NO ₆ ⁺	18.80	1	248.1123	0.5	14.3	0.0334	1.7
malate-1MC-2E	C ₈ H ₁₃ O ₇ ⁺	14.90	-1	221.0643	1.3	22.2	0.0442	1.8

^a Difference between fatty acids-based RI in a standard and sample. ^b Difference between predicted and measured mass. ^c False discovery rate according to Benjamini and Hochberg [89].

^d Calculation after manual integration of respective [M+H]⁺ as quant mass that exhibited a peak-to-peak S/N >20. ^e Not calculable.

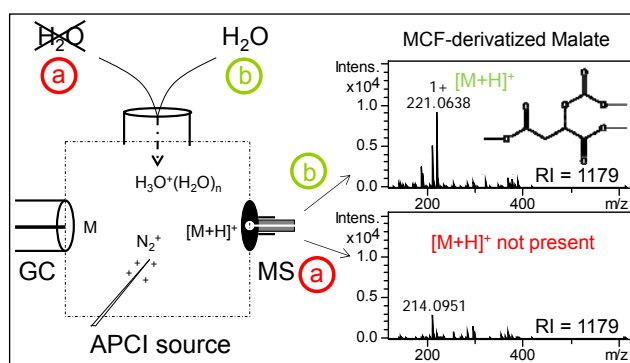


Figure 6.7 Influence of water infusion on detectability of $[M+H]^+$ of MCF-derivatized malate in a cell extract sample from the control group acquired by GC-APCI-TOFMS (a) without and (b) with 0.4 mL/h water infusion. $[M+H]^+$ was not formed in the absence of water infusion, thus impeding positive identification of this metabolite. Reprinted from [14].

6.3.7 Cross-validation of discriminating amino acids

To confirm that amino acids identified by metabolite fingerprinting to discriminate between control and 17-DMAG-treated cells reflect true differences in abundance of those amino acids, targeted analysis by means of HPLC-ESI(+)-MS/MS was performed using SIL-ISs. Concentration levels were normalized by the protein content of the respective cell pellet, which was 3.1 ± 0.2 mg and 1.5 ± 0.4 mg for the control and the 17-DMAG group, respectively. As displayed in **Table 6.3**, significant amino acids in metabolic fingerprinting were confirmed in all cases by HPLC-MS/MS, except for valine and cysteine, whose content could not be determined by the targeted approach. In addition to the discriminating amino acids already identified by the fingerprinting approach, HPLC-MS/MS revealed arginine and asparagine as significantly regulated. In the case of arginine this came as no surprise, as it is thermally unstable and, therefore, not amenable to GC-MS. Furthermore, the 1.3-fold difference in abundance of asparagine between the two groups detected by the untargeted approach was too small to reach significance due to the lower repeatability of GC-APCI/+H₂O-TOFMS in the absence of SIL-ISs. The FCs obtained by GC-APCI/+H₂O-TOFMS deviated less than 20% from HPLC-MS/MS results for 15 out of 17 amino acids (**Table 6.3**), emphasizing that the presented GC-MS approach was capable of reproducing fold changes in a quantitative manner.

Table 6.3 Comparison of fold changes in free amino acid abundance in 17-DMAG -treated cell extracts relative to controls determined both by GC-APCI/+H₂O-TOFMS and HPLC-ESI-MS/MS, respectively. Protein levels in cell pellets for normalization of concentration levels were determined by means of the FluoroProfile kit from Sigma-Aldrich. Amino acids marked with an asterisk (*) were found to discriminate control from treated cells irrespectively whether GC-APCI-TOFMS was performed without or with water infusion. Both approaches were well in accordance with each other. Reprinted from [14].

Amino acid	GC-APCI/+H ₂ O-TOFMS		HPLC-ESI-MS/MS		Fold GC-MS/ Fold HPLC- MS/MS
	Significant (Adjusted p-value ^a <0.05)	Fold ^b	Significant (Adjusted p-value ^a <0.05)	Fold	
Arg	Not detected		x	1.4	-
Gln	Not detected		Not determined ^c		-
Val*	x	2.5	Not determined ^c		-
Cys*	x	7.5	Not determined ^c		-
Gly*	x	3.9	x	2.8	1.4
Thr*	x	3.7	x	3.0	1.2
Ala*	x	3.5	x	2.6	1.4
Pro*	x	3.5	x	3.0	1.2
His*	x	2.7	x	2.4	1.1
Trp*	x	1.9	x	2.1	0.9
Leu*	x	2.4	x	2.3	1.0
Phe*	x	2.8	x	2.5	1.1
Ile*	x	2.0	x	2.4	0.8
Tyr	x	2.6	x	2.5	1.0
Ser	x	1.3	x	1.4	1.1
Met	x	1.9	x	2.0	1.0
Lys	x	2.5	x	2.3	1.1
Orn	x	1.9	x	1.8	1.1
Asn		1.3	x	1.1	1.2
Asp		1.0		1.1	0.9
Glu		1.0		1.1	1.1

^a False discovery rate according to Benjamini and Hochberg [89]; ^b Calculation after manual integration of respective [M+H]⁺ as quant mass that exhibited a peak-to-peak S/N >20 and normalization of peak areas to the protein content of the respective cell pellet; ^c Co-eluting compound present in case of Gln and Val, Cys was not detected in 17-DMAG treated samples and therefore a p value was not calculated.

6.4 Conclusions

The present study showed that the continuous infusion of water enhanced not only the efficiency but also the repeatability of APCI of MCF-derivatized metabolites, while exerting almost no or at times even a deleterious effect on APCI of MeOx-TMS derivatives. A likely explanation for the different behavior of the MCF and MeOx-TMS derivatives is the more polar nature of the former, which facilitates the formation of analyte-water ion

clusters. Overall, the beneficial effect of water infusion was most pronounced for carboxylic acids, less so for amino acids, and the least for (hetero) aromatic compounds. Water infusion was also found to improve the ability of GC-APCI-TOFMS to detect, determine, and identify metabolites, whose abundance changed significantly upon treatment of cancer cells with the Hsp90-inhibitor 17-DMAG. This was due to reduced in-source fragmentation of MCF-derivatized metabolites and a concomitant increase in $[M+H]^+$ ions, that improved not only detection sensitivity but also facilitated elucidation of elemental composition of discriminating metabolites, which also benefitted from the improved isotope patterns seen upon water infusion. Finally, the ability of metabolite fingerprinting based on MCF derivatization and GC-APCI/+H₂O-TOFMS to detect true differences in metabolite abundance could be confirmed by targeted HPLC-MS/MS analysis employing SIL-ISs. In summary, APCI/+H₂O provides softer ionization than conventional APCI, with reproducible formation of $[M+H]^+$ being favored over fragment ions for MCF derivatives, especially in the case of organic acids, thus extending the range of compounds amenable to metabolic fingerprinting.

7 Assessment of matrix effects in GC-APCI-MS

7.1 Introduction

As concluded in my master thesis, GC-APCI-TOFMS can be used for the quantitative analysis of metabolites. However, a major obstacle in the development of accurate and reliable quantitative methods for complex samples is the matrix effect, i.e., “the combined effect of all components of the sample other than the analyte on the measurement of the quantity” [137].

Matrix effects were examined in detail for LC-MS. Although less discussed, matrix effects are also found in GC-MS, e.g. in pesticide analysis [138,139]. Either suppression or enhancement of the analyte signal is possible. Electrospray ionization (ESI) and APCI are common ionization techniques in LC-MS, and both were found to be susceptible to matrix effects, although to different degrees [80,81,140]. Different mechanisms are currently discussed for the occurrence of matrix effects. Ion suppression may result in the liquid phase due to an altered efficiency of droplet formation or droplet desolvation caused by less volatile compounds or matrix co-eluting with the analyte of interest [141]. In turn, this affects the number of ions in the gas phase that enter the MS. Furthermore, other possible mechanisms are proposed that take place in the gas phase such as depletion of charge of the analytes by interfering species. In summary, as pointed out by several authors [80,141], many potential causes rather than only one may lead to the observed matrix effect. The effects may vary greatly among different sources and ionization modes [142].

Matrix effects occurring solely during ionization have been less of a concern in GC-MS compared to LC-MS. Since the mobile phase is a gas, mobile phase additives and the analyte transfer in the gas phase during ionization are irrelevant. Furthermore, GC provides enhanced analyte resolution, and, hence, a lower number of co-eluting analytes and matrix compounds are expected. Nevertheless, biological samples are highly com-

plex, and co-elution is commonly observed even in GC-MS. In addition to that, APCI for LC-MS and GC-MS is based on similar principles, because analyte ionization in the gas phase is initiated by corona discharge, which might be a source of ion suppression in the case of co-eluting species. So far, only few authors evaluated matrix effects in GC-MS with EI [37,143] and APCI [8,82] sources. For instance, Portoles et al. [82] found decreased signals of pesticides spiked into various food matrices in comparison to pure standard solutions. The authors concluded that this was due to a combined effect of the ionization and the injection [82]. Indeed, blocking of active sites in the liner by matrix compounds resulting in ionization enhancement of the analyte has been already described in the field of pesticide analysis [144].

The aim of this project was to investigate matrix effects in GC-APCI-TOFMS. It should be noted that the study only aimed at revealing potential matrix effect but the underlying mechanisms were not further studied. Three different matrices that are typically encountered in metabolomics were spiked with a mixture of 15 metabolites and recovery rates were determined for these analytes. In addition to that, possible interferences between three pairs of co-eluting analytes in standards were studied, which not only allowed the evaluation of defined amounts of possibly interfering compounds but also a closer look into whether APCI might be an actual source of ion suppression or enhancement.

The results of this chapter will be part of a manuscript in preparation.

7.2 Materials and methods

7.2.1 Spike-in experiment

Preparation of E. coli, serum and urine pool samples

The aqueous-methanolic *E. coli* BL21 extract was the same as described in a previous study by Almstetter et al. [12]. Sample volumes of 10 and 20 μL of serum and urine, respectively, were used according to established protocols in our laboratory [132]. Forty aliquots of 10 μL of human serum were precipitated with 50 μL of cold methanol and

then centrifuged at $3375 \times g$ for 5 min. The supernatants were collected and pellets re-extracted twice with 50 μL of methanol. All supernatants were combined in a pool sample. Forty aliquots of 20 μL of human urine were incubated with 20 μL of a 2U/ μL solution of urease in PBS at 37°C for 15 min. The reaction was stopped with 200 μL of ethanol, and all aliquots were pooled.

Preparation of standards and matrix samples

Since endogenous metabolite concentrations varied among the three matrices, individual spike-in mixtures were prepared in methanol from individual aqueous or methanolic stock solutions of the spike-in compounds (N=15). Their composition and the respective concentrations of the spike-in compounds are given in **Table 12.7**. Aliquots of 10 and 40 μL were taken for further preparation of spike-in standards to determine the amounts of spiked metabolites via the same calibration curves (see below) as for the matrix samples for calculating recovery rates. For the generation of the spiked matrix samples, 100 μL of *E. coli* extract, 150 μL of pooled serum extract and 240 μL of pooled pre-treated urine, respectively, were used, followed by the addition of 0, 10, 20, 30, 40, 50, and 60 μL of the respective spike-in mixture. The matrix samples contained amounts between 0.08-24.24 nmol absolute of the spike-in metabolites. Each matrix sample and spike-in standard was prepared in five replicates in a 2-mL glass vial with a 400- μL glass insert. They were then fortified with 10 μL of internal standard surrogate solution, vortexed, and evaporated to complete dryness, followed by manual MeOx-TMS derivatization including the odd-numbered, saturated fatty acids (C9-C19) prior to silylation. To check accuracy of the calibration curves, two standard samples were prepared as previously described [17] at a concentration of 125 and 15.63 μM , respectively.

Determination of endogenous and spiked levels of spike-in metabolites, and recovery

GC-APCI-TOFMS calibration curves from an existing quantitative method [17] were used for quantification of spike-in metabolites in spiked and unspiked matrix samples, and spike-in standards. For quantification the internal standards specified in **Table 12.7** were used. Recovery rates were determined in Excel for each spiked matrix sample

according to the following equation, which includes absolute amounts of spiked analytes that were calculated from those of the spike-in standards:

$$\text{Recovery(\%)} = \frac{[\text{Analyte absolute amount}]_{\text{spiked matrix sample}}}{[\text{Analyte absolute amount}]_{\text{spiked}} + [\text{Analyte absolute amount}]_{\text{endogenous}}} \times 100$$

Values higher or lower than 100% indicate ion enhancement or suppression, respectively.

Spiked and unspiked matrix samples as well as spike-in standards were measured in random order by GC-APCI-TOFMS.

7.2.2 Co-eluting analyte pairs

Two standard mixtures were prepared from separate stock solutions. Mixture A contained glycerol, glycine and phenyllactate, whereas mixture B comprised phosphate, succinate and phenylpyruvate at a concentration of 1 mM each in methanol. Both were serially diluted over a concentration range of 0.98-1000 μM .

According to the expected linear range of the calibration curves of the six analytes, one set of calibrators consisted of eight and six calibration samples that were prepared in triplicate from dilutions of mixtures A and B, respectively. Hundred μL or 75 μL of the respective diluted standard were transferred into a 2-mL glass vial with a 400- μL glass insert. Absolute amounts of 0.195, 0.391, 0.781, 1.172, 1.563, 3.125, 4.688 and 6.25 nmol of analytes in mixture A, respectively, and 0.098, 0.781, 1.563, 3.125, 6.25 and 12.5 nmol of analytes in mixture B were used. To study the slopes of the analytes' calibration curves in the absence of co-eluting compounds and their presence at three different concentration levels, four sets of calibrators were prepared in total. Then absolute amounts of 0, 1, 10, and 100 nmol of mixture B were added to the calibration samples from mixture A and vice versa. Further, all calibration samples were fortified with 20 μL of a 0.5 mM mixture of [$^2\text{H}_7$]trans-cinnamate and [$\text{U-}^{13}\text{C}$]fumarate, vortexed, evaporated to complete dryness, and manual MeOx-TMS derivatization was performed, without the addition of fatty acids.

For calibration, glycerol, phenyllactate and phenylpyruvate were normalized by [$^2\text{H}_7$]trans-cinnamate, whereas [$\text{U-}^{13}\text{C}$]fumarate was used for normalization of glycine, phosphate and succinate. The $[\text{M}+\text{H}]^+$ ion was chosen as quantifier for all analytes. Retention times of the analytes and internal standards are shown in **Figure 7.4**.

7.3 Results and discussion

7.3.1 Selection of internal standards for quantification of spike-in metabolites

At the initial stage of the evaluation of matrix effects an existing quantitative method including 43 analytes from different chemical classes [17] was used to determine the endogenous concentrations of these compounds in *E. coli*, serum, and urine. For the selection of spike-in compounds the ULOQs of the calibration curves were considered in addition to the determined endogenous levels. Concentration fold changes (FCs) of 1.5- to 4-fold of the endogenous levels were thought to be suitable to be spiked into the matrix samples. According to Annesley [145], analyte concentrations similar to those under real conditions should be used to evaluate ion suppression. Therefore, for instance, amino acids were not chosen because of their endogenous concentrations in the high micromolar range in *E. coli*. On the other hand, FCs lower than 1.5-fold might be too small for a correct evaluation. Furthermore, for analytes with endogenous concentrations below the LLOQ, 0.75 nmol absolute were used for spike-in level 1, and FCs between 2 and 6 for the higher spike-in levels. **Table 12.7** lists all the compounds included in the spike-in mixtures and their concentration levels.

Evaluation of matrix samples initially resulted in surprisingly high recoveries of up to 174% for glycerate and phenyllactate in *E. coli* extracts, and for homovanillate, phosphoglycerate and hydroxyphenylpyruvate both in serum and urine. This was caused by an internal standard recovery ($[\text{U-}^{13}\text{C}]$ lactate, $[\text{U-}^2\text{H}]$ succinate and $[\text{2,2,4,4-}^2\text{H}_4]$ citrate) of less than 75% in these matrices (**Figure 7.1**). A decrease of 25% or more in peak area has been previously regarded as indicator for ion suppression [81]. The other 4 internal

standards deviated less than 25%, but recoveries ranged between 76.1-113.4% (**Table 12.7**). Matrix pre-treatment steps are ruled out as a possible source for the observed analyte losses, because matrix aliquots were spiked post-extraction with internal standards. Hence, it was assumed that the poor IS recovery is due to the high endogenous concentrations of lactate and succinate in *E. coli*, and citrate in serum and urine, which might have suppressed their co-eluting stable isotope-labeled analogues. In fact, the actual concentration levels of these three endogenous metabolites could not be determined in these matrices because the ULOQ of the respective calibration curves had been exceeded. Another explanation in the case of labeled citrate in a urine matrix might be complexation of labeled citrate with bivalent calcium cations, because no precipitation step was included in urine pre-treatment. Similarly for LC-MS with ESI and APCI sources, Remane et al. [140] reported enhanced ion suppression for 14 ISs with increasing concentration levels of the native analogues, which was evaluated in the context of multi-analyte procedures to quantify drugs and poisons in clinical and forensic toxicology.

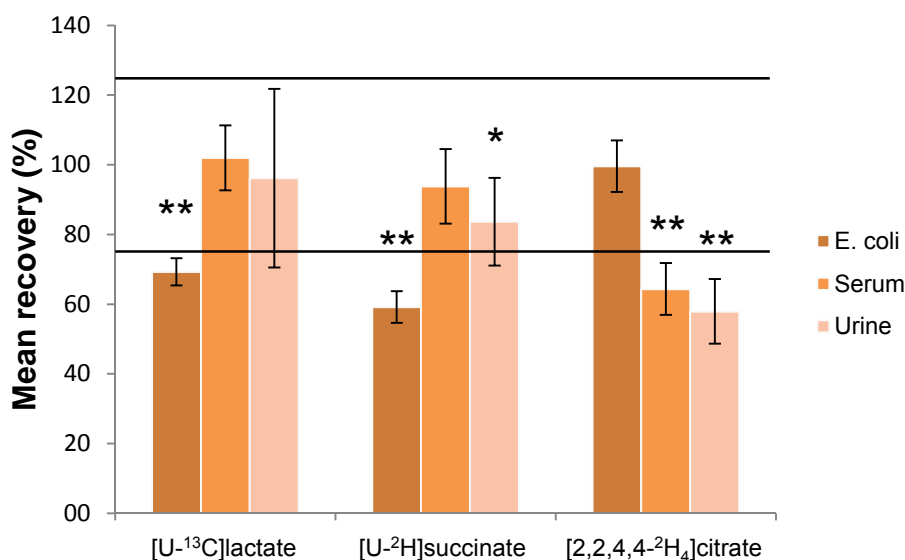


Figure 7.1 Mean recoveries \pm SD of critical internal standards in *E. coli*, serum, and urine in the spike-in experiment. These internal standards exhibited more than 25% deviation in peak area in at least one of the biological matrices (N=30 spiked matrix samples each) in comparison to standard samples (N=12). The black solid lines indicate mean recoveries of 75% and 125%, respectively. Student's t test was performed on the area integrals of $[M+H]^+$ in the spiked matrix samples (N=30) versus those in standards (N=12) *P value <0.05; **P value <0.001.

A decreased IS signal caused by co-eluting endogenous analytes might cause severe problems especially if only a limited number of ISs are available in a study. Hence, use of SIL-ISs for quantification of analytes other than their native analogues may result in inaccurate quantification.

As a consequence of the poor recoveries of [U-¹³C]lactate, [U-²H]succinate and [2,2,4,4-²H₄]citrate, endogenous concentrations of glycerate, phenyllactate, homovanillate, phosphoglycerate and hydroxyphenylpyruvate were recalculated using calibration curves based on different SIL-ISs (listed in bold in **Table 12.7**). The recalculated values differed from the original data. Since the calculation of the spike-in amounts was based on the original data, the resulting actual fold changes are slightly different and therefore fold changes are not consistent for all spike-in metabolites.

7.3.2 Recovery rates of 15 spiked metabolites in three different matrices

Mean recoveries over all spike-in levels of the 15 spike-in compounds were determined in the three different matrices. Results are presented in **Figure 7.2**. Most metabolites yielded recovery rates between 75-125%. One of the exceptions was methylmalonate in *E. coli* with a mean recovery of 68%. However, this was probably due to incorrect peak integration since valine (m/z [M+H]⁺: 262.1653 Da) was partially co-eluting (~ 25% peak overlap) with methylmalonate (m/z [M+H]⁺: 263.1129 Da), and an extraction width of ±0.1 Da for the quantifier turned out as too wide to distinguish methylmalonate and the first isotope of valine based on their mass. High recoveries up to 142% were seen for phenylacetate in serum and urine in the presence of threonine as co-eluting matrix compound. Among the matrices no obvious differences were recognized. For most spike-in compounds high mean RSD values of up to 30% were obtained. This is attributed to missing corresponding SIL-ISs for 13 out of 15 spike-in analytes. RSD values for fumarate and pyruvate, for which a corresponding SIL-IS was included, were distinctly lower (below 9% in all cases).

Next, mean recoveries of individual spike-in levels were compared, because it has been pointed out in LC-MS that the degree of ion suppression can depend not only on the

total amount of co-eluting matrix but also the matrix/analyte ratio [145]. Even though no consistent fold changes had been generated for all analytes, as described in the previous section, a total of 12 spike-in metabolite – matrix combinations could be evaluated. Only combinations with comparable spiked FCs in the specified matrix were considered. As shown in **Figure 7.3**, no big differences were seen among the different spike-in levels, except for methylmalonate in *E. coli* and urine, and pyruvate in urine (ANOVA, $p < 0.05$), which exhibited significantly decreased recoveries for the lowest spike-in level. In the case of pyruvate this was surprising, as no co-eluting matrix compound was recognized in the chromatogram. Hence, this might be caused more likely due to a slightly over-estimated endogenous pyruvate level in urine that had the largest impact on the calculation of recovery rates for low spike-in levels.

Previously, Garcia-Villalba et al. [8] investigated matrix effects in GC-APCI-TOFMS for phenolic compounds in extra virgin olive oil. The authors compared peak areas of analytes spiked post-extraction into the oil matrix to those in a neat solution, but did not report any significant differences [8]. In contrast, Portoles et al. [82] studied insecticides in various food matrices by means of GC-APCI-MS/MS, and found a distinct decrease in the analyte signal in most of the investigated matrices, which was on average 55% of that in standard solutions. The authors concluded that such effects resulted from combined effects of the ionization and the injection [82]. The latter is well known as a source for ion enhancement in the analysis of pesticides by GC [144]. Accordingly, the outcome of matrix evaluation experiments may strongly depend on the type of matrix.

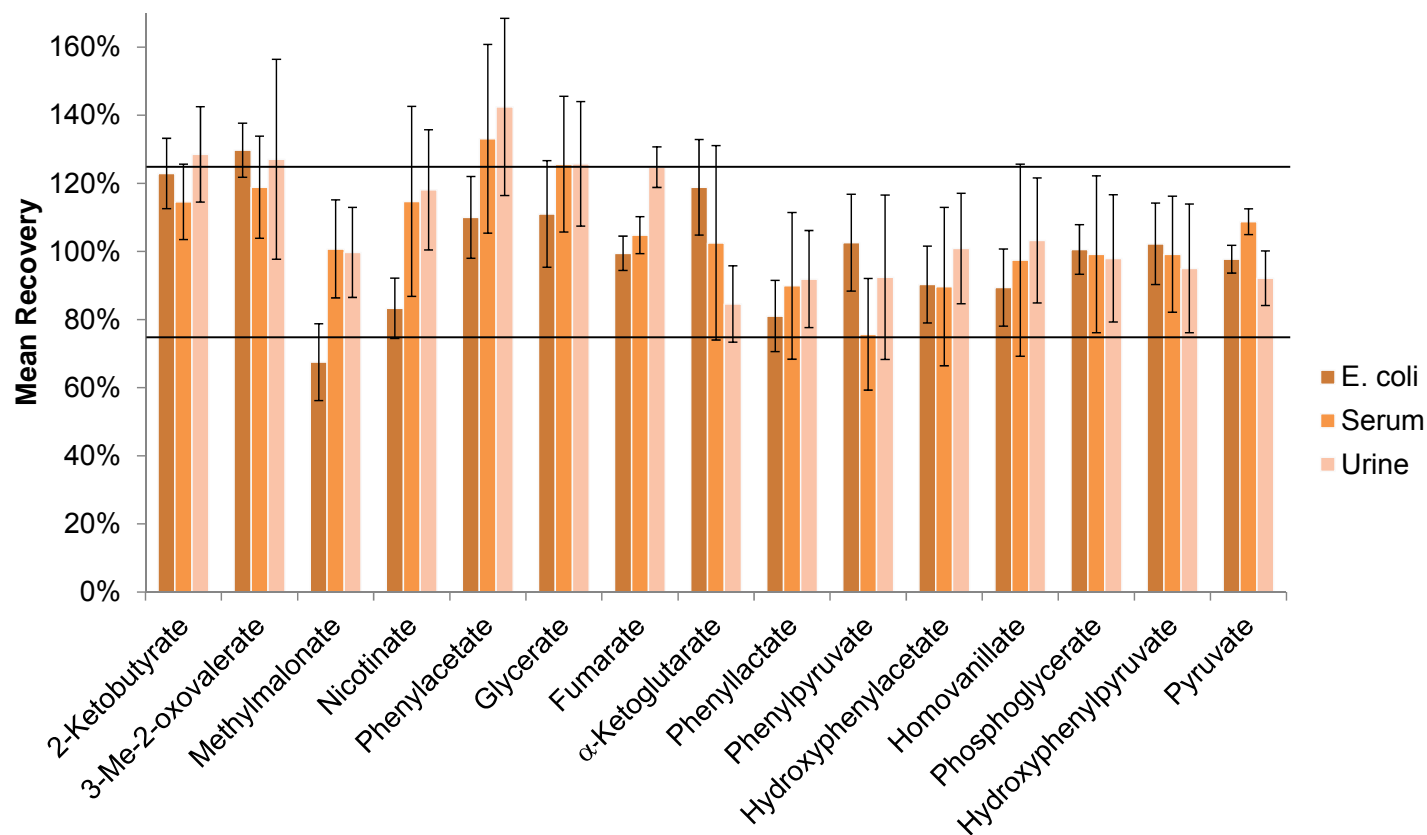


Figure 7.2 Mean recoveries over all spike-in levels \pm SD of the 15 metabolites spiked into *E. coli*, serum, and urine. Six spike-in levels (N=5 derivatization replicates each) were generated using absolute spiked amounts between 0.08-24.24 nmol. Spike-in levels were based on the endogenous concentration of the spike-in metabolites determined from unspiked matrix samples and the ULOQ of the respective calibration curves. The black solid lines indicate mean recoveries of 75% and 125%, respectively. For the majority of spike-in compounds mean recovery rates ranged between 75-125% indicating that matrix compounds hardly interfered.

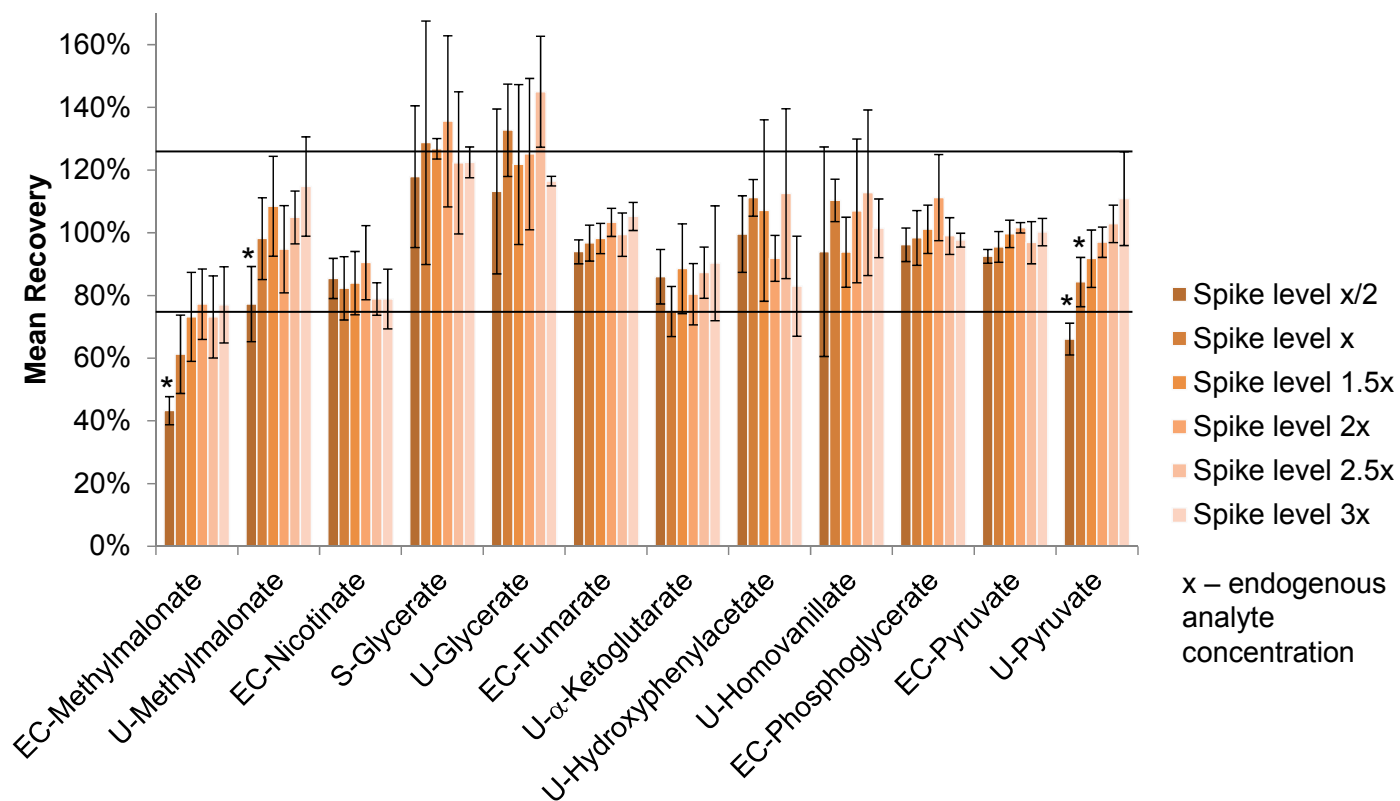


Figure 7.3 Mean recoveries of individual spike-in levels \pm SD of 12 spike-in metabolite – matrix combinations with comparable fold changes based on the endogenous analyte concentration in the specified matrix. The black solid lines indicate mean recoveries of 75% and 125%, respectively. Differences among individual spike-in levels were small indicating that the matrix-to-analyte ratio is rather not crucial for the onset of matrix effects. EC, *E. coli*; S, serum; U, urine, *, ANOVA, $p < 0.05$ in at least one comparison to a higher spike-in level.

Another explanation for the rather minor effects seen for the 15 spike-in compounds might be insufficient matrix amounts. Heller [146] by means of API-LC/MS studied varying amounts of co-injected matrix in the analysis of drug residues in animal tissue extracts. It was concluded that a critical matrix concentration exists for the onset of matrix effects [146]. This might also be valid in GC-MS. Nevertheless, typical volumes of sample were used for the experiments described above, according to established protocols in our laboratory. Furthermore, MeOx-TMS derivatization was used, which is considered as most versatile among common derivatization strategies in metabolomics; therefore, a comparably high number of matrix compounds was expected to be derivatized. On the other hand, urease pre-treatment in the case of urine removed urea, which would have yielded a broad peak and might have caused alterations in the ionization efficiency of co-eluting spike-in compounds.

7.3.3 Slopes of calibration curves of six analytes in the presence of co-eluting compounds

Apart from interferences by matrix compounds, ion suppression and enhancement tests should be performed for co-eluting analytes. Using calibrators that include at least two of the latter might hamper the quantitative determination of them in real samples in the absence of co-elution, and vice versa. Therefore, the effect of co-eluting analytes on the slopes of calibration curves of five metabolites and phosphate was studied by addition of defined amounts of co-eluting analytes to the calibrator samples. **Figure 7.4** depicts the TIC chromatogram of a standard containing the six investigated compounds and two internal standards included in the experiment. The three co-eluting pairs, namely glycerol/phosphate, glycine/succinate and phenyllactate/phenylpyruvate, gave rise to mixed APCI-TOFMS mass spectra as shown at the bottom of **Figure 7.4**. These compounds might co-exist in real samples as well, e.g., complete removal of phosphate in cell culture samples is virtually not possible; hence, phosphate might interfere with the quantitative analysis of glycerol.

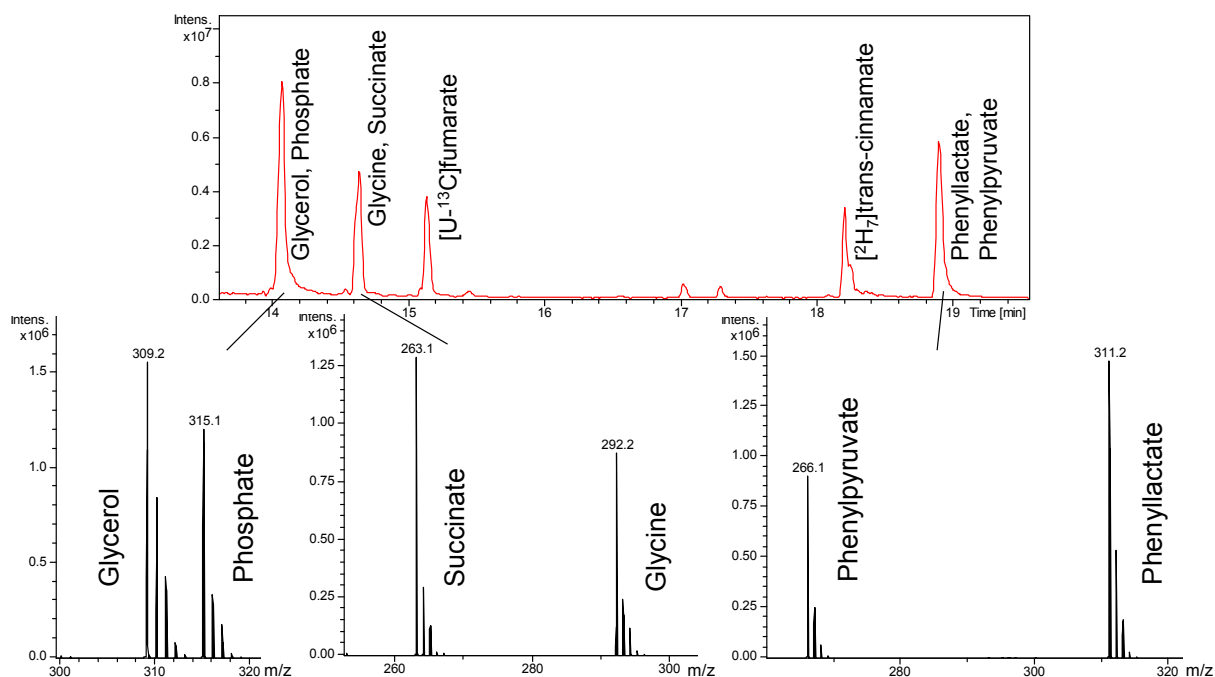


Figure 7.4 Gas chromatographic separation of the six analytes and two internal standards included in the co-elution experiment as well as APCI-TOFMS mass spectra of the three pairs of co-eluting compounds. Experimental details are given in chapter 5.

According to the linear range of their calibration curves [17], analytes were combined in two mixtures. Calibration curves were acquired in the absence of co-eluting analytes and their presence at three different amounts (1, 10, and 100 nmol absolute). As example **Figure 7.5** shows a glycerol calibration without and with phosphate at 100 nmol absolute being present in the calibrators.

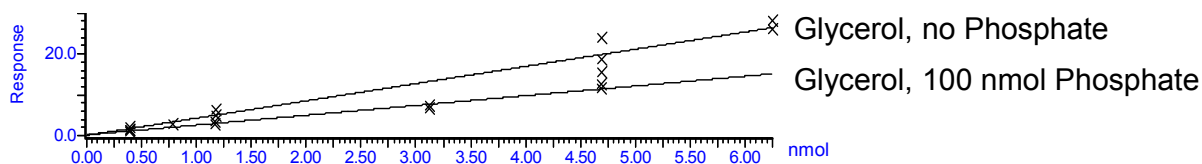


Figure 7.5 Calibration curves for the example of glycerol in the absence of co-eluting phosphate and its presence at 100 nmol absolute. The slope of the latter curve was distinctly decreased.

Further, the slopes of all calibration curves were normalized by those obtained in the absence of a co-eluting compound. The data is illustrated in **Figure 7.6** including averaged figures of merit of the four calibration curves of each analyte. As indicated by the two black solid lines spanning from 75% to 125% of normalized slopes, slopes of cali-

bration curves in the presence of 100 nmol of co-eluting analytes were distinctly decreased in all cases. Whereas those of glycine (80%) and phenylpyruvate (77%) were slightly above the established threshold of 75%, they were much lower for the other four analytes ranging between 29-57%. Furthermore, slopes of phosphate (81%) and succinate (76%) were remarkably decreased even in the presence of intermediate (10 nmol) levels of co-eluting analytes. In general, due to the presence of co-eluting analytes, mean RSDs over all triplicates and calibration points were quite high (**Figure 7.6** at the bottom), especially in the case of glycine (39%).

