

**Development and Optimization of Targeted/Untargeted
Lipidomic Screening Methodologies in Pharmaceutical and
Clinical Bioanalysis by Liquid Chromatography-High
Resolution-Mass Spectrometry**

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät
der Eberhard Karls Universität Tübingen
zur Erlangung des Grades eines
Doktors der Naturwissenschaften
(Dr. rer. nat.)

vorgelegt von
Bernhard Drotleff
aus Mediasch / Siebenbürgen (Rumänien)

Tübingen
2019

The research described in this thesis was conducted between April 15th 2015 and July 15th 2019 at the Institute of Pharmaceutical Sciences, Division Pharmaceutical (Bio-)Analysis, Eberhard Karls Universität Tübingen under the supervision of Prof. Dr. Michael Lämmerhofer.

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation: 10.12.2019

Dekan: Prof. Dr. Wolfgang Rosenstiel

1. Berichterstatter: Prof. Dr. Michael Lämmerhofer
2. Berichterstatter: Prof. Dr. Stefan Laufer
3. Berichterstatter: Prof. Dr. Gerhard Scriba

Table of Contents

I.	Summary of the Dissertation	V
II.	Zusammenfassung der Dissertation	VII
III.	List of publications	X
IV.	Author contributions	XI
V.	Poster Presentations	XVI
VI.	Oral Presentations	XVII
VII.	Abbreviations	XVIII
1.	Introduction	1
1.1.	Lipidomics	1
1.1.1.	Sample Preparation	3
1.1.2.	Analytical Strategies	5
1.2.	Mass Spectrometry	7
1.2.1.	Ionization	8
1.2.2.	Mass Analyzers	10
1.2.3.	Detector System	15
1.2.4.	Scan Modes	17
1.2.4.1.	Product Ion Scan	17
1.2.4.2.	Data-dependent Acquisition	19
1.2.4.3.	Data-independent Acquisition	20
1.2.5.	Method Optimization	22
1.2.6.	Data Processing in Untargeted Lipidomics	25
1.2.7.	Statistical Evaluation	28
1.3.	Bioanalytical Method Validation	29
1.3.1.	Calibration of Endogenous Compounds	33
1.4.	References	36
2.	List of Figures	42
3.	List of Tables	43
4.	Objectives of the Thesis	44
5.	Results and Discussion	46
5.1.	Quantification of Steroid Hormones in Plasma Using a Surrogate Calibrant Approach and UHPLC-ESI-QTOF-MS/MS with SWATH-Acquisition Combined with Untargeted Profiling	46
5.1.1.	Abstract	47
5.1.2.	Introduction	48
5.1.3.	Experimental Section	50

5.1.4.	Results and Discussion.....	53
5.1.5.	Conclusions	66
5.1.6.	References	67
5.1.7.	Supplementary Data	69
5.2.	Guidelines for Selection of Internal Standard-based Normalization Strategies in Untargeted Lipidomic Profiling by LC-HR-MS/MS	107
5.2.1.	Abstract	108
5.2.2.	Introduction.....	109
5.2.3.	Methods.....	110
5.2.4.	Results and Discussion.....	114
5.2.5.	Conclusion.....	122
5.2.6.	References	124
5.2.7.	Supporting Information.....	126
5.3.	Comprehensive Lipidomics of Mouse Plasma using Class-Specific Surrogate Calibrants and SWATH Acquisition for Untargeted Quantification.	151
5.3.1.	Abstract	152
5.3.2.	Introduction.....	153
5.3.3.	Experimental Section.....	154
5.3.4.	Results and Discussion.....	160
5.3.5.	Conclusions	171
5.3.6.	References	174
5.3.7.	Supplementary Material.....	177
6.	Appendix	202
6.1.	Supplementary Accepted Manuscripts	202
6.1.1.	Insulin and Estrogen Independently and Differentially Reduce Macronutrient Intake in Healthy Men.....	202
6.1.2.	Comparison of Simple Monophasic versus Classical Biphasic Extraction Protocols for Comprehensive UHPLC-MS/MS Lipidomic Analysis of HeLa Cells	212
6.2.	Data Processing Script for R.....	222
6.2.1.	Selection of Internal Standard-based Normalization for MS-Data	222
6.2.1.1.	Readme.md	222
6.2.1.2.	Script Code	224
7.	Acknowledgements	236

I. Summary of the Dissertation

Throughout the last decades, substantial technological innovations in mass spectrometry (MS) fueled a continuous progress in analytical chemistry and enabled the accelerated exploration of various “omics” branches. Lipidomics, a subset of metabolomics, has recently attracted increasing attention, since lipids, with their diverse physiological functions as structural components of membranes, energy depots and signaling molecules, have been recognized as significant factors in disease onset and progression, e.g. for central health concerns like cardiovascular disease and cancer. Besides the improved instrument hardware, also analytical strategies and bioinformatics have advanced to yield reliable qualification and quantification, which aid in the ongoing decryption of the lipidome and its pathways and networks.

To obtain sufficient sensitivity, selectivity and analyte coverage, hybrid quadrupole time-of-flight (QTOF) mass spectrometers are often utilized, as their rapid acquisition rates allow to combine the recording of high resolution mass spectra with the hyphenation to ultra-high performance liquid chromatography (UHPLC). Along with the introduction of sequential window acquisition of all theoretical fragment ion mass spectra (SWATH), a powerful tool for true comprehensive analysis was made available.

In this thesis, the potential of a UHPLC-QTOF platform was exploited to achieve a maximum yield of extractable information from biological samples in the context of lipidomic workflows. Besides the comparison of different lipid extraction strategies for HeLa cells, also the development of a specialized sample preparation protocol for plasma steroids was conducted. Together with SWATH acquisition and a thorough optimization of MS parameters, absolute quantification of low picomolar levels of testosterone and estradiol was attained. An obstacle for the quantification of endogenous compounds is typically the lack of a true blank matrix, which was compensated by surrogate calibration via $^{13}\text{C}_3$ -labeled target analyte analogues. The established method was validated according to international guidelines and the accuracy and precision were additionally verified by the analysis of external quality control (QC) samples. Later, over 300 clinical samples were analyzed and the obtained results were utilized to monitor and interpret the influence of estradiol treatment on food intake in healthy men. The observed lowered protein consumption was shown to be independent from alterations in macronutrient ingestion induced by insulin administration. Furthermore, the merged targeted/untargeted study design enabled simultaneous screening of additional steroids and revealed a significant reduction in epitestosterone, dihydrotestosterone and hydroxyprogesterone levels after estradiol intake.

Moreover, in the framework of an exclusively untargeted lipidomic study, different normalization strategies, which are usually required to control for unwanted variation, were assessed. After compiling a QC-based workflow to effectively compare the performance of

normalization methods, novel guidelines for selecting the best suitable strategy, while maintaining data integrity, were proposed. In addition, a script-based statistical and bioinformatical tool was provided as an open source solution to facilitate the implementation of these guidelines into the existing data processing workflows. Eventually, a contribution towards the necessary harmonization of data handling approaches in untargeted lipidomics was supplied for the scientific community.

In spite of the increasing efforts in untargeted analysis, the majority of studies is still only able to report results as relative foldchanges. Without absolute quantification, though, comparability to other studies or databases is limited and follow-up experiments have to be conducted to estimate reference or abnormal levels of potential biomarkers. To overcome this issue, a study with deuterated, lipid class-specific standards as class-wide surrogate calibrants was designed. Matrix-matched calibrants and QCs were incorporated into the analytical sequence of an untargeted plasma study and precision and accuracy for representative surrogate lipids was validated. Lipid species were separated with a reversed-phase UHPLC method and data was acquired using SWATH. Due to different ionization efficiencies and instrument responses between lipid species, which depend on carbon chain length, degree of saturation, matrix effects and solvent composition during elution, response factors had to be considered for class-wide extrapolation of absolute concentrations. It could be demonstrated that surrogate calibration resulted in more accurate quantification of lipid levels than one-point calibration. With post-acquisition re-calibration, which is describing the experimental determination of response factors after analysis and data processing for lipids of interest, a workflow, that is capable to estimate lipid levels in untargeted assays, was suggested.

II. Zusammenfassung der Dissertation

In den vergangenen Jahrzehnten haben beachtliche technologische Innovationen in der Massenspektrometrie (MS) einen kontinuierlichen Fortschritt der analytischen Chemie vorangetrieben und die Erforschung verschiedener „Omik“ Bereiche beschleunigt. Die Lipidomik, ein Teilgebiet der Metabolomik, hat in letzter Zeit zunehmend Aufmerksamkeit erregt, da Lipiden ein signifikanter Einfluss beim Ausbruch und Verlauf von weitläufigen Krankheiten wie Krebs oder Herz-Kreislauf-Erkrankungen zugesprochen wird. Dies ist vor allem auf ihre vielfältigen physiologischen Aufgaben, z.B. als Membranbausteine, Energiespeicher und Signalmoleküle, zurückzuführen. Neben der verbesserten Geräteleistung haben sich auch die analytische Herangehensweise und die Methoden der Bioinformatik weiterentwickelt. Hierdurch wurden Grundlagen für eine zuverlässige Qualifizierung und Quantifizierung geschaffen, die bei der stetigen Entschlüsselung des Lipidoms sowie der damit verbundenen Stoffwechselwege und metabolischen Netzwerke, hilfreich sind.

Damit eine ausreichende Empfindlichkeit, Selektivität und Analytabdeckung erreicht werden kann, setzt man häufig Quadrupol-Flugzeit-Massenspektrometer (QTOFs) ein, da ihre schnellen Aufnahmegeschwindigkeiten die Generierung von hochauflösenden Massenspektren während der Kopplung an Ultra-Hochleistungs-Flüssig-Chromatographie (UHPLC) erlauben. Mit der Einführung einer ergänzenden Akquisitionstechnik namens SWATH (engl.: sequential window acquisition of all theoretical fragment-ion spectra) wurde ein leistungsstarkes Hilfsmittel für eine umfassende Analytik zur Verfügung gestellt.

In der folgenden Arbeit wurde das Potenzial eines UHPLC-QTOF Instruments genutzt, um bei der Untersuchung des Lipidoms in biologischen Proben den maximalen Informationsgewinn zu erzielen. Neben dem Vergleich verschiedener Extraktionsprotokolle für HeLa-Zellen wurde zudem die Entwicklung einer spezialisierten Probenaufbereitungsmethode für Plasma-Steroide durchgeführt. Mit Hilfe von SWATH und einer sorgfältigen Optimierung der massenspektrometrischen Parameter, wurden Bestimmungsgrenzen im niedrigen picomolaren Bereich für Testosteron und Estradiol erreicht. Ein Hindernis bei der Quantifizierung endogener Verbindungen ist üblicherweise das Fehlen einer echten Leermatrix. Dies konnte jedoch durch den Einsatz einer sogenannten Surrogat-Kalibrierung mittels $^{13}\text{C}_3$ -markierten Analoga der Zielanalyte kompensiert werden. Die fertige Methode wurde nach internationalen Richtlinien validiert und Präzision und Richtigkeit wurden zusätzlich über externe Qualitätskontrollproben bestätigt. Anschließend wurden über 300 klinische Proben vermessen. Anhand der erhaltenen Ergebnisse konnte der Einfluss einer Estradiol-Behandlung auf die Nahrungsaufnahme bei männlichen Probanden überwacht und interpretiert werden. Dabei wurde ein verringerter Proteinkonsum festgestellt, der unabhängig von Veränderungen der Makronährstoffaufnahme ist, welche durch Insulingabe induziert

werden. Durch das kombinierte Studiendesign, welches neben den Zielanalyten eine ungezielte Detektion von Analyten erlaubt, war es zudem möglich nach weiteren Steroiden in den Proben zu suchen. Hierdurch konnte gezeigt werden, dass eine signifikante Abnahme der Epitestosteron-, Dihydrotestosteron- sowie Hydroxyprogesteron-Level durch die Einnahme von Estradiol verursacht wurde.