These results are in good accordance with those in the previous subsection on the suppression of stable isotopically labeled lactate, succinate, and citrate in the presence of high concentrations of their native analogues. The effects observed in the co-elution experiment, however, are more likely solely due to ion suppression because only standards were used and no biological matrix was present. Furthermore, only calibrators well in the linear range of the analytes were used, so that saturation of the TOFMS could be excluded. In summary, alterations in ionization efficiency were found within all three co-eluting analyte pairs, emphasizing the importance of corresponding SIL-ISs for quantitative analysis.

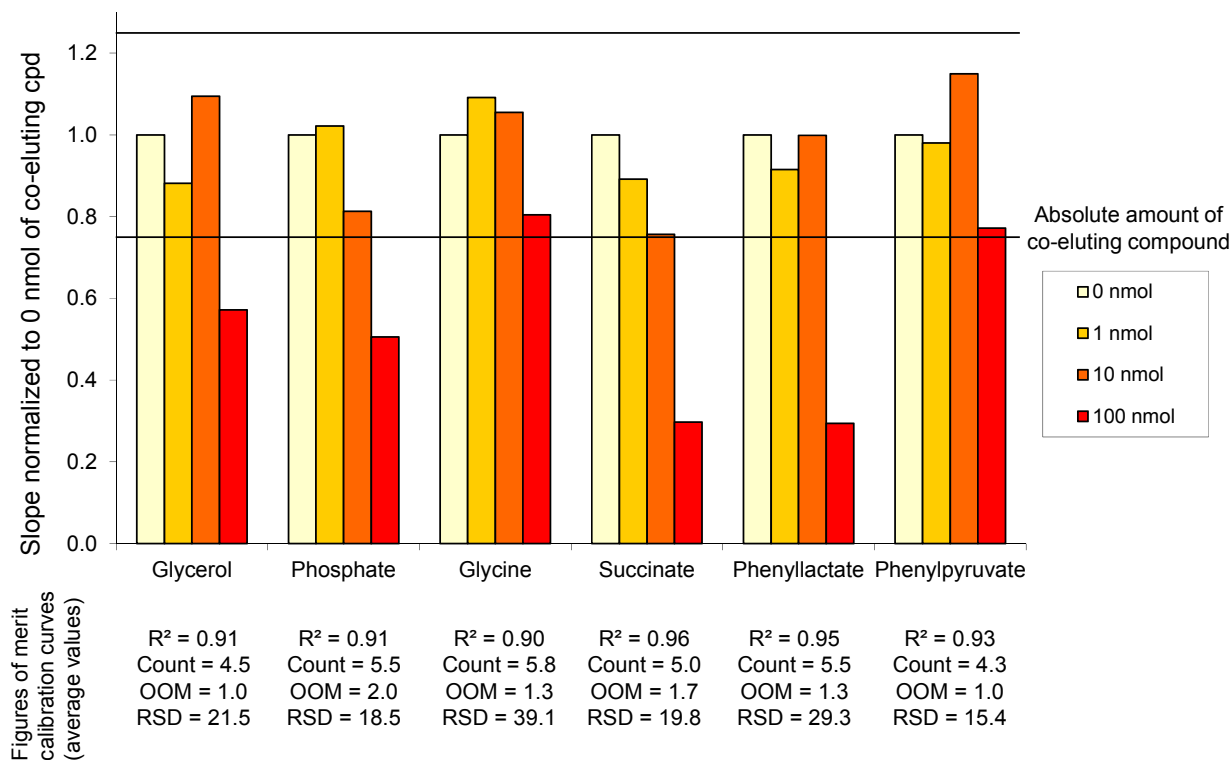


Figure 7.6 Normalized slopes of calibration curves of six analytes in the presence of three different levels of their respective co-eluting compound. Depending on the analyte, the calibration curves were based on calibrator levels between 0.98-125 μ M (N=3 derivatization replicates per calibration level) with 0, 1, 10, or 100 nmol absolute of respective co-eluting compound added to each calibration sample. The three co-eluting analyte pairs are arranged adjacent to each other in the figure. Normalization was based on the slope obtained in the absence of co-elution. Figures of merit of the calibration curves listed at the bottom correspond to the averaged values from the four curves of each analyte. The black solid lines indicate normalized slopes of 75% and 125%, respectively. While slopes of calibration curves were distinctly decreased in all cases in the presence of 100 nmol of co-eluting compounds, this was also seen for phosphate and succinate in the presence of intermediate (10 nmol) levels of co-eluting analytes. Counts, number of included calibration points; OOM, order of magnitude of the linear range on the basis of a log-10 scale; RSD, average mean relative standard deviation in peak area over all individual RSDs of calibration triplicates.

7.4 Conclusions

The experiments presented in this chapter clearly indicate the presence of matrix effects or interferences due to co-eluting analytes, which result at least in part from the ionization in the APCI source. In the spike-in experiment, recovery of three out of the seven internal standards was below 75% in at least one of the three biological matrices in comparison to standard samples. This was attributed to suppression by their respective

unlabeled analogue that was present at a high concentration, which was reinforced by the fact that co-eluting analyte pairs in standards suppressed each other at concentrations of 1000 μM of the co-eluting compound. It emphasizes the need for careful selection of internal standards used for several analytes with respect to possibly occurring ionization interferences.

Recoveries of the 15 spike-in compounds in the three biological matrices were within 75-125% in most cases, which does not indicate severe matrix effects. Nevertheless, the situation might be different if higher concentrated samples or other matrices are analyzed, or different derivatization strategies such as MCF derivatization are employed. Furthermore, the matrix-to-analyte ratio did not prove as critical for the onset of matrix effects, albeit only a small number of matrix – spike-in compound combinations were evaluated. A clear dependence on the concentration level of the interfering analyte for the onset of ion suppression was observed for three analyte pairs. In fact, depending on the analyte, 100 or 1000 μM of interfering compound led to at least 20% suppression of the slope of the analyte's calibration curve. This points to the ionization in the APCI source as cause for suppression since no matrix was present in that experiment.

In summary, during quantitative method development matrix effects should always be evaluated. They may be reduced by sample cleanup. Alternatively, an analyte's altered ionization may be corrected by its corresponding SIL-IS, or selectivity may be tuned. The latter can be achieved by choosing a GC column of different polarity or modification of the temperature program in such a way that co-elution with the target analytes becomes negligible.

8 Performance evaluation of a redesigned APCI source

8.1 Introduction

Following APCI source commercialization for GC-MS, an increasing number of successful applications has been realized in various fields, as summarized in section 4.3.2. Nevertheless, weaknesses such as unstable ionization conditions remained and spurred further development and modifications of the source design. In 2006, Östman et al. introduced an APCI microchip that exhibited good repeatability [135]. Klee et al. [71] reported the development of a novel capillary APCI (cAPCI) ionization source for GC making use of liquid instead of metal corona discharge point electrodes. The authors observed controlled and very stable conditions during source operation, which was partly attributed to the continuously replenished surface of the point electrode by the liquid flow. Recent applications with a conventional source design relied more and more on the addition of water as a modifier to provide a well-defined humidity inside the APCI source for controlled ionization [14,82,131]. Continuous infusion of water into the APCI source clearly enhanced efficiency as well as repeatability of APCI in the case of MCF-derivatized metabolites, as presented in chapter 6, but this was not observed for MeOx-TMS derivatives. Hence, the second generation “GC-APCI II” design from Bruker Daltonics, which was introduced during my doctoral thesis and supposedly addressed shortcomings of the original GC-APCI I source, was further tested with regard to its performance for the latter class of derivatives.

The improved design claims to minimize gas turbulences in the source and to better shield the ionization chamber against uncontrolled gas exchange with the outside air, thereby reducing chemical background and providing better control over humidity in the source. In the present study, the novel GC-APCI II source coupled to HRTOFMS analysis was compared to its first-generation predecessor (“GC-APCI I”) for the application to

metabolomics. Its performance was initially compared over the entire chromatographic run using a homologous series of fatty acid methyl esters. Furthermore, calibration curves of 20 MeOx-TMS derivatized metabolite standards were used to compare figures of merit for metabolic profiling, such as linear range and derivatization repeatability. Finally, to comparatively assess the qualitative output, features were extracted and identified from APCI I and II mass spectra from analyses of cell culture supernatants.

Preparation of calibration and cell culture supernatant samples, as well as partial evaluation of calibration curves was carried out by Thomas Hahn as part of his bachelor thesis under my guidance. This chapter was published in [13].

8.2 Materials and methods

8.2.1 Materials

A 1:100 (v/v) working dilution of the FAME mixture in methylene chloride was prepared, yielding concentrations of 5-13 μM of the 19 individual compounds that were finally evaluated. The metabolite mix contained 20 metabolites (compare **Table 8.2**) in methanol at 1 mM each, prepared from separate stock solutions in water, methanol or a mixture thereof. The mix was serially diluted to generate 20 calibration points from 0.002 μM to 1000 μM . MeOx-TMS derivatized calibration standards were generated as described in subchapter 5.2.

8.2.2 Instrumentation

APCI II source setup

The “GC-APCI II” source (Bruker Daltonics GmbH, Bremen) consists of an atmospheric pressure chemical ionization source chamber equipped with a corona discharge needle assembly and a flexible heated GC transfer line. A scheme of the APCI II source is shown in **Figure 8.1**. In contrast to its predecessor, the vaporizing unit had been removed because it caused excessive turbulences in the APCI I source (compare **Figure**

4.7 and section 4.3.1 on the GC-APCI-TOFMS instrumental setup). Instead, the heat for the APCI ionization process is now generated at the head of the GC transfer line by an external heating device. Further reductions in gas turbulences are accomplished by generating a hot laminar flow from the head of the GC transfer line, illustrated in orange in **Figure 8.1**. This laminar gas flow guides the carrier gas towards the MS inlet. Opposite to the head of the GC transfer line a modified spray shield assembly is mounted at the entrance of the transfer capillary of the mass spectrometer. This spray shield arrangement has 2 functions. It produces an optimized electrical gradient to direct the ions into the MS inlet capillary and it generates an optimized laminar dry gas flow, illustrated in red in **Figure 8.1**, outside of the inner guiding gas flow. According to simulations performed by the manufacturer, both counter-directed laminar gas flows do not disturb each other and do not generate turbulences (personal communication, Dr. Thomas Arthen-Engeland, Bruker Daltonics Bremen, Germany).

Both gas flows shield and maintain the heat in the inner APCI ionization area. This prevents contamination, which reduces the chemical background and excludes uncontrolled gas exchange with the outside air thereby controlling the water content in the ion source. This is further supported by the airtight design of the source chamber. Finally, in contrast to the original design, the heat required for the APCI process is now generated in close vicinity to the ionization region and hence generates higher temperatures in the APCI ionization region.

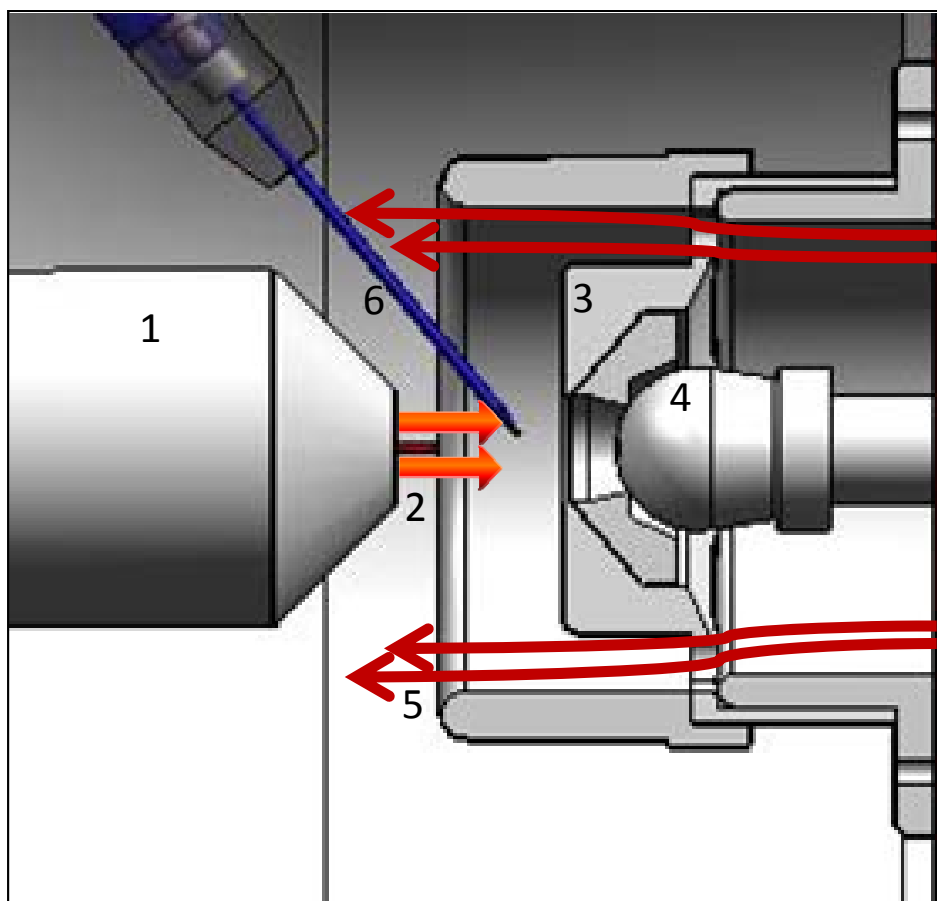


Figure 8.1 Scheme of the GC-APCI II ion source: 1 head of the GC transfer line; 2 GC capillary outlet and heated nitrogen gas guiding the eluting analytes towards the MS inlet; 3 APCI electrical shield with outer ring cover enclosing the APCI ionization area and guiding both gas flows into and from MS; 4 inlet mass spectrometer; 5 dry gas flow direction; 6 corona needle. The GC effluent is delivered by a flexible GC transfer line into the APCI region of the ion source. Mechanical design, gas flows, heat distribution and electrical fields were optimized for efficient AP chemical ionization and ion transfer into the MS inlet. The illustration is reprinted from [13] with permission from Bruker Daltonics GmbH, Bremen.

GCxGC-EI-TOFMS

Instrumental settings for the employed LECO Pegasus 4D GCxGC-EI-TOFMS instrument are described in detail elsewhere [17,147]. The first-dimension GC column was identical to the one used for GC-APCI-TOFMS in the current chapter and chapter 7. The temperature program was only slightly different with 50°C for 0.2 min, raised at 8°C/min to 265°C, and held for 10 min. Analyte fractions were transferred in 4-second intervals from the first- onto the second-dimension column by means of a dual-stage, quad-jet thermal modulator. An RTX-1701 (2 m × 0.1 mm i. d. × 0.1 µm film thickness, Restek)

was employed as second dimension column, and kept at a positive offset of 5°C relative to the first-dimension column. Splitless injection of 1.5 µL was performed by means of a PTV injector (set at 50°C for 0.5 min, ramped at 12°C/s to 250°C, held for 1 min). EI ionization under standard 70 eV electron ionization parameters at 200°C was used. Data were acquired 40-600 *m/z* at a rate of 100 spectra/s. The solvent delay was 8 min.

8.2.3 Data analysis

8.2.3.1 FAMES

For automatic peak integration, the “Find Compounds – Chromatogram” version 3.0 peak finder in DataAnalysis was applied to the extracted ion chromatograms (EICs) of $[M+H]^+$ ions of all investigated compounds (see Table 1) using an extraction width of ± 0.02 Da. Peak finder parameters were as follows: sensitivity: 90%; area threshold: off; intensity threshold: 1000; minimum peak valley: 30%; and mass spectrum calculation settings: default. Peak identification of investigated FAMES was based on the accurate mass of each compound, common elution rules [148], as well as peak abundance in case of C18:1n9c and C18:1n9t, as the concentration of the former is twice as high in the FAME mixture.

8.2.3.2 Cell culture supernatant samples

For compound identification, the “Find Compounds – Dissect” peak-finding algorithm implemented in DataAnalysis was used. This deconvolution algorithm combines all ions with a similar chromatographic profile into one compound. Dissect compounds were extracted over a chromatographic range of 9-37 min of recalibrated data files using the following parameters: internal S/N threshold: 20; maximum number of overlapping compounds: 3; mass spectrum calculation parameters: default; and advanced settings: enabled chromatographic resolving power of 0.3. Identification of metabolites was achieved in several ways: (1) *m/z* values of extracted dissect compounds were searched in the GC-APCI-HRMS online database established by Pacchiarotta et al. [74] and retention indices compared when available in the database. (2) Accurate masses and retention indices were matched against an in-house GC-APCI-HRTOFMS database

of 100 metabolite standards including sugars, fatty acids, organic acids, alcohols, amino acids and amino acid metabolites (compare **Table 12.8**). (3) GC×GC-EI-TOFMS raw data were processed using the ChromaTOF software as previously described [147] with a S/N threshold of 100 for peak detection in the first dimension and a S/N threshold of 30 for second-dimension subpeaks. Peaks were combined in the second-dimension separation using a spectral matching factor of 500. Annotation of deconvoluted mass spectra was carried out by mass spectral matching to the NIST 05 library (based on a match >500 in forward search mode) and an attempt was made for annotated metabolites to find the accurate masses of corresponding [M+H]⁺ ions in the APCI-HRTOFMS data. (4) For additional confidence in identified metabolites according to (1)-(3), elemental formulas were calculated using the SmartFormula tool implemented in DataAnalysis (compare section 5.4.4).

In addition to the Dissect algorithm, the FMF algorithm, which is recommended for quantitative analyses, was used to extract features from the chromatograms and compare cumulative distributions of RSDs of peak areas. The following settings were applied for extraction and bucketing of FMF features: S/N threshold, 20; correlation coefficient threshold, 0.6; minimum compound length, 15; smoothing width, 5; additional smoothing, enabled; proteomics, enabled; bucketing basis, M+H. Retention time range, 9-37 min; mass range, *m/z* 100-650; advanced bucketing tolerance parameters, 0.1 min and 10 mDa for retention time and mass, respectively; other parameters, always none or disabled. The bucket table was exported into Excel, features were normalized to the area integrals of [²H₇]trans-cinnamate, and the RSDs calculated for all APCI I and II FMF features.

For the qualitative assessment of technical variability, a single supernatant sample was injected trice and overlaid EICs were manually inspected.

8.3 Results and discussion

8.3.1 GC-APCI-HRTOFMS analysis of FAMES with the APCI I and II sources

The general performance of APCI II in combination with HRTOFMS was investigated with a commercial 37-component FAME mixture and compared to that of APCI I-HRTOFMS. The two data sets were acquired using identical chromatographic conditions (**Figure 8.2**).

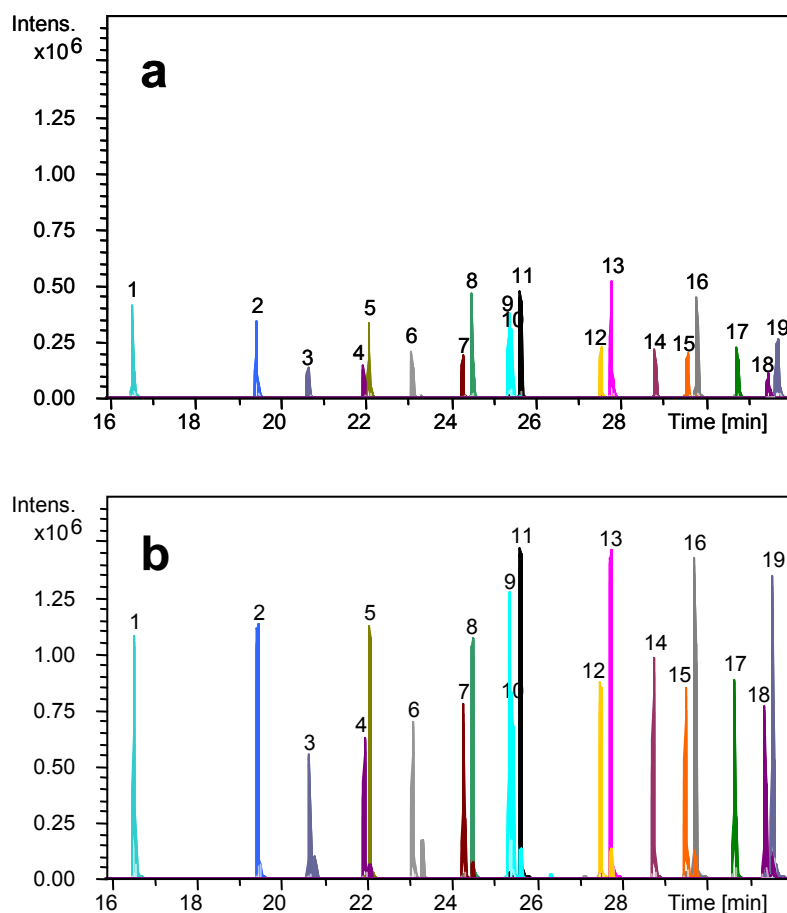


Figure 8.2 Overlaid extracted ion GC-APCI I and II-HRTOFMS chromatograms of FAMES. A subset of N=19 components of a commercial FAMES mixture were investigated in four subsequent replicate runs by means of GC-HRTOFMS equipped with an (a) APCI I and (b) APCI II ion source, respectively. Extracted masses of protonated molecules are listed in **Table 8.1**. An average 4.4-fold increase in peak areas was observed and peak apices aligned better in case of APCI II. (Concentration range: 5-13 μ M. Δ EIC = 20 mDa). Peak identification: 1, C11:0; 2, C13:0; 3, C14:1; 4, C15:1; 5, C15:0; 6, C16:1; 7, C17:1; 8, C17:0; 9, C18:1c; 10, C18:1t; 11, C18:0; 12, C20:1; 13, C20:0; 14, C21:0; 15, C22:1; 16, C22:0; 17, C23:0; 18, C24:1; 19, C24:0. Reprinted from [13].

The FAME mixture was used, because repeatability of retention time, peak width and ionization efficiency can be evaluated for compounds of widely different volatility. Moreover, FAMES are alternative reference compounds for the calculation of retention indices instead of alkanes, which evade ionization in APCI(+) mode [78]. Keeping the use of FAMES as reference compounds in mind, we opted not to use a dedicated FAME capillary GC column but rather the general-purpose Rxi-5MS column, though it is not well suited for the separation of complex FAME mixtures. Especially, unsaturated FAMES with C18 and C20 backbones are not baseline resolved [148]. However, separation of even- and odd-numbered, saturated and monounsaturated FAMES was sufficient. The FAMES C4:0 and C6:0 eluted with the solvent and C8:0, C10:0, C12:0, C14:0, and C16:0 were excluded from this subset because they were in detector saturation with GC-APCI II-HRTOFMS already at concentrations of 25 μ M in the diluted mixture. Consequently, 19 FAMES remained that were used for comparison. Very stable retention times across four injection replicates were obtained with an average relative standard deviation of 0.01% for all 19 FAMES. Signal intensities (peak areas) of the protonated molecules of the 19 FAMES were on average 4.4-fold greater with the APCI II than the APCI I source, most likely due to the more efficient ion transfer to the MS inlet in the APCI II source (**Table 8.1**). This is also reflected in the S/N ratios that were on average 2.6-fold higher (**Table 12.9**). Furthermore, RSDs of peak areas obtained with APCI II were drastically reduced from 24.3% with APCI I to 4.6%. This low value is the more impressive as peak areas had not been normalized to an IS. Equally impressive were the symmetric GC peaks with fairly uniform and highly repeatable (average RSD of 1.5%) peak widths at half height (w_h) of approximately 2 s that could be observed over the entire elution range with GC-APCI II-HRTOFMS (**Figure 8.2**, **Table 8.1**). In contrast, for APCI I, values of w_h increased for late-eluting FAMES (compare **Figure 8.3**), most likely due to adsorption of the higher boiling analytes to cold spots in the ionization chamber, and were generally less repeatable (13.0%). The excellent repeatability of the peak area and w_h observed with the APCI II source resulted from the generation of less ragged peak profiles as a consequence of the reduction in gas turbulences in the APCI II source accomplished by the modified spray shield and the removal of the vaporizing unit.

Table 8.1 Statistical results for quadruplicate analyses of FAMES (N=4) by GC-HRTOFMS with the APCI I and APCI II source, respectively. Given are the retention times and the m/z values of the protonated molecules of the nineteen FAMES examined. Student's t-tests with correction for multiple testing according to Benjamini and Hochberg [89] were performed on the peak areas and peak widths at half height (w_h). The average 4.4-fold increase in peak area obtained with APCI II was significant for all evaluated FAMES ($p < 0.05$). APCI II also outperformed APCI I in terms of repeatability of peak area and w_h and yielded similar peak widths over the entire retention time range. A paired t-test was performed to compare mean area RSDs. For comparison of the not normally distributed w_h RSDs, the Wilcoxon signed-rank test was used. Reprinted from [13].

FAME	GC-APCI II- HRTOFMS ret.time (min)	m/z APCI [M+H] ⁺ ion	Fold change mean peak area APCI II/I	GC-APCI I-HRTOFMS			GC-APCI II-HRTOFMS		
				Area RSD (%)	Mean w_h (s)	w_h RSD (%)	Area RSD (%)	Mean w_h (s)	w_h RSD (%)
C11:0	16.61	201.18	3.4	13.6	1.9	17.7	2.2	2.0	0.6
C13:0	19.50	229.22	3.8	17.4	2.1	9.7	3.5	2.0	0.8
C14:1	20.70	241.22	5.3	20.0	2.0	10.7	3.8	1.9	1.1
C15:1	22.00	255.23	5.0	21.7	2.0	6.6	4.0	2.0	2.5
C15:0	22.12	257.25	4.3	29.4	1.9	11.4	5.2	2.0	0.5
C16:1	23.13	269.25	5.0	40.9	1.9	12.9	4.5	1.9	0.7
C17:1	24.32	283.26	4.9	25.0	1.9	4.3	5.2	1.9	2.3
C17:0	24.53	285.28	3.4	22.7	1.8	15.1	4.3	1.9	1.2
C18:1c	25.39	297.28	4.3	16.6	1.8 ^a	5.2	4.5	2.1 ^a	0.6
C18:1t	25.45	297.28	4.6	43.3	2.0	31.6	4.9	2.0	2.3
C18:0	25.65	299.29	3.5	10.0	2.0	8.4	3.8	2.2	2.1
C20:1	27.54	325.31	4.8	23.0	1.9	11.4	4.5	2.0	1.6
C20:0	27.78	327.33	3.5	20.7	2.1	8.7	3.5	2.2	2.6

Table 8.1 continued.

FAME	GC-APCI II- HRTOFMS ret.time (min)	<i>m/z</i> APCI [M+H] ⁺ ion	Fold change mean peak area APCI II/I	GC-APCI I-HRTOFMS			GC-APCI II-HRTOFMS		
				Area RSD (%)	Mean <i>w_h</i> (s)	<i>w_h</i> RSD (%)	Area RSD (%)	Mean <i>w_h</i> (s)	<i>w_h</i> RSD (%)
C21:0	28.79	341.34	4.5	18.5	2.2	8.9	5.1	1.9	1.1
C22:1	29.55	353.34	5.7	37.3	2.4 ^a	8.7	5.6	2.0 ^a	0.8
C22:0	29.77	355.36	3.9	31.1	2.1	10.3	3.7	2.2	1.0
C23:0	30.71	369.37	4.7	26.5	2.4	32.8	7.1	2.0	2.3
C24:1	31.43	381.37	5.7	22.9	2.9	19.8	7.2	2.1	2.4
C24:0	31.64	383.39	4.0	20.6	3.4 ^a	12.5	4.7	2.3 ^a	2.8
Mean			4.4	24.3 ^b	2.1	13.0 ^c	4.6 ^b	2.0	1.5 ^c

^a $p_{adj} < 0.05$ (Student's t-test, corrected for multiple testing). ^b $p < 0.05$ (Students t-test). ^c $p < 0.05$ (Wilcoxon)

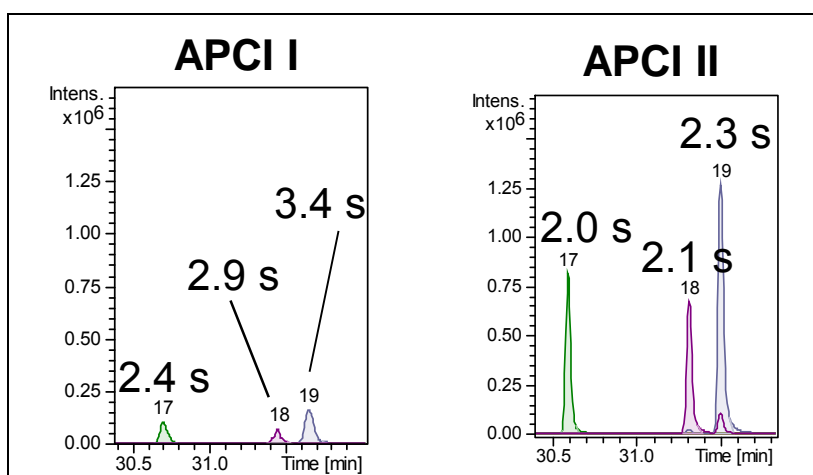


Figure 8.3 Peak widths (W_h) of late-eluting FAMEs obtained from GC-APCI-TOFMS measurements with APCI I (left) and APCI II (right) ion sources. APCI II yielded distinctly narrower peaks in comparison to APCI I. 17, C23:0; 18, C24:1; 19, C24:0.

8.3.2 Comparison of GC-APCI-HRTOFMS calibration curves

The quantitative performance of GC-APCI II-HRTOFMS was assessed by comparing the standard calibration curves of 20 metabolites to data acquired with an APCI I source. The chosen set of metabolites included compounds from various chemical classes, including amino acids; amines; mono-, di-, tri-, and aromatic carboxylic acids; 2-oxo-acids; and sugars. The linear range determined by the lower limit of quantification (LLOQ) and upper limit of quantification (ULOQ), its order of magnitude on a log-10 scale, and the RSD of the area response for derivatization replicates (N=5) of a single calibration point (31.25 μ M) are given in **Table 8.2** together with the GC-APCI II-HRTOFMS retention times, m/z values of $[M+H]^+$ ions, and the ISs used for normalization. Retention times, m/z values, and RSDs of the non-normalized peak areas of the ISs are listed in **Table 12.10**. Employing the APCI II source, the analytical performance for the MeOx-TMS derivatized metabolite standards was distinctly enhanced. The RSD of the area response of the 31.25 μ M calibration point decreased more than twofold compared to APCI I, from 7.5 to 3.4%. The calibration point was well within the linear range of all compounds for both the APCI I and the APCI II source, except for GABA-3TMS, which slightly exceeded the ULOQ of GC-APCI II-HRTOFMS. Similarly, RSD values determined for the non-normalized peak areas of the ISs were slightly lower for

the APCI II measurements (**Table 12.10**). LLOQs decreased on average more than threefold from 1.5 to 0.45 μM . Considering the median rather than the average LLOQ, a more than fourfold decrease from 0.61 to 0.15 μM was observed. The median LLOQ provides a more appropriate evaluation of the data, as the LLOQ of 4-methyl-2-oxovalerate (3.91 μM) was much higher than the LLOQs of the other analytes, which ranged from 0.002 – 0.73 μM . The linear concentration range was significantly increased from 2.4 to almost three orders of magnitude, with R^2 values above 0.99 in all cases. The increase was achieved through better repeatability and, thus, inclusion of lower calibration points with less than 20% deviation from the curve. Increased detectability also contributed to the broader linear range as indicated by a distinctly lower median LLOQ in comparison to the median limit of detection (LOD) previously obtained for GC-APCI I-HRTOFMS (0.15 versus 0.24 μM) [17]. A drastic improvement in LLOQ from 1.46 μM to 0.045 μM was observed for lactose, due to the generation of narrower peaks for high-boiling compounds using the APCI II source. Even greater improvements in LLOQ were achieved for GABA and glycine, with all amino acids yielding consistently LLOQs in the submicromolar range and a linear range over at least three orders of magnitude in five out of seven cases, including the non-proteinogenic GABA. As reported previously by our group [17] and others [37], amino acids are difficult to analyze by silylation and more suitable derivatization strategies have been recommended [37,38]. Using GC-APCI II-HRTOFMS, RSDs were below 5% for all amino acids except Ala-2TMS (8.6%), but in general lower than previously reported data obtained by GC-MS with EI and CI sources [17]. This constitutes a major improvement especially if one takes into account that derivatization was performed manually in the present study.

Table 8.2 Figures of merit for quantitative profiling of 20 metabolites after MeOx-TMS derivatization and GC-HRTOFMS analysis with the APCI I or APCI II source. RSDs of response values were determined for a concentration level of 31.25 μM (N = 5 derivatization replicates). APCI II outperformed APCI I in terms of lower LLOQs, broader linear ranges, and better repeatability. (TMS derivatization status and isomer elution number are given for metabolites yielding multiple peaks after derivatization). Reprinted from [13].

Compound	GC-APCI II-HRTOFMS ret.time (min)	m/z [M+H] ⁺ ion	Corresponding IS	GC-APCI I-HRTOFMS			GC-APCI II-HRTOFMS		
				LR (μM)	Order of Magnitude	RSD (%)	LR (μM)	Order of Magnitude	RSD (%)
Lactate	10.88	235.12	[U- ¹³ C]lactate	0.12 - 500	3.6	1.7	0.12-187.5	3.2	9.3
Ala-2TMS	11.56	234.13	Nval-2TMS	0.98 - 375	2.6	15.9	0.18-250 ^a	3.1 ^b	8.6
2-Hydroxybutyrate	12.02	249.13	[U- ¹³ C]3-hydroxybutyrate	0.24- 93.75	2.6	4.9	0.12-375	3.5 ^b	4.5
3-Hydroxybutyrate	12.55	249.13	[U- ¹³ C]3-hydroxybutyrate	0.12 - 500	3.6	4.3	0.12-250	3.3	2.9
4-Me-2-oxovalerate	13.46	232.14	[U- ¹³ C]3-hydroxybutyrate	7.81 - 500	1.8	2.0	3.91-500	2.1	3.7
Val-2TMS	13.53	262.17	Nval-2TMS	0.37 - 125	2.5	13.5	0.73-93.75	2.1	1.2
Leu-2TMS	14.44	276.18	Nval-2TMS	0.73 - 125	2.2	15.9	0.73-187.5 ^a	2.4 ^b	2.3
Ile-2TMS	14.81	276.18	Nval-2TMS	1.95 - 125	1.8	15.6	0.12-187.5 ^a	3.2 ^b	2.0
Pro-2TMS	14.89	260.15	Nval-2TMS	3.91 - 375	2.0	3.0	0.49-500 ^a	3.0 ^b	1.7
Gly-3TMS	15.04	292.16	[U- ² H ₄]succinate	0.12 - 62.5	2.7	9.0	≤ 0.002 - 46.88 ^a	4.4 ^b	2.5
Succinate	15.05	263.11	[U- ² H ₄]succinate	0.18 - 46.88	2.4	6.3	0.09-46.88 ^a	2.7	1.6

Table 8.2 continued.

Compound	GC-APCI II-HRTOFMS ret.time (min)	m/z [M+H] ⁺ ion	Corresponding IS	GC-APCI I-HRTOFMS			GC-APCI II-HRTOFMS		
				LR (μM)	Order of Magnitude	RSD (%)	LR (μM)	Order of Magnitude	RSD (%)
Glycerate	15.41	323.15	[U- ¹³ C]fumarate	0.06 - 46.88	2.9	14.7	0.12-46.88	2.6	1.5
Fumarate	15.55	261.10	[U- ¹³ C]fumarate	0.12 - 93.75	2.9	2.4	0.37-187.5	2.7	0.8
Adipate	17.97	291.14	[U- ² H ₄]succinate	2.93 - 187.5	1.8	10.4	0.37-46.88	2.1	2.2
GABA-3TMS	18.43	320.19	[U- ¹³ C]fumarate	0.37 - 46.88	2.1	2.5	≤ 0.002- 15.63 ^a	3.9 ^b	2.8 ^c
cis-Aconitate	21.35	391.14	[² H ₇]trans-cinnamate	0.49 - 62.5	2.1	6.5	0.09-62.5	2.8	4.4
Homogentisate	22.52	385.17	[U- ¹³ C]glucose-1	5.86 - 250	1.6	4.6	0.73-31.25	1.6	2.3
Hydroxyphenylpyruvate	23.36	354.16	[U- ¹³ C]glucose-1	0.73 - 500	2.8	11.6	0.24-375 ^a	3.2 ^b	6.8
5-HIAA-2TMS	26.79	336.15	[4,6,7- ² H ₃]5-HI-[² H ₂]AA	1.46 - 500	2.5	3.0	0.49-46.88	2.0	1.1
Lactose-1	31.86	948.47	[U- ¹³ C]lactose-1	1.46 - 93.75	1.8	2.5	0.045-375	3.9	5.0
Arithmetic mean				1.50-230 ^{d,e}	2.4 ^f	7.5 ^g	0.45-191 ^{d,e}	2.9 ^f	3.4 ^g

^a LLOQ lower than previously for GC×GC-EI-TOFMS [17]. ^b Broader linear range than previously for GC×GC-EI-TOFMS [17]. ^c Standard concentration above ULOQ. ^d p (Wilcoxon, LLOQ) = 0.002. ^e p (Wilcoxon, ULOQ) = 0.365. ^f p (t test) = 0.014. ^g p (Wilcoxon) = 0.005.