Des Weiteren wurden im Rahmen einer anderen Studie, welche ausschließlich auf die ungezielte Detektion von Analyten ausgerichtet war, verschiedene Methoden zur Normalisierung beurteilt. Diese werden für gewöhnlich angewendet um ungewollte Variationen in den Proben zu verringern. Nachdem ein auf Qualitätskontrollproben basierender Arbeitsablauf zur effektiven Beurteilung der Güte einer Normalisierung zusammengestellt wurde, konnten neue Richtlinien zur Auswahl der bestmöglichen Strategie vorgeschlagen werden. Ein wichtiger Aspekt war dabei auch, die Überprüfung der Plausibilität der Daten nach der Normalisierung. Darüber hinaus wurde ein bioinformatisches und statistisches Open Source Skript zur Verfügung gestellt, welches eine vereinfachte Implementierung der formulierten Richtlinien in bestehende Abläufe zur Datenprozessierung gewährleisten kann. Dadurch wurde letztendlich ein Beitrag für die wissenschaftliche Gemeinschaft erbracht, der zur erforderlichen Harmonisierung der Datenhandhabung bei ungerichteten Lipidomanalysen genutzt werden kann.

Trotz des erhöhten Aufwands bei ungerichteten Analysen, wird bei den meisten Studien nur eine relative Quantifizierung erzielt. Ohne absolute Konzentrationsangaben ist die Vergleichbarkeit mit anderen Studien oder Datenbanken jedoch eingeschränkt und Folgeversuche zur Abschätzung von abnormalen Konzentrationen oder Referenzbereichen potenzieller Biomarker müssen unternommen werden. Um diesem Problem entgegenzuwirken wurde eine Studie initiiert, bei der deuterierte, Lipidklassen-spezifische Standardsubstanzen als Surrogat-Kalibranten verwendet wurden. Hierdurch sollte eine klassenbasierte Quantifizierung realisiert werden. Kalibranten und Qualitätskontrollproben der Surrogat-Substanzen wurden in unbehandeltem Plasma hergestellt und in die Sequenz einer ungerichteten Plasmalipid-Studie eingebettet. Dabei wurde die Präzision und Richtigkeit der Surrogat-Kalibranten überprüft und validiert. Die einzelnen Lipidspezies wurden über eine Umkehrphasen-UHPLC Methode getrennt und mittels SWATH analysiert.

Aufgrund von abweichenden Ionisierungseffizienzen und Unterschieden in der Signalstärke zwischen Lipidspezies, welche von der Kettenlänge, vom Sättigungsgrad, von Matrixeffekten und von der Zusammensetzung der mobilen Phase während der Elution abhängen, mussten sogenannte Response-Faktoren beachtet werden um eine klassenweite Extrapolation der Ergebnisse zu gewährleisten. Es wurde aufgezeigt, dass die Surrogat-Kalibrierung eine genauere Quantifizierung von Lipiden erlaubt als eine Ein-Punkt-Kalibrierung.

Durch eine neu formulierte Strategie der „Post-Akquisition Re-Kalibrierung“, welche die experimentelle Bestimmung von Response-Faktoren für Lipide von besonderem Interesse nach der Messung und Datenprozessierung beschreibt, wurde ein möglicher Arbeitsablauf angedeutet, mit dem eine Abschätzung von Lipidgehalten in ungezielten Analysen vollzogen werden kann.

III. List of publications

Publication I

Drotleff, B.; Hallschmid, M.; Lämmerhofer, M. Quantification of steroid hormones in plasma using a surrogate calibrant approach and UHPLC-ESI-QTOF-MS/MS with SWATH-acquisition combined with untargeted profiling. *Anal. Chim. Acta* **2018**, 1022:70–80, DOI: 10.1016/j.aca.2018.03.040.

Publication II

Drotleff, B.; Lämmerhofer, M. Guidelines for Selection of Internal Standard-based Normalization Strategies in Untargeted Lipidomic Profiling by LC-HR-MS/MS. *Anal. Chem.* **2019**, in press, DOI: 10.1021/acs.analchem.9b01505.

Publication III (Submitted Manuscript; *Analytica Chimica Acta*)

Drotleff, B.; Illison, J.; Schlotterbeck, J.; Lukowski, R.; Lämmerhofer, M. Comprehensive Lipidomics of Mouse Plasma using Class-Specific Surrogate Calibrants and SWATH Acquisition for Untargeted Quantification. Status: Resubmitted after revision.

Publication IV

Krug, R.; Mohwinkel, L.; Drotleff, B.; Born, J.; Hallschmid, M. Insulin and Estrogen Independently and Differentially Reduce Macronutrient Intake in Healthy Men. *J. Clin. Endocrinol. Metab.* **2018**, 103:1393-1401, DOI: 10.1210/jc.2017-01835.

Publication V

Calderón, C.; Sanwald, C.; Schlotterbeck, J.; Drotleff, B.; Lämmerhofer, M. Comparison of simple monophasic versus classical biphasic extraction protocols for comprehensive UHPLC-MS/MS lipidomic analysis of HeLa cells. *Anal. Chim. Acta* **2018**, 1048:66-74, DOI: 10.1016/j.aca.2018.10.035.

IV. Author contributions

Publication I

Quantification of steroid hormones in plasma using a surrogate calibrant approach and UHPLC-ESI-QTOF-MS/MS with SWATH-acquisition combined with untargeted profiling.

Bernhard Drotleff:

General idea generation

Method development

Sample preparation and analysis

Data processing and interpretation

Main writing of the manuscript

Prof. Dr. Manfred Hallschmid:

Sample collection

Immunoassay measurements

Proofreading of the manuscript

Prof. Dr. Michael Lämmerhofer:

Generation, initiation, coordination and financing of the project

Discussion of results and interpretation

Partial writing and editing of the manuscript

Proofreading and final approval of the manuscript

Corresponding Author

Publication II

Guidelines for Selection of Internal Standard-based Normalization Strategies in Untargeted Lipidomic Profiling by LC-HR-MS/MS.

Bernhard Drotleff:

General idea generation

Sample preparation and analysis

Data processing and interpretation

Programming in R statistical language

Main writing of the manuscript

Prof. Dr. Michael Lämmerhofer:

Generation, initiation, coordination and financing of the project

Discussion of results and interpretation

Partial writing and editing of the manuscript

Proofreading and final approval of the manuscript

Corresponding Author

Publication III (Submitted Manuscript; Analytica Chimica Acta)

Comprehensive Lipidomics of Mouse Plasma using Class-Specific Surrogate Calibrants and SWATH Acquisition for Untargeted Quantification.

Bernhard Drotleff:

General idea generation

Sample preparation and analysis

Data processing and interpretation

Main writing of the manuscript

Julia Illison:

Handling of mice and sample collection

Jörg Schlotterbeck:

Proofreading of the manuscript

Mass spectrometer maintenance

Prof. Dr. Robert Lukowski:

Generation, initiation, coordination and financing of the biological part of the project

Proofreading of the manuscript

Prof. Dr. Michael Lämmerhofer:

Generation, initiation, coordination and financing of the project

Discussion of results and interpretation

Partial writing and editing of the manuscript

Proofreading and final approval of the manuscript

Corresponding Author

Publication IV

Insulin and Estrogen Independently and Differentially Reduce Macronutrient Intake in Healthy Men.

Rosemarie Krug:

Designed and conducted research

Data analysis

Writing of the manuscript

Linda Mohwinkel:

Conducted research

Data analysis

Bernhard Drotleff:

Conducted research/MS experiments

Data analysis (steroid analysis)

Proofreading of the manuscript

Prof. Dr. Jan Born:

Designed research

Prof. Dr. Manfred Hallschmid:

Designed and conducted research

Data analysis

Writing of the manuscript

Publication V

Comparison of simple monophasic versus classical biphasic extraction protocols for comprehensive UHPLC-MS/MS lipidomic analysis of HeLa cells.

Carlos Calderón:

General idea generation

Sample preparation and analysis

Data processing and interpretation

Main writing of the manuscript

Corinna Sanwald:

HeLa cell cultivation

Proofreading of the manuscript

Jörg Schlotterbeck:

Proofreading of the manuscript

Mass spectrometer maintenance

Bernhard Drotleff:

Advice on lipidomics profiling by SWATH-MS/MS

Proofreading of the manuscript

Mass spectrometer maintenance

Prof. Dr. Michael Lämmerhofer:

Generation, initiation, coordination and financing of the project

Discussion of results and interpretation

Partial writing and editing of the manuscript

Proofreading and final approval of the manuscript

Corresponding Author

V. Poster Presentations

Balaton Symposium on High-Performance Separation Methods 2015, Siófok, Hungary, September 4th – 6th.

Development of an LC-QToF-MS/MS Method for the Determination of Paxilline and its Metabolites in Mouse Plasma.

Bernhard Drotleff, Michael Lämmerhofer

International Symposium on Chromatography (ISC) 2016, Cork, Ireland, August 28th – September 1st.

Quantification of Steroid Hormones in Plasma using a Surrogate Analyte Approach and HPLC-HR-MS/MS with SWATH-Acquisition.

Bernhard Drotleff, Manfred Hallschmid, Michael Lämmerhofer

HPLC 2017, Prague, Czech Republic, June 18th – 22nd.

Quantification of Steroid Hormones in Plasma using a Surrogate Analyte Approach and UHPLC-HR-MS/MS with SWATH-Acquisition.

Bernhard Drotleff, Manfred Hallschmid, Michael Lämmerhofer

Awarded with an Agilent Technologies Best Poster Award

International Symposium on Separation Sciences (ISSS) 2017, Vienna, Austria, September 19th – 22nd.

Combination of Untargeted Lipidomic Profiling and Class-Specific Quantification.

Bernhard Drotleff, Michael Lämmerhofer

DPhG (German Pharmaceutical Society) Annual Meeting 2018, Hamburg, Germany, October 2nd – 5th.

Enhancing Inter-Batch Comparability in Untargeted Lipidomics by a Semi-Quantitative Study Design.

Bernhard Drotleff, Robert Lukowski, Michael Lämmerhofer

Awarded with a Lesmüller Best Poster Award

Berliner LC-MS/MS Symposium, Berlin, Germany, April 1st – 2nd.

Towards Absolute Quantification in Untargeted Lipidomics using Class-Specific Surrogate Calibrants.

Bernhard Drotleff, Michael Lämmerhofer

Awarded with a Poster Award

VI. Oral Presentations

PhD Meeting on Separation Science (Doktorandenseminar des AK Separation Science (GDCh, German Chemical Society)) 2018, Hohenroda, Germany, January 7th – 9th.

Simultaneous Targeted Quantification and Untargeted Profiling for Plasma Steroidomics using Data-Independent Acquisition.

Bernhard Drotleff, Michael Lämmerhofer

VII. Abbreviations

2D-LC	Two-dimensional liquid chromatography
AAPS	American Association of Pharmaceutical Scientists
ACN	Acetonitrile
ADC	Analog-to-digital converter
AIF	All-ion-fragmentation
APCI	Atmospheric pressure chemical ionization
CCS	Collision cross section
CE	Collision energy
CEM	Channel electron multiplier
Cer	Ceramide
CHCl ₃	Chloroform
cps	Counts per second
DDA	Data-dependent acquisition
DG	Diacylglycerol
DP	Declustering potential
EIC	Extracted ion chromatogram
ESI	Electrospray ionization
FDA	United States Food and Drug Administration
FFA	Free fatty acid
FT-ICR	Fourier transform ion cyclotron resonance
fwhm	Full width at half maximum
HILIC	Hydrophilic interaction liquid chromatography
HR	High resolution
Hz	Hertz
IDA	Information-dependent acquisition
IMS	Ion mobility spectrometry
IPA	Isopropanol
IRD	Ion release delays
IRW	Ion release width
IS	Internal standard
ITC	Ion transmission control
LC	Liquid chromatography
LLOQ	Lower limit of quantification
LPC	Lysophosphatidylcholine
LSMD	LIPID MAPS Structure Database

MCP	Microchannel plate
MeOH	Methanol
MRM	Multiple reaction monitoring
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
MS-DIAL	Mass Spectrometry Data-Independent Analysis
MTBE	Methyl-tert-butyl ether
NMR	Nuclear magnetic resonance
PA	Phosphatidic acid
PC	Phosphatidylcholine
PCA	Principal component analysis
PE	Phosphatidylethanolamine
PG	Phosphatidylglycerole
PIP	Phosphatidylinositol-phosphates
ppm	Parts per million
PS	Phosphatidylserine
q2	Collision cell (RF-only quadrupole)
QC	Quality control
QqQ	Triple quadrupole
QTOF	Quadrupole time-of-flight
ROC	Receiver operating characteristics
RP	Reversed-phase
S/N	Signal-to-noise
S1P	Sphingosine-1-phosphate
SD	Standard deviation
SFC	Supercritical fluid chromatography
SM	Sphingomyelin
SPE	Solid phase extraction
SRM	Selected reaction monitoring
TDC	Time-to-digital converter
TG	Triacylglycerol
TIC	Total ion current
TOF	Time-of-flight
t _R	Retention time
UHPLC	Ultra-high performance liquid chromatography
ULOQ	Upper limit of quantification
XIC	Extracted ion chromatogram

1. Introduction

1.1. Lipidomics

Following groundbreaking analytical developments in DNA-sequencing,^{1,2} genomics emerged as the first “omics” field and initiated the rise of various other disciplines like proteomics, transcriptomics and metabolomics. As one of the most recent research areas amongst omics, lipidomics was introduced in 2003 as a subset of metabolomics³ and has ever since attracted a continuously increasing scientific interest (Figure 1). This trend was also fueled by analytical advancements, in particular in instrumentation and acquisition techniques of mass spectrometry (MS) (see Chapter 1.2), which enabled the iterative construction of characteristics, pathways and networks of the lipidome.

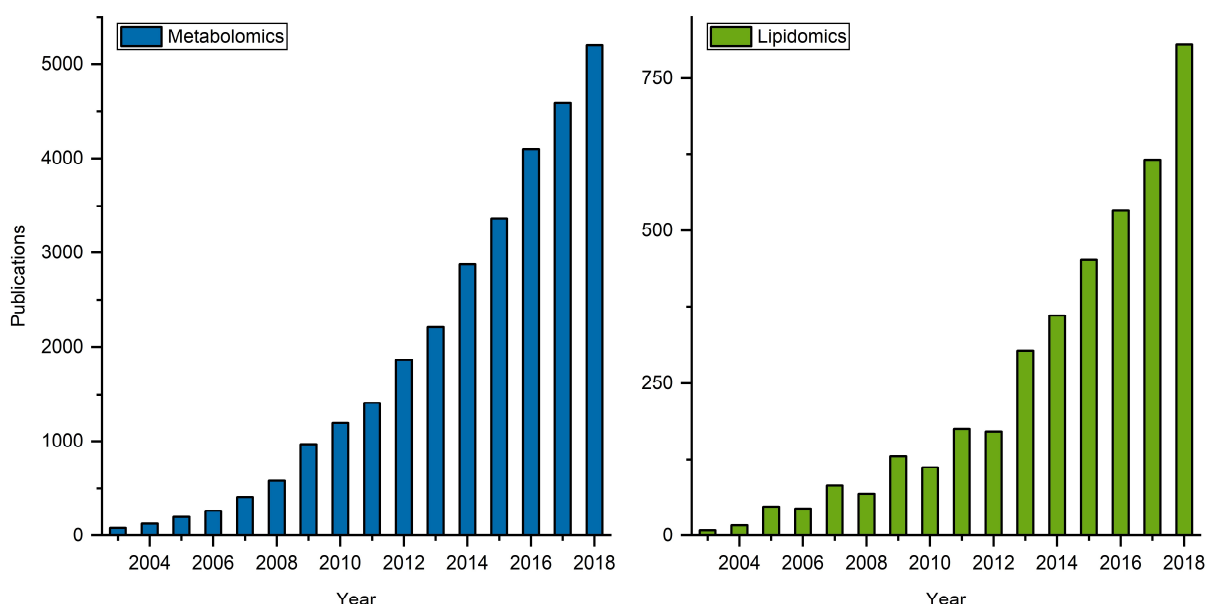


Figure 1. Published research in metabolomics and lipidomics. Data extracted from the PubMed database,⁴ showing the number of articles that include metabolomics and lipidomics as keywords.

In general, lipids are summarized as biological compounds that are soluble in nonpolar solvents.⁵ Given this simplified description, it is likely to underestimate the complexity and structural diversity of biological lipids, which is captured in the LIPID MAPS Structure Database (LMSD),⁶ the most comprehensive, public lipid database with 43,413 distinct entries (access date: 07/01/2019). For clarification and harmonization, a designated expert consortium^{7,8} established a classification system that divides biological lipids into eight categories: fatty acyls, glycerolipids, glycerophospholipids, polyketides, prenol lipids, saccharolipids, sphingolipids and sterol lipids (Figure 2). Moreover, a nomenclature system to precisely define

structures of acyl/alkyl chains including carbon count, degree of saturation, stereochemistry and double-bond geometry was introduced,⁶ which was also adopted for this work.

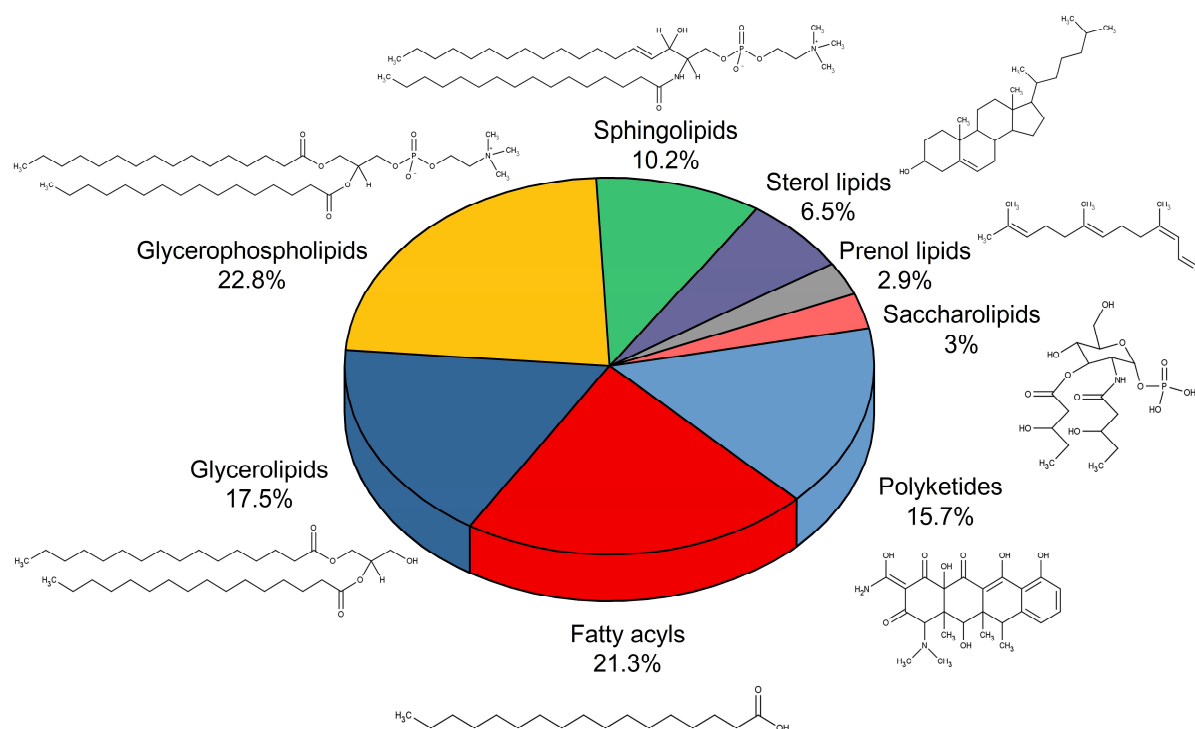


Figure 2. Distribution and exemplary structures of biological lipid categories. Data according to LMSD⁶. Lipid categories can be further separated into classes and subclasses⁶⁻⁸.

The vast complexity of the lipidome is also reflected in the diverse biological functions that are exerted by lipid compounds. Primarily, lipids create membranes by formation of bilayers,⁹ which represent the vital process of compartmentalization of cells and organelles. Besides its properties as molecular barriers, membrane lipids also influence signaling of embedded proteins via regulatory membrane-protein interactions.¹⁰ Many others, e.g. eicosanoids, phosphoinositides and sterol lipid hormones, act directly as messengers and control cellular functions via intra- and extracellular signaling.¹¹ Furthermore, key roles as energy depots, co-factors, pigments and vesicular transport units have been assigned to lipids.^{12,13}

The deep involvement in numerous physiological processes emphasizes the great potential to discover novel biomarker candidates amongst lipid compounds. Many clinical manifestations like Alzheimer's disease,¹⁴ atherosclerosis,¹⁵ breast cancer,¹⁶ cardiovascular disease,¹⁷ diabetes,¹⁸ liver cancer,¹⁹ obesity,²⁰ and prostate cancer²¹ were already shown to be associated to significant changes of the lipidome. These findings have put lipidomics on the verge to enter clinical application²² and will contribute to realize personalized medicine.²³

Overall, the strong relevance of lipid metabolism has been recognized by the scientific community, which is aiming for comprehensive exploration of the lipidome. Ultimately, by

interconnecting the findings of omics disciplines, a holistic understanding of physiological processes in health and disease is pursued. Continuous analytical innovation is mandatory and will play a central role as a pacemaker to achieve set goals.

1.1.1. Sample Preparation

The biggest contribution to measurement uncertainty in an analytical workflow is introduced in the pre-analytical phase²⁴ from where random or systematic errors will propagate and yield inaccurate results. Good analytical practices therefore need to be respected already during sample collection and sample storage. Sample type-specific considerations to obtain homogenous and representative portions of body fluids or tissues are obligatory, including pre-acquisition normalization procedures, e.g. adjustments according to DNA/protein concentration, cell count or dry weight.²⁵ Moreover, any contamination that affects analyte integrity as well as hemolysis have to be strictly avoided, as the obtained results will lead to misinterpretation and erroneous inter-study comparison.

As some lipids are sensitive compounds that can undergo post-sampling modifications by oxidation, peroxidation or hydrolysis,²⁶ instant quenching of enzymatic activities by immediate sample processing or low temperature (-80 °C) storage, optimally under inert gas (such as argon) and in appropriate, light-absorbing containers, is advised.²⁷ To further avoid degradation during storage or thawing, also degassing or addition of antioxidants is administered, in particular during sample preparation in oxylipin analysis or oxylipidomics in general.²⁸

Lipid extraction from biological samples involves a variety of organic solvents like chloroform (CHCl₃),^{29,30} methanol (MeOH),²⁹⁻³¹ butanol,³² methyl-tert-butyl ether (MTBE),^{33,34} hexane,³¹ and isopropanol (IPA),³⁵⁻³⁷ which provide an enhanced solubility for lipophilic compounds and dispatch proteins via precipitation. Depending on the miscibility of the used solvent with water, extraction approaches can be generally divided into monophasic and biphasic protocols. Traditional methods are based on biphasic extraction using CHCl₃/MeOH/H₂O mixtures according to the protocols of Bligh and Dyer²⁹ or Folch.³⁰ Although being most widely used over decades, certain drawbacks exist: (i) the desired organic phase, predominantly containing CHCl₃ ($\rho = 1.49 \text{ g/cm}^3$, 25 °C), is accumulated at the bottom, topped by the protein-rich interphase and the aqueous phase. Collection of the organic phase is aggravated as the upper phases have to be fully removed or penetrated, leading to enhanced risk for contamination; (ii) CHCl₃ has to be treated with stabilizing agents to prevent accumulation of phosgene and hydrochloric acid and consequential modification of susceptible lipids;³⁸ (iii) CHCl₃ toxicity as a carcinogenic substance increases environmental and personnel health risks;³⁹ (iv) CHCl₃ is incompatible with most plastic laboratory materials used for extraction tubes and pipette tips

(e.g. polypropylene)⁴⁰ and usage of glassware is strongly recommended; (v) high volumes are needed e.g. Folch's extraction requires a CHCl₃/MeOH/H₂O-ratio of 8:4:3 (v:v). Biphasic protocols are therefore trending towards alternative extraction solvents like MTBE. Its major benefit is the low density ($\rho = 0.74 \text{ g/cm}^3$, 20 °C), which enables formation of the organic phase as the upper layer, which is highly simplifying the collection process. Moreover, MTBE is non-halogenated and shows a favorable toxicity profile.³⁴ Nevertheless, (iv) and (v) also apply to MTBE, which is diminishing its usage in automated, high-throughput workflows like 96-well plates.⁴¹