The excellent overall quantitative performance of the APCI II source and, in particular, the improvements observed for analytes without a corresponding IS are striking even in comparison to GC×GC-EI-TOFMS, which had performed the best in a previous study [17]. GC-APCI II-HRTOFMS yielded lower LLOQs and broader linear ranges than GC×GC-EI-TOFMS [17] for eight of the investigated metabolites. The respective metabolites are marked with superscript *a* indicating a lower LLOQ and *b* indicating a broader linear range (**Table 8.2**). Overall, the median LLOQ was comparable to that of GC×GC-EI-TOFMS (0.15 versus 0.12 μM). The average linear range was slightly less wide for GC-APCI II-HRTOFMS (2.9 versus 3.2 orders of magnitude) because of detector saturation for higher calibration points.

8.3.3 Analysis of cell culture supernatants

To assess the performance of GC-APCI II-HRTOFMS in the analysis of complex biological matrices, we measured MeOx-TMS derivatized cell culture supernatants from three replicate MiaPaCa-2 pancreatic cancer cell cultures. In total, 275 peaks (dissect compounds) were extracted over a chromatographic window of 9-37 min from the APCI II data using an S/N threshold of 20. In comparison, only 127 peaks were extracted from the APCI I data applying the same “Find Compounds – Dissect” peak finder parameters. The difference in extracted peaks is illustrated in **Figure 8.4** showing the dissect compounds from the same supernatant sample for both APCI sources. Particularly impressive is the approximately 8-fold gain in extracted peaks for compounds eluting after 25 minutes. The increased number of extracted peaks with the APCI II source is mainly attributed to the improved detection sensitivity but also to the improved, less ragged peak shapes that facilitate peak detection by the peak finding algorithm. This is again discussed at the end of this section. Furthermore, late eluting peaks are narrower, which enhances peak detection. In addition, the improved heating of the source reduces potential adsorption of late eluting high boiling analytes, which may also contribute to the higher number of detected late eluting analytes.

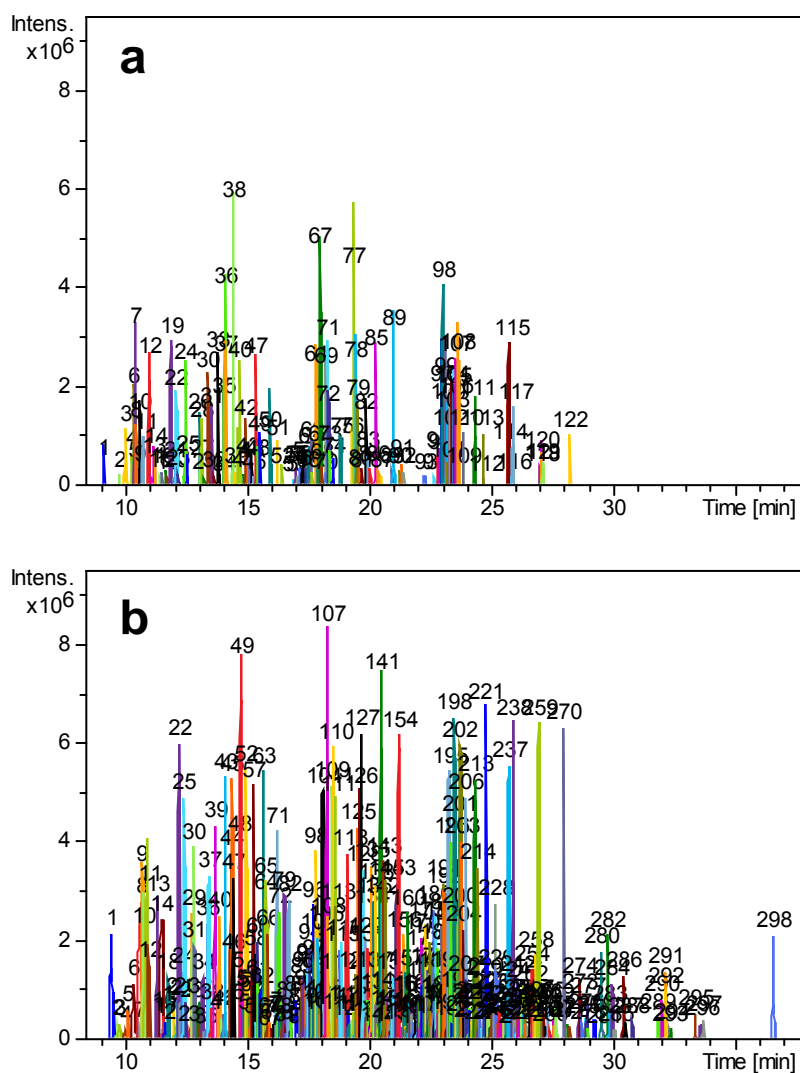


Figure 8.4 Features extracted from GC-APCI I-HRTOFMS (a) and GC-APCI II-HRTOFMS (b) base peak chromatograms of a representative supernatant sample of MiaPaCa-2 pancreatic cancer cells. The “Find Compounds – Dissect” peak finder was applied with an internal S/N threshold of ≥ 20 over the chromatographic range of 9-37 min. On average 127 and 275 features were extracted from the APCI I and APCI II data, respectively. The gain in extracted signals was the greatest for analytes eluting after 25 minutes; it was approximately eightfold higher for the APCI II than the APCI I source. Reprinted from [13].

Eliminating signals also present in blank samples and multiple signals found within a retention time and mass window of 0.1 min and 10 mDa, a revised total of 96 and 251 dissect compounds was obtained from the APCI I and APCI II data, respectively. It should be noted that these numbers do not equal individual metabolites. Certain metabolites yield multiple GC peaks such as partly silylated amino acids or cis-trans isomers of methoxime derivatives of carbonyl compounds (see **Table 8.3**).

For compound identification different strategies were pursued. Accurate masses were initially searched in the GC-APCI-HRMS online database established by Pacchiarotta et al. [74]. This database also contains retention indices based on n-alkanes used as reference. Further, GC×GC-EI-TOFMS data of the supernatant samples were used to support compound identification. Based on the mass of the $[M+H]^+$ ions of tentatively annotated metabolites, m/z values corresponding to the respective EI molecular ions and potential $[M-CH_3]^+$ fragment ions were calculated, searched in the GC×GC-EI-TOFMS data and then subjected to a NIST 05 library search. The $[M-CH_3]^+$ fragment ion was used, because silylated metabolites often form this fragment with EI ionization, while the molecular ion is low abundant or not detectable. To make use of the retention indices stored in the GC-APCI-HRMS online database, n-alkanes (C8-C20 and C21-C40) were run within the same batch as the supernatant samples on the GC×GC-EI-TOFMS, retention indices were calculated and then compared to those stored in the GC-APCI-HRMS online database. Based on the fatty acids, $[^2H_7]$ trans-cinnamate and 22 metabolites (marked in **Table 8.3**), a prediction model for GC-APCI-HRTOFMS retention times based on GC×GC-EI-TOFMS data and vice versa was established. The 22 selected metabolites were unambiguously identified by means of their accurate mass, retention index and NIST 05 match >800. The aim was to comprehensively exploit GC-APCI-HRTOFMS and GC×GC-EI-TOFMS data for compound identification, which is beneficial for metabolite annotation as previously reported by other groups [78,79]. Linear correlation of GC-APCI-HRTOFMS and GC×GC-EI-TOFMS retention times of the seven standards (fatty acids and $[^2H_7]$ trans-cinnamate) and 22 metabolites was obtained ($R^2=0.999$) and GC-APCI-HRTOFMS retention indices based on n-alkanes could be deduced from the corresponding fatty acids-based indices.

Twenty-two compounds that are not included in the online database could be assigned by mass spectral matching (NIST 05 match >700 in all cases) and/or by matching of accurate masses and retention indices with an in-house GC-APCI-HRTOFMS metabolite database. In total, 63 dissect compounds were annotated corresponding to 49 individual metabolites and phosphate (see **Table 8.3**).

Table 8.3 Figures of merit for identified features in cell culture supernatants of MiaPaCa-2 cells. Features were extracted with an internal S/N threshold of ≥ 20 . A total of 63 features could be assigned to 49 individual metabolites and phosphate representing approximately 46% and 24%, respectively, of all signals detected in GC-APCI I and II-HRTOFMS measurements. Twelve signals annotated with asterisks were identified solely by GC-APCI II-HRTOFMS in at least two out of three biological replicates. RI values based on fatty acids <900 were obtained through extrapolation. MeOx, group introduced by methoximation; std, analysis of an authentic reference standard; TMS, trimethylsilyl group; db, GC-APCI-HRMS metabolite database by Pacchiarotta et al. [74]. Reprinted from [13].

No. ^a	Dissect compound	Elemental formula [M+H] ⁺	Ret.time (min)	Ref. RI ^{b,c}	Δ RI ^d (%)	m/z expt (Da)	Δ m ^e (mDa)	Identification basis
1	Pyruvate+1MeOx+1TMS+H	C ₇ H ₁₆ NO ₃ Si ⁺	10.73	458 ^b	-0.9	190.0920	-2.6	NIST; std
2 ^a	Lactate+2TMS+H	C ₉ H ₂₃ O ₃ Si ₂ ⁺	10.82	475 ^b	1.2	235.1179	0.2	db; NIST; std
3	Valine+1TMS+H	C ₈ H ₂₀ NO ₂ Si ⁺	11.39	514 ^b	-0.4	190.1304	-4.6	std
4 ^a	Alanine+2TMS+H	C ₉ H ₂₄ NO ₂ Si ₂ ⁺	11.58	536 ^b	1.0	234.1354	-1.3	db; NIST; std
5	Leucine+1TMS+H	C ₉ H ₂₂ NO ₂ Si ⁺	12.49	1160 ^c	-0.5	204.1416	-0.1	db; std
6	3-Hydroxybutyrate+2TMS+H	C ₁₀ H ₂₅ O ₃ Si ₂ ⁺	12.57	617 ^b	0.7	249.1313	2.4	NIST; std
7	Proline+1TMS+H	C ₈ H ₁₈ NO ₂ Si ⁺	12.85	1180 ^c	-0.6	188.1124	-2.2	db; std ^g
8	Isoleucine+1TMS+H	C ₉ H ₂₂ NO ₂ Si ⁺	12.88	1182 ^c	-0.5	204.1389	2.6	db; std
9	3-Methyl-2-oxovalerate+1MeOx+1TMS+H	C ₁₀ H ₂₂ NO ₃ Si ⁺	12.92	644 ^b	0.5	232.1360	0.3	NIST; std
10	4-Methyl-2-oxovalerate+1MeOx+1TMS+H	C ₁₀ H ₂₂ NO ₃ Si ⁺	13.47	690 ^b	0.6	232.1354	0.9	NIST; std
11	Valine+2TMS+H	C ₁₁ H ₂₈ NO ₂ Si ₂ ⁺	13.53	694 ^b	0.4	262.1685	-3.2	NIST; std

Table 8.3 continued.

No. ^a	Dissect compound	Elemental formula [M+H] ⁺	Ret.time (min)	Ref. RI ^{b,c}	Δ RI ^d (%)	m/z expt (Da)	Δ m ^e (mDa)	Identification basis
12 ^a	Urea+2TMS+H	C ₇ H ₂₁ N ₂ O ₂ Si ₂ ⁺	13.78	1246 ^c	0.3	205.1143	4.4	db; NIST
13 ^a	Benzoate+1TMS+H	C ₁₀ H ₁₅ O ₂ Si ⁺	14.05	1254 ^c	-1.0	195.0851	-1.5	db; NIST
14	Serine+2TMS+H	C ₉ H ₂₄ NO ₃ Si ₂ ⁺	14.19	744 ^b	-0.2	250.1292	-0.2	NIST; std
15 ^a	Leucine+2TMS+H	C ₁₂ H ₃₀ NO ₂ Si ₂ ⁺	14.43	768 ^b	0.5	276.1782	2.7	db; NIST; std
16	Glycerol+3TMS+H	C ₁₂ H ₃₃ O ₃ Si ₃ ⁺	14.47	773 ^b	0.7	309.1737	-0.5	NIST; std
17 ^a	Phosphate+3TMS+H	C ₉ H ₂₈ PO ₄ Si ₃ ⁺	14.50	1284 ^c	-0.1	315.1006	2.1	db; NIST
18 ^a	Isoleucine+2TMS+H	C ₁₂ H ₃₀ NO ₂ Si ₂ ⁺	14.79	797 ^b	0.5	276.1786	2.4	db; NIST; std ^g
19	Threonine+2TMS+H	C ₁₀ H ₂₆ NO ₃ Si ₂ ⁺	14.82	795 ^b	-0.2	264.1471	-2.5	NIST; std ^f
20	Proline+2TMS+H*	C ₁₁ H ₂₆ NO ₂ Si ₂ ⁺	14.89	1308 ^c	-0.1	260.1492	0.4	db; NIST; std ^g
21 ^a	Glycine+3TMS+H	C ₁₁ H ₃₀ NO ₂ Si ₃ ⁺	15.04	815 ^b	0.1	292.1571	0.8	db; NIST; std
22	Glycerate+3TMS+H	C ₁₂ H ₃₁ O ₄ Si ₃ ⁺	15.39	847 ^b	0.5	323.1495	2.9	NIST; std
23	Fumarate+2TMS+H*	C ₁₀ H ₂₁ O ₄ Si ₂ ⁺	15.53	858 ^b	0.4	261.0955	1.8	db; NIST; std ^g
24	Serine+3TMS+H	C ₁₂ H ₃₂ NO ₃ Si ₃ ⁺	15.86	910 ^b	0.2	322.1653	3.2	NIST; std
25	Threonine+3TMS+H	C ₁₃ H ₃₄ NO ₃ Si ₃ ⁺	16.30	940 ^b	0.2	336.1826	1.5	NIST; std
26 ^a	Thymine+2TMS+H*	C ₁₁ H ₂₃ N ₂ O ₂ Si ₂ ⁺	16.50	1412 ^c	0.0	271.1268	2.5	db; NIST; std ^g
27	Methionine+1TMS+H	C ₈ H ₂₀ NO ₂ SSi ⁺	16.60	1416 ^c	-0.2	222.0955	2.4	db
28	Aspartate+2TMS+H	C ₁₀ H ₂₄ NO ₄ Si ₂ ⁺	16.77	1430 ^c	-0.1	278.1209	2.9	db; std
29	Nicotinamide+1TMS+H	C ₉ H ₁₅ N ₂ O ₂ Si ⁺	17.63	1029 ^b	-0.3	195.0905	4.3	NIST; std
30	Malate+3TMS+H	C ₁₃ H ₃₁ O ₅ Si ₃ ⁺	17.79	1044 ^b	0.2	351.1433	4.0	NIST; std

Table 8.3 continued.

No. ^a	Dissect compound	Elemental formula [M+H] ⁺	Ret.time (min)	Ref. RI ^{b,c}	Δ RI ^d (%)	<i>m/z</i> expt (Da)	Δ m ^e (mDa)	Identification basis
31	Adipate+2TMS+H	C ₁₂ H ₂₇ O ₄ Si ₂ ⁺	17.95	1054 ^b	0.1	291.1415	2.7	std ^f
32	Glutamate+2TMS+H*	C ₁₁ H ₂₆ NO ₄ Si ₂ ⁺	18.12	1536 ^c	1.0	292.1371	2.4	db ^g
33	Erythritol+4TMS+H	C ₁₆ H ₄₃ O ₄ Si ₄ ⁺	18.15	1070 ^b	0.3	411.2219	1.4	NIST; std
34 ^a	Methionine+2TMS+H	C ₁₁ H ₂₈ NO ₂ SSi ₂ ⁺	18.26	1532 ^c	0.0	294.1352	2.2	db; NIST
35	Phenylalanine+1TMS+H	C ₁₂ H ₂₀ NO ₂ Si ⁺	18.66	1556 ^c	-0.3	238.1270	-1.2	db; std
36	α -ketoglutarate+1MeOx+2TMS+H	C ₁₂ H ₂₆ NO ₅ Si ₂ ⁺	19.01	1132 ^b	0.3	320.1316	2.8	NIST; std ^g
37	Phenylpyruvate+1MeOx+1TMS+H*	C ₁₃ H ₂₀ NO ₃ Si ⁺	19.22	1145 ^b	0.0	266.1178	2.9	NIST; std ^g
38 ^a	Phenylalanine+2TMS+H	C ₁₅ H ₂₈ NO ₂ Si ₂ ⁺	19.81	1644 ^c	0.0	310.1660	-0.7	db; NIST; std
39	Laurate+1TMS+H	C ₁₅ H ₃₃ O ₂ Si ⁺	19.97	1201 ^b	-0.1	273.2230	1.4	NIST; std
40 ^a	Ribose+1MeOx+4TMS+H	C ₁₈ H ₄₆ NO ₅ Si ₄ ⁺	20.60	1700 ^c	-0.2	468.2409	3.9	db; NIST
41 ^a	Xylitol+5TMS+H	C ₂₀ H ₅₃ O ₅ Si ₅ ⁺	21.18	1734 ^c	-0.8	513.2738	-0.4	db; NIST
42	Ornithine+3TMS+H	C ₁₄ H ₃₇ N ₂ O ₂ Si ₃ ⁺	21.46	1767 ^c	-0.2	349.2116	4.2	db; std
43	Glycerol-1-phosphate+4TMS+H	C ₁₅ H ₄₂ PO ₆ Si ₄ ⁺	21.64	1333 ^b	0.2	461.1748	4.3	std
44 ^a	Ornithine+4TMS+H	C ₁₇ H ₄₅ N ₂ O ₂ Si ₄ ⁺	22.32	1387 ^b	0.0	421.2545	0.8	db; NIST; std
45	Myristate+1TMS+H*	C ₁₇ H ₃₇ O ₂ Si ⁺	22.49	1838 ^c	-0.8	301.2545	1.2	db; NIST ^g
46 ^a	Hippurate+1TMS+H	C ₁₂ H ₁₈ NO ₃ Si ⁺	22.51	1414 ^b	-0.9	252.1014	3.6	db; NIST; std ^g
47	Lysine+3TMS+H	C ₁₅ H ₃₉ N ₂ O ₂ Si ₃ ⁺	22.70	1418 ^b	0.0	363.2304	1.0	db; std
48 ^a	Tyrosine+2TMS+H	C ₁₅ H ₂₈ NO ₃ Si ₂ ⁺	23.11	1907 ^c	0.2	326.1592	1.1	db; NIST; std
49	Fructose+1MeOx+5TMS+H	C ₂₂ H ₅₆ NO ₆ Si ₅ ⁺	23.27	1457 ^b	-0.5	570.2966	-1.8	NIST; std

Table 8.3 continued.

No. ^a	Dissect compound	Elemental formula [M+H] ⁺	Ret.time (min)	Ref. RI ^{b,c}	Δ RI ^d (%)	<i>m/z</i> expt (Da)	Δ m ^e (mDa)	Identification basis
50	Lysine+4TMS+H	C ₁₈ H ₄₇ N ₂ O ₂ Si ₄ ⁺	23.60	1490 ^b	-0.1	435.2673	3.6	db; NIST; std ^g
51 ^a	Tyrosine+3TMS+H	C ₁₈ H ₃₆ NO ₃ Si ₃ ⁺	23.82	1509 ^b	-0.1	398.1949	4.8	db; NIST; std ^g
52	Sorbitol+6TMS+H	C ₂₄ H ₆₃ O ₆ Si ₆ ⁺	23.99	1974 ^c	-0.3	615.3266	-3.1	db; std
53 ^a	Pantothenate+3TMS+H	C ₁₈ H ₄₂ NO ₅ Si ₃ ⁺	24.44	2016 ^c	-0.2	436.2326	3.9	db; NIST; std
54	Palmitoleate+1TMS+H*	C ₁₉ H ₃₉ O ₂ Si ⁺	24.61	2030 ^c	-0.3	327.2702	1.2	db ^g
55 ^a	Palmitate+1TMS+H	C ₁₉ H ₄₁ O ₂ Si ⁺	24.82	2049 ^c	-0.3	329.2845	2.6	db; NIST
56	Hippurate+2TMS+H*	C ₁₅ H ₂₆ NO ₃ Si ₂ ⁺	25.24	1628 ^b	-0.8	324.1402	4.4	std ^g
57 ^a	Myo-inositol+6TMS+H	C ₂₄ H ₆₁ O ₆ Si ₆ ⁺	25.74	1681 ^b	-0.2	613.3059	2.0	db; NIST; std
58	Linoleate+1TMS+H*	C ₂₁ H ₄₁ O ₂ Si ⁺	26.66	2218 ^c	-0.5	353.2838	3.2	db ^g
59	Oleate+1TMS+H*	C ₂₁ H ₄₃ O ₂ Si ⁺	26.71	2222 ^c	-0.5	355.2990	3.7	db ^g
60	Cis-Vaccenate+1TMS+H*	C ₂₁ H ₄₃ O ₂ Si ⁺	26.77	2229 ^c	-0.5	355.2989	3.8	db ^g
61 ^a	Tryptophan+2TMS+H*	C ₁₇ H ₂₉ N ₂ O ₂ Si ₂ ⁺	26.85	1808 ^b	0.7	349.1753	1.0	db; NIST; std ^g
62 ^a	Stearate+1TMS+H	C ₂₁ H ₄₅ O ₂ Si ⁺	26.96	2246 ^c	-0.6	357.3154	2.9	db; NIST
63	Tryptophan+3TMS+H	C ₂₀ H ₃₇ N ₂ O ₂ Si ₃ ⁺	26.99	2250 ^c	-0.5	421.2118	4.0	db; std ^g

^a Compound used for GC-APCI II-HRTOFMS and GC×GC-EI-TOFMS retention time correlation; ^b RI calculation based on fatty acids; ^c RI calculation based on alkanes; ^d Difference between RI in a standard and sample based on either fatty acids or alkanes relative to the determined RI in the sample; ^e Difference between predicted and measured mass; ^f Compound only extracted from GC-APCI I-HRTOFMS data; ^g Compound only extracted from GC-APCI II-HRTOFMS data.

Metabolite annotation was confirmed by the correct elemental formula using the Smart-Formula tool and, if available, by analysis of a reference standard. Using the minimum reporting standards proposed by Sumner et al. for metabolomics data [26], 48 level 1 and 15 level 2 identifications were obtained. Experimental masses deviated less than 5 mDa from their calculated m/z values. Relative differences in retention indices did not exceed the threshold of 1.0% proposed by Strehmel et al. [128], except for lactate due to its early elution as previously reported [79]. Nevertheless, lactate identification was confirmed by a standard. The applied identification strategy is specified in **Table 8.3** for all annotated compounds including information on retention time and accurate mass. Assigned compounds represented 46% and 24% of the total number of compounds extracted from APCI I and APCI II data, respectively. Twelve dissect compounds annotated with asterisks in **Table 8.3** were only extracted from the APCI II data of at least two supernatant replicates due to the improved signal-to-noise ratio obtained with the redesigned source. This could be confirmed by lowering the S/N threshold to 2, which then enabled the extraction of these twelve signals also from the APCI I data. Three additional low abundant compounds, namely α -ketoglutarate-MeOx-2TMS, hippurate-1TMS, and tryptophan-3TMS were only extracted from the GC-APCI II-HRTOFMS data of a single supernatant replicate but not found in the APCI I data set. Further, threonine-2TMS and adipate-2TMS were exclusively extracted from APCI I data and proline-1TMS, isoleucine-2TMS, lysine-4TMS and tyrosine-3TMS from APCI II data, although their respective peaks were fairly abundant. Since co-eluting compounds were present, this was attributed to failure of deconvolution by the dissect algorithm in these cases.

For determination of repeatability, the “Find Molecular Feature” algorithm was used to extract features from the chromatograms of the three biological replicates, which were then aligned in a single data matrix using the Profile Analysis Tool to compare RSD values of the area integrals obtained with the two APCI sources. Extraction and bucketing of features was performed separately for APCI I and APCI II data. Area integrals of a total of 412 (APCI I) and 1462 (APCI II) features were obtained and normalized to [$^2\text{H}_7$]trans-cinnamate. Next, RSDs were calculated for all features in both data sets. Cumulative distributions of RSD values for data acquired with both sources are shown in **Figure 8.5**. RSD values were distinctly ($p < 0.05$, Wilcoxon signed rank test) lower for

APCI II measurements, as indicated by the steeper gradient over the first part of the corresponding curve. The median RSDs decreased from 33% to 24% using APCI II.

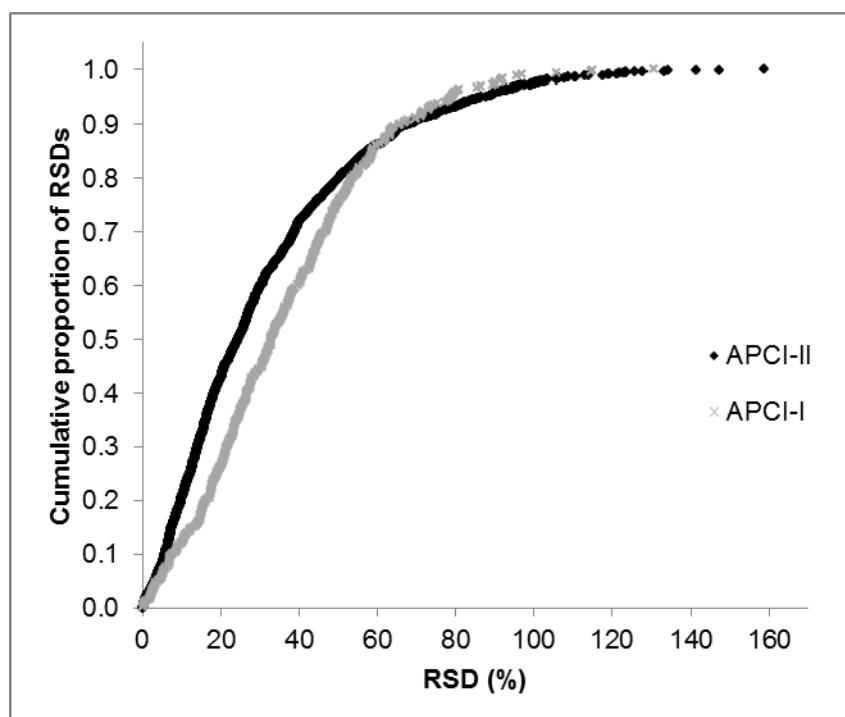


Figure 8.5 Cumulative distributions of area RSDs of extracted GC-APCI I and II-HRTOFMS features from cell culture supernatant of MiaPaCa-2 cells. A total of 412 (APCI I) and 1462 (APCI II) features were extracted over 9-37 min using the “Find Compounds – Molecular Features” algorithm and normalized to [$^2\text{H}_7$]trans-cinnamate. RSD values were distinctly ($p < 0.05$, Wilcoxon signed rank test) lower in case of APCI II measurements. Reprinted from [13].

Overall, improved peak abundances and shapes, as well as better ionization repeatability were attained by APCI II. This is exemplified in **Figure 8.6**, in which the extracted ion chromatograms of fumarate-2TMS and nicotinamide-1TMS from three injection replicates of a supernatant sample using either the APCI I or the APCI II source are overlaid.

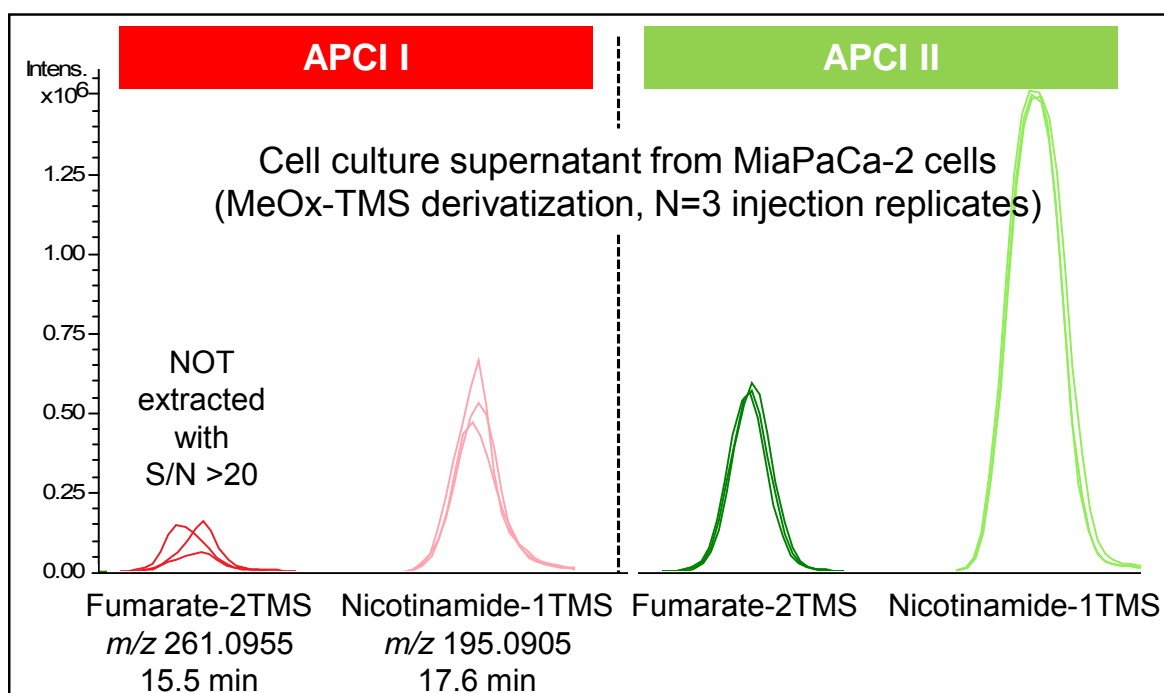


Figure 8.6 Influence of APCI source type on peak abundance and repeatability illustrated by the example of two metabolites identified in the supernatant of MiaPaCa-2 cells. Three injection replicates were analyzed and m/z values of $[M+H]^+$ of fumarate-2TMS and nicotinamide-1TMS were extracted. APCI II in comparison to APCI I resulted in more abundant peaks and yielded better repeatability of the ionization. Reprinted from [13].

8.4 Conclusions

The redesigned APCI II ion source provides lower quantification limits, broader linear ranges, and improved repeatability due to less ragged peak profiles in comparison to its predecessor. This is particularly obvious for late-eluting compounds due to an optimized heat distribution in the ionization chamber, which minimizes adsorption of high-boiling analytes to cold spots. Applied to a complex biological matrix, the APCI II source increased the number of identified metabolites by one third, making GC-APCI II-HRTOFMS a true alternative to GC \times GC-EI-TOFMS, especially given easier data analysis and the feasibility of coupling GC via the APCI II source to ultra-high-performance FTICR MS.

9 Development of an enantioselective quantitative profiling method for the oncometabolite D-2-hydroxyglutarate

9.1 Introduction

Analysis of 2-hydroxyglutarate (2-HG) enantiomers has gained in importance in recent years. Apart from metabolic diseases such as D- or L-2-hydroxyglutaric aciduria [149,150], the accumulation of D-2-HG in several types of cancers harboring neomorphic mutations in the isocitrate dehydrogenase 1 and 2 genes has become a hot topic in cancer biology [151]. IDH1 and IDH2 are enzymes that catalyze the decarboxylation of isocitrate to form α -ketoglutarate using NADP⁺ and producing NADPH. IDH1 is a cytosolic and peroxisomal enzyme while IDH2 is found in the mitochondria. In different types of human cancers, somatic point mutations have been found that result in a loss of normal catalytic activity and concomitantly a gain of function, which is the formation of D-2-HG from α -ketoglutarate. Mutations in IDH1 or IDH2 are found frequently in grade 2 – 3 gliomas and secondary glioblastomas and in 5-20% of acute myeloid leukemia (AML) patients, as well as other tumor types [152]. **Figure 9.1** gives an overview over normal functions of the IDH1 and IDH2 enzymes in the cytoplasm and mitochondrion, respectively, together with other related cellular metabolic processes (top picture) and depicts reactions catalyzed by wild type and mutant IDH1 and IDH2 enzymes (lower picture) in the interconversion of isocitrate, 2-oxoglutarate and (R)-2-HG, respectively.

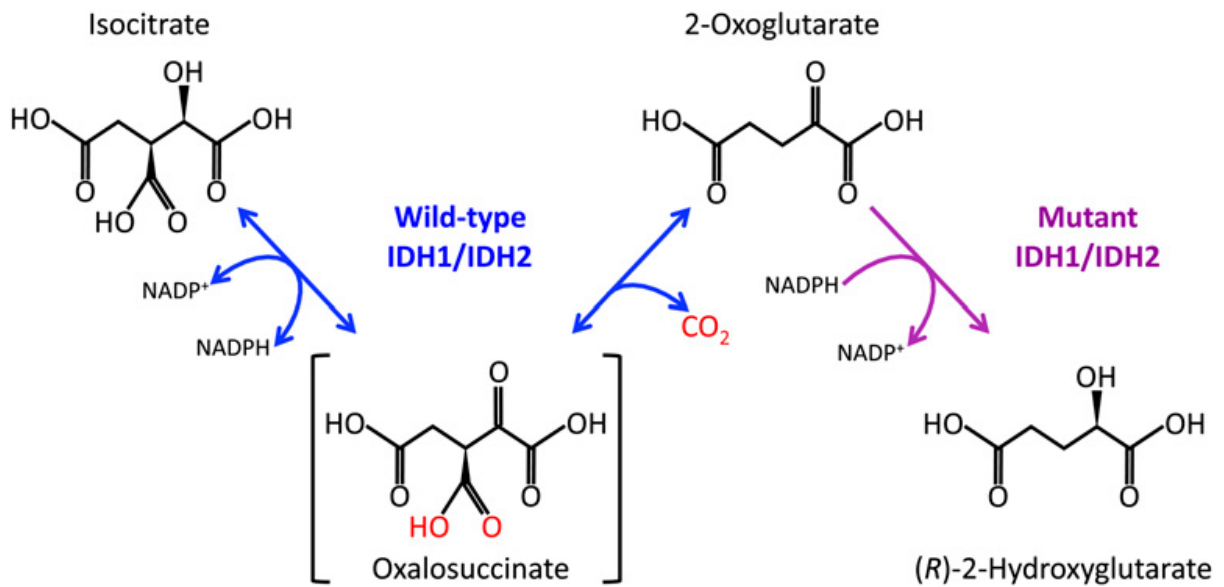
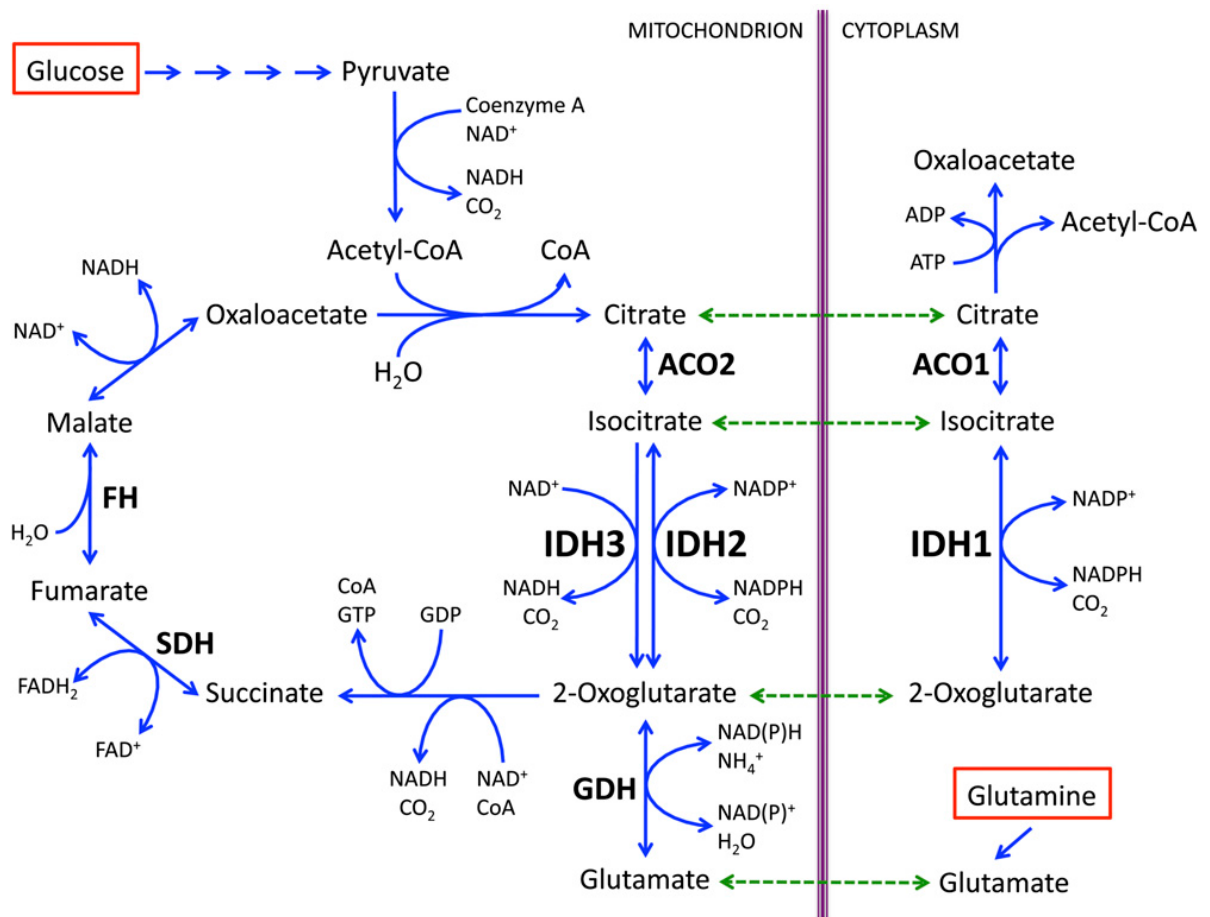


Figure 9.1 Normal enzymatic functions of IDH 1 and IDH2 together with related cellular metabolic processes in the cytoplasm and mitochondrion (top picture) as well as reactions catalyzed by wild type and mutant IDH1 and IDH2 enzymes (lower picture) in the interconversion of isocitrate, 2-oxoglutarate and (R)-2-HG, respectively. Reprinted from [153].