Monophasic extraction is carried out via addition of water-miscible organic solvents (i.e. IPA, MeOH, acetonitrile (ACN)). Here, precipitated proteins are easily removed by subsequent centrifugation of the mixture (sometimes after storage at reduced temperatures to complete protein precipitation) and the resulting supernatant can be directly analyzed, if no supplementary processing is conducted. Added volumes are much lower than for biphasic extraction as a minimum ratio of 2.5:1 (organic solvent:sample, v:v) has been shown to achieve sufficient removal of proteins for plasma samples.⁴² These minimum steps in preparation not only provide a time effective sample handling but also have less potential to introduce variance, which is ultimately resulting in improved precision.³⁶ IPA-based protocols have been shown to yield comparable or superior results regarding lipid coverage,^{35,36} since also polar lipid classes like lysophosphatidylcholines (LPCs) or free fatty acids (FFAs) are efficiently extracted (see Figure 3). On the other hand, also polar, non-lipid compounds are obtained by IPA-extraction, which can cause interference or matrix effects e.g. in direct-infusion mass spectrometry (i.e. shotgun MS). When using reversed-phase (RP) liquid chromatography (LC) in conjunction with MS, this effect can be neglected, as polar impurities are separated from major lipid classes

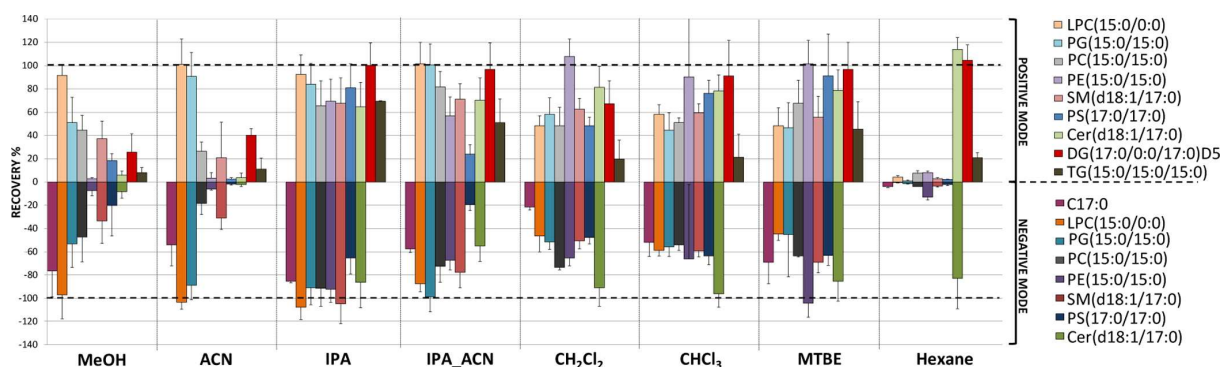


Figure 3. Comparison of the lipid extraction profiles of typically used organic solvent systems.

Nine individual lipid classes were assessed via a representative lipid species (PG, phosphatidylglycerole; PC, phosphatidylcholine; PE, phosphatidylethanolamine; SM, sphingomyelin; PS, phosphatidylserine; Cer, Ceramide; DG, diacylglycerol; TG, triacylglycerol; C17:0, FFA 17:0). Extractions were compared in both polarities of the ESI mode. Error bars indicate standard deviation (SD) of the recovery. Reprinted with permission from Sarafian et. al³⁶ (Copyright 2014, American Chemical Society).

during early gradient progression.⁴³ In conclusion, additionally regarding its cost effectiveness and acceptable toxicity profile, IPA is an excellent choice for global lipid extraction and high-throughput application.

Lipid extraction yields can be shifted towards specific classes by pH-adjustments, i.e. it has been reported that acidification increases recovery for phosphatidic acids (PAs), sphingosine-1-phosphates (S1Ps) and phosphatidylinositol-phosphates (PIP).^{44,45} However, acidic conditions can also lead to hydrolysis. Accordingly, the duration of the exposure should be kept at a minimum when global lipid profiling is the goal.⁴⁶

If a more specialized extraction of lipid targets is demanded, like in the case for low abundant steroid or eicosanoid species, solid phase extraction (SPE) is the method of choice. Various class-specific extraction protocols have been already developed to regulate recovery, purification, and enrichment of analytes.^{26,28,47}

At last, analysis results are also highly dependent on the composition of the injection solvent, which ideally shows maximum solubility for all lipids but has no detrimental effects on chromatography i.e. peak shape of the analytes including early eluted ones. Micro- or nano-LC systems in RP mode are particularly prone to peak distortion effects of injection solvents with an exceeding elution power, since the injection volume accounts for a relatively high proportion of the system volume. Improperly chosen injection solvents (e.g. when drying and redissolving is conducted) can in consequence result in premature elution, peak broadening and loss in sensitivity.⁴⁸

1.1.2. Analytical Strategies

To qualitatively and quantitatively encompass the full lipidome with its immense physicochemical diversity is a challenging approach that requires a broad spectrum of analytical methodologies. Furthermore, the thousands of distinct lipids in various biological matrices show concentrations ranging from picomolar to micromolar levels and are under constant dynamic change.⁴⁹ To date, prominent contributions in lipidomics derive from spectroscopy, chromatography and mass spectrometry.⁵⁰

Amongst spectroscopic methods, which are capable to observe interactions between electromagnetic radiation and matter, nuclear magnetic resonance (NMR) spectroscopy is the most noteworthy in the context of lipidomics, as it offers some unprecedented analytical properties: (i) non-destructive detection; (ii) high instrumental robustness and analytical reproducibility; (iii) elucidation of structural information and molecular dynamics; (iv) direct quantitative data.⁵¹ Still, its relatively low sensitivity and the convoluted spectra that are obtained for complex mixtures are limiting its potential impact on the field.²⁶

Currently, main advancements and discoveries in lipidomics are driven by MS analysis, either utilizing shotgun methods, hyphenation to LC systems or imaging techniques.²⁶ Highlights of shotgun analysis are short run times and the robust quantitative performance due to consistent matrix effects and co-detection of internal standards (ISs).⁵² However, the large quantity of analytes that are concurring for ionization can cause significant ion suppression⁵³ and low abundant lipids are likely to remain uncaptured. Another issue is in-source fragmentation, which can lead to an apparent lipid conversion and ultimately to feature misannotation when no orthogonal information like retention time (t_R) is present.⁵⁴ In addition, signal interferences, originating from the presence of numerous isobars and isomers,⁵⁵ are to be expected and require sophisticated deisotoping/deconvolution algorithms. Many of these difficulties are approached by additional sample preparation efforts,⁵² intra-source separation and selective ionization,⁵⁶ high-resolution MS instruments⁵⁷ or accessory ion mobility spectrometry (IMS).⁵⁸ In conjunction with chromatographic separation, matrix effects and potential signal interferences are drastically reduced, which is making less concentrated compounds accessible and improves identification.²⁶ The most widely used chromatographic approach for lipidomics is based on RPLC. It provides the separation of complex lipid mixtures and resolves lipid species according to their chain length and degree of saturation. Retention time increases with number of carbon atoms, while it decreases with number of double bonds.⁵⁹⁻⁶¹ In contrast, lipid class separation can be achieved by hydrophilic interaction liquid chromatography (HILIC), since the separation mechanism mainly relies on the polar head groups.⁶² HILIC, currently replacing normal phase LC, is preferably used for quantitative purposes due to the co-elution of analytes and class-specific ISs. However, complications, owing to peak interferences (as above mentioned for shotgun lipidomics), limit its application. For efficient lipid class separation, also supercritical fluid chromatography (SFC) is eligible, which has been drawing attention due to rapid run times.⁶³ SFC also allows retention of apolar, neutral lipids, like normal-phase LC, which is difficult to achieve with HILIC for what reason neutral lipids elute unresolved with or close after t_0 in HILIC lipid class separation methods. Another promising separation technique for lipidomics is two-dimensional LC (2D-LC), as it can deliver increased peak capacity by highly orthogonal retention and elution, e.g. via combination of HILIC in the first dimension and RPLC in the second dimension.⁶⁴ An overview about frequently applied workflows is provided in Figure 4.

The above-mentioned methods allow the analysis of representative biofluids and cell bundles or tissue regions. However, it is of great scientific interest to be able to characterize specific cell types and its metabolomic pathways in order to decipher mechanisms of disease progression. Suitable methods to pursue this goal are predominantly MS-based as well, and include single-cell analysis or MS imaging.⁶⁹⁻⁷³

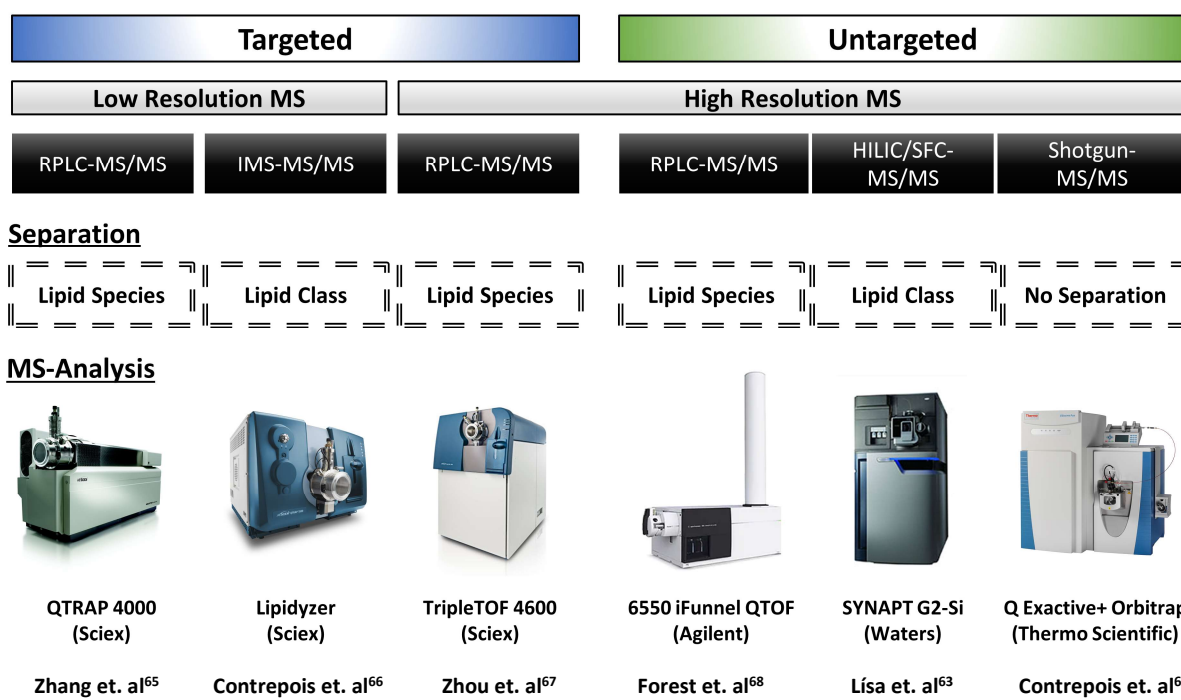


Figure 4. Summary of common workflows. The Lipidyzer platform consists of a QTRAP 5500 and is equipped with a SelexION unit for differential ion mobility.

Powerful analytical tools and workflows have been developed in the recent years, though, each platform has its specific advantages and disadvantages so that a gold-standard method for lipidomic analysis has not yet been proclaimed. Accordingly, further research is simultaneously progressing in multiple directions, distributed across the available techniques. Together with the steady development of relatively novel technologies, like 2D-LC or IMS, and hyphenation of subsidiary instrumentation, the lipidomics community is advancing to solve some of the currently existing challenges.

1.2. Mass Spectrometry

Throughout the last century, MS has evolved to become an indispensable and pervasive technology that brought innovation to many scientific fields. Since its first major impact in 1920, when the existence of stable isotopes could be verified,⁷⁴ striking instrumental advancements have been made, and many being recognized with a Nobel Prize.⁷⁵

The principle of MS relies on the determination of the mass-to-charge (m/z) ratio of ionized atoms or molecules by using magnetic and/or electric fields. As effective ionization is still one of the biggest challenges of MS, historical improvements of this process have iteratively led to broader application. For example, the invention of electrospray ionization (ESI) greatly simplified interfacing of LC and MS and directed subsequent developments.⁷⁶

Nowadays, the analytical chemist can access various ionization techniques (see Chapter 1.2.1) and many technologically advanced instruments to address specific demands. Especially LC-MS, owing to its unmatched sensitivity, selectivity, analyte coverage, and reliable performance in qualification and quantification, has become the gold-standard method in the analysis of biomolecules in complex matrices. Fundamental aspects of MS will be discussed in the next chapters.