Several analytical methods have been used to study the role of D-2-HG in the formation and malignant progression of cancers, such as an enzymatic assay [154], NMR spectroscopy [155] as well as LC [156] or GC [155,157,158] coupled to MS. GC-MS has been most widely applied owing to its high robustness and ability to achieve low quantification limits. However, existing weaknesses among current procedures include unintended [158,159] or poor [157] chromatographic separation of D and L enantiomers of 2-HG, thus limiting the analysis to total 2-HG levels. Furthermore, indirect methods separating diastereomeric 2-HG derivatives on a non-chiral column have been reported [149], but inadvertent racemization and side reactions caused by contaminants in the derivatizing reagent may lead to erroneous quantitative results. On the other hand, direct enantioselective analysis of D/L-2-HG has been successfully applied [160]; however, applied derivatization strategy was neither simple nor fast and, according to the authors, had to be performed very carefully, which would render its automation difficult. Irrespective of whether D/L-2-HG enantiomers were chromatographically resolved, quantitative data based on MS measurements is provided to a very limited extent [159], completely lacking [155], or a corresponding SIL-IS of 2-HG was not included for quantification [155,158], which may severely compromise reproducibility and accuracy. On top of that, determination of peak areas or concentration levels using GC-MS has been performed with electron ionization (EI) [157-160], which results in extensive fragmentation of molecular ions in general. Consequently, a fragment of lower mass has to be selected as quantification trace for the target analyte, which might be less specific and common to co-eluting compounds across various matrices hampering accurate quantification or restricting general applicability of the method.

Derivatization of 2-HG with methylchloroformate/methanol (MCF/MeOH) is simple and fast and lends itself to further automation as established for propyl chloroformate-derivatized amino acids [38]. Further, in a previous study of our group, MCF/MeOH derivatization and separation on a chiral CD-based Rt- γ DEXsa capillary column yielded excellent resolution of amino acid enantiomers [39]. Soft ionization techniques, such as CI and APCI, meanwhile have a proven track record in metabolomics as an alternative to conventional GC-EI-MS due to higher selectivity and abundance of $[M+H]^+$ ions over fragment ions [9,17,86,136].

In this chapter, the development of an enantioselective approach for the quantitative determination of D/L-2-HG using GC-APCI II coupled to a TOF-MS is presented based on the separation of MCF-derivatized D- and L-2-HG enantiomers on a chiral Rt- γ DEXsa column. D/L-[2,3,3-²H₃]2-HG was included as internal standard for accurate and robust quantification. The general applicability of the method was shown in different matrices including human serum, urine, cell extracts and cell culture supernatants.

Sample preparation was done together with Claudia Samol and Nadine Nürnberger. Wolfram Gronwald and Magdalena Waldhier carried out NMR analysis and initial GC-EI-qMS measurements, respectively. Katja Dettmer set up the quantitative HPLC-ESI-MS/MS method for the determination of 2-HG in urine specimens.

The results of this chapter will be part of a manuscript in preparation.

9.2 Methods

9.2.1 Sample preparation

An aqueous stock solution of racemic D/L-2-HG at a concentration of 1 mM was prepared and geometrically diluted over 17 points (0.02-1000 μ M) for calibration purposes. Deuterium-labeled racemic D/L-2-HG was used as internal standard and an aqueous working solution with a concentration of 100 μ M was prepared. For preparation of standard samples, 150 μ L of stock solution or a dilution of racemic D/L-2-HG, according to the desired final standard concentration, were transferred with 10 μ L of internal standard to a 2-mL glass vial for derivatization. Addition of standard and IS to supernatant or serum specimens was carried out prior to the precipitation step unless otherwise stated under Results and Discussion. To check for racemization, separate aqueous stock solutions of individual D and L enantiomers of 2-HG were prepared and 30 μ L of a 1 mM dilution of them were derivatized without IS.

The entire set of biological samples analyzed in this project including serum, urine, and cell culture samples was provided by collaborators at the University Hospital of Re-

gensburg. More details are provided under Results and Discussion. The samples were stored at -80°C until preparation. For extraction of cell pellets, 10 μL of IS (100 μM) were added to the pellet followed by 600 μL cold 80% MeOH (80/20 MeOH/H₂O, v/v). The sample was vortexed and centrifuged at 4°C and $9500 \times g$ for 5 min. The supernatant was collected and the pellet was washed with 200 μL cold 80% MeOH, centrifuged at 4°C and $9500 \times g$ for 5 min and finally washed with 200 μL H₂O and centrifuged at 4°C and $11200 \times g$ for 5 min. The supernatants were combined and evaporated to dryness using a vacuum evaporator. For preparation of serum and cell culture supernatants, 150 μL of sample were spiked with 10 μL of internal standard and 600 μL of cold MeOH was added for protein precipitation. Then, the procedure described above for cell pellets was followed with the exception that the second wash of the protein pellet was also performed with 80% MeOH. Finally, derivatization of serum specimens with MCF/MeOH was performed directly in the aqueous phase using 150 μL of serum and 10 μL of IS without prior protein precipitation. The latter protocol was also applied to urine samples. Manual derivatization with MCF/MeOH was carried out as described under 5.2.3.

9.2.2 Analytical approaches

GC-APCI-TOFMS

As denoted in Results and Discussion, calibration was repeated several times in the course of method development using varying instrumental or operational settings. In fact, an APCI I source with water infusion was used, an injection volume of 2 μL was used, or no water was infused instead of the conventional settings described in subchapter 5.3.

D/L-2-HG calibration curves based on either one of the two derivatives for each enantiomer were obtained as described in more detail under 5.4.2. Method #1 in **Table 9.1** was employed in the spike-in experiment, whereas method #4 (**Table 9.1**) was used for quantification of urine and serum samples in section 9.3.4.

HPLC-ESI-MS/MS

High-performance liquid chromatography – electrospray ionization – tandem mass spectrometry (HPLC-ESI-MS/MS) in negative ionization mode was employed as reference method for the determination of total levels of (D+L)-2-HG. The optimized method was capable of distinguishing 2-HG from 5-oxotetrahydro-2-furancarboxylate and also allowed quantification of the latter. A mixture containing 2-HG and 5-oxotetrahydro-2-furancarboxylate at a concentration of 2 mM in water was serially diluted covering 16 points over 0.06-2000 μM . For calibration, 90 μL of standard were transferred into a glass vial with a glass insert along with 10 μL of internal standard. The latter was obtained from 500 μL of the original aqueous stock solution of deuterium-labeled D/L-2-HG that was incubated with 200 μL of formic acid at 100°C for 30 min, evaporated to complete dryness, re-suspended in 500 μL of water, and diluted to a final concentration of 100 μM . To determine 2-HG levels in urine and serum, 10 μL of urine were added to 10 μL of internal standard and 80 μL of water, and 50 μL of serum were precipitated with 250 μL of 100% MeOH and washed twice with 100 μL of 80% MeOH as described above. Dried extracts were reconstituted in 50 μL of water.

Using an Agilent 1200 SL HPLC system (Böblingen, Germany), separation of analytes was achieved on a Supelco Discovery HS F5 (150 \times 2.1 mm i.d., 3 μm) PFPP column. Elution with mobile phases A (0.1% formic acid in water, v/v) and B (acetonitrile) was performed as follows: 0-6.5 min, held at 100% A at a flow rate of 0.2 mL/min; 6.5-8 min, 0-100% B at 0.35 mL/min; 8-10 min, held at 100% B at 0.35 mL/min; 10-10.1 min, 0-100% A at 0.35 mL/min; 10.1-17 min, held at 100% A at 0.35 mL/min; 17-18 min, held at 100% A at 0.2 mL/min. The injection volume was 5 μL , and the column temperature was 30°C. MS/MS analysis was performed in MRM mode on a 4000 QTRAP from AB Sciex using one selective transition each for the unlabeled and labeled analytes as follows: m/z 147.1 (M-H)⁻ to m/z 84.8 (product ion) for 2-HG, m/z 150.1 (M-H)⁻ to m/z 87.8 (product ion) for D₃-labeled 2-HG, m/z 128.8 (M-H)⁻ to m/z 100.8 (product ion) for 5-Oxotetrahydro-2-furancarboxylate, and m/z 131.8 (M-H)⁻ to m/z 88.0 (product ion) for D₃-labeled 5-Oxotetrahydro-2-furancarboxylate. To obtain the calibration curves from the calibrator samples, nominal concentration ratio (analyte/internal standard) was plot-

ted against peak area ratio (analyte/internal standard), whereby the corresponding labeled analogue was used for normalization for each target analyte as internal standard. Acquisition and analysis of data was carried out in Analyst version 1.5.1 from AB Sciex.

Creatinine values were determined as previously described [161].

NMR spectroscopy

Serum specimens as well as aqueous and MCF derivatized standards of 2-HG and 5-oxotetrahydro-2-furancarboxylate (1 mM each) were analyzed. Five hundred μL of serum were filtered using Nanosep centrifugal devices (Pall Corporation, Port Washington, NY) with a 3K molecular weight cutoff. Four hundred μL of filtrate, aqueous standard, or derivatized standard that had been carefully evaporated and reconstituted in water, were transferred into an NMR tube, followed by the addition of 200 μL of potassium phosphate buffer, pH 7.4, and 50 μL of 29.02 mM TSP in D_2O as internal standard. Following established protocols [162], 1D ^1H and 2D ^1H - ^{13}C TOCSY spectra were acquired on a 600MHz Bruker Avance III (Bruker BioSpin GmbH, Rheinstetten, Germany) employing a triple resonance (^1H , ^{13}C ^{31}P , ^2H lock) cryogenic probe equipped with z-gradients and an automatic cooled sample changer. Amix viewer V3.9.13 (Bruker BioSpin) was used for manual inspection of spectra. Signals were assigned based on the predicted ^1H chemical shifts by ChemDraw and assuming similar ^1H intensities for methyl groups of each of the three derivatives, whereas the identity of the methylchloroformate signal was confirmed by comparison to a blank sample.

9.3 Results and discussion

9.3.1 Selection of a specific quantifier ion trace from APCI-TOFMS versus EI-qMS mass spectra

2-HG is a polar molecule that requires derivatization prior to GC analysis. Based on the experiences gained with the enantioselective GC analysis of amino acids in our laboratory, preliminary experiments were performed using derivatization with methyl chloroformate/methanol (MCF/MeOH) and separation on an Rt- γ DEXsa column [39].

GC-MS analysis yielded four different peaks for a derivatized 2-HG racemate, of which the late-eluting pair was more abundant and, hence, initially used for quantification. It corresponded to the D and L enantiomers of the methyl ester of 5-oxotetrahydro-2-furancarboxylate formed as the condensation product from racemic D/L-2-HG. The example of a GC-APCI-TOFMS base peak chromatogram of a derivatized racemic standard of D/L-2-HG (100 μ M each for the two enantiomers) is given in **Figure 9.5**, with structural formulas provided for the two pairs of derivatives (denoted as D/L-2-HG-1MC-2E and D/L-2-HG-1E according to **Figure 9.5**). In the case of the more abundant methyl esters of D/L-5-oxotetrahydro-2-furancarboxylate, extensive fragmentation was observed in EI mass spectra, which were dominated by a fragment mass of m/z 85, while the relative intensity of the molecular ion with an m/z of 144 Da was <2% of the base peak in the absence of other ion species. However, m/z 85 as quantifier trace proved nonspecific for accurate D-2-HG quantification in the initial analysis of cell culture supernatants of DLD-1 colon cancer cells, as a highly abundant matrix compound present in cell culture media supplemented with bovine serum albumin yielded the same mass. Furthermore, such a low mass is probably common to several additional metabolites.

APCI as a soft ionization technique on the other hand preserved the $[M+H]^+$ ion of the methyl esters of D/L-5-oxotetrahydro-2-furancarboxylate (m/z $[M+H]^+$: 145.050 Da) by suppression of in-source fragmentation, while TOFMS analysis enabled the extraction of quantifier ions with a narrow width of ± 10 mDa. With such an enhanced specificity, it

became clearly obvious that a matrix compound with a mass of m/z 220.086 was co-eluting with D-2-HG in the initially analyzed cell culture supernatant sample, as shown in **Figure 9.2B**, which had not been obvious from the EI mass spectrum of the same peak (**Figure 9.2A**). Hence, further method development was solely carried out on the GC-APCI-TOFMS instrument. Similar to these observations, such a benefit of soft ionization in terms of specificity has also been shown for the quantitative profiling of amino and nonamino organic acids by GC-MS employing CI coupled to MS/MS analysis [86] as well as in the case of pyrethroid insecticides by means of APCI-MS/MS [82].

Despite the possibility to distinguish between D-2-HG-1E and the co-eluting matrix compound based on unique masses, the temperature ramp was modified to chromatographically separate these two compounds to minimize possible matrix effects. As discussed in chapter 7, high amounts of co-eluting analytes or matrix compounds may suppress the signal of an analyte in APCI. Given that the signal of D-2-HG-1E that was spiked at a concentration of 50 μM into the supernatant sample was approximately as abundant as that of the co-eluting compound (**Figure 9.2**), comparably high amounts of the latter could be assumed to be present, possibly deteriorating lower quantification limits of the method in these samples due to suppression of the signal of D-2-HG-1E. Among the three different heating rates tested (4, 5, and 6°C/min), no co-elution occurred solely in the case of 5°C/min for both 2-HG-1E enantiomers in the analysis of a randomly selected urine, serum, cell extract and supernatant samples. Hence, further experiments were carried out using 5°C/min as heating rate. As depicted in **Figure 9.3**, huge concentration differences of D- and L-2-HG could be initially detected in cell extracts and supernatants of dendritic cells from an IDH2 R140Q mutation transfected and untransfected cell line, as well as serum of an AML patient with an IDH2 R140Q mutation versus healthy control. The L enantiomer was not detected. The IS of D/L-2-HG was poorly recovered (~30%) due to extraction losses for these biological samples. Hence, the corresponding peaks of D/L-[2,3,3- $^2\text{H}_3$]2-HG are small or not visible in **Figure 9.3** due to the highly abundant endogenous metabolites and the resulting scaling of the chromatogram. It should be noted that the peaks for D/L-[2,3,3- $^2\text{H}_3$]2-HG still exceeded a S/N ratio of 10.

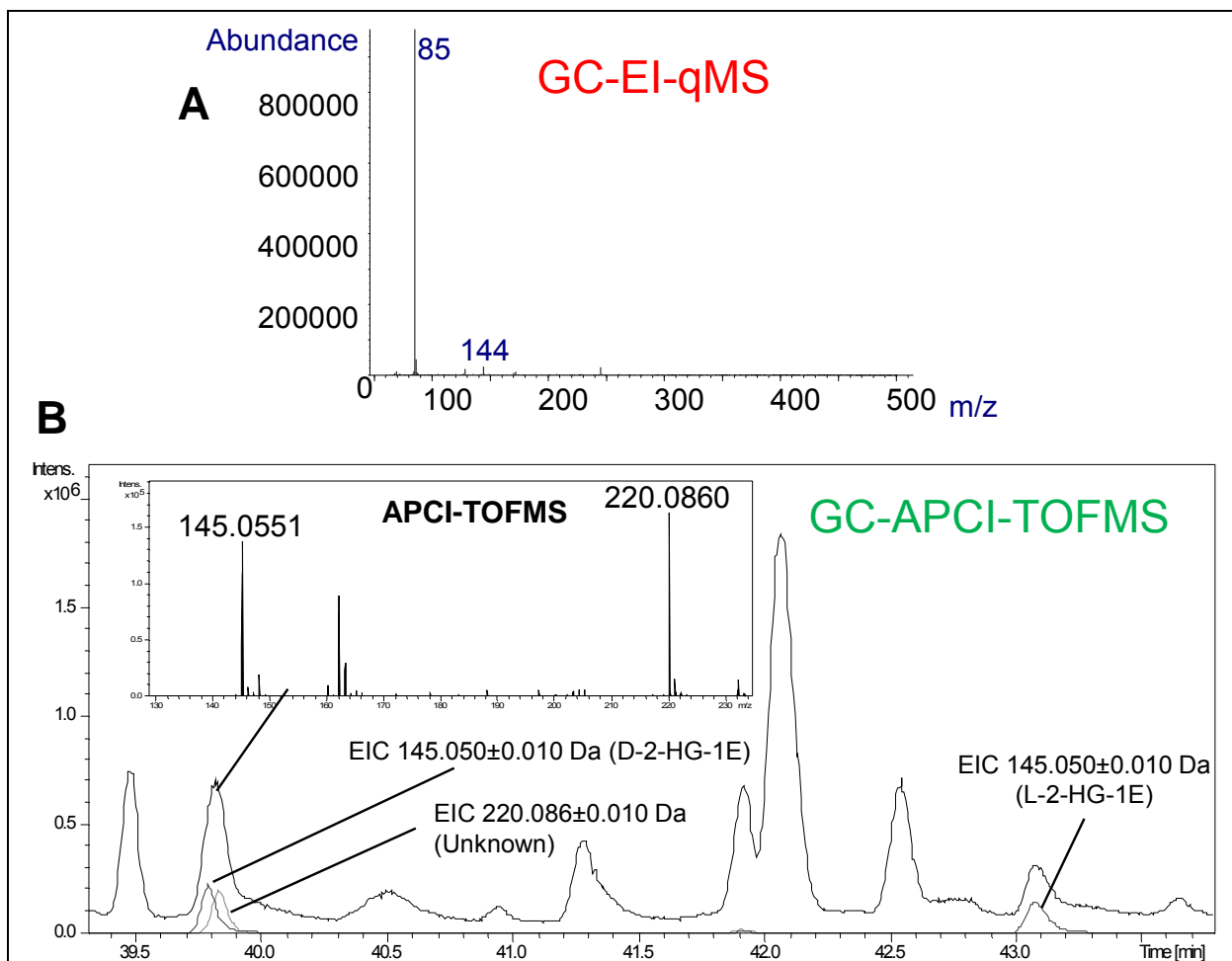


Figure 9.2 Ionization of D-2-HG-1E and a co-eluting matrix component by APCI versus EI. For illustration, a cell culture supernatant sample containing huge amounts of the co-eluting compound, from DLD-1 colon cancer cells grown in DMEM medium supplemented with 10% fetal calf serum, was spiked with racemic D/L-2-HG (50 μ M of each enantiomer), fortified with MeOH for protein precipitation, MCF-derivatized and analyzed by GC-APCI-TOFMS and GC-EI-qMS. GC-EI-qMS method parameters were as previously described in [39] (method E) and adopted for GC-APCI-TOFMS. (B) Extracted ion chromatograms based on unique masses from the APCI-TOFMS mass spectrum revealed that D-2-HG-1E was co-eluting with an unknown compound, whereas (A) the EI mass spectrum for the same chromatographic peak only yielded m/z 85 as base peak in the virtual absence of other characteristic ions. E, methyl ester group.

Similar peak profiles in extracted ion chromatograms of D/L-2-HG-1E obtained from a standard and the latter biological samples demonstrate the formation of a selective mass trace for this compound by APCI (**Figure 9.3**). Furthermore, GC-APCI-TOFMS with MCF derivatization and separation on the chiral CD-based column provided high chromatographic resolution of the two enantiomers of 2-HG-1E, which were separated by more than 2.5 min.

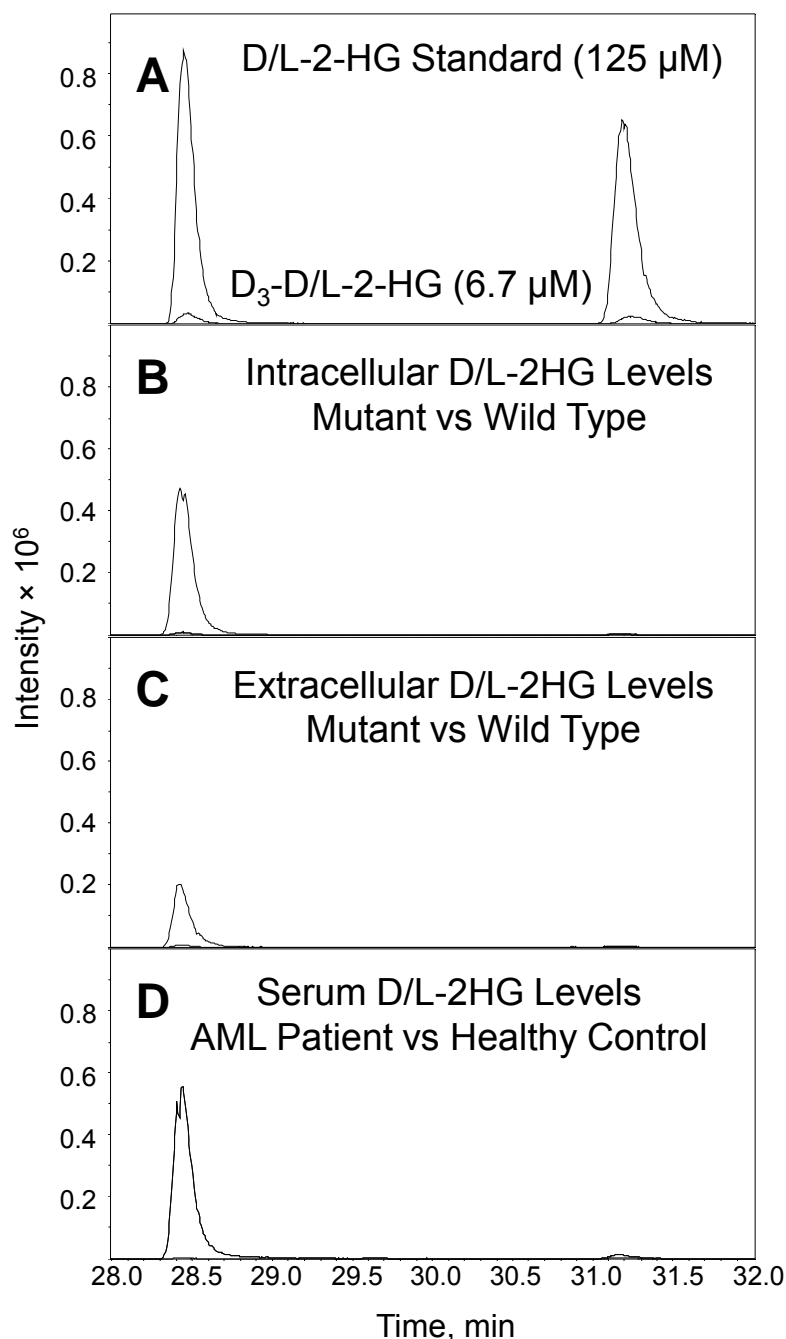


Figure 9.3 Extracted ion chromatograms of D/L-2-HG-1E from a standard and biological samples. (A) Standard sample containing both unlabeled (total (D+L) concentration of 125 μM) and D₃-labeled (minor peak, total (D+L) concentration of 6.7 μM) racemic D/L-2-HG. (B) Methanolic cell extracts of IDH2 R140Q mutation transfected and untransfected dendritic cells and (C) the corresponding cell culture supernatant samples. (D) Serum samples of an AML patient with an IDH2 R140Q mutation and a healthy control. Peak profiles of D-2-HG-1E (first peak) in the different biological matrices were very similar to the standard demonstrating the formation of a selective mass trace for this compound by APCI. This enabled the detection of huge concentration differences of D- and L-2-HG in cell culture and serum samples, respectively. Experimental details and method parameters are described under 9.2.

9.3.2 Ionization behavior of D/L-2-HG in GC-APCI-TOFMS infusing water as modifier

APCI requires traces of water vapor in the source for protonation of analytes [45]. As described in chapter 6, a tremendous effect of continuous water infusion on APCI of methyl chloroformate derivatives was observed, in particular for organic acids. Upon water infusion at 0.5 mL/h a more than 4-fold increase in abundance of the $[M+H]^+$ ion of D/L-2-HG-1E was found, in comparison to no water infusion (APCI/-H₂O), which proved optimum over the range of investigated infusion rates (0.1-0.6 mL/h in 0.1 mL/h increments, compare **Figure 9.4**). Water infusion exerted no effect on the ion species formed for that derivative as the $[M+H]^+$ ion represented the base peak of the APCI spectrum and no fragments with a relative intensity >1% were observed. Interestingly, for the other pair of peaks in the chromatogram of a derivatized racemic standard of D/L-2-HG corresponding to the D/L O-methoxycarbonyldimethyl ester of 2-HG (D/L-2-HG-1MC-2E), a $[M+H-CO_2-CH_3OH]^+$ fragment ion (m/z 159.065) dominated the APCI/-H₂O mass spectrum of both enantiomers, while the corresponding $[M+H]^+$ ion (m/z 235.081) was almost absent (relative intensity <1%). In contrast, the $[M+H]^+$ ion represented the most abundant ion in the APCI/+H₂O spectra. Nevertheless, its abundance was only one fourth of that of D/L-2-HG-1E, which was initially used for quantification of D/L-2-HG. **Figure 9.5** depicts the overlaid base peak chromatograms of a derivatized racemic D/L-2-HG standard acquired w/o water infusion, respectively, demonstrating that APCI ionization efficiency was distinctly promoted by water infusion (upper chromatogram in grey).

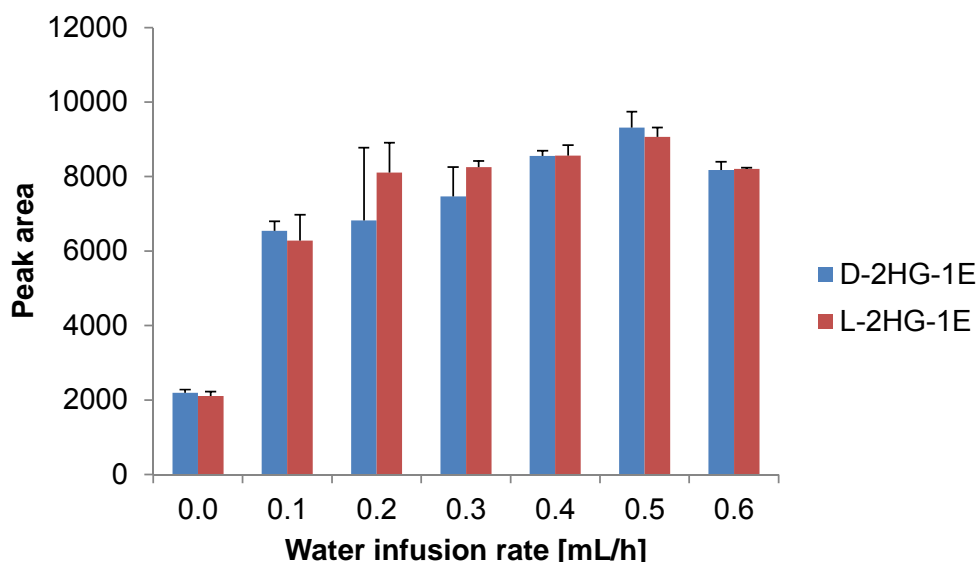


Figure 9.4 Impact of water infusion at different rates on APCI efficiency of D/L-2-HG-1E. Arithmetic means \pm SD of $[M+H]^+$ peak areas of D/L-2-HG-1E are given that were obtained by GC-APCI-TOFMS analysis of a 7.81 μ M racemic D/L-2HG standard (N=3 derivatization replicates). Water was infused in 0.1 mL/h increments over the range of 0.1-0.6 mL/h in consecutive runs and compared to no water infusion demonstrating that an infusion rate of 0.5 mL/h proved optimum yielding a more than 4-fold increase in peak areas. E, methyl ester group.

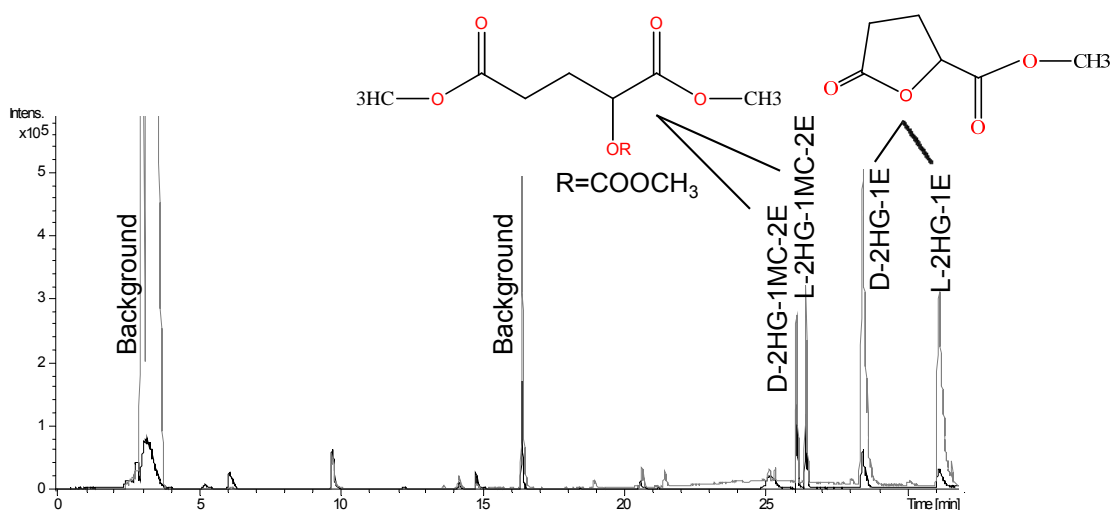


Figure 9.5 Impact of water infusion on APCI efficiency of an MCF-derivatized racemic D/L-2-HG standard. Water infusion distinctly increased peak abundances as demonstrated by the overlaid base peak chromatograms of the same standard acquired by GC-APCI-TOFMS without (chromatogram in black) and with (0.5 mL/h) water infusion (chromatogram in grey). The total (D+L)-2-HG concentration in the standard was 200 μ M. MC, methoxycarbonyl group; E, methyl ester group.

9.3.3 Method validation

Since D/L-2-HG-1E eluted as symmetric peaks with no interferences in the case of the biological samples, calibration was carried out over a concentration range of 0.02-1000 μM . **Table 9.1** presents figures of merit for the calibration curves.

Table 9.1 Figures of merit for the quantitative profiling of D/L-2-HG by GC-APCI-TOFMS. Calibration was carried out over 0.02-1000 μM . Racemic D/L-[2,3,3- $^2\text{H}_3$]2-HG was added as internal standard prior to MCF derivatization. Several calibration curves were acquired in the course of method development that were based on either of the two derivatives and acquired with (0.5 mL/h) or without water infusion, with an APCI I or II source, and with different injection volumes. Water infusion and APCI II source yielded the lowest LLOQs for both enantiomers of 2-HG-1E, thus enabling quantification of 2-HG levels that are well below expected concentrations in physiological specimens, while higher LLOQs were obtained for D/L-2-HG-1MC-2E in comparison to D/L-2-HG-1E. E, methyl ester group; MC, methoxycarbonyl group; IS, internal standard; LR, linear range; RSD, relative standard deviation.

#	Compound	Water inf. (mL/h)	Source	Inj. Vol. (μL)	RT (min)	m/z Quantifier ion (Da)	IS	R ²	LR (μM)	Order of magnitude	RSD deriv. repl. [†] (%)
1	D-2-HG-1E	0.5	APCI II	1	28.45	145.05 \pm 0.01	D ₃ -D-2-HG-1E	0.9997	0.06-188	3.5	3.8
1	L-2-HG-1E	0.5	APCI II	1	31.20	145.05 \pm 0.01	D ₃ -L-2-HG-1E	0.9988	0.06-188	3.5	2.2
2	D-2-HG-1E	0.0	APCI II	1	28.45	145.05 \pm 0.01	D ₃ -D-2-HG-1E	0.9995	0.12-500	3.6	4.8
2	L-2-HG-1E	0.0	APCI II	1	31.20	145.05 \pm 0.01	D ₃ -L-2-HG-1E	0.9972	0.49-500	3.0	6.8
3	D-2-HG-1E	0.5	APCI I	1	28.33	145.05 \pm 0.01	D ₃ -D-2-HG-1E	0.9998	3.91-1000	2.4	10.9
3	L-2-HG-1E	0.5	APCI I	1	31.03	145.05 \pm 0.01	D ₃ -L-2-HG-1E	0.9985	1.95-1000	2.7	2.5
4	D-2-HG-1MC-2E	0.5	APCI II	2	26.05	235.08 \pm 0.01	D ₃ -D-2-HG-1MC-2E	0.9995	0.49-250	2.7	4.8
4	L-2-HG-1MC-2E	0.5	APCI II	2	26.36	235.08 \pm 0.01	D ₃ -L-2-HG-1MC-2E	0.9986	0.24-250	3.0	9.5

[†] N=5, conc.=7.81 μM

Three different calibrations were evaluated based on D/L-2-HG-1E (#1-#3 in **Table 9.1**). Calibration #1, which was performed while infusing 0.5 mL/h water infusion into the APCI II source, yielded the lowest LLOQ for the two enantiomers (0.06 μM), a broad linear concentration range over three and a half orders of magnitude, and good repeatability with RSD of response of a calibration triplicate below 4%. The method benefitted from improvements in the APCI II source design over APCI I as well as the continuous infusion of water as modifier into the APCI II source, which is clearly obvious from the comparably poor performance of the methods that correspond to calibrations #2 and #3 (**Table 9.1**). Such a low LLOQ for the quantitative determination of both enantiomers of 2-HG in method #1 was well below expected concentration levels in human serum. In fact, Janin et al. [157] reported serum (D+L)-2-HG levels between 1-2 μM for controls and at least 10 times higher concentrations for patients with AML and an IDH mutation,

and suggested a diagnostic cutoff of 2 μM as a predictor of the presence of IDH mutations and outcome in these patients.

To further elaborate on the accuracy of method #1, a spike-in experiment in control serum was performed, using three different concentration levels of racemic D/L-2-HG (5 μM , 50 μM , and 200 μM , N=3 derivatization replicates per spike-level). Concentrations in the derivatized samples were determined by means of the calibration curves, endogenous levels of D- and L-2-HG obtained from unspiked serum samples were subtracted, and recoveries then calculated. **Figure 9.6** depicts arithmetic means \pm SD of recovery in those experiments based on D/L-2HG-1E, which ranged well between 89-100% for the two enantiomers.