1.2.1. Ionization

In order to separate and detect atoms or molecules with MS, gas-phase ions must be present. The ionizing conversion is achieved either by electron ejection, electron capture, protonation, deprotonation or adduct formation with charged ions.⁷⁷ The majority of currently used ionization methods operate under atmospheric pressure to allow LC coupling or solvent injection, e.g. ESI or atmospheric pressure chemical ionization (APCI). Another widely applied method is matrix-assisted laser desorption/ionization (MALDI), which is particularly favorable for MS imaging.⁷²

ESI, also utilized for the experimental work in this dissertation, is regarded as the standard ion source and is installed in most commercialized MS instruments. Here, a liquid flow (1 – 1000 $\mu\text{L}/\text{min}$) is directed through a narrow needle or capillary to which a high voltage potential of up to ≈ 5 kV is applied.⁷⁸ Every solvent shows a specific minimum onset voltage for spray formation, which depends on its surface tension. With increasing voltage, initially spherical drops elongate until the so-called Rayleigh limit is reached (see Figure 5). Here, the liquid flow is transformed into a spray by Coulombic explosion due to the accumulated electric repulsion on the surface.⁷⁷ A “Taylor cone”, that is constantly emitting charged droplets, can now be observed on the tip. Supported by a heated dry gas (commonly N_2 or synthetic air), the solvent in the droplets evaporates, again resulting in an elevated charge-to-surface ratio, increasing repulsion and ultimately Coulombic explosion into even finer droplets. This process is repeated until ions are desorbed from the successional droplets or, in the case of large molecules, until the solvent is completely evaporated.^{77,79} An applied nebulizer gas guides the spray to the MS orifice, where ions are electrostatically drawn into an ion path and towards the mass analyzer. The initially axial installation of the spray and the orifice is nowadays mainly replaced by an orthogonal setup. This way excessive contamination is avoided and the diameter of the orifice can be increased, leading to enhanced instrument robustness and sensitivity.⁷⁷



Figure 5. Formation of the Taylor cone. Left: No or low voltage with a spherical drop; Middle: Elongated drop, close to the onset voltage; Right: Formation of the Taylor cone and release of charged droplets after the Rayleigh limit has been surpassed.

ESI usually yields $[M+H]^+$, $[M-H]^-$ or adduct ions depending on the polarity, pH and modifier content in the solution. For large molecules (e.g. proteins) also multiply charged ions are generated, which can shift the measured m/z ratio to observable values when mass analyzers with limited mass range are utilized. Another characteristic of ESI is its sensitivity to concentration rather than to total mass.⁸⁰ Accordingly, in case of preceded chromatographic separation, an increased sensitivity can be achieved at lower flow rates due to the decreased sample dilution factor. However, lower flow rates also have to be combined with smaller diameters of chromatographic columns to avoid cancellation of the reached gain in sensitivity by peak broadening due to increased longitudinal diffusion.⁸¹ It should also be noted that typical electrochemical processes are occurring at the tip of the ESI electrode.⁸² When operated, the ESI aperture functions as an electrolytic cell in which reduction and oxidation reactions produce an electric current, which is limiting the total amount of extractable ions.⁸³

Ionization efficiency is further dependent on the presence of co-eluting compounds that concur for ionization or affect droplet formation.⁸⁴ This is known as the matrix effect, which can result in ion enhancement or, more commonly, in ion suppression.

APCI is reported to be less prone to matrix effects due to its differing ionization mechanism.⁸⁵ Here, a gas is used to nebulize the liquid flow into droplets that pass through a heated ceramic tube. The evaporated gas-phase is then directed towards a high voltage corona discharge needle where an ionization plasma is formed.⁷⁷ By a transfer cascade via ambient gas molecules and nebulized solvent molecules, a charge is transferred to the analyte. APCI can yield better ionization efficiency for specific analytes and is particularly suited for nonpolar lipids and compounds in the lower molecular range up to 1500 Da.

With MALDI, an ionization technique that is exceptionally suited for large molecules like DNA or proteins was introduced. It shows a good robustness to contamination and does not require exceeding sample preparation. In MALDI, the analyte is dissolved in a liquid matrix that contains molecules with a high absorption coefficient for the wavelength of a laser. After drying, the analyte is embedded in this matrix and exposed to strong laser pulses. During this process, the absorbing molecules are extremely heated and evaporated, and subsequent ionization takes place via proton transfer.⁸⁶ Due to its unique ionization procedure, it is the predominantly used technique for MS imaging.

1.2.2. Mass Analyzers

After ionization in the source, ions pass the MS orifice through which they are directed by an electromagnetic field. Subsequent ion trajectories and unwanted fragmentation can only be preserved in high vacuum conditions, thus, powerful turbopumps are essential to sustain a vacuum of 10^{-3} - 10^{-7} Torr. Also, increased requirements for the interface are needed when atmospheric pressure ionization methods are used.

To date, analytical chemists have to choose between a selection of diverse mass analyzers, most of them with complementary properties regarding (i) the observable mass range, (ii) the resolving power, (iii) the sensitivity (depending on duty cycle, transmission and detector efficiency), (iv) the acquisition speed and (v) the mass accuracy.^{77,87} Commercialized instruments usually employ quadrupole, ion trap, Fourier transform ion cyclotron resonance (FT-ICR), Orbitrap or time-of-flight (TOF) mass analyzers. Various combinations of two or more mass analyzers are also available, e.g. triple quadrupoles (QqQs) or in case of different mass analyzers as hybrid instruments, for tandem mass spectrometry (MS/MS), which provides a repertory of scan modes to improve qualification and quantification. An overview about the characteristics of currently available high resolution (HR) hybrid MS instruments is shown in Table 1.

The instrument that was utilized for the experimental work in this dissertation was a hybrid quadrupole time-of-flight (QTOF) MS by Sciex (TripleTOF 5600+), operated with a DuoSpray source (consisting of an ESI and APCI probe). Its domain is the recording of HR TOF-MS spectra and the additional acquisition of MS/MS fragments to gain structural information.

Table 1. Performance parameters of exemplary HR hybrid MS instruments.

Instrument	Resolving Power	Acquisition Rate [Hz]
Agilent 6550 iFunnel QTOF	40 000	50
Bruker impact II (QTOF)	50 000	50
JEOL SpiralTOF-TOF	60 000	10
SCIEX TripleTOF 6600 (QTOF)	35 000	100
SCIEX TOF/TOF 5800	26 000	10
SCIEX X500R (QTOF)	35 000	100
ThermoFisher Q Exactive HF	240 000	18
ThermoFisher Orbitrap Fusion Tribrid	450 000	15
Waters Xevo G2-XS QTOF	40 000	30

Resolving power was calculated for specified m/z values at full width at half maximum (fwhm). Adapted with permission from Kind et. al⁸⁸ (Copyright 2017, Wiley Periodicals, Inc).

The first element of a QTOF is a quadrupole, which generally consists of four perfectly parallel, hyperbolic rods in square formation to which a radio frequency current (RF) and a

superimposed direct current (DC) are applied (see Figure 6).⁸⁹ Arriving ions traverse through the quadrupole, since they are attracted by a low voltage potential at its opposite end. During their passage, they interact with the electromagnetic fields, which are induced by the currents, and follow helical trajectories. Depending on the amplitude and frequency of the RF, the offset potential of the DC, the initial position and the traversing velocity, only ions with a specific m/z ratio will obey trajectories that ensure axial transmission through the quadrupole.⁹⁰ Other ions experience increasing oscillation and are ultimately neutralized when striking the rods.⁹²

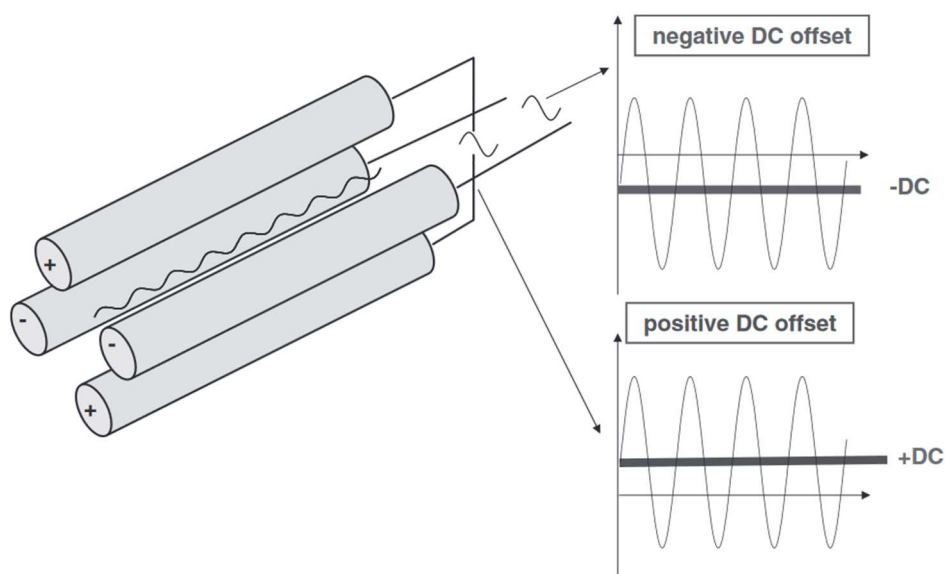


Figure 6. Voltage profile of the rods in a quadrupole analyzer. Depending on the amplitude and frequency of the RF and the DC offset, dynamically changing electromagnetic fields are created, which force ions into m/z ratio specific trajectories. Reprinted with permission from Somogyi et. al.⁹¹ (Copyright 2008, Elsevier B.V.).

Accordingly, an efficient m/z filter is obtained by selected modulation of RF and DC. In Figure 7, an exemplary stability diagram for different m/z ratios is shown. The stability areas below given m/z ratios represent possible combinations of RF and DC values that will result in stable trajectories through the quadrupole. Many m/z values have overlapping stability areas and RF and DC have to be set carefully to achieve sufficient resolving power to discriminate m/z values. Thus, during a scan, RF and DC are linearly increased, resulting in a scan line (see Figure 7). The intersection of the scan line and the stability areas is directly correlated to the spectral peak width and therefore also to resolving power. To attain uniform resolving power for a given mass range, quadrupoles have to be tuned and are usually set to yield spectral peak widths of 0.7 – 1.0 units at fwhm on the m/z ratio scale (see Figure 7). As often observed in MS, higher resolution is compromised with a loss in sensitivity and vice versa.^{93,94}

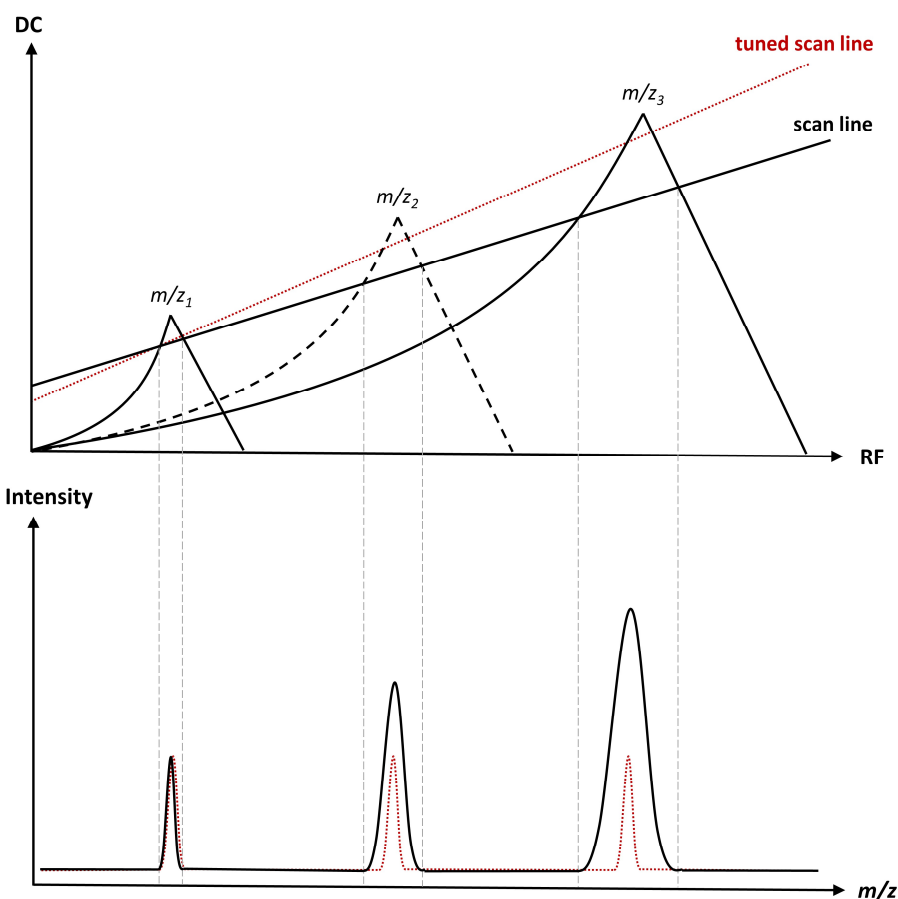


Figure 7. Stability diagram of a quadrupole mass analyzer. $m/z_1 < m/z_2 < m/z_3$. Heavier ions with higher m/z ratios show bigger stability regions since their reaction to electromagnetic field changes is slower than that of lighter ions due to the higher inertia. For RF scanning, most quadrupoles use amplitude scans at fixed frequencies rather than frequency scans at fixed amplitudes.⁹⁵ The initial DC offset is represented by the y-intercept of the scan line. Units are omitted in the diagrams for reasons of simplicity.

Quadrupoles are also regularly employed as ion guides or collision cells when operated in “RF-only” mode. With the absence of DC in the rods, wider m/z transmission windows can be controlled via RF settings to yield an “open” quadrupole or a focused ion beam for ion optics. Traversing ions can be fragmented in MS/MS by introduction of an inert gas into an RF-only quadrupole and subsequent collision-induced dissociation (CID). The degree of fragmentation is dependent on the kinetic energy of the ions, which can be modulated via an additional potential termed collision energy (CE).

QTOF instruments consist of two quadrupoles (disregarding additional quadrupoles used as ion guides), whereas the first one (Q1) is utilized as a mass filter and the second one (q2) functions as a collision cell (see Figure 8). However, both can also be operated in open state to forward unfiltered precursor ions to the successional TOF analyzer. The underlying principle of a TOF is based on the determination of ion flight times (t) in a field-free region after acceleration by a potential (V).⁹⁶

According to the basic equation:

$$E_{kin} = \frac{1}{2}mv^2 = qV = zeV \quad \text{Equation 1}$$

ions with a mass m and a total charge q (defined as the product of the charge number z and the elementary charge e) show distinct velocities (v) after uniform acceleration in a field (V) to reach a defined kinetic energy (E_{kin}).⁷⁷ Velocities are calculated by rearrangement of the equation to:

$$v = \sqrt{\frac{2zeV}{m}} \quad \text{Equation 2}$$

The flight time of the ions is then determined via the length of the flight tube (L) and the velocity (v):

$$t = \frac{L}{v} \quad \text{Equation 3}$$

Combination of Equation 2 and Equation 3 yield:

$$\frac{m}{z} = 2eV\left(\frac{t}{L}\right)^2 \quad \text{Equation 4}$$

Equation 4 shows that the m/z ratio of an ion can be directly calculated from the measured drift time t when the length of the flight tube L and the amplitude of the potential V are known. Accordingly, heavier ions with an increasing m/z value will show prolonged flight times compared to lighter ions with lower m/z values.⁹¹ Moreover, uncontrolled changes in L or V can affect the measurement accuracy. The main factors are temperature fluctuations that affect the length of the flight tube or the power supply output. For this reason, TOF instruments require temperature controlled laboratories and denser mass calibration intervals compared to quadrupoles.⁹⁷

TOF resolution can be enhanced with increasing flight duration, yet maximizing the flight tube or decreasing the acceleration voltage were shown to be impractical.⁷⁷ A major factor for insufficient resolution are the unequal starting conditions of the ions caused by kinetic and spatial dispersion.^{96,98} Both deteriorating factors were substantially minimized by the introduction of the reflectron⁹⁹ and delayed pulsed extraction^{100,101} elements.

A reflectron acts as an ion mirror (see Figure 8) and corrects kinetic spread of ions with the same m/z ratio. It relies on the elongated flight path of ions with higher initial velocity, as they permeate deeper into the ion mirror until they are reflected. The detector should be positioned in a suitable location to capture ions of identical m/z values in the moment when the ions of higher velocity reach the ions of slower velocity.

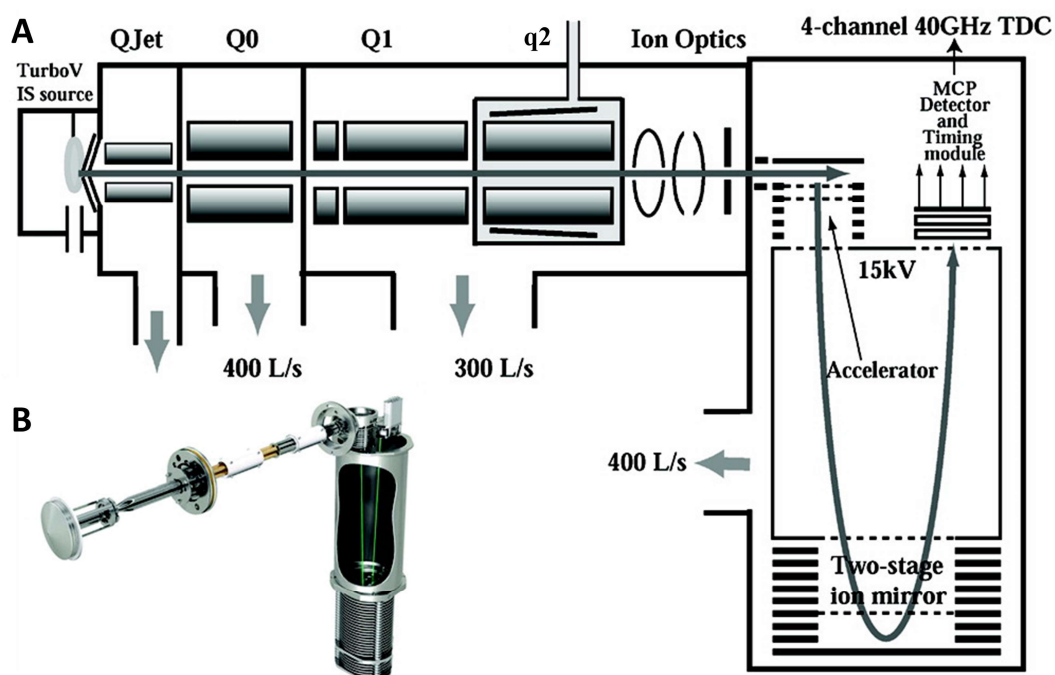


Figure 8. Design of a QTOF instrument (TripleTOF 5600). A: Scheme of the instrument. The ion course is indicated with the arrows from left to right. B: 3D orientation of the instrument ion path. Adapted with permission from Andrews et. al¹⁰² (Copyright 2011, American Chemical Society).

In a continuous ion beam, ions with identical m/z values can show differing spatial distribution due to disparate velocities or nonuniform angles of the beam. To regain the lost resolution, arriving ions are gathered in a field-less region before they are accelerated by a short pulse. This region is therefore denominated as the accelerator or ion pulser. Ions closer to the origin of the pulse (and initially further from the detector) are more intensely exposed to the pulse and experience higher acceleration. On the way to the detector they ultimately join the previously separated ions, which received less accelerating energy, and are recorded with identical flight times. In a QTOF this delayed pulsing is usually done in an orthogonal direction after the ions have passed the collision cell.¹⁰³ This way the continuous ion beam is transformed into an ion pulse, which is preferred for TOF applications. Moreover, the axis of detection is independent from the ion beam direction and thus acquisition speed, resolution and sensitivity is improved.¹⁰⁴

QTOF instruments are now widely used for various applications as they combine HR spectral acquisition with fast acquisition speed. This allows comprehensive analysis, even in conjunction with ultra-high performance liquid chromatography (UHPLC). Technological improvements have also provided enhanced sensitivity and wider linear dynamic range, enabling quantitative performance comparable to QqQs.^{47,105,106}

1.2.3. Detector System

The detector system represents the final bottleneck during MS data acquisition, as previous efforts in ionization and mass filtering are degraded without the proper conversion of ions into interpretable signals. Direct detection of few single ions with sufficient signal-to-noise (S/N) ratio is challenging due to the low electric current they induce.⁷⁷ Thus, with distinct exceptions,¹⁰⁷ incoming ions require amplification mechanisms, mainly in the form of electron multipliers.¹⁰⁸

In a TOF, after fully passing the flight tube, ions are accelerated towards the electron multiplier dynode using a high voltage of the opposite polarity. This way, ions strike the electron multiplier with enhanced kinetic energy and release several charged secondary particles (negative mode: positive ions; positive mode: negative ions and electrons).⁷⁷ By a subsequent impact with the dynode, the secondary particles are converted into electrons, which release additional electrons in the following cascade of impacts. Depending on the applied voltage and the geometry and surface of the dynode, signal multiplication factors of $>10^6$ can be reached.

The predominant electron multiplier type used in QTOF instruments are microchannel plates (MCPs). These are small plates that are perforated with many cylindrical holes, which all function as single amplifying dynodes. The valuable properties for QTOF application are the fast response time, due to a rather short electron path, and the increased detection area, ideal for large incoming ion beams.

The next element of a QTOF detector system is a converter that digitizes the ejected electron signals of the MCP. This can either be an analog-to-digital converter (ADC), that show a wide linear dynamic range and can also record signal amplitudes, or a time-to-digital converter (TDC). Most frequently, TDCs are employed in QTOF instruments, as they are typically more sensitive, owing to a beneficial S/N ratio.⁹⁷

During acquisition, the TDC is synced to the pulsing of the accelerator (Figure 9). After amplification, incoming ion signals are forwarded by a discriminator if a certain threshold is reached. The TDC then registers the arrival time that has passed since the last pulse.⁹⁷ This measurement result can be converted into an m/z value (see Equation 4). A major drawback of the TDC is its inability to register an additional ion event during its dead time (usually few nanoseconds).^{97,109} In consequence, if multiple ions with an identical m/z value arrive at the detector at the same time during a pulse cycle, they are only counted as one single ion.

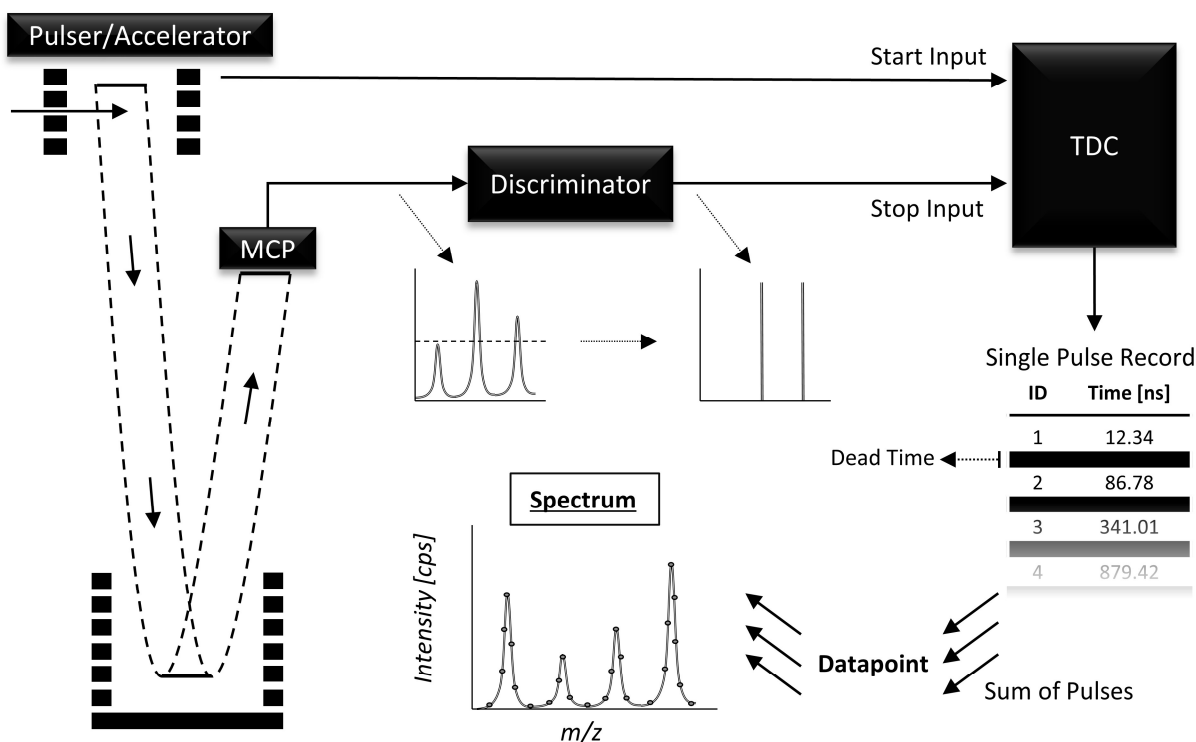


Figure 9. Scheme of the spectrum acquisition with a TDC. The pulser frequency is limited by the highest mass that is transmitted by the q2, since heavier ions have longer flight times and a new pulse can only be initiated after all ions have arrived at the detector. With the sum of pulses an arrival time histogram is created. As the TDC is operated at 40 GHz, the minimum time bin for the histogram (\cong the maximum achievable resolution) is 25 ps.