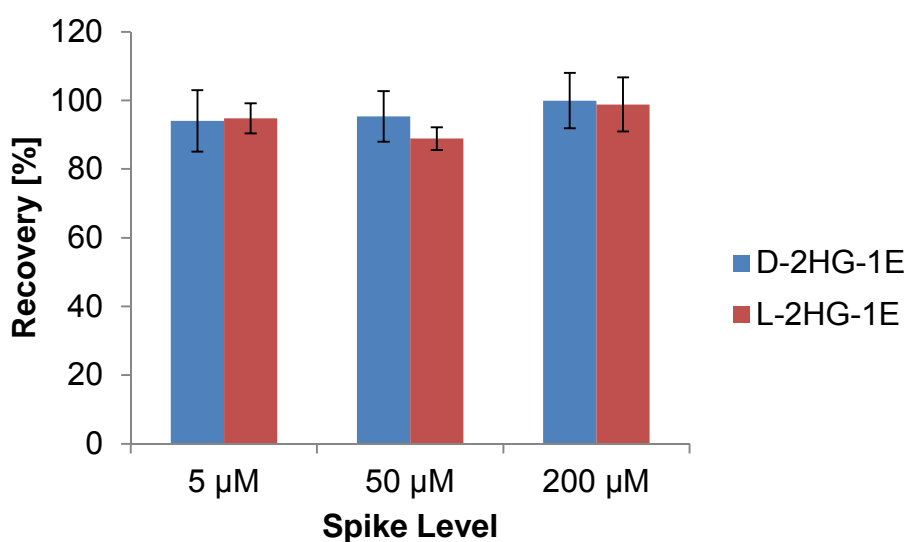


Figure 9.6 Arithmetic means \pm SD of recovery of D/L-2HG in matrix spike-in experiments. Racemic D/L-2-HG was spiked at three different levels (5 μM , 50 μM , and 200 μM , N=3 derivatization replicates per spike level) into human control serum, MCF derivatization was performed after precipitation of proteins with 100% MeOH, and concentration levels were determined by means of the calibration curves based on D/L-2HG-1E (#1 in **Table 9.1**). Endogenous D/L-2-HG levels were obtained from three unspiked serum samples and subtracted from spiked ones for calculation of recovery. As a result, good recoveries between 89-100% were found for the two enantiomers. E, methyl ester group.

Moreover, to check for racemization, individual standards of D and L enantiomers were analyzed (at a concentration of 200 μM , N=3 derivatization replicates each). Mean D- and L-2-HG-1E area ratios, i.e., the ratio of areas of D- or L-enantiomer divided by the sum of (D+L) enantiomers, were 0.6 and 1.8% in derivatized pure standards of L- and

D-2-HG, respectively. Hence, relevant falsely high concentrations of D- or L-2-HG due to racemization were only expected in the presence of very high concentrations of approximately 100 μM or higher of the other enantiomer.

Although good accuracy of method #1 in matrix spike-in experiments could be demonstrated, significant losses of 2-HG and its IS occurred during sample preparation as evidenced by the poor IS recovery of $\sim 30\%$ for serum specimens compared to standards, which entails higher LLOQ values in serum samples than in the standards. In addition to that, signal-to-noise of the analyte or IS may be compromised to a point where precision or accuracy becomes negatively affected in patient samples, which deteriorates a discriminant analysis between groups of them. To reveal the cause of analyte loss, samples containing solely IS or serum and IS were prepared in five different ways and compared to each other in terms of IS recovery. Mean peak area $\pm\text{SD}$ of IS for the five different sets of samples ($N=3$ derivatization replicates) are presented in **Figure 9.7**, which were prepared as described below the bars. As a result, analyte loss could be clearly attributed to the precipitation step in the protocol, because ion suppression did not occur (fourth bar from left) and loss of IS due to evaporation was not observed (second bar from left). As opposed to that, direct MCF derivatization of spiked serum without precipitation (fifth bar from left) yielded an almost full recovery of $\text{D}_3\text{-D/L-2-HG-1E}$. Hence, the latter protocol was applied for further analysis of serum, urine, and cell culture supernatants.

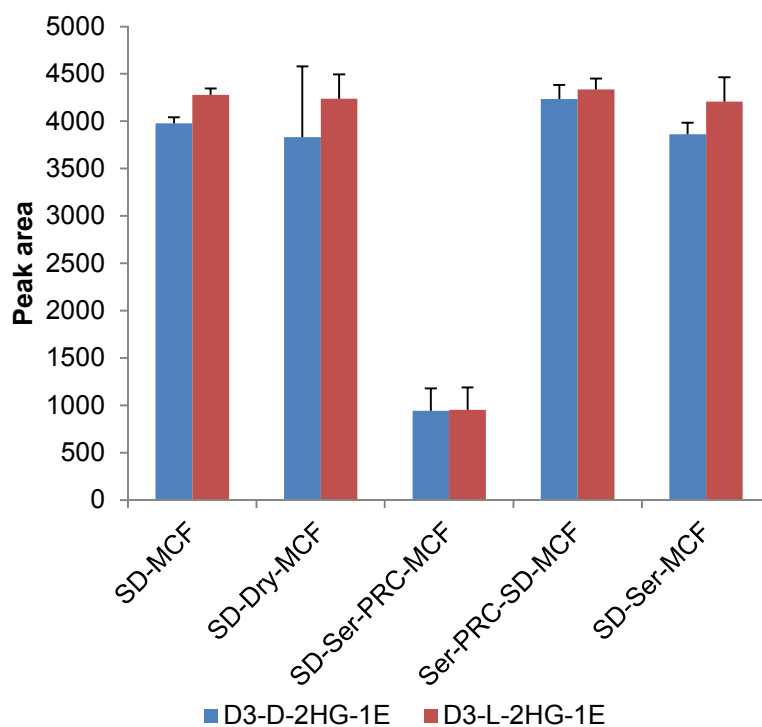


Figure 9.7 Arithmetic means \pm SD of $[M+H]^+$ peak areas of D_3 -D/L-2-HG-1E from MCF-derivatized IS and serum samples spiked with IS. Different sample preparation strategies were tested and compared with regard to increasing IS recovery in serum. Sample preparation was performed in order from left to right according to the labeling below the bars (N=3 derivatization replicates). Direct MCF derivatization and no precipitation of spiked serum yielded an almost full recovery of D_3 -D/L-2-HG-1E. SD, Deuterium-labeled standard was added (total (D+L) concentration of 6.7 μ M); MCF, MCF derivatization; Dry, evaporation of solvent; Ser, 150 μ L of control serum was added; PRC, precipitation with 100 % MeOH and two washing steps; E, methyl ester group.

9.3.4 Lactonization of D/L-2-HG in derivatized standards and biological samples

As pointed out earlier, two peaks for each enantiomer of 2-HG were obtained in the analysis of an MCF-derivatized racemic D/L-2-HG standard (**Figure 9.5**). A constant area ratio of 4:1 was found for each enantiomer (peak #2/peak #1) in samples analyzed in the course of method validation. As the same was true for the IS, (peak #2/peak #2-IS)-to-(peak #1/peak #1-IS) ratio of approximately 1 was obtained, e.g., 0.92 ± 0.09 for the D enantiomer in 31 standards and spiked serum samples analyzed in the course of method validation. As further revealed by NMR, lactonization of 2-HG, which has been reported by others as well [149,160], occurred due to MCF derivatization. **Figure 12.3**

depicts overlaid ^1H NMR spectra of 1mM MCF-derivatized standards of 2-HG and 5-oxotetrahydro-2-furancarboxylate (B) showing that the latter yielded one signal of its corresponding methyl ester while 2-HG formed three different derivatives including the lactone species. Among these three derivatives, the two-fold derivatized 2-HG species did probably not elute from the GC column because of its free hydroxyl group. As opposed to that, aqueous standards of the same compounds were pure with no signal overlap of their methylene protons was observed in the corresponding ^1H NMR spectra (**Figure 12.3C**).

In contrast to the constant (peak #2/peak #2-IS)-to-(peak #1/peak #1-IS) ratio of about 1 obtained so far, the analysis of serum of an AML patient with an IDH2 R140Q mutation (the same sample as in **Figure 9.3D**) resulted in a higher ratio of 1.2, while another sample of this type yielded an even higher ratio of 2.9. Therefore, it was hypothesized that lactone is present in these biological samples prior to derivatization, which could be proven by HPLC-ESI-MS/MS and NMR afterwards. **Figure 12.4** shows a TIC chromatogram of a standard mixture of 2-HG and lactone (31.25 μM each) acquired by HPLC-ESI-MS/MS on a Discovery HS F5 PFPP column. Both compounds were well separated and also present in the serum specimens. Accordingly, all signals in 2D ^1H - ^{13}C TOCSY spectra of aqueous 2-HG and lactone standards overlapped with those from the serum samples (see **Figure 12.3D**). In addition to that, lactone was also found in urine specimens that had not been subjected to derivatization. Consequently, quantification of 2-HG by GC-APCI-TOFMS based on the lactone peak did no longer appear feasible and peak #1 corresponding to the three-fold derivatized open-chain 2-HG species was used instead.

Peaks #1 of D and L enantiomers of 2-HG were well separated by 0.3 min and calibration was feasible. Figures of merit are given in **Table 9.1** (method #4). Calibration and further quantification of 2-HG in biological samples was performed using 2 μL as injection volume since a 1.8-fold increase in peak area was obtained compared to previous 1 μL , whereas peak area did not increase with a longer splitless time of 1.5 min. However, LLOQ values of 0.49 and 0.24 μL for D- and L-2-HG in standards were close to expected concentrations in human serum.

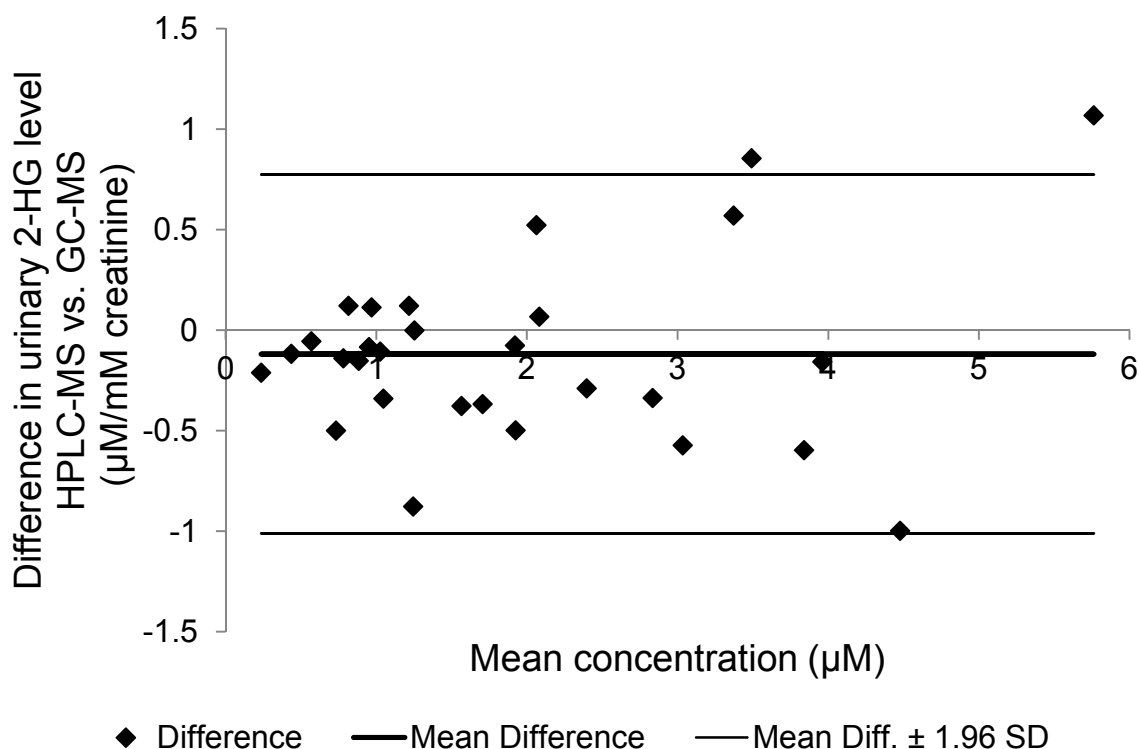


Figure 9.8 Bland-Altman plot to assess the agreement between GC-APCI-TOFMS and HPLC-ESI-MS/MS in the quantification of total (D+L) urinary levels of 2-HG. A total of 23 urine samples from AML patients and 6 controls were analyzed by both methods. They were well in accordance with each other as indicated by an absolute and relative mean difference of $-0.12 \mu\text{M}$ 2-HG/mM creatinine (represented by the centerline) and 6%, respectively. The outer lines correspond to mean $\pm 1.96 \times \text{SD}$ of the differences.

Nevertheless, cross-validation of the present GC-APCI-TOFMS approach with HPLC-MS/MS using urine specimens from 23 wild-type IDH AML patients and 6 healthy controls was possible, because normal urinary levels of 2-HG are higher than in serum. The two approaches yielded comparable (D+L)-2-HG levels as illustrated by the Bland-Altman plot in **Figure 9.8**; in fact, absolute and relative mean difference of $-0.12 \mu\text{M}/\text{mM}$ creatinine and 6%, respectively, were obtained over all samples. Only three out of the 29 samples exhibited more than 30% deviation, but this was attributed to the particular sample or analysis rather than a general disagreement between the two methods. For 23 samples less than 20% deviation was observed. Interestingly, an average enantiomeric excess of 80% of the L enantiomer was found in these urine samples, which could solely be revealed by the GC-MS approach, but this was not significantly different between wild-type IDH AML patients and controls, as was the case for (D+L)-2-HG concentration levels normalized to creatinine. As opposed to that, large concentration

differences could be detected in qualitative and quantitative terms in cell extracts of IDH2 R140Q mutation transfected and untransfected dendritic cell lines and corresponding cell culture supernatants, as well as serum specimens of a healthy control versus an AML patient carrying an IDH2 R140Q mutation (compare **Figure 9.3**). In the latter serum sample, (D+L)-2-HG concentration levels of 57 and 52 μM were determined by HPLC-MS/MS and GC-APCI-TOFMS, respectively, which were well in accordance with each other. However, the latter approach additionally revealed an enantiomeric excess of 98% of the D enantiomer. Such concentration and enantiomeric excess values for D/L-2HG are well comparable to other reports for the same type of sample [157,163].

9.4 Conclusions

A quantitative approach by GC-APCI-TOFMS was established for the enantioselective profiling of the oncometabolite D-2-HG and applied to various matrices including human serum, urine, cell extracts and cell culture supernatants. Both D and L enantiomers of 2-HG were readily separable as their MCF derivatives on the employed chiral Rt- γ DEXsa column. Further, the method benefitted from highly specific quantifier ions due to soft ionization and high-resolution TOFMS. Moreover, the infusion of water as modifier improved LLOQ and repeatability for both enantiomers, with the improved APCI II source yielding about 50-fold lower LLOQs than the original APCI I source. Omission of the protein precipitation step in the preparation of serum specimens increased recovery of the IS from 30% to almost 100%.

The natural presence of the 5-membered lactone of 2-HG in biological samples represented an interesting finding, which might be of significance in the elucidation of the molecular mechanism, by which 2-HG potentially promotes tumor progression. According to this, quantification of 2-HG had to be based on its open-chain three-fold derivative resulting in LLOQ values of 0.49 and 0.24 μM for the D and L enantiomer, respectively, which are close to the expected serum concentrations. Hence, further optimization in terms of lower limits of quantification of the method may be required. The general ap-

plicability of the method in distinguishing 2-HG concentration differences in serum of diseased and healthy individuals and the accurate quantification of urinary 2-HG levels could be shown, with GC-APCI-TOFMS concentration levels well in agreement with those determined by HPLC-MS/MS analysis. In addition to the latter approach, the developed GC-APCI-TOFMS method enabled the determination of the enantiomeric excess in the biological samples, which can be expected to be very helpful not only in the further elucidation of biological effects of 2-HG in tumors but also the utility of D-2-HG as a diagnostic and/or prognostic tumor marker.

10 Conclusions and perspectives

This thesis exemplifies the great potential of GC-APCI-TOFMS in metabolic fingerprinting and profiling providing both the means for identification of unknown metabolites by accurate mass measurements and the sensitive determination of metabolites [9,10]. Significant gains in detection sensitivity could be realized for MCF derivatives by the continuous infusion of water into the APCI source. Further gains in reliability and accuracy were achieved by means of the recognition and minimization of matrix effects and a redesigned APCI source developed by Bruker Daltonics.

The benefit of continuous water infusion during APCI of MCF-derivatized metabolites became particularly obvious in the comparative metabolic fingerprinting of control and 17-DMAG treated cancer cell extracts. The doubling of identified, significantly regulated metabolites to a total of 23 was due to a significant increase in abundance of $[M+H]^+$ ions as well as a concomitant decrease in technical variability. That the peak areas of significantly differentiating amino acids reflected true differences in concentration between the groups was corroborated through cross-validation by targeted HPLC-MS/MS analysis, yielding concordant results for 14 out of 15 amino acids that could be analyzed by both GC-APCI-TOFMS and HPLC-MS/MS. Further improvements in detection sensitivity and repeatability for both MCF and MeOx-TMS could be realized employing a redesigned APCI source commercialized recently by Bruker Daltonics. Due to a tighter source construction, which reduces chemical background, and the minimization of gas turbulences, the absolute number of identified compounds increased by 33% compared to the original APCI source and was accompanied by a distinct decrease in RSD of area integrals of features extracted from the biological replicates.

There remain areas for further improvement, such as the high level of manual intervention still required in peak integration and metabolite identification, which might be facilitated by the introduction of comprehensive GC-APCI-MS libraries. Moreover, the lack of reference compounds impedes unambiguous assignment of unknowns. Recent years

have seen a flurry of novel computational approaches for identifying metabolites, but systematic evaluations of their performance are still lacking, making a definitive judgment on their routine usefulness difficult. Overall, validated standard operating procedures for the annotation of novel metabolites and open-access metabolite libraries have to be extended. The introduction of International Chemical Identifier (InChI) strings for consistent naming of identified metabolites as well as supporting global data exchange formats, e.g. the netCDF format, for sharing mass spectra and associated information constitute promising steps in this regard. Still, higher community efforts are necessary.

GC-APCI-TOFMS proved to be a useful quantitative tool in metabolic profiling. The redesigned APCI II source, in comparison to APCI I, increased the linear range to almost three orders of magnitude mostly due to lower limits of quantification in the sub-micromolar range. Hence, GC-APCI(II)-TOFMS appeared promising for the quantification of potential biomarkers as exemplified here for the enantioselective profiling of the oncometabolite D-2-hydroxyglutarate, which accumulates in tumors harboring neomorphic mutations in IDH1 and IDH2 genes, respectively. The developed method relied on baseline separation of D and L enantiomers of 2-HG as their MCF derivatives on a chiral Rt- γ DEXsa column, highly specific quantifier ions due to soft ionization and high-resolution TOFMS, and the infusion of water as modifier improved lower quantification limits and repeatability. During initial method development D-2HG was separated from a highly abundant matrix compound by modifying the temperature ramp to prevent matrix effects. In the further course of method development, quantification of D/L-2-HG based on the 5-membered lactone peak of D/L-2-HG proved unsuitable, because the lactone was naturally present in several of the investigated biological specimens. Quantitative analysis was thus based on the less abundant open-chain three-fold derivative of 2-HG that yielded LLOQ values of 0.49 and 0.24 μ M for the D and L enantiomer, respectively, which are close to the expected serum concentrations. Hence, further optimization of the method may be needed, which might be achieved by increasing the yield of the three-fold derivative through modification of the derivatization protocol as well as maximizing ion transmission in the low mass range by specifically tuning the mass spectrometer. The method was successfully applied to various biological matrices including human serum and urine. The urinary 2-HG levels of 23 wild-type IDH AML patients and

6 healthy controls were well in accordance with those determined by HPLC-MS/MS. The natural presence of the 5-membered lactone of 2-HG in several tumor samples was of special interest as this finding might be helpful in elucidating the molecular mechanism, by which D-2-HG potentially promotes tumor progression. Whether the D-lactone is biologically even more active than D-2-HG, is the focus of future investigations.

11 References

- [1] E.C. Horning, M.G. Horning, D.I. Carroll, I. Dzidic, R.N. Stillwell, *Anal Chem* 45 (1973) 936.
- [2] E.C. Horning, D.I. Carroll, I. Dzidic, K.D. Haegele, S. Lin, C.U. Oertli, R.N. Stillwell, *Clin Chem* 23 (1977) 13.
- [3] R. Schiewek, M. Lorenz, R. Giese, K. Brockmann, T. Benter, S. Gab, O.J. Schmitz, *Anal Bioanal Chem* 392 (2008) 87.
- [4] C.N. McEwen, R.G. McKay, *J Am Soc Mass Spectrom* 16 (2005) 1730.
- [5] S. Davis, J. Hughes, A. Makarov, Google Patents (2001) WO2001018846 A2. Publication date: 15th March 2001.
- [6] R.K. Mitchum, G.F. Moler, W.A. Korfmacher, *Anal Chem* 52 (1980) 2278.
- [7] T. Bristow, M. Harrison, M. Sims, *Rapid Commun Mass Spectrom* 24 (2010) 1673.
- [8] R. Garcia-Villalba, T. Pacchiarotta, A. Carrasco-Pancorbo, A. Segura-Carretero, A. Fernandez-Gutierrez, A.M. Deelder, O.A. Mayboroda, *J Chromatogr A* 1218 (2011) 959.
- [9] A. Carrasco-Pancorbo, E. Nevedomskaya, T. Arthen-Engeland, T. Zey, G. Zurek, C. Baessmann, A.M. Deelder, O.A. Mayboroda, *Anal Chem* 81 (2009) 10071.
- [10] T. Pacchiarotta, E. Nevedomskaya, A. Carrasco-Pancorbo, A.M. Deelder, O.A. Mayboroda, *J Biomol Tech* 21 (2010) 205.
- [11] M. Ishimaru, M. Yamada, I. Nakagawa, S. Sugano, *J Breath Res* 2 (2008) 037021.
- [12] M.F. Almstetter, I.J. Appel, M.A. Gruber, C. Lottaz, B. Timischl, R. Spang, K. Dettmer, P.J. Oefner, *Anal Chem* 81 (2009) 5731.
- [13] C.J. Wachsmuth, T.A. Hahn, P.J. Oefner, K. Dettmer, *Anal Bioanal Chem*, Epub ahead of print on June 21.
- [14] C.J. Wachsmuth, K. Dettmer, S.A. Lang, M.E. Mycielska, P.J. Oefner, *Anal Chem* 86 (2014) 9186.
- [15] C.J. Wachsmuth, P.J. Oefner, K. Dettmer, in B.C. Weimer, C. Slupsky (Editors), *Metabolomics in food and nutrition*, Woodhead Publishing, Philadelphia, 2013, p. 3.
- [16] C.J. Wachsmuth, F.C. Vogl, P.J. Oefner, K. Dettmer, in T. Hyotylainen, S. Wiedmer (Editors), *Chromatographic Methods in Metabolomics*, RSC Publishing, Cambridge 2013, p. 87.
- [17] C.J. Wachsmuth, M.F. Almstetter, M.C. Waldhier, M.A. Gruber, N. Nurnberger, P.J. Oefner, K. Dettmer, *Anal Chem* 83 (2011) 7514.
- [18] K. Dettmer, B.D. Hammock, *Environ Health Perspect* 112 (2004) A396.
- [19] S.P. Gygi, Y. Rochon, B.R. Franza, R. Aebersold, *Mol Cell Biol* 19 (1999) 1720.
- [20] L.W. Sumner, P. Mendes, R.A. Dixon, *Phytochemistry* 62 (2003) 817.
- [21] W.B. Dunn, D.I. Broadhurst, H.J. Atherton, R. Goodacre, J.L. Griffin, *Chem Soc Rev* 40 (2011) 387.

- [22] K. Dettmer, P.A. Aronov, B.D. Hammock, *Mass Spectrom Rev* 26 (2007) 51.
- [23] D.S. Wishart, C. Knox, A.C. Guo, R. Eisner, N. Young, B. Gautam, D.D. Hau, N. Psychogios, E. Dong, S. Bouatra, R. Mandal, I. Sinelnikov, J. Xia, L. Jia, J.A. Cruz, E. Lim, C.A. Sobsey, S. Shrivastava, P. Huang, P. Liu, L. Fang, J. Peng, R. Fradette, D. Cheng, D. Tzur, M. Clements, A. Lewis, A. De Souza, A. Zuniga, M. Dawe, Y. Xiong, D. Clive, R. Greiner, A. Nazyrova, R. Shaykhutdinov, L. Li, H.J. Vogel, I. Forsythe, *Nucleic Acids Res* 37 (2009) D603.
- [24] M. Sud, E. Fahy, D. Cotter, A. Brown, E.A. Dennis, C.K. Glass, A.H. Merrill, Jr., R.C. Murphy, C.R. Raetz, D.W. Russell, S. Subramaniam, *Nucleic Acids Res* 35 (2007) D527.
- [25] C.A. Smith, G. O'Maille, E.J. Want, C. Qin, S.A. Trauger, T.R. Brandon, D.E. Custodio, R. Abagyan, G. Siuzdak, *Ther Drug Monit* 27 (2005) 747.
- [26] L.W. Sumner, A. Amberg, D. Barrett, M.H. Beale, R. Beger, C.A. Daykin, T.W.-M. Fan, O. Fiehn, R. Goodacre, J.L. Griffin, T. Hankemeier, N. Hardy, J. Harnly, R. Higashi, J. Kopka, A.N. Lane, J.C. Lindon, P. Marriott, A.W. Nicholls, M.D. Reily, J.J. Thaden, M.R. Viant, *Metabolomics* 3 (2007) 211.
- [27] O. Fiehn, *Plant Mol Biol* 48 (2002) 155.
- [28] F.V. Filipp, D.A. Scott, Z.A. Ronai, A.L. Osterman, J.W. Smith, *Pigment Cell Melanoma Res* 25 (2012) 375.
- [29] J.W. Locasale, A.R. Grassian, T. Melman, C.A. Lyssiotis, K.R. Mattaini, A.J. Bass, G. Heffron, C.M. Metallo, T. Muranen, H. Sharfi, A.T. Sasaki, D. Anastasiou, E. Mullarky, N.I. Vokes, M. Sasaki, R. Beroukhim, G. Stephanopoulos, A.H. Ligon, M. Meyerson, A.L. Richardson, L. Chin, G. Wagner, J.M. Asara, J.S. Brugge, L.C. Cantley, M.G. Vander Heiden, *Nat Genet* 43 (2011) 869.
- [30] J.M. Halket, V.G. Zaikin, *Eur J Mass Spectrom (Chichester, Eng)* 9 (2003) 1.
- [31] M.M. Koek, R.H. Jellema, J. van der Greef, A.C. Tas, T. Hankemeier, *Metabolomics* 7 (2011) 307.
- [32] C. Birkemeyer, A. Kolasa, J. Kopka, *J Chromatogr A* 993 (2003) 89.
- [33] O. Fiehn, J. Kopka, R.N. Trethewey, L. Willmitzer, *Anal Chem* 72 (2000) 3573.
- [34] J.M. Halket, D. Waterman, A.M. Przyborowska, R.K. Patel, P.D. Fraser, P.M. Bramley, *J Exp Bot* 56 (2005) 219.
- [35] H.H. Kanani, M.I. Klapa, *Metab Eng* 9 (2007) 39.
- [36] G. Noctor, G. Bergot, C. Mauve, D. Thominet, C. Lelarge-Trouverie, J.-L. Prioul, *Metabolomics* 3 (2007) 161.
- [37] S.G. Villas-Bôas, K.F. Smart, S. Sivakumaran, G.A. Lane, *Metabolites* 1 (2011) 3.
- [38] H. Kaspar, K. Dettmer, W. Gronwald, P.J. Oefner, *J Chromatogr B Analyt Technol Biomed Life Sci* 870 (2008) 222.
- [39] M.C. Waldhier, K. Dettmer, M.A. Gruber, P.J. Oefner, *J Chromatogr B Analyt Technol Biomed Life Sci* 878 (2010) 1103.
- [40] S.G. Villas-Boas, D.G. Delicado, M. Akesson, J. Nielsen, *Anal Biochem* 322 (2003) 134.
- [41] M.M. Koek, B. Muilwijk, M.J. van der Werf, T. Hankemeier, *Anal Chem* 78 (2006) 1272.
- [42] M.F. Almstetter, P.J. Oefner, K. Dettmer, *Anal Bioanal Chem* 402 (2012) 1993.
- [43] M.C. Waldhier, M.F. Almstetter, N. Nurnberger, M.A. Gruber, K. Dettmer, P.J. Oefner, *J Chromatogr A* 1218 (2011) 4537.

- [44] C.J. Wachsmuth, Master Thesis, University of Regensburg, Regensburg 2011.
- [45] I. Dzidic, D.I. Carroll, R.N. Stillwell, E.C. Horning, *Anal Chem* 48 (1976) 1763.
- [46] R.K. Mitchum, W.A. Korfmacher, G.F. Mohler, *Anal Chem* 54 (1982) 719.
- [47] M.W. Siegel, W.I. Fite, *J Phys Chem* 80 (1976) 2871.
- [48] M. Yamashita, J.B. Fenn, *J Phys Chem* 88 (1984) 4451.
- [49] E.D. Erickson, C.G. Enke, J.F. Holland, J.T. Watson, *Anal Chem* 62 (1990) 1079.
- [50] S. Abate, Y.G. Ahn, T. Kind, T.R. Cataldi, O. Fiehn, *Rapid Commun Mass Spectrom* 24 (2010) 1172.
- [51] T. Cajka, J. Hajslova, R. Kazda, J. Poustka, *J Sep Sci* 28 (2005) 601.
- [52] K. Qian, G.J. Dechert, *Anal Chem* 74 (2002) 3977.
- [53] R. Zimmermann, W. Welthagen, T. Groger, *J Chromatogr A* 1184 (2008) 296.
- [54] S. Klee, S. Albrecht, V. Derpmann, H. Kersten, T. Benter, *Anal Bioanal Chem* 405 (2013) 6933.
- [55] S. Klee, V. Derpmann, W. Wissdorf, S. Klopotoski, H. Kersten, K.J. Brockmann, T. Benter, S. Albrecht, A.P. Bruins, F. Dousty, T.J. Kauppila, R. Kostianen, R. O'Brien, D.B. Robb, J.A. Syage, *J Am Soc Mass Spectrom* 25 (2014) 1310.
- [56] G.J. Van Berkel, S.P. Pasilis, O. Ovchinnikova, *J Mass Spectrom* 43 (2008) 1161.
- [57] S. Krieger, A. von Trotha, K.S. Leung, O.J. Schmitz, *Anal Bioanal Chem* 405 (2013) 1373.
- [58] T.J. Kauppila, H. Kersten, T. Benter, *J Am Soc Mass Spectrom* 25 (2014) 1870.
- [59] T. Portoles, J.V. Sancho, F. Hernandez, A. Newton, P. Hancock, *J Mass Spectrom* 45 (2010) 926.
- [60] C.N. McEwen, *International Journal of Mass Spectrometry* 259 (2007) 57.
- [61] S. Klee, A. Brockhaus, W. Wissdorf, M. Thinius, N. Hartmann, T. Benter, *Rapid Commun Mass Spectrom* 29 (2015) 143.
- [62] F.J. Andrade, J.T. Shelley, W.C. Wetzel, M.R. Webb, G. Gamez, S.J. Ray, G.M. Hieftje, *Anal Chem* 80 (2008) 2646.
- [63] D. Carroll, I. Dzidic, E. Horning, R. Stillwell, *Applied Spectroscopy Reviews* 17 (1981) 337.
- [64] W. Wissdorf, L. Pohler, S. Klee, D. Muller, T. Benter, *J Am Soc Mass Spectrom* 23 (2012) 397.
- [65] F. Dousty, R.T. O'Brien, R. Gahler, H. Kersten, T. Benter, *Rapid Commun Mass Spectrom* 27 (2013) 1969.
- [66] A. Good, D. Durden, P. Kebarle, *The Journal of Chemical Physics* 52 (1970) 212.
- [67] P. Kebarle, R.M. Haynes, J.G. Collins, *Journal of the American Chemical Society* 89 (1967) 5753.
- [68] P. Kebarle, S.K. Searles, A. Zolla, J. Scarborough, M. Arshadi, *Journal of the American Chemical Society* 89 (1967) 6393.
- [69] J. Sunner, G. Nicol, P. Kebarle, *Anal Chem* 60 (1988) 1300.
- [70] W. Wissdorf, L. Seifert, V. Derpmann, S. Klee, W. Vautz, T. Benter, *J Am Soc Mass Spectrom* 24 (2013) 632.
- [71] S. Klee, M. Thinius, K.J. Brockmann, T. Benter, *Rapid Commun Mass Spectrom* 28 (2014) 1591.
- [72] T.J. Kauppila, R. Kostianen, A.P. Bruins, *Rapid Commun Mass Spectrom* 18 (2004) 808.

- [73] N. Ishii, K. Nakahigashi, T. Baba, M. Robert, T. Soga, A. Kanai, T. Hirasawa, M. Naba, K. Hirai, A. Hoque, P.Y. Ho, Y. Kakazu, K. Sugawara, S. Igarashi, S. Harada, T. Masuda, N. Sugiyama, T. Togashi, M. Hasegawa, Y. Takai, K. Yugi, K. Arakawa, N. Iwata, Y. Toya, Y. Nakayama, T. Nishioka, K. Shimizu, H. Mori, M. Tomita, *Science* 316 (2007) 593.
- [74] T. Pacchiarotta, R.J. Derks, E. Hurtado-Fernandez, P. van Bezooijen, A. Henneman, R. Schiewek, A. Fernandez-Gutierrez, A. Carrasco-Pancorbo, A.M. Deelder, O.A. Mayboroda, *Bioanalysis* 5 (2013) 1515.
- [75] E. Hurtado-Fernandez, T. Pacchiarotta, E. Longueira-Suarez, O.A. Mayboroda, A. Fernandez-Gutierrez, A. Carrasco-Pancorbo, *J Chromatogr A* 1313 (2013) 228.
- [76] E. Hurtado-Fernandez, T. Pacchiarotta, O.A. Mayboroda, A. Fernandez-Gutierrez, A. Carrasco-Pancorbo, *Anal Bioanal Chem* 407 (2015) 547.
- [77] T. Pacchiarotta, R.J. Derks, E. Nevedomskaya, W. van der Starre, J. van Dissel, A. Deelder, O.A. Mayboroda, *Analyst* 140 (2015) 2834.
- [78] N. Strehmel, J. Kopka, D. Scheel, C. Böttcher, *Metabolomics* 10 (2014) 324.
- [79] C. Jaeger, V. Tellström, G. Zurek, S. König, S. Eimer, B. Kammerer, *Metabolomics* 10 (2014) 859.
- [80] B.K. Matuszewski, M.L. Constanzer, C.M. Chavez-Eng, *Anal Chem* 75 (2003) 3019.
- [81] D. Remane, M.R. Meyer, D.K. Wissenbach, H.H. Maurer, *Rapid Commun Mass Spectrom* 24 (2010) 3103.
- [82] T. Portolés, J.G. Mol, J.V. Sancho, F. Hernández, *Anal Chem* 84 (2012) 9802.
- [83] W. Zhu, A.P. Stevens, K. Dettmer, E. Gottfried, S. Hoves, M. Kreutz, E. Holler, A.B. Canelas, I. Kema, P.J. Oefner, *Anal Bioanal Chem* 401 (2011) 3249.
- [84] J. Gullberg, P. Jonsson, A. Nordstrom, M. Sjostrom, T. Moritz, *Anal Biochem* 331 (2004) 283.
- [85] S.K. Lien, H.F. Kvitvang, P. Bruheim, *J Chromatogr A* 1247 (2012) 118.
- [86] H.F. Kvitvang, T. Andreassen, T. Adam, S.G. Villas-Boas, P. Bruheim, *Anal Chem* 83 (2011) 2705.
- [87] C. Birkemeyer, A. Luedemann, C. Wagner, A. Erban, J. Kopka, *Trends Biotechnol* 23 (2005) 28.
- [88] A. Barladi, C. Enders, *Journal of School Psychology* 48 (2010) 5.
- [89] Y. Benjamini, Y. Hochberg, *J Roy Statist Soc Ser B (Methodological)* 57 (1995) 289.
- [90] S. Dudoit, J. Fridlyand, T.P. Speed, *Journal of the American Statistical Association* 97 (2002) 77.
- [91] J. Hochrein, M.S. Klein, H.Zacharias, J. Li, G. Wijffels, H.J. Schirra, R. Spang, P.J. Oefner, W. Gronwald, *J Proteome Res* 11 (2012) 6242.
- [92] Q. Cui, I.A. Lewis, A.D. Hegeman, M.E. Anderson, J. Li, C.F. Schulte, W.M. Westler, H.R. Eghbalnia, M.R. Sussman, J.L. Markley, *Nat Biotechnol* 26 (2008) 162.
- [93] B. Zhou, J. Wang, H.W. Ransom, *PLoS One* 7 (2012) e40096.
- [94] D.P. Leader, K. Burgess, D. Creek, M.P. Barrett, *Rapid Commun Mass Spectrom* 25 (2011) 3422.
- [95] T. Kind, O. Fiehn, *BMC Bioinformatics* 7 (2006) 234.
- [96] T. Kind, O. Fiehn, *BMC Bioinformatics* 8 (2007) 105.