This saturation effect limits the linear dynamic range of the detector and requires counteraction, like mathematical correction via probability statistics. As the output signal is generated by summation of several thousand individual pulses (with mass range dependent accelerator frequencies of about 10 – 30 kHz), Poisson statistics can be used to predict the amount of ions that were missed during dead times.⁹⁷ This automated signal correction enhances the linear dynamic range by 10-fold. Another approach to overcome TDC limitations is multi-channeling. The TripleTOF5600+ contains four TDC channels,¹⁰² which theoretically increase ion counting capacity by 4-fold (if a uniform spread of the ion beam is assumed). Moreover, the system offers automated ion gating, i.e. ion transmission control (ITC).¹¹⁰ Here, a lens in front of the Q0 region (see Figure 8) dynamically modulates the total ion transmission depending on the intensity of the total ion current (TIC). Accordingly, the ion load is decreased and TDC saturation is less likely. Prior to the data output, a correction factor is applied to the signal intensities that were recorded with reduced ion transmission. Due to the vastly reduced ion load in MS/MS experiments, ITC is only activated in TOF-MS by default, but not in MS/MS scan modes.

1.2.4. Scan Modes

In general, MS data acquisition can be separated into targeted and untargeted approaches. By targeted acquisition, most commonly conducted with QqQ instruments, high quality data of preselected analytes is obtained. It is regarded as the gold-standard of absolute quantification and is perfectly suited to verify predefined hypotheses. In contrast, the goal of untargeted MS is to capture all detectable analytes, including unknowns, to stimulate the discovery of novel metabolites and potential biomarkers. Rather than absolute quantification, it focusses on relative quantification and qualification of found features. In order to obtain reasonable and interpretable results from untargeted analysis of complex sample matrices, HR-MS instruments, like QTOFs or Orbitraps, are mandatory. The most prominently utilized scan modes of the TripleTOF 5600+ are briefly described in the following sections.

The central scan mode of a QTOF is the obligatory TOF-MS survey scan, where a full HR precursor ion spectrum is recorded. Both quadrupoles are operated in the open state with low CE to avoid unwanted fragmentation. The subsequent TOF-analysis provides a specified resolving power of $\geq 30,000_{\text{fwhm}}$ (@ m/z 829.5393 in ESI⁺ and @ m/z 933.6363 in ESI⁻). Via the obtained accurate masses and isotopic patterns, potential sum formulas for the detected ions can be estimated. However, this data alone does not yield high levels of confidence for compound identification.¹¹¹⁻¹¹⁴ Thus, additional data acquisition is necessary to enhance the extracted information content.

1.2.4.1. Product Ion Scan

The experiment cycle can be extended by addition of one or several product ions scans to merge untargeted and targeted acquisition.¹¹⁵ Here, narrow Q1 transmission ranges (typically with an m/z width of 0.7; see Figure 7) are used to isolate precursor molecules, that are subsequently fragmented in the q2 using elevated CE values. This scan mode is similar to the product ion scan in QqQs, but an HR spectrum of all fragments is obtained. Although the duty cycle is much lower than for QqQs, the enhanced selectivity can result in beneficial S/N ratios. The obtained fragmentation results provide the highest achievable MS selectivity for the instrument and can be utilized for structure elucidation and compound identification. Moreover, the continuous acquisition of MS and MS/MS data enables the generation of extracted ion chromatograms (EICs), which can be exploited for relative or absolute quantification on both MS levels.

When MS is used in conjunction to prior chromatographic separation, reasonable considerations regarding total cycle time, which is the sum of experiment accumulation times and system times (e.g. settling time etc.), need to be considered. As a minimum of 10 data

points per peak is typically required to accurately describe a peak with sufficient precision,¹¹⁶ the distributable accumulation time is limited by the chromatographic peak width. Given that increasing accumulation time yields an improved S/N ratio,⁴⁷ it should always be maximized for the least sensitive analyte. In order to multiplex the acquisition of target precursors, the analytical run can be split up into several periods with individual MS and MS/MS settings.⁶⁷ Period experiments are ideally designed to monitor target analytes only during their respective retention time intervals. This way, comparable to scheduled multiple reaction monitoring (MRM) in QqQs, accumulation times can be optimally distributed between all target analytes. In contrast to a QqQ, the investigation of fragment ratios is automatically enabled without any additional MS experiments. These ratios are suitable for (cross-)validation of selectivity, since they should remain stable in the absence of interferences.⁴⁷

The product ion scan (as well as the other MS/MS scans) can be operated in high sensitivity or in high resolution mode. During high sensitivity acquisition, the ion beam is focused after ion ejection from the q2 (see Figure 10A). This way the yield of accumulating ions in the accelerator, and consequently the duty cycle, is increased. However, the linear correlation between the initial ion velocity and the ion position in the pulser, on which the compensation by delayed pulsed extraction and reflection is based (see Chapter 1.2.2), is hereby distorted. The 2- to 5-fold gain in sensitivity thus comes along at expense of a loss in resolution (resolving power $\geq 15,000_{\text{fwhm}}$). In the high resolution mode, ion optics are turned off and ions are partially lost at the skimmer to the entrance of the accelerator (see Figure 10B). In consequence, sensitivity is decreased owing to the reduced duty cycle. Yet, an equivalent resolution as in TOF-MS (resolving power $\geq 30,000_{\text{fwhm}}$) is reached due to the feasible correction of velocity and spatial spreads.¹¹⁷

Beyond these two options for acquisition, signal intensity can be further increased by enabling the enhancement of a specific m/z region in the high sensitivity mode. This can be achieved via pre-tuned ion optics potentials that induce temporary ion trapping in the q2, subsequent rapid ion gating into the accelerator and an optimized accelerator timing of the TOF pulse.¹¹⁸ Depending on the m/z value of the targeted ion, the system calculates ideal ion release delays (IRDs, i.e. the time between ion gating and the TOF pulse) and ion release widths (IRWs, i.e. the duration of the gating pulse). The result is a ≥ 3 -fold increase in sensitivity for a limited m/z region of about 400 units around the set value. With the achieved duty cycle of up to 100 %, a sensitivity comparable to QqQ instruments is reached.

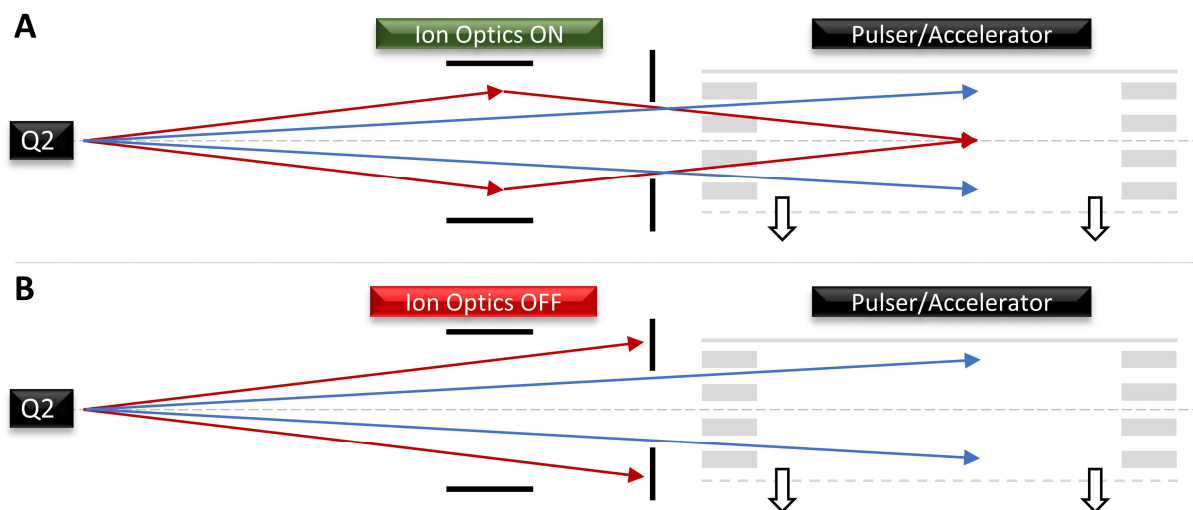


Figure 10. Principle of the high sensitivity and high resolution mode. In high sensitivity mode (A), correlation of initial ion velocities and ion positions before acceleration are partially lost. The red arrows indicate the path of two ions (of identical m/z value) that ultimately have the same position in the accelerator but their velocity shows different directions. In high resolution mode (B), ions are not deflected by ion optics. The positions of the ions entering the accelerator (blue arrows) are correlated to their velocity and the initial spread can be corrected.

1.2.4.2. Data-dependent Acquisition

Although the preceding TOF-MS survey scan generates untargeted MS data, product ion scanning is far from comprehensive on the MS/MS level. To increase the amount of MS/MS data in order to improve the confidence level for compound annotation, information-dependent acquisition (IDA, i.e. data-dependent acquisition (DDA)) was introduced.^{119,120} It can be briefly described as a dynamic product ion scan, which is triggered after preset criteria are fulfilled in the TOF-MS survey scan. In truly untargeted acquisition, IDA is adjusted to acquire product ion spectra of the most abundant ions found in TOF-MS. Furthermore, IDA-methods can be further specified by addition of inclusion and/or exclusion lists.

The number of recordable product ion scans per cycle is limited by the acquisition frequency of the instrument (100 Hz for the TripleTOF5600+). In (LC-MS) practice, usually the top 10 – 20 ions per cycle for QTOF¹²¹ (in case of slow acquisition instruments like orbitraps usually top 4 – 5) are selected for further fragmentation, as accumulation times need to be adjusted to achieve 10 data points per peak (in TOF-MS) with adequate S/N ratio (in IDA-MS/MS). IDA drastically increases the number of processable and selective spectra for compound annotation compared to regular product ion scanning. Yet, the increased selectivity on the MS/MS level cannot be exploited for quantitative purposes as most peaks are not multiply triggered to reach a sufficient number of data points per peak. Relative or absolute quantification, therefore, is restricted to TOF-MS results, which are more prone to interference. Since there is still a risk to miss relevant precursors, as it is the case when analytes are of low

abundance, co-eluting, or not efficiently triggered, IDA does also not provide full comprehensiveness on the MS/MS level.

1.2.4.3. Data-independent Acquisition

Global fragment ion data for all detected precursors can only be captured by data-independent acquisition (DIA) techniques. A simple approach to achieve full comprehensive data acquisition is MS^E (first reported on QTOF instrument of Waters) or also termed All-ion-fragmentation (AIF; QTOF of Agilent, Orbitraps of Thermo Fisher).^{122,123} Here, two (TOF)-MS scans, one with low CE for precursor detection, and one with high CE for registration of all fragments, are exerted. If the chromatographic separation is not efficient, the obtained spectra are highly complex and require deconvolution,¹²⁴ which is aggravated for perfectly co-eluting peaks. Another method, termed MS/MS^{ALL},¹²⁵ approached DIA by discrete stepping of unit mass resolution Q1 precursor windows in small intervals to achieve sequential fragmentation (see Figure 11). As this workflow basically represents full comprehensive product ion scanning (see Chapter 1.2.4.1), total cycle times are incompatible with chromatographic hyphenation. Thus, MS/MS^{ALL} is strictly limited to shotgun analysis.