- [97] T. Kind, M. Scholz, O. Fiehn, PLoS One 4 (2009) e5440.
- [98] D.S. Wishart, Bioanalysis 1 (2009) 1579.
- [99] D.S. Wishart, Bioanalysis 3 (2011) 1769.
- [100] W.B. Dunn, A. Erban, R.J.M. Weber, D.J. Creek, M. Brown, R. Breitling, T. Hankemeier, R. Goodacre, S. Neumann, J. Kopka, M.R. Viant, Metabolomics 9 (2013) 44.
- [101] T. Kind, O. Fiehn, Bioanal Rev 2 (2010) 23.
- [102] E. Kováts, Helvetica Chimica Acta 41 (1958) 1915.
- [103] H. Van Den Dool, P.D. Kratz, J Chromatogr A 11 (1963) 463.
- [104] E. Malvoisin, E. Evrard, M. Roberfroid, M. Mercier, J Chromatogr 186 (1979) 81.
- [105] J. Liseč, N. Schauer, J. Kopka, L. Willmitzer, A.R. Fernie, Nat Protoc 1 (2006) 387.
- [106] T. Kind, G. Wohlgemuth, Y. Lee do, Y. Lu, M. Palazoglu, S. Shahbaz, O. Fiehn, Anal Chem 81 (2009) 10038.
- [107] J. Kopka, N. Schauer, S. Krueger, C. Birkemeyer, B. Usadel, E. Bergmuller, P. Dormann, W. Weckwerth, Y. Gibon, M. Stitt, L. Willmitzer, A.R. Fernie, D. Steinhauser, Bioinformatics 21 (2005) 1635.
- [108] G.G. Desbrosses, J. Kopka, M.K. Udvardi, Plant Physiol 137 (2005) 1302.
- [109] J. Kopka, J Biotechnol 124 (2006) 312.
- [110] S. Kumari, D. Stevens, T. Kind, C. Denkert, O. Fiehn, Anal Chem 83 (2011) 5895.
- [111] S.E. Stein, V.I. Babushok, R.L. Brown, P.J. Linstrom, J Chem Inf Model 47 (2007) 975.
- [112] V.V. Mihaleva, H.A. Verhoeven, R.C. de Vos, R.D. Hall, R.C. van Ham, Bioinformatics 25 (2009) 787.
- [113] S. Bieri, P.J. Marriott, Anal Chem 78 (2006) 8089.
- [114] K. Héberger, J Chromatogr A 1158 (2007) 273.
- [115] H. Lu, Y. Liang, W.B. Dunn, H. Shen, D.B. Kell, Trends Anal Chem 27 (2008) 215.
- [116] O. David Sparkman, J Am Soc Mass Spectrom 7 (1996) 313.
- [117] S.E. Stein, D.R. Scott, J Am Soc Mass Spectrom 5 (1994) 859.
- [118] K.F. Smart, R.B. Aggio, J.R. Van Houtte, S.G. Villas-Boas, Nat Protoc 5 (2010) 1709.
- [119] E.L. Schymanski, M. Meringer, W. Brack, Anal Chem 81 (2009) 3608.
- [120] M. Kanehisa, M. Araki, S. Goto, M. Hattori, M. Hirakawa, M. Itoh, T. Katayama, S. Kawashima, S. Okuda, T. Tokimatsu, Y. Yamanishi, Nucleic Acids Res 36 (2008) D480.
- [121] K. Dettmer, N. Nurnberger, H. Kaspar, M.A. Gruber, M.F. Almstetter, P.J. Oefner, Anal Bioanal Chem 399 (2011) 1127.
- [122] R-DevelopmentCore Team. R: A language and environment for statistical computing; R Foundation for Statistical Computing: Vienna, Austria 2007.
- [123] G.K. Smyth, Stat Appl Genet Mol Biol 3 (2004) Article3.
- [124] J. Josse, F. Husson, J SFdS 153 (2012) 79.
- [125] K.S. Pollard, S. Dudoit, M.J. van der Laan, MULTTEST Multiple Testing Procedures and Applications to Genomics; Division of Biostatistics.

- [126] H. Kaspar, K. Dettmer, Q. Chan, S. Daniels, S. Nimkar, M.L. Daviglus, J. Stamler, P. Elliott, P.J. Oefner, *J Chromatogr B Analyt Technol Biomed Life Sci* 877 (2009) 1838.
- [127] T. Portolés, E. Pitarch, F.J. López, F. Hernández, W.M.A. Niessen, *Rapid Commun Mass Spectrom* 25 (2011) 1589.
- [128] N. Strehmel, J. Hummel, A. Erban, K. Strassburg, J. Kopka, *J Chromatogr B Analyt Technol Biomed Life Sci* 871 (2008) 182.
- [129] D.G. Altman, J.M. Bland, *The statistician* 32 (1983) 307.
- [130] J.M. Bland, D.G. Altman, *Lancet* 1 (1986) 307.
- [131] S. Matysik, G. Schmitz, S. Bauer, J. Kiermaier, F.M. Matysik, *Biochem Biophys Res Commun* 446 (2014) 751.
- [132] K. Dettmer, M.F. Almstetter, I.J. Appel, N. Nurnberger, G. Schlamberger, W. Gronwald, H.H.D. Meyer, P.J. Oefner, *Electrophoresis* 31 (2010) 2365.
- [133] A.T. van der Goot, W. Zhu, R.P. Vázquez-Manrique, R.I. Seinstra, K. Dettmer, H. Michels, F. Farina, J. Krijnen, R. Melki, R.C. Buijsman, M. Ruiz Silva, K.L. Thijssen, I.P. Kema, C. Neri, P.J. Oefner, E.A.A. Nollen, *Proc Natl Acad Sci U S A* 109 (2012) 14912.
- [134] M.M. Shahin, *J Chem Phys* 45 (1966) 2600.
- [135] P. Östman, L. Luosujärvi, M. Haapala, K. Grigoras, R.A. Ketola, T. Kotiaho, S. Franssila, R. Kostianen, *Anal Chem* 78 (2006) 3027.
- [136] C.R. Warren, *Metabolomics* 9 (2013) S110.
- [137] G. Guilbault, M. Hjelm, *Pure and applied chemistry* 61 (1989) 1657.
- [138] J. Hajšlová, K. Holadova, V. Kocourek, J. Poustka, M. Godula, P. Cuhra, M. Kempný, *Journal of Chromatography A* 800 (1998) 283.
- [139] J. Hajšlová, J. Zrostlikova, *Journal of Chromatography A* 1000 (2003) 181.
- [140] D. Remane, D.K. Wissenbach, M.R. Meyer, H.H. Maurer, *Rapid Commun Mass Spectrom* 24 (2010) 859.
- [141] F. Gosetti, E. Mazzucco, D. Zampieri, M.C. Gennaro, *J Chromatogr A* 1217 (2010) 3929.
- [142] J.M. Marin, E. Gracia-Lor, J.V. Sancho, F.J. Lopez, F. Hernandez, *J Chromatogr A* 1216 (2009) 1410.
- [143] D. Ostheimer, M. Cremese, A.H. Wu, D.W. Hill, *J Anal Toxicol* 21 (1997) 17.
- [144] C.F. Poole, *J Chromatogr A* 1158 (2007) 241.
- [145] T.M. Annesley, *Clin Chem* 49 (2003) 1041.
- [146] D.N. Heller, *Rapid Commun Mass Spectrom* 21 (2007) 644.
- [147] M.F. Almstetter, I.J. Appel, K. Dettmer, M.A. Gruber, P.J. Oefner, *J Chromatogr A* 1218 (2011) 7031.
- [148] K. Dettmer, *Anal Bioanal Chem* 406 (2014) 4931.
- [149] M. Duran, J.P. Kamerling, H.D. Bakker, A.H. van Gennip, S.K. Wadman, *J Inherit Metab Dis* 3 (1980) 109.
- [150] B. Wilcken, J. Pitt, D. Heath, P. Walsh, G. Wilson, N. Buchanan, *J Inherit Metab Dis* 16 (1993) 501.
- [151] L. Dang, D.W. White, S. Gross, B.D. Bennett, M.A. Bittinger, E.M. Driggers, V.R. Fantin, H.G. Jang, S. Jin, M.C. Keenan, K.M. Marks, R.M. Prins, P.S. Ward, K.E. Yen, L.M. Liao, J.D. Rabinowitz, L.C. Cantley, C.B. Thompson, M.G. Vander Heiden, S.M. Su, *Nature* 462 (2009) 739.
- [152] H. Yang, D. Ye, K.-L. Guan, Y. Xiong, *Clinical Cancer Research* 18 (2012) 5562.

- [153] J.A. Losman, W.G. Kaelin, Jr., *Genes Dev* 27 (2013) 836.
- [154] J. Balss, S. Pusch, A.C. Beck, C. Herold-Mende, A. Kramer, C. Thiede, W. Buckel, C.D. Langhans, J.G. Okun, A. von Deimling, *Acta Neuropathol* 124 (2012) 883.
- [155] C. Chesnelong, M.M. Chaumeil, M.D. Blough, M. Al-Najjar, O.D. Stechishin, J.A. Chan, R.O. Pieper, S.M. Ronen, S. Weiss, H.A. Luchman, J.G. Cairncross, *Neuro Oncol* 16 (2014) 686.
- [156] P. Koivunen, S. Lee, C.G. Duncan, G. Lopez, G. Lu, S. Ramkissoon, J.A. Losman, P. Joensuu, U. Bergmann, S. Gross, J. Travins, S. Weiss, R. Looper, K.L. Ligon, R.G. Verhaak, H. Yan, W.G. Kaelin, Jr., *Nature* 483 (2012) 484.
- [157] M. Janin, E. Mylonas, V. Saada, J.B. Micol, A. Renneville, C. Quivoron, S. Koscielny, L. Scourzic, S. Forget, C. Pautas, D. Caillot, C. Preudhomme, H. Dombret, C. Berthon, R. Barouki, D. Rabier, N. Auger, F. Griscelli, E. Chachaty, E. Leclercq, M.H. Courtier, A. Bennaceur-Griscelli, E. Solary, O.A. Bernard, V. Penard-Lacronique, C. Ottolenghi, S. de Botton, *J Clin Oncol* 32 (2014) 297.
- [158] J.H. Wang, W.L. Chen, J.M. Li, S.F. Wu, T.L. Chen, Y.M. Zhu, W.N. Zhang, Y. Li, Y.P. Qiu, A.H. Zhao, J.Q. Mi, J. Jin, Y.G. Wang, Q.L. Ma, H. Huang, D.P. Wu, Q.R. Wang, Y. Li, X.J. Yan, J.S. Yan, J.Y. Li, S. Wang, X.J. Huang, B.S. Wang, W. Jia, Y. Shen, Z. Chen, S.J. Chen, *Proc Natl Acad Sci U S A* 110 (2013) 17017.
- [159] A. Terunuma, N. Putluri, P. Mishra, E.A. Mathe, T.H. Dorsey, M. Yi, T.A. Wallace, H.J. Issaq, M. Zhou, J.K. Killian, H.S. Stevenson, E.D. Karoly, K. Chan, S. Samanta, D. Prieto, T.Y. Hsu, S.J. Kurley, V. Putluri, R. Sonavane, D.C. Edelman, J. Wulff, A.M. Starks, Y. Yang, R.A. Kittles, H.G. Yfantis, D.H. Lee, O.B. Ioffe, R. Schiff, R.M. Stephens, P.S. Meltzer, T.D. Veenstra, T.F. Westbrook, A. Sreekumar, S. Ambis, *J Clin Invest* 124 (2014) 398.
- [160] H.J.C. das Neves, J.P. Noronha, H. Rufino, *J High Resol Chromatogr* 19 (1996) 161.
- [161] K. Landfried, W. Zhu, M.C. Waldhler, U. Schulz, J. Ammer, B. Holler, D. Wolff, M. Edinger, K. Peter, M. Kreutz, R. Andreesen, P.J. Oefner, E. Holler, *Blood* 118 (2011) 6971.
- [162] W. Gronwald, M.S. Klein, H. Kaspar, S.R. Fagerer, N. Nuernberger, K. Dettmer, T. Bertsch, P.J. Oefner, *Anal Chem* 80 (2008) 9288.
- [163] L. Sellner, D. Capper, J. Meyer, C.D. Langhans, C.M. Hartog, H. Pfeifer, H. Serve, A.D. Ho, J.G. Okun, A. Kramer, A. Von Deimling, *Eur J Haematol* 85 (2010) 457.
- [164] A. Ballesteros-Gomez, J. de Boer, P.E. Leonards, *Anal Chem* 85 (2013) 9572.
- [165] M.P. Barrow, K.M. Peru, J.V. Headley, *Anal Chem* 86 (2014) 8281.
- [166] E. Canellas, P. Vera, C. Domeno, P. Alfaro, C. Nerin, *J Chromatogr A* 1235 (2012) 141.
- [167] E. Canellas, P. Vera, C. Nerin, *J Mass Spectrom* 49 (2014) 1181.
- [168] L. Cherta, T. Portoles, J. Beltran, E. Pitarch, J.G. Mol, F. Hernandez, *J Chromatogr A* 1314 (2013) 224.
- [169] F. David, P. Sandra, P. Hancock, *LC-GC Eur* 24 (2011) 16.
- [170] C. Domeno, E. Canellas, P. Alfaro, A. Rodriguez-Lafuente, C. Nerin, *J Chromatogr A* 1252 (2012) 146.

- [171] L.C. Garratt, R. Linforth, A.J. Taylor, K.C. Lowe, J.B. Power, M.R. Davey, *Plant Biotechnol J* 3 (2005) 165.
- [172] M.L. Gomez-Perez, P. Plaza-Bolanos, R. Romero-Gonzalez, J.L. Martinez Vidal, A. Garrido Frenich, *J Am Soc Mass Spectrom* 25 (2014) 899.
- [173] J. Nacher-Mestre, R. Serrano, T. Portoles, M.H. Berntssen, J. Perez-Sanchez, F. Hernandez, *J Agric Food Chem* 62 (2014) 2165.
- [174] T. Portolés, J.G. Mol, J.V. Sancho, F. Hernández, *Journal of Chromatography A* 1339 (2014) 145.
- [175] M. Raro, T. Portoles, J.V. Sancho, E. Pitarch, F. Hernandez, J. Marcos, R. Ventura, C. Gomez, J. Segura, O.J. Pozo, *J Mass Spectrom* 49 (2014) 509.

12 Appendix

Table 12.1 Overview on recent GC(\times GC)-APCI(+)-MS studies of environmental samples, foodstuffs, pharmaceutical impurities, pesticides, steroids and volatile organic compounds (Focus: 1=Ionization behavior, 2=Screening/ identification, 3=Quantification, 4=Cross-platform analysis, 5=Application).

Authors/ Technique/ Field	Samples	Focus	Highlights
Ballesteroz-Gómez et al. (2013) [164] Direct probe/ GC/ GC \times GC coupled to APCI-HRTOFMS, LC-ESI/APCI/APPI-HRTOFMS Environmental	Flame retardant standards/ electronic waste and car interior samples	3, 5	Direct probe, GC, and GC \times GC coupled to APCI-HRTOFMS were implemented for targeted quantification of flame retardants in environmental samples and compared to common LC-HRTOFMS approaches. Targeted screening and quantification by GC-APCI-HRTOFMS exhibiting excellent detection capabilities revealed several previously unreported flame retardants in electronic waste and car interior samples.

Table 12.1 continued.

Authors/ Technique/ Field	Samples	Focus	Highlights
Barrow et al. (2014) [165] GC-APCI-FTICRMS Environmental	Water samples related to oil sands	2, 5	GC-APCI-FTICRMS resolved nominal masses for a wide range of compounds in the analysis of oil sands process water. The resolution of isomers was seen as a huge added value given the rising demand for elaboration on component structure in environmental toxicity studies.
Bristow et al. (2010) [7] GC-APCI-TOFMS, GC-qMS with EI and CI sources Pharmaceutical	Standards of organic compounds/ synthesized cyanoamide sample	2, 4	High mass accuracy of GC-APCI-TOFMS (Δm : 1-2 mDa) allowed the identification of impurities in organic reaction mixtures based on their elemental formulas. During optimization, compounds presenting a broad range of polarity could be detected, but RSDs were reported as fairly high.

Table 12.1 continued.

Authors/ Technique/ Field	Samples	Focus	Highlights
Canellas et al. (2012) [166] GC-APCI-QTOFMS, GC-EI-qMS Organic contaminants/ foodstuffs	Acrylic adhesives/ standards of organic compounds/ Tenax [®] for migration studies	2, 4	GC-APCI-QTOFMS appeared as complementary tool to GC-EI-qMS in the analysis of non-intentionally added substances in acrylic adhesives used in food packaging materials. The low degree of overlap between the two techniques was not further elucidated; nevertheless, GC-APCI-QTOFMS revealed three new compounds of relevance. However, they did not diffuse into the packaged food and, hence, were considered as safe constituents of the packaging material.
Canellas et al. (2014) [167] GC-APCI-QTOFMS, GC-EI-qMS Organic contaminants/ foodstuffs	Labels with adhesive/ standards of organic compounds/ Tenax [®] and natural pork intestine for migration studies	2, 4	Unknown, toxic non-intentionally added substances from autoadhesive labels used for direct food contact were identified by GC-APCI-QTOFMS. Compared to GC-EI-qMS, calculation of elemental formulas, extensive database search and acquisition of MS/MS mass spectra more than doubled the number of confirmed compounds, for which migration to the food simulants was shown in migration studies.

Table 12.1 continued.

Authors/ Technique/ Field	Samples	Focus	Highlights
Cherta et al. (2013) [168] GC-APCI-QqQMS Pesticides/ foodstuffs	Pesticide stand- ards/ food samples	3	A quantitative GC-APCI-MS/MS method was developed and fully validated for 142 pesticides followed by its application to six different food matrices. Highly abundant $[M+H]^+$ ions as precursor ions favoring high selectivity and detection sensitivity were available in APCI as opposed to the extensive fragmentation in EI for these compounds.
David et al. (2011) [169] GC-APCI-TOFMS Environmental	Reference phthalate samples/ sediment extract	1	Simultaneous quantification of two high-molecular-weight phthalates without comprising detection sensitivity was achieved by GC-APCI-TOFMS in sedi- ments. Whereas EI yielded abundant, but unspecific fragment ions and only minor $[M-CH_3]^+$ ions that were specific for the two co-eluting compounds, the $[M+H]^+$ ion represented the base peak of the APCI spectrum in both cases.

Table 12.1 continued.

Authors/ Technique/ Field	Samples	Focus	Highlights
Domeño et al. (2012) [170] GC-APCI-QTOFMS Environmental	PAH and NPAH standards/ moss samples	3	Quantification of 15 PAHs and 8 nitrated PAHs by GC-APCI-QTOFMS was achieved at very low concentration levels in mosses. Acquired data can be used for pattern recognition of pollutants pointing to the different locations of sampling points and likely emissions.
García-Villalba et al. (2011) [8] GC-APCI-TOFMS Foodstuffs	Phenolic standard compounds/ Spanish extra-virgin olive oil samples	2, 3	The use of GC-APCI-TOFMS for screening, structural assignment and quantitative determination of phenolic compounds in virgin olive oil is demonstrated. Optimization of chromatographic and MS conditions and a full validation of the method (selectivity, linear range, LOD, precision, accuracy, matrix effect) was carried out.

Table 12.1 continued.

Authors/ Technique/ Field	Samples	Focus	Highlights
Garratt et al. (2005) [171] APCI-MS/ linked GC- APCI/EI-MS Foodstuffs/ VOCs	Transgenic lettuce plants and azy- gous controls	5	Six VOCs were studied by means of combined GC-APCI/EI-MS in the head-space from transgenic lettuce plants expressing an IPT gene that exhibited a senescence-suppressing effect by autoregulated cytokinin biosynthesis. Significant increases in acetaldehyde, ethanol and dimethyl sulphide in transgenic lettuce could be related to other indices of cellular senescence, demonstrating the general applicability of the approach.
Gómez-Pérez et al. (2014) [172] GC-APCI-QTOFMS Pesticides	Pesticide stand- ards/ meat sam- ples	1, 3	A total of 71 pesticides were evaluated by GC-APCI-QTOFMS in terms of their ionization behavior and quantitative performance for the application to meat samples. Quantitative analysis was carried out for fifty-one analytes that had shown acceptable performance in the validation study including the determination of linear quantification range, LOD, recovery, intra- and inter-day precision, and matrix effect.

Table 12.1 continued.

Authors/ Technique/ Field	Samples	Focus	Highlights
Matysik et al. (2014) [131] GC-APCI-TOFMS Steroids	Standards for sterols and related analytes/ human EDTA-plasma	5	The general applicability of GC-APCI-TOFMS for simultaneous screening and determination of concentration levels of sterols in human plasma was shown. Well-controlled humidity was established in the modified APCI source with regard to reproducible ionization.
Nácher-Mestre et al. (2014) [173] GC-APCI-QTOFMS Pesticides/ aquacultural	Pesticide and polycyclic aromatic hydrocarbon standards/ feed ingredients in aquafeeds/ aquafeeds/ fish samples	2	A reliable screening method based on GC-APCI-QTOFMS was developed and validated for 133 pesticides and 24 PAHs in aquacultural and fish samples. Screening detection limit of a compound, i.e. the threshold concentration at which detection becomes trustworthy, was set at the lowest spike-in concentration level that yielded a positive identification hit in at least 19 out of 20 replicate matrix samples. Several matches for the validated compounds were reported in fish samples.

Table 12.1 continued.

Authors/ Technique/ Field	Samples	Focus	Highlights
Östman et al. (2006) [135] GC-microchip APCI- QqQMS VOCs	Standards of volatile organic compounds/ human urine	5	The general suitability of GC-microchip APCI-MS for qualitative and quantitative analysis of volatile and semi-volatile compounds was demonstrated. Mass spectra produced by microchip APCI in this study yielded the same type of information as conventional APCI.
Portolés et al. (2010) [59] GC-APCI-QTOFMS, GC- EI-TOFMS Pesticides/ foodstuffs	Pesticide standards/ food samples	1, 2	GC-APCI-QTOFMS was evaluated for wide-scope pesticide screening purposes in three varieties of food samples. Water as modifier led to the predictable presence of $[M+H]^+$ ions for 90 out of 100 model pesticides favoring rapid screening, whereas the EI molecular ion is typically absent. MS/MS experiments increased confidence in the identity of several detected compounds in real food samples.

Table 12.1 continued.

Authors/ Technique/ Field	Samples	Focus	Highlights
Portolés et al. (2012) [82] GC-APCI-QqQMS, GC-EI-QqQMS Pesticides/ foodstuffs	Pyrethroid standards/ fruit and vegetables samples	1, 4	<p>The beneficial impact of APCI with water as modifier over EI was demonstrated for GC-QqQMS analysis of pyrethroids, which are highly fragmented in EI.</p> <p>More specific MRM transitions based on $[M+H]^+$ ions as precursor ions and lower detection limits in food samples were obtained.</p> <p>Strong matrix effects were observed for several pyrethroid/ food matrix combinations.</p>
Portolés et al. (2014) [174] GC-APCI-QTOFMS/ GC-EI-TOFMS Environmental	Standards of organic contaminants/ groundwater samples	2, 4	<p>A wide-scope screening method for initially 170 organic contaminants was established based on GC-APCI-QTOFMS and compared to GC-EI-TOFMS.</p> <p>Detection of these compounds based on molecular-ion searching was more effective than matching deconvoluted EI mass spectra to reference libraries.</p> <p>Number of compounds included in GC-APCI-QTOFMS screening approach could be extended by 85 because of the predictable presence of $[M+H]^+$ ions and the simultaneous acquisition of high and low collision energy spectra.</p>

Table 12.1 continued.

Authors/ Technique/ Field	Samples	Focus	Highlights
Raro et al. (2014) [175] GC-APCI-QTOFMS Steroids	Steroid standards	1	<p>Ionization behaviour of 60 underivatized and TMS-derivatized anabolic androgenic steroids and steroid metabolites was studied.</p> <p>Water as modifier proved beneficial for the formation of dominating $[M+H]^+$ ions that may serve as selective and sensitive precursor ions in tandem MS.</p> <p>Relationship between steroid structure and observed ions was established that might be helpful in the elucidation of unknown metabolites of steroids.</p>

Table 12.2 Gas chromatographic and mass spectrometric characteristics of the MCF and MeOx-TMS derivatives of the 20 standard compounds and fatty acids selected for method evaluation. Retention times, quantifier ions, prominent ions (marked in bold) present in APCI(+) mass spectra acquired at water infusion rates of 0.0 mL/h and 0.4 mL/h, respectively, and calculated $\log P_{\text{octanol/water}}$ values, as a measure of lipophilicity, for both MCF and MeOx-TMS derivatives, including the number of trimethylsilyl groups of the latter derivatives, are presented. Reprinted from [14].

compound	MCF derivatives					MeOx-TMS derivatives					
	retention time [min]	<i>m/z</i> Quantifier [M+H] ⁺ ion	APCI(+) mass spectrum, 0.0 mL/h, prominent ions (<i>m/z</i> (rel. Intensity [%])) ^a	APCI(+) mass spectrum, 0.4 mL/h, prominent ions (<i>m/z</i> (rel. Intensity [%])) ^a	$\log P^b$	derivatization status	retention time [min]	<i>m/z</i> Quantifier [M+H] ⁺ ion	APCI(+) mass spectrum, 0.0 mL/h, prominent ions (<i>m/z</i> (rel. Intensity [%])) ^a	APCI(+) mass spectrum, 0.4 mL/h, prominent ions (<i>m/z</i> (rel. Intensity [%])) ^a	$\log P^b$
Adipate	12.49	175.0965	115.0796 (100), 111.0484 (92.5), 143.0719 (68.1), 129.0577 (10.2), 175.0965 (3.3)	175.0971 (100), 161.0814 (86.7), 143.0708 (55.9), 129.0554 (21.7), 115.0766 (18.6), 111.0448 (11.6)	0.95	2TMS	14.55	291.1442	201.0970 (100), 291.1492 (80.5), 111.0450 (25.6)	291.1506 (100), 201.0973 (44.2)	2.74
Ala	11.08	162.0761	102.0597 (100), 162.0766 (41.3), 130.0531 (10.2)	162.0770 (100), 102.0557 (27.1), 130.0507 (25.2)	-0.16	2TMS	8.50	234.1340	234.1373 (100), 144.0848 (7.0)	234.1377 (100), 144.0858 (5.6)	2.25
4-Aminobutyrate	14.65	176.0917	144.0659 (100), 130.0526 (95.9), 112.0425 (51.0), 176.0958 (25.8)	176.0915 (100), 144.0656 (74.9), 130.0501 (37.8)	0.12	3TMS	13.83	320.1892	320.1970 (100), -	320.1961 (100), -	3.52
Asp	16.33	220.0816	160.0613 (100), 128.0359 (23.2), 220.0818 (14.3)	220.0820 (100), 160.0612 (38.9), 188.0558 (17.2)	-0.36	3TMS	14.18	350.1634	350.1712 (100), 278.1287 (5.0)	350.1698 (100), -	3.07
Benzoate	10.33	137.0597	137.0618 (100), -	137.0606 (100), -	2.12	1TMS	11.74	195.0836	195.0865 (100), 179.0506 (88.1), 105.0357 (35.8)	195.0860 (100), 105.0350 (9.9)	2.60
Fumarate	9.03	145.0495	145.0508 (100), 113.0272 (7.6)	145.0505 (100), 163.0606 (72.7), 113.0237 (10.2)	0.43	2TMS	11.98	261.0980	261.1021 (100), 171.0489 (9.9)	261.1020 (100), 189.0605 (8.5)	1.76
Glu	17.92	234.0972	202.0712 (100), 174.0772 (65.0), 142.0527 (54.0), 234.0983 (26.2), 114.0590 (20.9)	234.0970 (100), 202.0707 (34.6), 174.0758 (13.8)	-0.19	3TMS	15.43	364.1790	364.1867 (100), 274.1327 (8.5)	364.1849 (100), 292.1439 (10.8)	3.20
Hippurate	19.92	194.0812	194.0814 (100), 105.0401 (52.5)	194.0822 (100), 105.0352 (9.6)	0.80	1TMS	20.67	252.1050	252.1086 (100), 162.0570 (13.9)	252.1090 (100), 162.0553 (8.3)	1.88
Ile	13.89	204.1230	144.1036 (100), 204.1229 (13.2)	204.1238 (100), 144.1025 (63.7)	1.21	2TMS	10.77	276.1810	276.1868 (100), 158.1378 (5.4)	276.1857 (100), 204.1436 (5.3)	3.62
Leu	13.73	204.1230	144.1038 (100), 204.1238 (16.2)	204.1235 (100), 144.1024 (53.2)	1.21	2TMS	10.49	276.1810	276.1864 (100), 186.1330 (5.3)	276.1857 (100), 204.1436 (6.5)	3.62
Malonate	8.18	133.0495	133.0507 (100), 101.0284 (72.9), 119.0359 (43.9)	133.0504 (100), 119.0351 (54.2)	-0.12	2TMS	10.65	249.0973	249.1024 (100), 178.1104 (47.2), 303.1225 (28.8), 159.0482 (18.3)	249.1002 (100), 303.1204 (67.3), 178.1099 (25.7)	1.73
Met	18.10	222.0795	162.0589 (100), 174.0762 (64.6), 222.0797 (20.6), 114.0594 (14.1)	222.0803 (100), 162.0594 (44.5), 190.0552 (20.3), 174.0770 (11.8)	0.46	2TMS	14.64	294.1374	294.1433 (100), -	294.1427 (100), 222.1008 (7.0)	2.98
Methyl-malonate	8.50	147.0652	115.0443 (100), 133.0523 (93.6), 147.0672 (54.2)	147.0662 (100), 133.0503 (73.8)	0.24	2TMS	10.56	263.1129	263.1175 (100), 251.1493 (90.1), 117.0743 (37.7), 161.1008 (14.8)	251.1540 (100), 263.1176 (67.6), 117.0745 (10.8)	2.09

Table 12.2 continued.

compound	MCF derivatives					MeOx-TMS derivatives					
	retention time [min]	<i>m/z</i> Quantifier [M+H] ⁺ ion	APC(+), mass spectrum, 0.0 mL/h, prominent ions (<i>m/z</i> (rel. intensity [%])) ^a	APC(+), mass spectrum, 0.4 mL/h, prominent ions (<i>m/z</i> (rel. intensity [%])) ^a	logP ^b	derivatization status	retention time [min]	<i>m/z</i> Quantifier [M+H] ⁺ ion	APC(+), mass spectrum, 0.0 mL/h, prominent ions (<i>m/z</i> (rel. intensity [%])) ^a	APC(+), mass spectrum, 0.4 mL/h, prominent ions (<i>m/z</i> (rel. intensity [%])) ^a	logP ^b
Nicotinate	11.35	138.0550	138.0567 (100), -	138.0558 (100), -	0.79	1TMS	12.65	196.0788	196.0826 (100), -	196.0827 (100), -	1.66
Nval	13.23	190.1074	130.0887 (100), 190.1074 (9.4)	190.1082 (100), 130.0874 (64.1), 158.0817 (10.5)	0.86	2TMS	10.15	262.1653	262.1697 (100), 172.1164 (6.6)	262.1703 (100), 190.1287 (5.7)	3.27
Phe	19.55	238.1074	178.0872 (100), 190.0506 (42.3), 238.1078 (14.7), 146.0616 (13.1)	238.1085 (100), 178.0873 (49.7), 190.0509 (27.4)	1.19	2TMS	16.04	310.1653	310.1715 (100), -	310.1698 (100), 238.1285 (5.7)	3.76
Phenylacetate	11.74	151.0754	151.0764 (100), 91.0612 (14.2)	151.0760 (100), 91.0548 (6.4)	1.91	1TMS	12.62	209.0992	209.0992 (100), 193.0701 (18.1)	209.1024 (100), -	2.54
Pro	15.45	188.0917	128.0742 (100), 188.0930 (14.7)	188.0924 (100), 128.0718 (49.0)	0.27	2TMS	11.29	260.1497	260.1540 (100), 170.1012 (6.1)	260.1543 (100), 188.1127 (11.4)	2.17
Suberate	15.13	203.1278	171.1024 (100), 189.1128 (53.9), 143.1085 (46.2), 139.0772 (28.5), 111.0849 (27.1), 203.1283 (8.1)	203.1287 (100), 189.1132 (90.7)	1.88	2TMS	16.92	319.1755	319.1819 (100), 229.1287 (65.9), 139.0763 (12.9)	319.1818 (100), 247.1398 (35.1), 229.1244 (11.5)	3.64
Val	12.64	190.1074	130.0889 (100), 190.1082 (8.4)	190.1083 (100), 130.0873 (70.6), 158.0816 (12.8)	0.70	2TMS	9.76	262.1653	262.1698 (100), 144.1219 (5.8)	262.1693 (100), 190.1275 (6.7)	3.11
Nonanoate ^c	10.90	173.1536	173.1541 (100), -	173.1548 (100), 191.1648 (17.8)	3.84	1TMS	12.01	231.1775	231.1779 (100), 215.1371 (10.2)	231.1779 (100), 141.0390 (5.4)	4.72
Undecanoate ^c	13.48	201.1849	201.1854 (100), -	201.1857 (100), 219.1959 (15.2)	4.86	1TMS	14.46	259.2088	259.2084 (100), 243.1745 (6.9)	259.2090 (100), -	5.74
Tridecanoate ^c	15.96	229.2162	229.2166 (100), -	229.2171 (100), 247.2269 (15.0)	5.88	1TMS	16.76	287.2401	287.2405 (100), -	287.2403 (100), -	6.76
Pentadecanoate ^c	18.34	257.2475	257.2478 (100), -	257.2481 (100), 275.2580 (13.7)	6.90	1TMS	18.98	315.2714	315.2721 (100), -	315.2719 (100), -	7.78
Heptadecanoate ^c	20.56	285.2788	285.2802 (100), -	285.2800 (100), 303.2909 (7.6)	7.92	1TMS	21.10	343.3027	343.3040 (100), -	343.3026 (100), 271.2628 (5.5)	8.80
Nonadecanoate ^c	22.81	313.3101	313.3114 (100), -	313.3105 (100), 331.3173 (5.5)	8.94	1TMS	23.15	371.3340	371.3347 (100), -	371.3350 (100), 299.2929 (7.4)	9.82

^a *m/z* of [M+H]⁺ marked in bold. Masses of fragment and adduct ions exceed a S/N of 20 and a relative intensity of 10%, except the most intensive fragment ion in a spectrum exceeds only 5%. ^b Calculated by ACD/Labs V12.01. ^c Concs. = 67 μM and 100 μM for MCF and MO-TMS derivatives, respectively.

Table 12.3 Impact of different rates of continuous water infusion on APCI of MCF and MeOx-TMS derivatized standard compounds. Mean FCs in peak area of [M+H]⁺ ions in relation to no water infusion and RSDs of peak areas for N=5 derivatization replicates are given for both MCF (200 μM each) and MeOx-TMS (100 μM each) derivatives (A=amide group, E=methyl ester group). ANOVA and post hoc tests were performed based on the [M+H]⁺ peak areas.^{a,b} Reprinted from [14].

MCF derivative	0.1 mL/h		0.2 mL/h		0.3 mL/h		0.4 mL/h		0.5 mL/h	
	FC	RSD (%)	FC	RSD (%)	FC	RSD (%)	FC	RSD (%)	FC	RSD (%)
Adipate-2E ^c	23.0 ^j	19.3	71.8 ^g	30.0	133.7 ^g	17.0	163.9 ^g	4.2	137.8 ^g	24.7
Ala-1A-1E ^c	9.8 ⁱ	25.6	30.4 ^g	6.3	29.8 ^g	40.0	36.7 ^g	16.7	27.3 ^g	9.8
4-Aminobutyrate-1A-1E ^c	5.0 ^j	3.1	16.3 ⁱ	30.1	30.8 ^g	12.0	33.7 ^g	30.9	32.4 ^g	22.0
Asp-1A-2E ^c	4.7 ⁱ	42.0	13.9 ^g	32.8	17.9 ^g	30.1	15.6 ^g	22.6	12.0 ^g	23.1
Benzoate-1E ^c	2.0 ^h	34.6	2.6 ^g	13.6	2.7 ^g	13.1	1.8 ^h	14.7	2.1 ^g	17.1
Fumarate-2E ^c	1.7 ⁱ	12.0	2.0 ^g	30.4	1.8 ^g	28.7	1.2 ^j	15.0	1.3 ^j	21.7
Glu-1A-2E ^c	5.3 ^j	32.8	14.9 ^g	27.9	14.7 ^g	27.5	16.8 ^g	24.9	14.4 ^g	24.5
Hippurate-1E ^c	4.5 ^h	17.2	9.5 ^g	13.5	8.0 ^g	16.3	7.1 ^g	32.4	5.2 ^g	13.4
Ile-1A-1E ^c	6.0 ⁱ	18.6	15.1 ^g	18.0	17.4 ^g	24.7	15.9 ^g	32.3	16.6 ^g	12.7
Leu-1A-1E ^c	7.3 ⁱ	22.7	15.2 ^g	7.9	15.1 ^g	42.7	14.6 ^g	33.2	13.1 ^g	33.9
Malonate-2E ^c	57.1 ⁱ	33.5	93.6 ^g	7.7	137.2 ^g	25.7	163.9 ^g	3.5	126.9 ^g	12.7
Met-1A-1E ^c	6.8 ⁱ	48.4	15.3 ^g	22.7	20.9 ^g	34.0	16.3 ^g	32.0	13.7 ^g	26.2
Me-malonate-2E ^c	59.8 ^h	22.7	98.8 ^g	14.4	125.8 ^g	27.8	142.0 ^g	15.7	121.3 ^g	24.0
Nicotinate-1E ^c	1.0 ^j	13.1	2.1 ^g	31.2	1.7 ⁱ	23.6	1.7 ⁱ	10.4	1.2 ^j	48.2
Nval-1A-1E ^c	4.3 ^j	42.4	14.6 ^g	7.8	15.4 ^g	17.9	17.0 ^g	27.3	13.3 ^g	19.1
Phe-1A-1E ^c	4.4 ⁱ	30.4	10.5 ^g	18.0	12.0 ^g	18.7	8.6 ^g	28.0	10.0 ^g	6.3
Phenylacetate-1E ^c	5.8 ^g	31.2	9.0 ^g	18.7	10.1 ^g	14.4	9.7 ^g	10.8	6.9 ^g	9.7
Pro-1A-1E ^c	3.1 ⁱ	29.2	7.0 ^g	9.6	7.6 ^g	9.2	10.8 ^g	7.6	7.2 ^g	21.4
Suberate-2E ^c	10.5 ^j	30.7	38.4 ^g	34.4	47.7 ^g	14.9	60.5 ^g	24.7	35.0 ^g	27.5
Val-1A-1E ^d	4.4 ^j	18.2	13.3 ^g	39.9	18.8 ^g	9.7	21.6 ^g	6.9	15.4 ^g	27.1
Geometric mean	6.2	23.2	14.3	17.9	16.6	20.4	16.6	16.3	13.8	19.2

Table 12.3 continued.

MeOx-TMS derivative	0.1 mL/h		0.2 mL/h		0.3 mL/h		0.4 mL/h		0.5 mL/h	
	FC	RSD (%)	FC	RSD (%)	FC	RSD (%)	FC	RSD (%)	FC	RSD (%)
Adipate-2TMS ^d	1.2 ^j	35.1	1.7 ^j	54.2	2.2 ⁱ	31.8	2.4 ^h	13.6	2.2 ^j	24.2
Ala-2TMS ^f	1.1	31.9	1.0	24.4	1.2	30.2	1.3	29.5	1.0	20.6
4-Aminobutyrate-3TMS ^f	1.2	19.1	1.2	25.8	1.3	36.5	1.3	21.1	1.5	19.3
Asp-3TMS ^f	0.9	31.6	1.1	44.0	1.0	20.7	1.6	41.5	1.3	33.1
Benzoate-1TMS ^d	1.5 ^j	50.5	2.6 ^h	9.7	2.4 ^h	26.0	2.3 ^h	31.9	1.9 ^j	30.0
Fumarate-2TMS ^f	1.3	16.1	1.4	25.4	1.1	24.5	1.3	35.6	1.5	19.9
Glu-3TMS ^f	1.6	37.6	0.8	21.2	1.3	24.8	1.5	45.1	1.1	39.1
Hippurate-1TMS ^c	1.1 ^j	35.5	1.7 ⁱ	13.9	1.3 ^j	37.9	0.9 ^j	37.9	2.4 ^g	13.9
Ile-2TMS ^f	1.2	23.4	1.3	32.7	1.2	30.9	1.6	16.1	1.0	55.2
Leu-2TMS ^f	1.1	37.3	1.4	39.6	0.9	32.6	1.3	31.4	1.0	51.3
Malonate-2TMS ^f	1.8	43.2	2.1	15.7	2.0	27.2	1.7	20.5	2.2	29.1
Met-2TMS ^f	1.0	38.9	0.8	35.6	0.7	47.8	0.9	45.6	1.1	21.2
Me-malonate-2TMS ^f	1.1	41.0	1.4	30.2	1.4	36.4	1.4	26.0	1.3	17.6
Nicotinate-1TMS ^e	1.5 ^j	12.8	1.3 ^j	27.6	1.6 ⁱ	20.1	1.7 ^h	12.8	1.6 ⁱ	20.8
Nval-2TMS ^f	0.9	30.3	0.8	25.4	0.7	41.1	1.3	41.4	0.7	32.3
Phe-2TMS ^f	0.9	28.6	1.3	30.6	1.2	14.6	1.6	30.3	1.3	22.0
Phenylacetate-1TMS ^d	1.3 ^j	23.1	1.8 ⁱ	17.0	2.1 ^h	28.8	2.3 ^g	16.9	1.7 ^j	24.4
Pro-2TMS ^e	1.1 ^j	34.4	1.1 ^j	17.5	0.8 ^j	13.4	1.6 ⁱ	29.6	0.8 ^j	26.2
Suberate-2TMS ^f	1.1	13.9	1.3	29.3	1.5	15.8	1.6	25.0	1.4	18.1
Val-2TMS ^f	0.8	17.8	0.7	26.5	0.7	31.6	1.1	40.1	0.9	29.5
Geometric mean	1.2	28.1	1.3	25.3	1.2	27.2	1.5	27.6	1.3	25.7

^a Ranges of p values in ANOVA indicated as follows: ^c p <0.001; ^d p <0.01; ^e p <0.05; ^f p >0.05.

^b Ranges of p values in post hoc test indicated as follows: ^g p <0.001; ^h p <0.01; ⁱ p <0.05; ^j p >0.05.

Table 12.4 Area under the curve values for the corresponding ROC curves of MCF and MeOx-TMS derivatives for all pairwise fold changes. ROC curves were generated from 20 compounds spiked at six different concentrations into aliquots of a serum extract, followed by derivatization (N=5 derivatization replicates per spike level) and GC-APCI-TOFMS analysis without and at a water infusion rate of 0.4 mL/h. Reprinted from [14].

FC	AUC MCF derivatives		AUC MeOx-TMS derivatives	
	0.0 mL/h	0.4 mL/h	0.0 mL/h	0.4 mL/h
1.17	0.36	0.64	0.60	0.50
1.20	0.65	0.46	0.64	0.64
1.25	0.57	0.67	0.58	0.66
1.33	0.42	0.82	0.44	0.51
1.40	0.77	0.73	0.74	0.63
1.50	0.66	0.77	0.73	0.76
1.50	0.77	0.77	0.81	0.77
1.67	0.79	0.86	0.81	0.85
1.75	0.79	0.84	0.80	0.77
2.00	0.81	0.95	0.85	0.87
2.00	0.84	0.92	0.92	0.84
2.33	0.85	0.87	0.87	0.88
2.50	0.90	0.93	0.91	0.90
3.00	0.93	0.96	0.94	0.89
3.50	0.95	0.94	0.94	0.90

Table 12.5 Figures of merits of linear regression analysis of expected and observed fold changes in the MCF spike-in experiment as a function of water infusion. In the spike-in bucket table, blank feature area values were subtracted as “background” from the peak areas of investigated features, i.e., features that could be extracted from both, APCI/-H₂O and APCI/+H₂O spectra. All possible fold changes were plotted and respective slopes and intercepts of linear regression equations, regression coefficients and mean RSD values over all fold changes are listed. In the absence of [M+H]⁺ ions, a fragment ion was monitored instead, with the respective mass given after the underscore character. For alanine, 4-aminobutyrate, malonate, and methylmalonate, no ions could be extracted across all spike-in levels in the absence of water infusion. Reprinted from [14].

feature	0.0 mL/h				0.4 mL/h			
	slope	intercept	r	mean RSD/%	slope	intercept	r	mean RSD/%
Adipate_m/z 115 ^a	1.56	-0.41	0.832	32.7	1.33	-0.26	0.922	21.0
Asp_m/z 160 ^a	1.14	-0.11	0.945	19.5	1.20	-0.16	0.944	20.1
Benzoate	1.29	-0.21	0.884	25.0	1.11	-0.07	0.935	19.5
Fumarate	1.29	-0.20	0.884	31.3	1.11	-0.06	0.877	25.0
Glu_m/z 202 ^a	0.29	0.73	0.498	32.1	0.52	0.49	0.709	24.2
Hippurate	1.08	0.01	0.816	38.6	0.80	0.21	0.883	21.6
Ile	1.04	0.13	0.655	51.4	1.30	-0.22	0.856	29.4
Leu	1.37	0.04	0.350	67.2	0.70	0.29	0.873	22.5
Met_m/z 162 ^a	1.18	-0.05	0.761	44.1	1.00	0.03	0.891	25.8
Nicotinate	1.62	-0.43	0.792	40.7	1.37	-0.26	0.883	22.5
Nval_m/z 130 ^a	1.90	-0.70	0.876	32.8	1.06	-0.02	0.919	17.8
Phe	0.67	0.39	0.683	35.1	0.68	0.33	0.800	24.0
Phenylacetate	1.77	-0.53	0.629	47.3	0.98	0.05	0.916	16.1
Pro	1.18	-0.01	0.727	44.0	0.81	0.22	0.826	28.4
Suberate_m/z 189 ^a	0.37	0.61	0.735	20.4	0.53	0.45	0.887	15.9
Val_m/z 130 ^a	1.00	0.06	0.787	33.5	0.89	0.15	0.841	29.9

^a Mass of investigated feature corresponding to a fragment.

Table 12.6 Determination of elemental compositions of known MCF derivatized spike-in compounds. Compounds from the spike-in mix (concn=200 μ M) were spiked into a serum extract (N=5 derivatization replicates), followed by MCF derivatization and GC-APCI-TOFMS analysis without and with water infusion at 0.4 mL/h. Mass accuracy was determined after internal recalibration of mass spectra. In addition, isotope abundance accuracy (referred to as percentage error of the first isotope as well as mSigma value obtained from DataAnalysis, Bruker Daltonics) and average rank of correct elemental formula with total average number of calculated formulas are presented (A=amide group, E=methyl ester group). Reprinted from [14].

MCF derivative	elemental composition [M+H] ⁺	m/z calcd	0.0 mL/h				0.4 mL/h			
			$\Delta m \pm SD$ (mDa) ^a	isotope ratio error M+H+1/M+H (%) ^a	mSigma \pm SD	avg. rank of correct formula ^b avg. total number of formulas	$\Delta m \pm SD$ (mDa) ^a	isotope ratio error M+H+1/M+H (%) ^a	mSigma \pm SD	avg. rank of correct formula ^b avg. total number of formulas
Adipate-2E	C ₈ H ₁₅ O ₄ ⁺	175.0965	-0.3 \pm 1.8	^c	52.3 \pm 0.0	2 4	-0.6 \pm 0.1	-0.1 \pm 0.4	7.1 \pm 0.3	2 4
Ala-1A-1E	C ₆ H ₁₂ NO ₄ ⁺	162.0761	-0.7 \pm 0.5	0.4 \pm 1.3	9.0 \pm 2.0	2 4	-0.2 \pm 0.5	0.4 \pm 0.5	6.8 \pm 0.7	2 4
4-Aminobutyrate-1A-1E	C ₇ H ₁₄ NO ₄ ⁺	176.0917	^c	^c	^c	^c ^c	0.0 \pm 0.4	0.3 \pm 1.4	9.5 \pm 2.9	2 4
Asp-1A-2E	C ₈ H ₁₄ NO ₆ ⁺	220.0816	-0.6 \pm 0.5	^c	55.1 \pm 0.0	4 8	-0.8 \pm 0.4	^c	55.1 \pm 0.0	4 8
Benzoate-1E	C ₈ H ₉ O ₂ ⁺	137.0597	-0.8 \pm 1.0	0.3 \pm 0.2	4.9 \pm 0.6	1 3	1.0 \pm 1.2	0.0 \pm 0.3	4.6 \pm 0.4	1 4
Fumarate-2E	C ₆ H ₉ O ₄ ⁺	145.0495	-0.6 \pm 0.8	0.5 \pm 0.4	6.9 \pm 0.8	2 4	1.4 \pm 3.0	-0.6 \pm 0.6	15.4 \pm 3.5	1 3
Glu-1A-2E	C ₉ H ₁₆ NO ₆ ⁺	234.0972	-1.1 \pm 0.9	-2.7 \pm 2.2	21.6 \pm 11.7	3 8	-0.4 \pm 0.4	0.0 \pm 0.7	9.3 \pm 2.5	2 7
Hippurate-1E	C ₁₀ H ₁₂ NO ₃ ⁺	194.0812	-0.9 \pm 0.5	-0.3 \pm 1.0	8.5 \pm 2.1	2 5	-0.8 \pm 0.5	0.3 \pm 0.5	7.6 \pm 0.7	2 4
Ile-1A-1E	C ₉ H ₁₈ NO ₄ ⁺	204.1230	-0.6 \pm 0.7	^c	60.9 \pm 0.0	3 4	-0.9 \pm 0.3	0.4 \pm 0.4	8.1 \pm 0.8	2 4
Leu-1A-1E	C ₉ H ₁₈ NO ₄ ⁺	204.1230	-0.5 \pm 0.3	^c	60.9 \pm 0.0	3 4	-0.9 \pm 0.4	0.4 \pm 0.2	7.9 \pm 0.2	2 4

Table 12.6 continued.

MCF derivative	elemental composition [M+H] ⁺	m/z calcd	0.0 mL/h					0.4 mL/h				
			$\Delta m \pm SD$ (mDa) ^a	isotope ratio error M+H+1/M+H (%) ^a	mSigma \pm SD	avg. rank of correct formula ^b	avg. total number of formulas	$\Delta m \pm SD$ (mDa) ^a	isotope ratio error M+H+1/M+H (%) ^a	mSigma \pm SD	avg. rank of correct formula ^b	avg. total number of formulas
Malonate-2E	C ₅ H ₉ O ₄ ⁺	133.0495	-0.4 \pm 0.8	^c	33.2 \pm 0.0	2	4	1.2 \pm 1.4	0.1 \pm 0.2	5.7 \pm 0.3	2	3
Met-1A-1E	C ₈ H ₁₆ NO ₄ S ⁺	222.0795	-0.2 \pm 0.8	-2.9 \pm 1.9	28.2 \pm 13.7	5	8	-0.8 \pm 0.2	0.5 \pm 0.3	4.0 \pm 1.4	2	8
Me-malonate-2E	C ₆ H ₁₁ O ₄ ⁺	147.0652	-1.0 \pm 1.4	-3.0 \pm 2.0	28.1 \pm 12.9	2	4	0.2 \pm 0.8	0.4 \pm 0.4	6.6 \pm 0.9	2	3
Nicotinate-1E	C ₇ H ₈ NO ₂ ⁺	138.0550	-0.5 \pm 1.0	-0.6 \pm 0.5	5.9 \pm 1.7	1	2	1.1 \pm 1.2	0.2 \pm 0.5	4.9 \pm 1.5	1	2
Nval-1A-1E	C ₈ H ₁₆ NO ₄ ⁺	190.1074	-0.6 \pm 0.4	^c	54.5 \pm 0.0	3	4	-0.4 \pm 0.3	0.5 \pm 0.2	7.7 \pm 0.5	2	4
Phe-1A-1E	C ₁₂ H ₁₆ NO ₄ ⁺	238.1074	-0.7 \pm 0.5	^c	79.6 \pm 0.0	7	8	-1.3 \pm 0.5	0.5 \pm 0.3	10.3 \pm 0.5	2	7
Phenylacetate-1E	C ₉ H ₁₁ O ₂ ⁺	151.0754	-0.6 \pm 0.4	-0.6 \pm 1.0	7.7 \pm 2.2	1	3	0.3 \pm 0.8	0.2 \pm 0.4	5.5 \pm 0.9	1	4
Pro-1A-1E	C ₈ H ₁₄ NO ₄ ⁺	188.0917	-0.8 \pm 0.3	^c	54.3 \pm 0.0	3	4	-0.6 \pm 0.2	-0.2 \pm 0.8	8.0 \pm 1.8	2	4
Suberate-2E	C ₁₀ H ₁₉ O ₄ ⁺	203.1278	-0.4 \pm 0.7	^c	65.1 \pm 0.0	3	4	-1.1 \pm 0.6	0.0 \pm 0.5	8.4 \pm 0.5	2	4
Val-1A-1E	C ₈ H ₁₆ NO ₄ ⁺	190.1074	-1.0 \pm 0.2	^c	54.5 \pm 0.0	3	4	-0.8 \pm 0.5	0.6 \pm 0.3	7.9 \pm 0.5	2	4

^a Difference between predicted and measured variable; ^b Sorting according to mSigma value; ^c Not detected.

Calculation parameter: positive adducts, M+H; $\Delta m \leq \pm 5$ mDa; charge, 1; even electron number; filter H/C element ratio ≤ 3 ; check ring plus double bonds from -0.5 to 40; $0 \leq n_N \leq 5$; $0 \leq n_P \leq 1$; $0 \leq n_S \leq 2$.

Table 12.7 Figures of merit for the study of matrix effects via spike-in experiment in chapter 7. Concentrations of spike-in compounds in spike-in mixtures for *E. coli*, serum and urine are given, the internal standard used for quantification of each spike-in compound is listed, and mean recoveries \pm SD for the internal standards in the three different matrices are presented. Since endogenous metabolite concentrations varied among the three matrices, separate spike-in mixtures were used. Volumes of 0, 10, 20, 30, 40, 50, and 60 μ L for spike-in levels 1-6 were added to aliquots of pre-processed matrix, which corresponded to absolute amounts between 0.08-24.24 nmol of the spike-in metabolites. The final volume of MeOx-TMS derivatized matrix sample was 110 μ L.

Compound	<i>E. coli</i> spike-in mix Conc. (μ M)	Serum spike-in mix Conc. (μ M)	Urine spike-in mix Conc. (μ M)	Corresponding internal standard ^a
<i>Spike-in compound</i>				
2-Ketobutyrate	75.0	75.0	75.0	[U- ¹³ C]pyruvate
3-Methyl-2-oxovalerate	75.0	75.0	69.9	[U- ¹³ C]pyruvate
Methylmalonate	138.6	9.8	47.9	[U- ¹³ C]fumarate
Nicotinate	98.1	75.0	75.0	[² H ₇]trans-cinnamate
Phenylacetate	75.0	75.0	75.0	[² H ₇]trans-cinnamate
Glycerate	64.9	10.6	35.3	[U- ² H]succinate/ [U- ¹³ C]fumarate
Fumarate	67.9	75.0	29.3	[U- ¹³ C]fumarate
α -Ketoglutarate	7.8	75.0	113.8	[2,3,3- ² H ₃]malate
Phenyllactate	75.0	75.0	75.0	[U- ¹³ C]lactate/ [² H ₇]trans-cinnamate
Phenylpyruvate	8.1	75.0	75.0	[U- ¹³ C]pyruvate
Hydroxyphenylacetate	75.0	75.0	79.3	[² H ₇]trans-cinnamate
Homovanillate	75.0	75.0	20.9	[2,2,4,4- ² H ₄]citrate/ [² H ₇]trans-cinnamate
Phosphoglycerate	24.1	75.0	36.6	[2,2,4,4- ² H ₄]citrate/ [² H ₇]trans-cinnamate
Hydroxyphenylpyruvate	75.0	75.0	17.2	[2,2,4,4- ² H ₄]citrate/ [² H ₇]trans-cinnamate
Pyruvate	410.6	75.0	194.2	[U- ¹³ C]pyruvate

Table 12.7 continued.

	Mean recovery in <i>E. coli</i> ^b (%)	Mean recovery in serum ^b (%)	Mean recovery in urine ^b (%)
<i>Internal standard</i>			
[U- ¹³ C]pyruvate	107.9±11.2	113.4±22.4	83.6±33.5
[U- ¹³ C]lactate	69.3±3.9	102±9.3	96.2±25.6
[U- ² H]succinate	59.2±4.6	93.8±10.7	83.7±12.6
[U- ¹³ C]fumarate	94.5±11.9	97±9.6	76.1±12.1
[2,3,3- ² H ₃]malate	99.7±6.9	99.8±8.1	80.4±11
[² H ₇]trans-cinnamate	109.8±7.2	103.8±18.2	90.6±19.1
[2,2,4,4- ² H ₄]citrate	99.6±7.4	64.3±7.4	57.9±9.3

^a Internal standards in bold were chosen instead of those with poor recovery. ^b Calculation based on area integrals of [M+H]⁺ in the spiked matrix samples (N=30) divided by those in standards (N=12).

Table 12.8 List of compounds contained in our in-house GC-APCI-HRTOFMS database of 100 MeOx-TMS derivatized metabolite standards. Mostly sugars, fatty acids, amines, organic acids, alcohols, amino acids and amino acid metabolites are included. Isomer elution numbers are given for metabolites yielding multiple peaks after derivatization. MeOx, group introduced by methoximation; TMS, trimethylsilyl group.

No.	Metabolite		
1	N-Acetyl-aspartate, 2TMS	44	Lactate, 2TMS
2	N-Acetyl-D-glucosamine, MeOx, 4TMS	45	Lactose-1, MeOx, 8TMS
3	N-Acetyl-neuraminate, MeOx, 6TMS	46	Lactose-2, MeOx, 8TMS
4	Acetylputrescine, 2TMS	47	Laurate, TMS
5	cis-Aconitate, 3TMS	48	Leucine, TMS
6	Adipate, 2TMS	49	Leucine, 2TMS
7	Alanine, 2TMS	50	Lysine, 3TMS
8	Aspartate, 2TMS	51	Lysine, 4TMS
9	Aspartate, 3TMS	52	Malate, 3TMS
10	Citrate, 4TMS	53	Malonate, 2TMS
11	Creatinine, 3TMS	54	Mandelate, 2TMS
12	Dihydroorotate, 3TMS	55	Mannitol, 6TMS
13	Dimethylsuccinate, 2TMS	56	Methylmalonate, 2TMS
14	Erythritol, 4TMS	57	3-Methyl-2-oxovalerate, MeOx, TMS
15	Ethanolamine, 2TMS	58	4-Methyl-2-oxovalerate, MeOx, TMS
16	Fructose-1, MeOx, 5TMS	59	Myo-inositol, 6TMS
17	Fructose-2, MeOx, 5TMS	60	Nicotinamide, TMS
18	Fumarate, 2TMS	61	Nicotinate, TMS
19	GABA, 3TMS	62	Nonadecanoate, TMS
20	Glucose-1, MeOx, 5TMS	63	Nonanoate, TMS
21	Glucose-2, MeOx, 5TMS	64	Norvaline, TMS
22	Glucose-6-phosphate-1, MeOx, 6TMS	65	Norvaline, 2TMS
23	Glucose-6-phosphate-2, MeOx, 6TMS	66	Ornithine, 3TMS
24	Glycerate, 3TMS	67	Ornithine, 4TMS
25	Glycerol, 3TMS	68	Orotate, 3TMS
26	Glycerol-1-phosphate, 4TMS	69	Pantothenate, 3TMS
27	Glycine, 3TMS	70	Pentadecanoate, TMS
28	Heptadecanoate, TMS	71	Phenylacetate, TMS
29	Hippurate, TMS	72	Phenylalanine, TMS
30	Hippurate, 2TMS	73	Phenylalanine, 2TMS
31	Homogentisate, 3TMS	74	Phenyllactate, 2TMS
32	Homovanillate, 2TMS	75	Phenylpyruvate, MeOx, TMS
33	2-Hydroxybutyrate, 2TMS	76	Phosphoenolpyruvate, 3TMS
34	3-Hydroxybutyrate, 2TMS	77	3-Phosphoglycerate, 4TMS
35	5-Hydroxy-indoleacetate, 2TMS	78	O-Phosphorylethanolamine, 4TMS
36	5-Hydroxy-indoleacetate, 3TMS	79	Proline, TMS
37	Hydroxyphenylacetate, 2TMS	80	Proline, 2TMS
38	Hydroxyphenylpyruvate, MeOx, 2TMS	81	Pyruvate, MeOx, TMS
39	Hydroxyproline, 3TMS	82	Quinolate, 2TMS
40	Isoleucine, TMS	83	Ribitol, 5TMS
41	Isoleucine, 2TMS	84	Serine, 2TMS
42	Itaconate, 2TMS	85	Serine, 3TMS
43	α -Ketoglutarate, MeOx, 2TMS	86	Sorbitol, 6TMS
		87	Suberate, 2TMS

Table 12.8 continued.

No.	Metabolite
88	Succinate, 2TMS
89	Threonine, 2TMS
90	Threonine, 3TMS
91	Thymine, 2TMS
92	Tridecanoate, TMS
93	Tryptophan, 2TMS
94	Tryptophan, 3TMS
95	Tyrosine, 2TMS
96	Tyrosine, 3TMS
97	Undecanoate, TMS
98	Valine, TMS
99	Valine, 2TMS
100	Vanillate, 2TMS

Table 12.9 Mean signal-to-noise ratios of FAMEs computed from four subsequent replicate runs by means of GC-APCI I or APCI II-HRTOFMS. The “Find Compounds – Chromatogram” peak finder was applied on EICs of protonated molecules (concentration range: 5-13 μ M). Student’s t-tests with correction for multiple testing according to Benjamini and Hochberg [89] were performed on signal-to-noise ratios. For comparison of the not normally distributed mean S/N, the Wilcoxon signed-rank test was used. Reprinted from [13].

FAME	Mean S/N	Mean S/N
	[M+H] ⁺ APCI I	[M+H] ⁺ APCI II
C11:0 ^a	3100	10389
C13:0 ^a	6708	20694
C14:1 ^a	4718	17974
C15:1 ^a	1619	6040
C15:0 ^a	1268	7038
C16:1	17392	19306
C17:1 ^a	3503	10269
C17:0 ^a	2988	11120
C18:1c ^a	12543	22875
C18:1t ^a	6903	14504
C18:0 ^a	18250	37792
C20:1 ^a	5174	11573
C20:0 ^a	16589	22813
C21:0 ^a	6636	15339
C22:1 ^a	2387	14289
C22:0 ^a	6813	29035
C23:0 ^a	3455	17251
C24:1 ^a	1468	11356
C24:0 ^a	3164	23060
Arithmetic mean ^b	6562	16985

^a $p_{\text{adj}} < 0.05$ (Student’s t-test, corrected for multiple testing). ^b $p < 0.05$ (Wilcoxon)

Table 12.10 Figures of merit for ISs used in metabolite profiling by GC-HRTOFMS with either the APCI I or the APCI II source. Calibration curves of twenty MeOx-TMS derivatized metabolites were analyzed. (TMS derivatization status and isomer elution number are given for metabolites yielding multiple peaks after derivatization). Reprinted from [13].

Compound	GC-APCI II- HRTOFMS ret.time (min)	m/z ⁺ [M+H] ion	GC-APCI I- HRTOFMS RSD ^a (%)	GC-APCI II- HRTOFMS RSD ^a (%)
[U- ¹³ C]lactate	10.87	238.13	2.0	5.8
[U- ¹³ C]3-hydroxybutyrate	12.55	253.15	7.3	3.2
Nval-2TMS	13.90	262.17	8.0	4.5
[U- ² H ₄]succinate	15.00	267.14	6.0	3.7
[U- ¹³ C]fumarate	15.54	265.11	2.6	2.4
[² H ₇]trans-cinnamate	18.56	228.14	4.7	1.4
[U- ¹³ C]glucose-1	23.50	576.32	5.8	2.4
[4,6,7- ² H ₃]5-HI-[² H ₂]AA	26.75	341.18	7.6	9.9
[U- ¹³ C]lactose-1	31.86	960.51	8.1	9.5
Arithmetic mean			5.8	4.8

^a Concentration = 100 μM, N = 5 derivatization replicates.

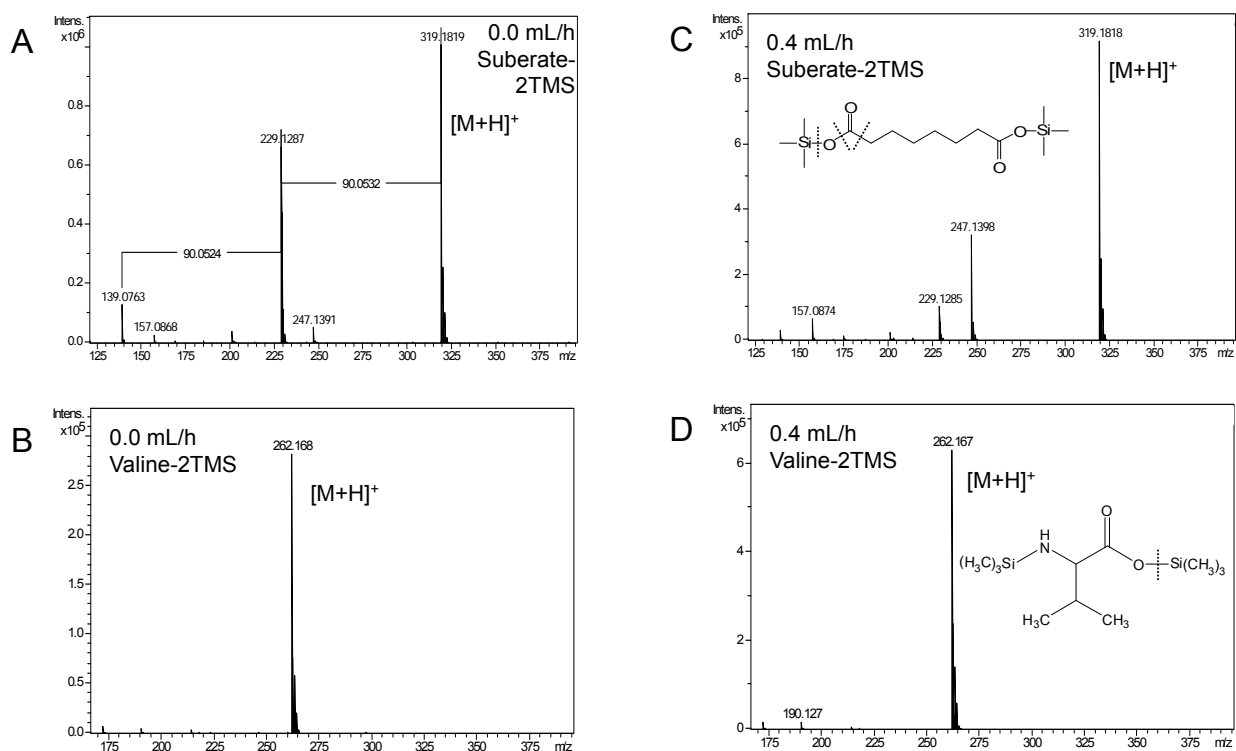


Figure 12.1 GC-APCI-TOFMS mass spectra of the MeOx-TMS derivatives of suberate and valine. Spectra A and B were acquired without, spectra C and D at 0.4 mL/h of continuous water infusion into the APCI source (Standard concn=100 μ M, TMS=trimethylsilyl group). Reprinted from [14].

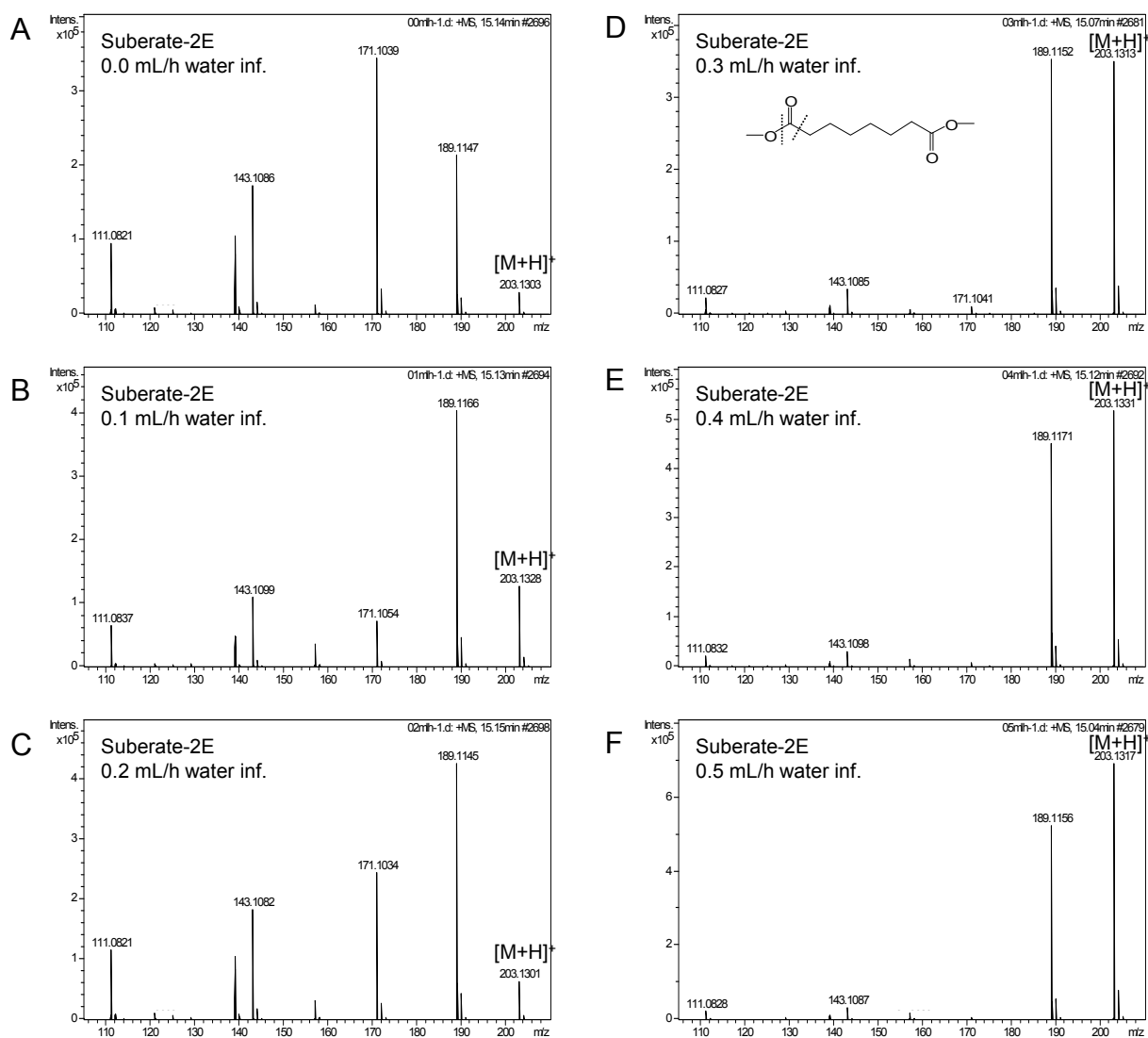


Figure 12.2 GC-APCI-TOFMS mass spectra of the MCF-derivatives of suberate and valine acquired without and at five different rates of water infusion. (A-F) refer to suberate-2E and (G-L) to valine-1E-1A, respectively. Water-infusion rate was increased in 0.1 mL/h increments from 0.1 to 0.5 mL/h (standard concn=200 μ M, E=methyl ester group, A=amide group). Reprinted from [14].

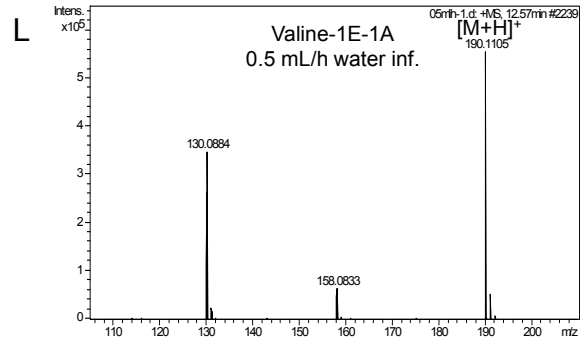
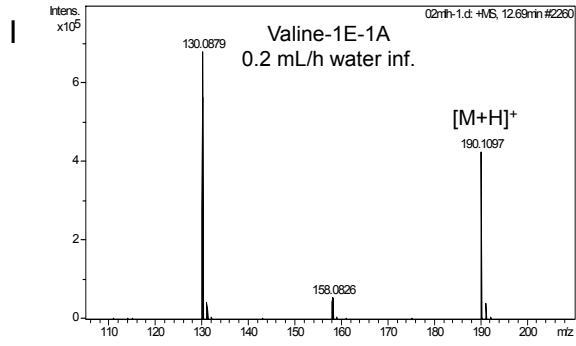
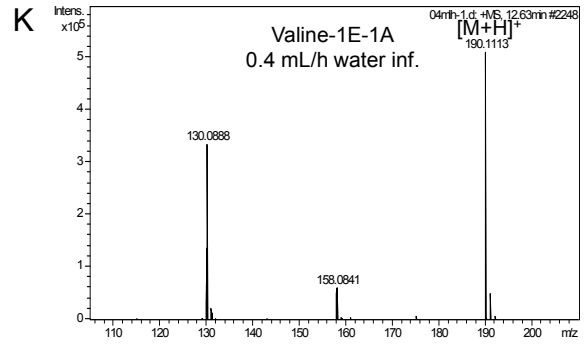
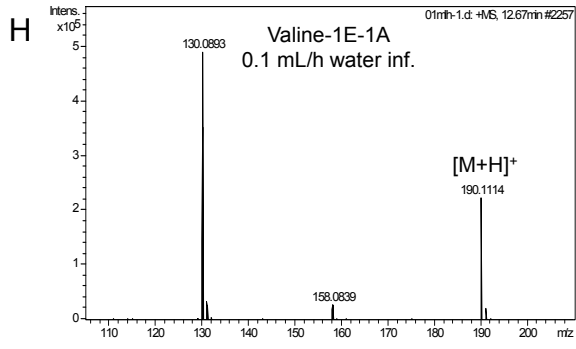
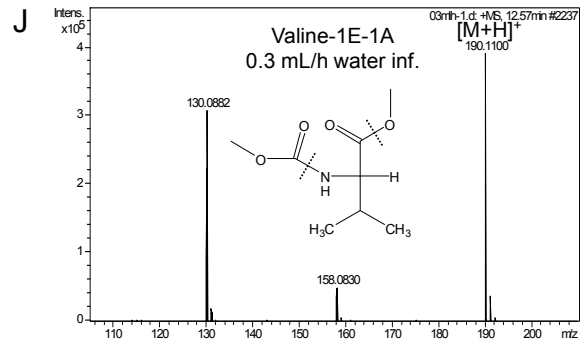
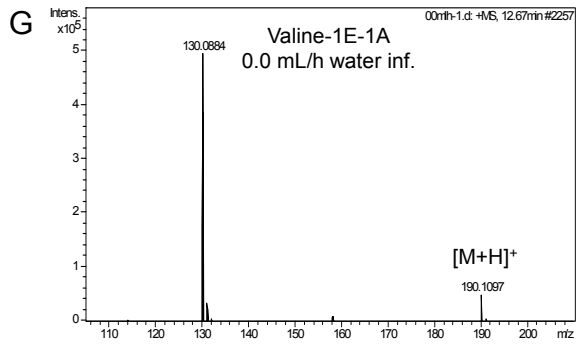


Figure 12.2 continued.

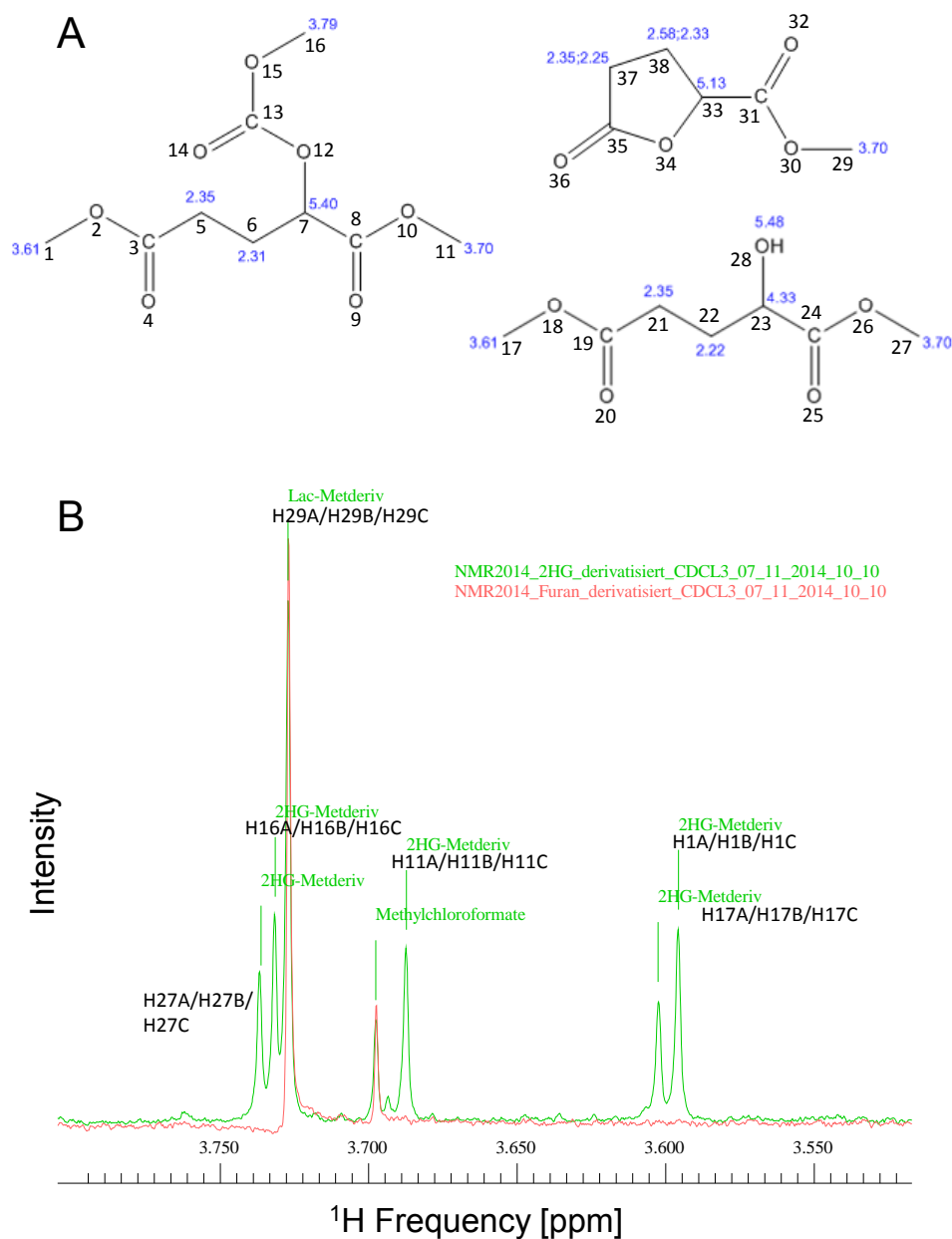


Figure 12.3 NMR analysis of aqueous and MCF derivatized standards of 2-HG and 5-Oxotetrahydro-2-furancarboxylate as well as serum of an AML patient with an IDH2 R140Q mutation. (A) Predicted ^1H -NMR chemical shifts in ppm (blue numbers) of the investigated derivatives of standard analytes by ChemDraw. (B, C) Overlaid 1D ^1H spectra of (B) MCF derivatized and (C) aqueous standards of 2-HG (green in (B), yellow in (C)) and 5-oxotetrahydro-2-furancarboxylate (orange in (B), green in (C)) at a concentration of 1 mM each. (D) Overlaid 2D ^1H - ^{13}C TOCSY spectra of serum of an AML patient with an IDH2 R140Q mutation (green) and aqueous standards of 2-HG (red) and 5-oxotetrahydro-2-furancarboxylate (orange). Experimental details are given in the experimental section of chapter 9. NMR peak assignments in (B) were based on predicted ^1H chemical shifts and similar ^1H intensities for methyl groups of each of the three derivatives (numbering is shown in (A)). Lac, 5-Oxotetrahydro-2-furancarboxylate; Metderiv, methyl ester.

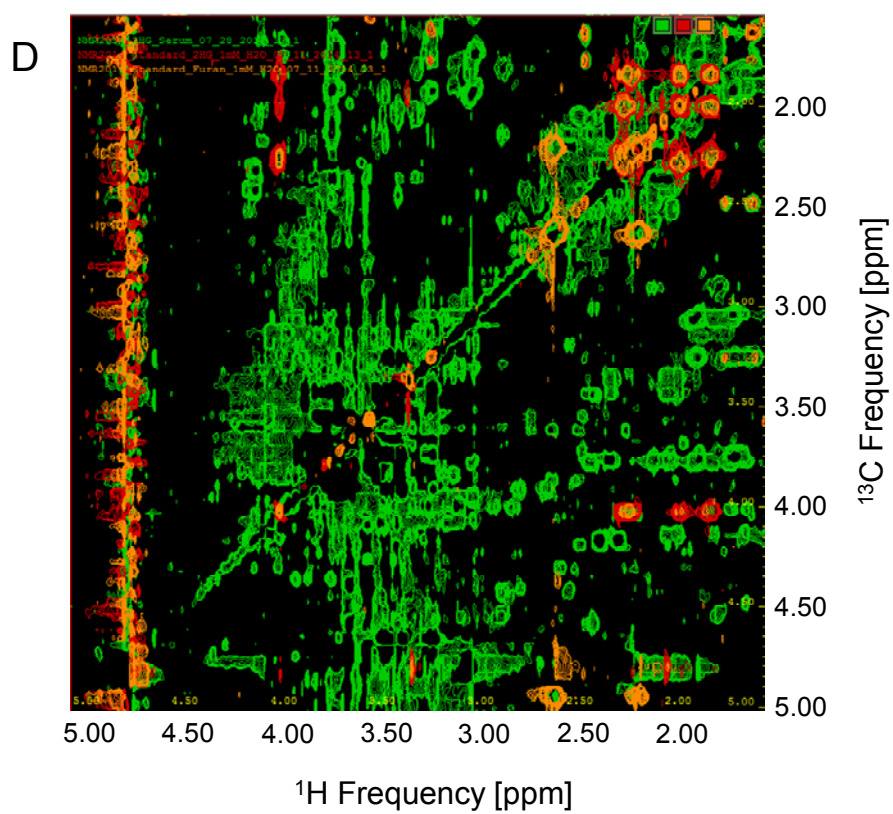
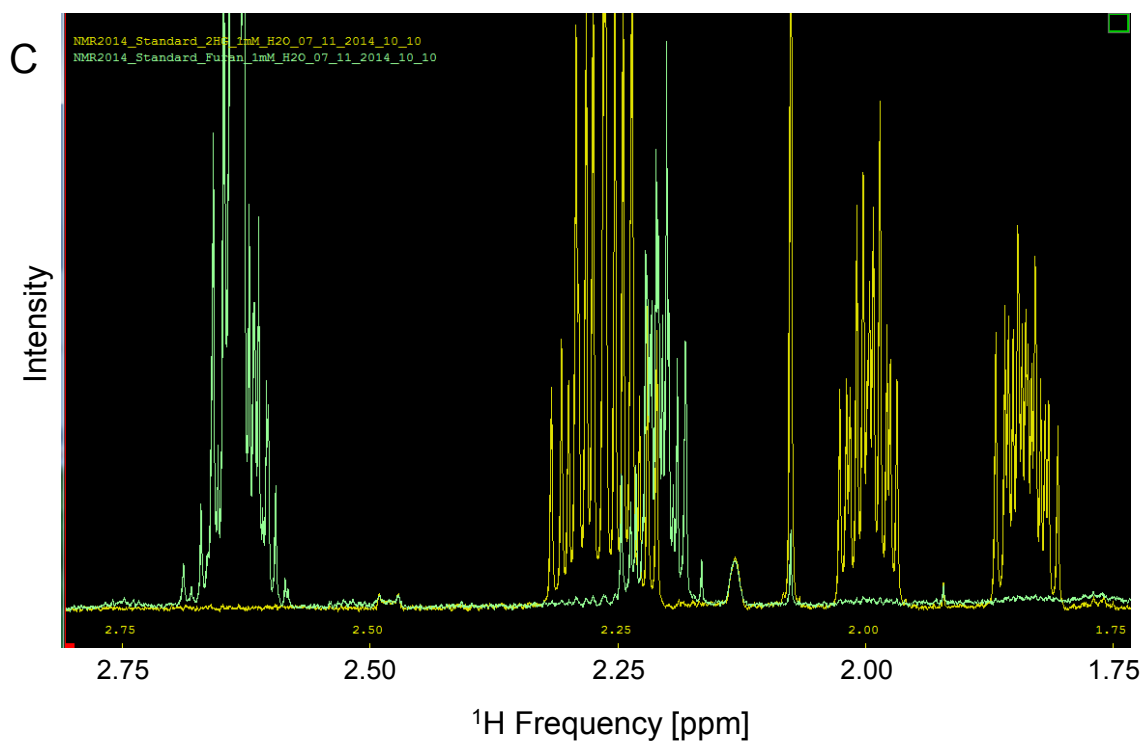


Figure 12.3 continued.

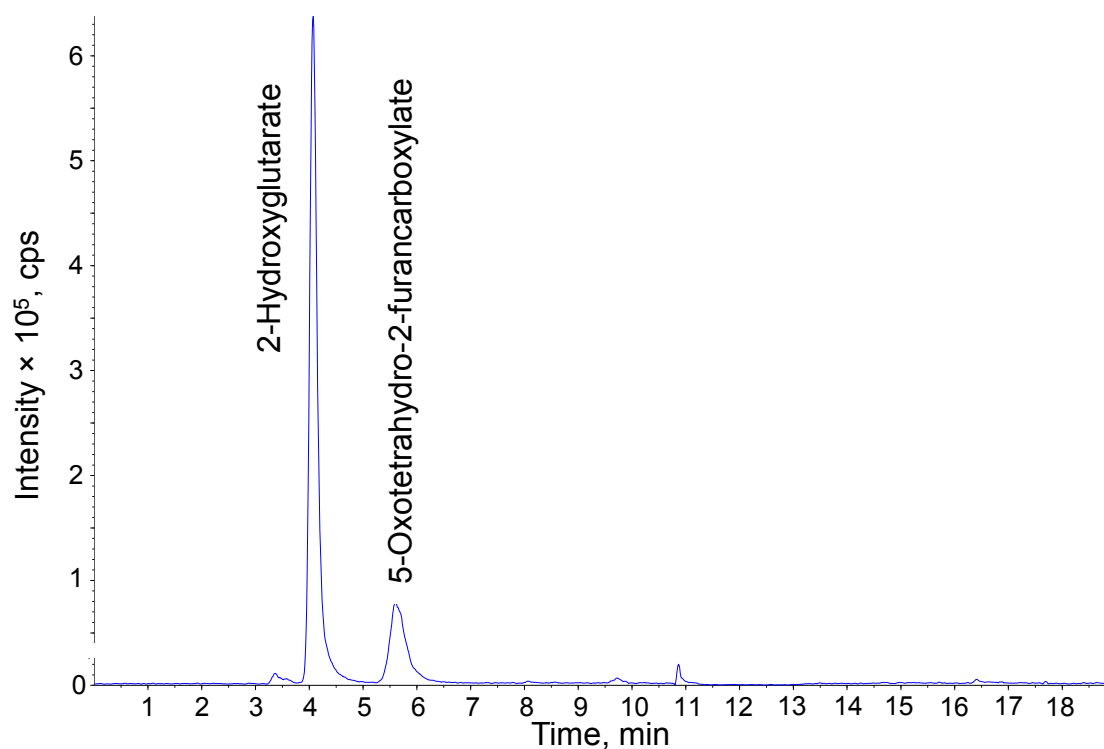


Figure 12.4 HPLC-ESI-MS/MS analysis of a standard mixture of 2-HG and 5-oxotetrahydro-2-furancarboxylate (31.25 μ M each) on a Discovery HS F5 PFPP column. 150 \times 2.1 mm, 3 μ m, 120 \AA , Discovery HS F5 (Supelco), P=134 bar. Elution with mobile phases A (0.1% formic acid in water, v/v) and B (acetonitrile) was performed using the following linear gradient: 0-6.5 min, hold at 100% A at a flow rate of 0.2 mL/min; 6.5-8 min, 0-100% B at 0.35 mL/min; 8-10 min, hold at 100% B at 0.35 mL/min; 10-10.1 min, 0-100% A at 0.35 mL/min; 10.1-17 min, hold at 100% A at 0.35 mL/min; 17-18 min, hold at 100% A at 0.2 mL/min. Detection was carried out using electrospray ionization in negative mode and multiple reaction monitoring on a 4000 QTRAP from AB Sciex. Further details are given in the experimental section of chapter 9.

13 Curriculum Vitae

Personal Details

Name	Christian J. Wachsmuth
Date of Birth	8th March 1986
Nationality	German

Education

04/2011 – 01/2015	PhD studies at the Institute of Functional Genomics/Metabolomics Department, University of Regensburg, Regensburg, Germany (Advisor: Prof. Dr. Peter J. Oefner)
10/2008 – 03/2011	MSc. of Chemistry, University of Regensburg Major: Analytical Chemistry
05/2010 – 03/2011	Master thesis at the IFG on "Evaluation of Gas Chromatography/Atmospheric Pressure Chemical Ionization Time-of-Flight Mass Spectrometry for Qualitative and Quantitative Metabolomics"
10/2005 – 09/2008	BSc. of Chemistry, University of Regensburg
09/1996 – 06/2005	A-Level, Grammar school in Mainburg

Further Qualifications

03/2013 – 06/2014	Intensive Course on Business Administration and Management Distance Education University of Hagen > Organization and leadership > Project and personnel management
02/2013 – 02/2014	Professional programme on "Examined Quality Expert GxP" German Chemists Society > GMP and GLP > Method validation in different quality systems
04/2012 – 03/2013	IT training programme, University of Regensburg
04/2012 – 02/2013	Series of seminars on "International Rhetorical Competence" University of Regensburg

Miscellaneous

Languages	English (business fluent oral and written) French (good oral and written)
Commitment	Pressman TSV Siegenburg, Division Table Tennis (2005-2011)
Activities	Swimming, Table Tennis, Reading

14 Publications and Presentations

14.1 Publications

Peer-Reviewed Journal Articles

- [1] Waldhier MC, **Wachsmuth CJ**, Gruber MA, Hellerbrand C, Zhu W, Nürnberger N, Dettmer K, Oefner PJ.
Chiralomics: adding a new dimension to metabolomics. (In preparation)
- [2] **Wachsmuth CJ**, Hahn TA, Oefner PJ, Dettmer K (2015).
Enhanced metabolite profiling using a redesigned atmospheric pressure chemical ionization source for gas chromatography coupled to high-resolution time-of-flight mass spectrometry. *Analytical Bioanalytical Chemistry*, Epub ahead of print on June 21.
- [3] Mycielska ME, **Wachsmuth CJ**, Dettmer K, Wagner C, Schlitt HJ, Oefner PJ, Geisler EK, Lang SA (2014).
Hsp90 inhibition affects cell metabolism by disrupting mitochondrial protein insertion. *Journal of Cell Biology and Cell Metabolism* 1(1):1-7.

- [4] **Wachsmuth CJ**, Dettmer K, Lang SA, Mycielska ME, Oefner PJ (2014).
Continuous water infusion enhances atmospheric pressure chemical ionization of methyl chloroformate derivatives in gas chromatography coupled to time-of-flight mass spectrometry based metabolomics.
Analytical Chemistry 86(18):9186-9195.
- [5] **Wachsmuth CJ**, Almstetter MF, Waldhier MC, Gruber MA, Nürnberger N, Oefner PJ, Dettmer K (2011).
Performance evaluation of gas chromatography – atmospheric pressure chemical ionization – time-of-flight mass spectrometry for metabolic fingerprinting and profiling.
Analytical Chemistry 83(19):7514-7522.

Book Chapters

- [1] **Wachsmuth CJ**, Oefner PJ, Dettmer K (2013).
Equipment and metabolite identification (ID) strategies for mass-based metabolomic analysis.
In B. C. Weimer & C. Slupsky (Eds.), *Metabolomics in food and nutrition* (pp. 3-28), Philadelphia: Woodhead Publishing.
- [2] **Wachsmuth CJ**, Vogl FC, Oefner PJ, Dettmer K (2013).
Gas chromatographic methods in metabolomics.
In T. Hyötyläinen & S. Wiedmer (Eds.), *Chromatographic methods in metabolomics* (pp. 87-113), Cambridge: RSC Publishing.
- [3] Dettmer K, Almstetter MF, **Wachsmuth CJ**, Oefner PJ, (2013).
Comprehensive two-dimensional gas chromatography for metabolomics. In W. Weckwerth & G. Kahl (Eds.), *The Handbook of Plant Metabolomics* (pp. 77-91), Weinheim: Wiley-VCH.

14.2 Oral Presentations

Continuous water infusion enhances atmospheric pressure chemical ionization for derivatization gas chromatography coupled to time-of-flight mass spectrometry in metabolomics.

30th International Symposium on Chromatography, Salzburg, Austria, 15 September 2014.

Gas chromatography – atmospheric pressure chemical ionization – time-of-flight mass spectrometry for qualitative and quantitative metabolomics.

Wissenschaftsforum Chemie 2013, Darmstadt, Germany, 2 September 2013. (Invited talk)

Gas chromatography – atmospheric pressure chemical ionization – time-of-flight mass spectrometry for qualitative and quantitative metabolomics.

13th Group Meeting of the Bavarian Genome Research Network, Martinsried, Germany, 29 January 2013.

Gas chromatography – atmospheric pressure chemical ionization – time-of-flight mass spectrometry for qualitative and quantitative metabolomics.

36th International Symposium on Capillary Chromatography and 9th GC×GC Symposium, Riva del Garda, Italy, 30 May 2012.

Evaluation of gas chromatography – atmospheric pressure chemical ionization – time-of-flight mass spectrometry for metabolic fingerprinting and profiling.

Bruker Daltonics Users' Meeting, Kassel, Germany, 22 March 2012. (Invited talk)

Gas chromatography – atmospheric pressure chemical ionization – time-of-flight mass spectrometry for qualitative and quantitative metabolomics.

22nd Postgraduate Seminar of the GDCh's AK Separation Science, Hohenroda, Germany, 9 January 2012.

14.3 Poster Presentations

Improvements in quantitative and qualitative metabolic profiling using a novel atmospheric pressure GC source coupled to high resolution TOF-MS analysis.

30th International Symposium on Chromatography, Salzburg, Austria, 14-18 September 2014.

Continuous water infusion enhances atmospheric pressure chemical ionization for derivatization gas chromatography coupled to time-of-flight mass spectrometry in untargeted metabolomics.

38th International Symposium on Capillary Chromatography and 11th GC×GC Symposium, Riva del Garda, Italy, 18-23 May 2014.

HPLC-MS/MS flux analysis reveals decreased mitochondrial activity in heat-shock protein 90 inhibited cancer cells.

39th International Symposium on High Performance Liquid Phase Separations and Related Techniques, Amsterdam, The Netherlands, 16-20 June 2013.

Analytical tools based on mass spectrometry for study of metabolism in heat-shock protein 90 inhibited cancer cells.

Symposium on Tumor Metabolism meets Immunology, Regensburg, Germany, 25-27 April 2013.

15 Summary

With the commercial introduction of atmospheric pressure chemical ionization for gas chromatography in 2008, GC-APCI coupled to high-resolution time-of-flight mass spectrometry (GC-APCI-HRTOFMS) became an interesting addition to the metabolomics toolbox. APCI is a soft ionization technique and its application to hyphenate GC to high resolution MS opens up promising new means for the identification of unknown signals in complex matrices. The actual utility of GC-APCI-HRTOFMS in metabolic fingerprinting and profiling applications to biological matrices is the topic of this doctoral thesis.

During comparison of GC-APCI-HRTOFMS with GC×GC-EI-TOFMS, GC-EI-TOFMS, GC-CI-qMS, and GC-EI-qMS in my master thesis, it was noticed that reproducibility of APCI was affected greatly by differences in humidity in the laboratory. Therefore, the impact of humidity in the APCI source on ionization efficiency and repeatability was systematically studied in the initial project of this doctoral thesis. Water was continuously infused to ensure a constant humidity during APCI in the analysis of methylchloroformate (MCF)- and methoxime-trimethylsilyl (MeOx-TMS) derivatized metabolites. These two different derivatization strategies are most commonly pursued in GC-MS based metabolome analyses. Several infusion rates were tested and a rate of 0.4 mL/h yielded an average 16.6-fold increase in intensity of the protonated molecules ($[M+H]^+$) of 20 MCF-derivatized metabolites through suppression of in-source fragmentation. Water infusion, however, did not improve efficiency and repeatability of APCI of methoxime-trimethylsilyl (MeOx-TMS) derivatives of metabolite standards. Then, the impact of water infusion on metabolic fingerprinting of biological specimens was investigated. Water infusion led to a marked increase in the number of metabolites identified in MCF-derivatized cancer cell extracts via their $[M+H]^+$ ions and improved repeatability of peak areas, almost doubling the number (N=23) of identified, significantly regulated metabolites (false discovery rate <0.05) between controls and cancer cells treated with the heat-shock protein 90 (Hsp90) inhibitor 17-DMAG.

Next, matrix effects caused by co-eluting compounds were investigated that might influence ionization. Strikingly, recovery of three out of seven internal standards used in a spike-in experiment was below 75% in one or several of the three biological matrices tested, namely *Escherichia coli* extract, serum, and urine. This was due to suppression by their respective endogenous metabolite that was present at a high concentration. Ion suppression caused by a co-eluting compound was further shown for three pairs of co-eluting standards employing standard mixtures with increasing concentrations of the co-eluting compounds. Overall these experiments demonstrated that matrix effects have to be taken into consideration in GC-APCI-MS.

In the course of my doctoral thesis, Bruker Daltonics (Bremen, Germany) introduced a redesigned APCI source. To test the capabilities of this source, MeOx-TMS derivatized supernatants of untreated cancer cells were analyzed by GC-APCI-HRTOFMS using both the original APCI I and the redesigned APCI II source. The latter source almost doubled the number of spectral features with signal-to-noise ratios greater than 20 that could be extracted from metabolite fingerprints and increased the absolute number of identified metabolites by 33% from 36 to 48. In addition, the median area RSDs of extracted features decreased from 33% to 24%. These improvements further resulted in a more than fourfold median decrease in lower limits of quantification to 0.002 - 3.91 μM as evidenced for 20 MeOx-TMS derivatized metabolite standards and a concomitant increase in the linear range by 0.5 to almost three orders of magnitude.

Finally, GC-APCI(II)-HRTOFMS was applied to the enantioselective quantitative profiling of the oncometabolite D-2-hydroxyglutarate (D-2-HG). MCF derivatization and GC analysis on a chiral γ -cyclodextrin (Rt- γ DEXsa) column were used to separate the D from the L enantiomer of 2-HG. Separation was optimized to avoid co-elution of D-2-HG with a highly abundant matrix compound present in cell culture media supplemented with bovine serum albumin. The use of APCI-HRTOFMS yielded highly specific quantifier ions and the infusion of water enhanced lower limits of quantification and repeatability by factors of ten and two, respectively. Analysis of a racemic 2-HG standard after MCF derivatization yielded a total of four peaks instead of the expected two signals for the D- and L-enantiomer. It was then shown that in addition to an open-chain three-fold

derivative of 2-HG the methyl ester of the D/L-2-HG lactone is formed during derivatization. Since lactone of D/L-2-HG was found to be naturally present in biological specimens, the developed method was eventually based on the open-chain three-fold derivative of 2-HG yielding LLOQ values of 0.49 μM (D-2-HG) and 0.24 μM (L-2-HG). The GC-APCI(II)-HRTOFMS approach was successfully applied to the determination of D/L-2-HG concentration levels in urine specimens of 23 acute myeloid leukemia (AML) patients and 6 healthy controls, which were validated by HPLC-MS/MS. The yet to be discovered source of 2-HG lactone in sera of AML patients carrying neomorphic isocitrate dehydrogenase mutations promises to shed new light on the pathogenesis and progression of AML.

In summary this doctoral thesis demonstrates that GC-APCI-HRTOFMS is a useful addition to the established GC-MS approaches in metabolomics. Studies on factors potentially influencing the ionization, namely water infusion, matrix effects and source type, distinctly widened the applicability of GC-APCI-HRTOFMS for qualitative and quantitative analysis of MeOx-TMS and MCF derivatized metabolites. The ability of APCI along with water infusion to efficiently ionize a broad range of MCF metabolites was proven and in the following applied to comparative metabolic fingerprinting in extracts of treated cancer cells. Finally, the outstanding quantitative capabilities of GC-APCI(II)-HRTOFMS were used in a first enantioselective profiling application for quantitative determination of the oncometabolite D-2-HG. Altogether, this doctoral thesis contributed significantly to the excellent progress GC-APCI-MS has made towards becoming a routine tool in metabolomics.

16 Zusammenfassung

Mit der kommerziellen Einführung der chemischen Ionisation bei Atmosphärendruck für die Gaschromatographie im Jahr 2008 wurde GC-APCI in Kombination mit hochauflösender Massenspektrometrie (GC-APCI-HRTOFMS) als zusätzliches Verfahren für das Methodenspektrum in der Metabolomik interessant. Bei der APCI handelt es sich um eine weiche Ionisationstechnik. In Kombination mit GC und hochauflösender MS eröffnet APCI neue erfolgversprechende Möglichkeiten zur Identifizierung von unbekanntem Signalen in komplexen Matrices. Diese Doktorarbeit behandelt die Frage, inwiefern GC-APCI-HRTOFMS in „Metabolic Fingerprinting“ und „Metabolic Profiling“-Untersuchungen in biologischen Matrices nützlich ist.

Im Laufe des Vergleichs von GC-APCI-HRTOFMS mit GC×GC-EI-TOFMS, GC-EI-TOFMS, GC-CI-qMS und GC-EI-qMS wurde festgestellt, dass sich Schwankungen in der Luftfeuchtigkeit im Labor sehr stark auf die Reproduzierbarkeit der APCI auswirkten. Deshalb wurde im anfänglichen Projekt dieser Doktorarbeit systematisch untersucht, ob die Feuchtigkeit in der APCI-Quelle die Ausbeute und Wiederholbarkeit der Ionisierung beeinflusst. Die Zugabe von Wasser erfolgte kontinuierlich während der Analyse von mittels Chlorameisensäuremethylester (MCF)/MeOH und sequenzieller Methoximierung/Silylierung derivatisierten Metaboliten, um einen konstanten Feuchtigkeitsgehalt bei der APCI zu gewährleisten. Diese beiden unterschiedlichen Derivatisierungsstrategien werden in Metabolomics-Untersuchungen mittels GC-MS am häufigsten gebraucht. Verschiedene Infusionsraten von Wasser wurden getestet. Durch Zugabe von 0.4 mL Wasser pro Stunde konnten die Intensitäten der protonierten Moleküle ($[M+H]^+$) der verwendeten 20 MCF derivatisierten Metabolitenstandards durchschnittlich um das 16.6-fache erhöht werden. Dies wurde aufgrund der resultierenden geringeren Fragmentierung in der APCI Quelle erreicht. Die Zugabe von Wasser führte hingegen im Fall von mittels sequenzieller Methoximierung/Silylierung (MeOx-TMS) derivatisierten Metabolitenstandards zu keiner Verbesserung bezüglich Ausbeute und Wiederholbarkeit der

APCI. Anschließend wurde überprüft, inwiefern sich die Zugabe von Wasser bei der Untersuchung von biologischen Proben mittels „Metabolic Fingerprinting“ auswirkt. Die Zugabe von Wasser erhöhte dabei die Anzahl der über das jeweilige protonierte Molekül identifizierten Metabolite in MCF-derivatisierten Extrakten von Krebszellen deutlich und verbesserte die Wiederholpräzision der Peakflächen, was die Zahl der identifizierten und signifikant regulierten Metabolite (False Discovery Rate <0.05) zwischen Kontrollen und den mit dem Hitzeschockprotein 90 (Hsp90) Inhibitor 17-DMAG behandelten Krebszellen auf insgesamt 23 fast verdoppelt hat.

Des Weiteren wurden mögliche Ionisierungseffekte aufgrund von ko-eluierenden Matrixbestandteilen untersucht. Insbesondere lag die Wiederfindungsrate von drei der sieben verwendeten internen Standards im Spike-in Experiment unterhalb von 75% in einer oder mehreren der drei untersuchten biologischen Matrices (*Escherichia coli* Extrakt, Serum und Urin). Dies war auf Ionensuppression durch den jeweiligen endogenen Metaboliten zurückzuführen, der hochkonzentriert vorlag. Durch eine ko-eluierende Standardverbindung hervorgerufene Ionisierungseffekte konnten ferner experimentell für drei solcher Paare aufgezeigt werden. Dabei wurden die mit den Analyten ko-eluierenden Standardverbindungen in steigenden Konzentrationen eingesetzt. Insgesamt hat sich in allen Experimenten zu Matrixeffekten herausgestellt, dass diese in der GC-APCI-MS berücksichtigt werden müssen.

Im Verlauf meiner Doktorarbeit wurde eine neu entworfene APCI-Quelle von Bruker Daltonics (Bremen, Deutschland) vorgestellt. Um die Leistungsfähigkeit dieser Quelle zu überprüfen, wurden MeOx-TMS derivatisierte Zellkulturüberstände unbehandelter Krebszellen mittels GC-APCI-HR-TOFMS analysiert. Dies erfolgte sowohl unter Verwendung der ursprünglichen APCI I Quelle als auch der neu entworfenen APCI II Quelle. Letztere Quelle konnte die Zahl der aus den metabolischen Fingerabdrücken gewonnenen Signale mit Signal-zu-Rausch-Verhältnis >20 fast verdoppeln. Darüber hinaus gelang es, die absolute Anzahl identifizierter Metabolite in diesen Proben um 33% von 36 auf 48 zu steigern und die mediane relative Standardabweichung der Peakflächen obiger gewonnener Signale von 33% auf 24% zu verringern. Außerdem war es unter Verwendung der verbesserten APCI II Quelle möglich, die mediane untere Quanti-

fizierungsgrenze für 20 MeOx-TMS derivatisierte Metabolitenstandards um mehr als das Vierfache auf 0.002-3.91 μM zu erniedrigen und gleichzeitig den linearen Bereich um eine halbe Größenordnung auf insgesamt drei Konzentrationsgrößenordnungen zu erhöhen.

Schließlich wurde GC-APCI(II)-HRTOFMS zur enantioselektiven, quantitativen Bestimmung des Onko-Metaboliten D-2-Hydroxyglutarat (D-2-HG) verwendet. MCF Derivatisierung mit nachfolgender gaschromatographischer Analyse mittels einer chiralen γ -Cyclodextrin (R_t - γ DEXsa) Kapillarsäule wurden eingesetzt, um das D- vom L-Enantiomer des 2-HGs abzutrennen. Die Trennung wurde optimiert, um die Koelution von D-2-HG und einer hochabundanten Matrixkomponente in Zellkulturmedien mit zugegebenem Rinderserumalbumin zu vermeiden. APCI und TOFMS ergaben Ionen von hoher Spezifität für die Quantifizierung. Die Zugabe von Wasser verbesserte die unteren Quantifizierungsgrenzen und die Reproduzierbarkeit der Methode um das Zehn- bzw. Zweifache. Bei der Analyse eines racemischen 2-HG Standards nach MCF Derivatisierung konnten anders als vermutet insgesamt vier statt zwei Signale für das D- und L-Enantiomer detektiert werden. Daraufhin wurde gezeigt, dass zusätzlich zum offenkettigen Dreifachderivat der Methylester des Lactons von 2-HG während der Derivatisierung gebildet wird. Dadurch, dass das fünfgliedrige Lacton von D/L-2-HG natürlich in biologischen Proben vorkommt, wurde in der entwickelten Methode schließlich das offenkettige Dreifachderivat von 2-HG zur Quantifizierung herangezogen. Daraus resultierten untere Quantifizierungsgrenzen von 0.49 μM (D-2-HG) und 0.24 μM (L-2-HG). GC-APCI(II)-HRTOFMS konnte erfolgreich zur Bestimmung von D/L-2-HG Spiegeln in Urinproben von 23 Patienten mit akuter myeloischer Leukämie (AML) und 6 gesunden Probanden angewandt und die erzielten Ergebnisse mittels HPLC-MS/MS validiert werden. Die Aufklärung des Ursprungs des 2-HG-Lactons in Serumproben von AML Patienten mit mutierter Isocitrat-Dehydrogenase verspricht neue Einblicke in Entstehen und Fortschreiten von AML zu geben.

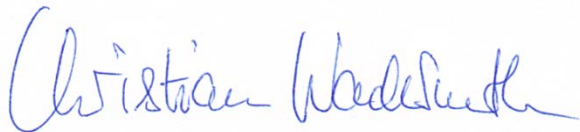
Zusammengefasst konnte in dieser Doktorarbeit aufgezeigt werden, dass GC-APCI-HRTOFMS ein nützliches zusätzliches Verfahren zu den etablierten GC-MS Methoden in der Metabolomik ist. Studien zu Faktoren, welche die Ionisierung möglicherweise

beeinflussen – nämlich die Zugabe von Wasser, Matrixeffekte und Quellentyp – konnten die Anwendbarkeit von GC-APCI-HRTOFMS für die qualitative und quantitative Analyse von MeOx-TMS und MCF derivatisierten Metaboliten erheblich steigern. Dabei konnte nachgewiesen werden, dass mittels APCI und Wasserzugabe eine große Bandbreite an MCF derivatisierten Metaboliten ionisiert werden kann. Dies wurde nachfolgend bei der vergleichenden Analyse metabolischer Fingerabdrücke von Extrakten behandelter Krebszellen angewandt. Schlussendlich wurde die ausgezeichnete quantitative Leistungsfähigkeit von GC-APCI(II)-HRTOFMS in einer ersten enantioselektiven „Metabolic Profiling“-Anwendung für die Quantifizierung des Onko-Metaboliten D-2-HG genutzt. Summa summarum hat diese Doktorarbeit einen wichtigen Beitrag zur zukünftigen, routinemäßigen Anwendung der GC-APCI-MS in der Metabolomik geleistet.

Erklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Regensburg, 29.06.2015



Christian Wachsmuth