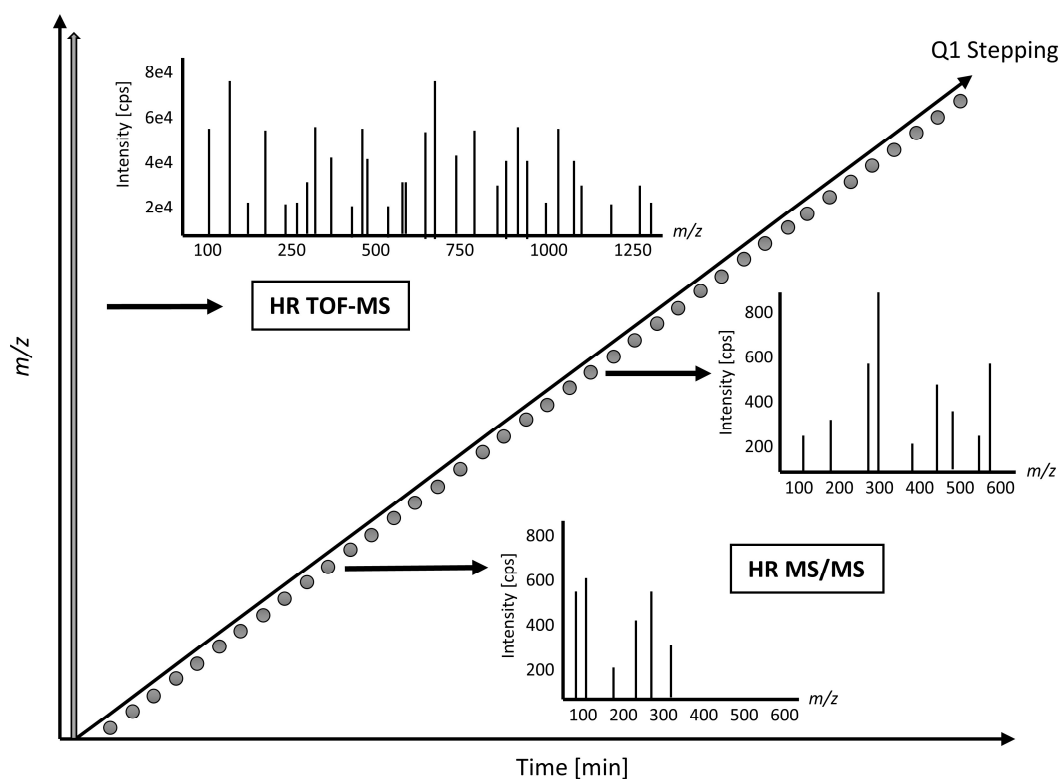


Figure 11. MS/MS^{ALL} acquisition. Due to the high number of sequential product ion scans, total run times for shotgun analysis with MS/MS^{ALL} can take up to six minutes. Recreated and adapted with permission from B. Simons et. al¹²⁵ (Copyright 2012 by the authors; open access).

Other groups experimented with sequential fragmentation of multiple precursors via intermediate Q1 isolation windows.¹²⁶ Each MS/MS experiment covered an m/z width of 10 units, but instruments were not yet fast enough to capture a broad precursor range when hyphenated to chromatography. Accordingly, at first instrumental limitations in resolution and acquisition speed had to be overcome until a novel DIA technique named SWATH (sequential window acquisition of all theoretical fragment-ion mass spectra) was introduced for commercialized instruments.^{122,127} Instead of unit mass precursor isolation, SWATH allows to set Q1 transmission windows of variable width (see Figure 12). Extracted precursors are then simultaneously fragmented in the q2, generating composite fragment spectra with significantly improved selectivity compared to MS^E (AIF).

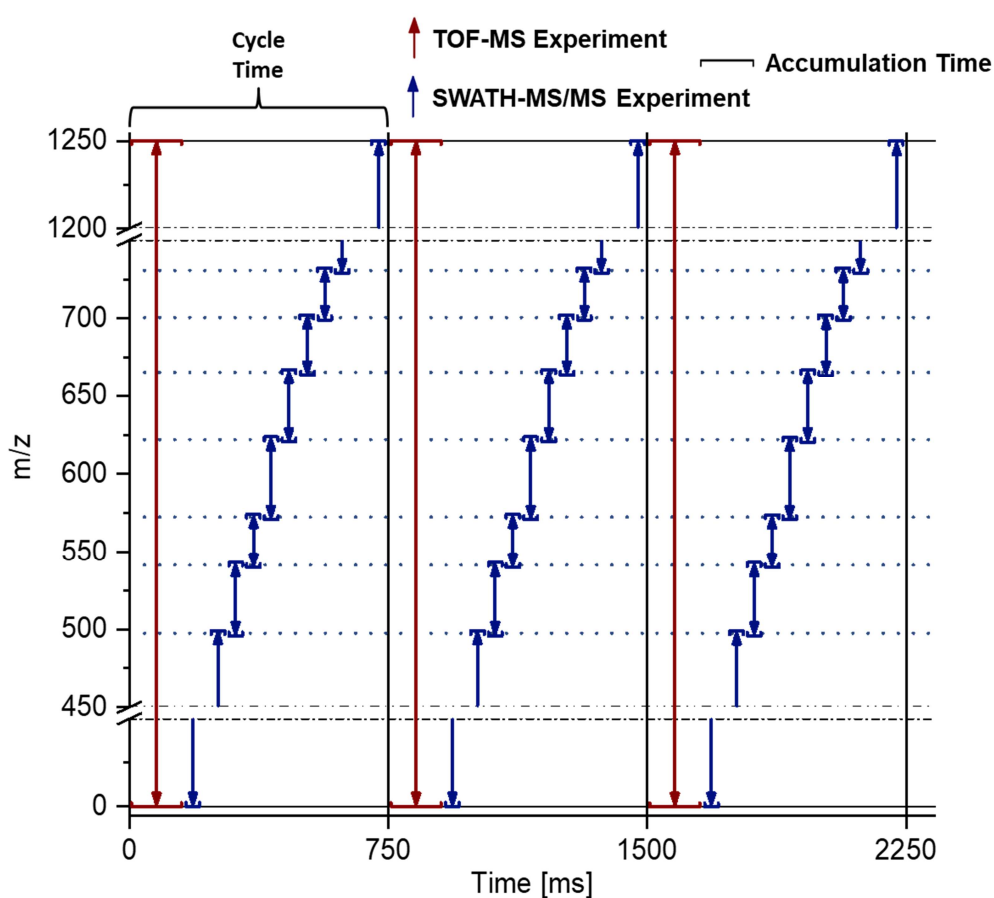


Figure 12. Exemplary scheme of a SWATH enabled acquisition cycle. The amount of SWATH-MS/MS experiments as well as the individual window width can be individually adjusted. Total cycle time has to be adjusted according to the chromatographic peak width.

Although the spectral quality is more susceptible to interference than IDA, numerous studies showed that SWATH results in higher identification rates and better analyte coverage.^{122,128,129} Moreover, SWATH also offers EIC generation on both MS and MS/MS level, which illustrates the excellent potential for quantification via the most sensitive and selective ion signals. In

addition, full comprehensive MS and MS/MS spectra facilitate the deconvolution process by matching of precursor and fragment retention times, which is exploited in the MS-DIAL (Mass Spectrometry - Data-Independent Analysis) software.¹³⁰

SWATH is further progressing by introduction of new bioinformatic tools like swathTUNER,¹³¹ which is used to calculate optimized SWATH window widths based on equalized precursor ion density or uniform TIC intensity. Also, merged experiment designs with narrow Q1 isolation windows for improved analysis of compounds of special interest are commonly reported.^{47,132,133} Additional advancements with novel techniques like scheduled SWATH^{131,133} (i.e. changing SWATH window settings throughout an analytical run) or scanning SWATH¹³⁴ (i.e. continuous scanning of the Q1 along the *m/z* range including fragment spectra deconvolution) can be expected in the future.

Ultimately, it can be noted that SWATH was proven to yield reliable and reproducible qualitative and quantitative data.¹³⁵ As an untargeted DIA technique it requires enhanced data processing efforts and demands bioinformatic solutions. Nevertheless, these obstacles have to be overcome to enable true comprehensive and retrospective analysis.

1.2.5. Method Optimization

Apart from sample preparation and chromatographic optimization, there are several MS parameters that can be modulated to obtain ideal results in terms of selectivity, sensitivity or linear dynamic range. In untargeted analysis, the main goal is to achieve a broad analyte coverage. Instrument settings are adjusted to comply with efficient ionization of many diverse compounds. Accordingly, optimization of MS parameters is based on finding the best compromise, which is often accomplished by selecting empirically derived standard settings. For targeted analysis, on the other hand, extensive optimization can be conducted in order to attain maximum method performance. This section will discuss the MS optimization parameters for the TripleTOF5600+ operated in the ESI configuration of a DuoSpray ion source.

Before any optimization efforts are initiated, the system should be thoroughly checked for errors, and tuning and calibration of all used acquisition modes (including Q1 settings, see Figure 7) must be properly verified. A major cause for a flawed performance is the ESI electrode, which needs to be free of contamination and corrosion to ensure a stable spray. Also, the protrusion of the electrode from the probe (i.e. the electrode holder) must be kept in the recommended limits between 0.5 – 1.0 mm, as a stable nebulizer gas flow around the tip can otherwise not be sustained. Ignoring these preparational steps can lead to high inter-sample variability and incorrect selection of optimized settings.

Assuming that the development of the chromatographic method has been completed, the target analyte should be dissolved in the mobile phase composition that is also expected during its elution. All subsequent parameter optimization steps are then conducted via continuous injection of this solution at the desired flow rate of the chromatographic method. Alternatively, if high flow rates are used, the defined mobile phase mixture can be provided by an LC instrument and the target analyte is added at minor flow rate via a T-piece. In both setups a reasonable analyte concentration should be administered to work outside any saturation conditions.

After preliminary tests have revealed the favorable polarity for analyte detection, the vertical and horizontal position of the probe can be adjusted. It is mainly dependent on the flow rate and typically yields best results if it is moved closer to the orifice with decreasing flow (and vice versa). Here, as well as for the adjustment of all other parameters, the settings should be optimized to obtain the best S/N ratio, as this will ultimately determine the analyte sensitivity. Nevertheless, settings that avoid exceeding contamination, implicitly relevant when adjusting the probe position, should be chosen to sustain a robust performance.

In the next step, the declustering potential (DP) and CE are to be optimized. The DP is a voltage that is applied to the MS orifice to extract analyte ions from the spray into the MS ion path. The voltage regulates the acceleration of ions towards the orifice, where they collide with residual gas molecules to reduce solvent clusters or unwanted adducts. At its optimized value it diminishes the division of analyte signals across different adducts and shifts the adduct distribution towards a preferred species. At highly elevated DP values, in-source fragmentation can occur, which decreases the precursor ion yield. Yet, in certain cases in-source fragments show enhanced sensitivity or can be used for structure elucidation.

Usually, MS/MS analysis results in increased selectivity and a beneficial S/N ratio. Fragmentation can be modulated by changing the CE, a difference of electric potentials between the Q0 and the q2 (see Figure 8). Transmitted ions are forwarded into the q2, where they undergo CID depending on the degree of acceleration. Eventually, optimum fragmentation is evaluated by acquiring data while ramping the CE voltage. Adjustments of DP and CE can also be utilized to balance instrument responses or to shift the linear dynamic range towards higher concentrations, if needed.

Ultimately, source and gas parameters are regulated. Besides their dependence on the flow rate and composition of the mobile phase, they are often also interdependent and require iterative fine tuning. In total, three different gas parameters are available for ESI: (i) the nebulizer gas, which flows around the electrode tip to support and direct the electrospray; (ii) the heater gas, that is originating from the heater to assist in solvent evaporation; (iii) the curtain gas, a gas flow between the orifice and the curtain plate that restrains contamination of the proximate ion optics by ambient air or solvent droplets. As a rule of thumb, the gases

should always be set to the highest possible value at which no loss in signal stability and sensitivity is observed. Especially the maximization of the curtain gas can lead to increased robustness and longevity of the system.

After the gas settings are assessed, the ion spray voltage is adjusted. It should be high enough to provide a stable spray but should also not forcibly operated at the maximum to avoid unwanted discharges and arcing.¹³⁶ Accordingly, its optimum setting is at the lowest value at which spray stability and sensitivity are not sacrificed. At the end, the heater temperature is adjusted to enable complete solvent evaporation and optimum ion desorption. A summary of optimization parameters and their operational ranges is given in Table 2.

There are also several advanced options to maximize the sensitivity. Exemplary measures are the alteration of the Q1 resolution (see Figure 7) to increase sensitivity while maintaining sufficient resolution and selectivity. Furthermore, the voltage of the detector (or MCP) can also be raised to yield higher signals. However, a constant increase of the voltage above its optimum value will drastically decrease its lifetime.

Table 2. Operational ranges for optimization parameters.

Parameter	Operational Range	Typical Value
Flow rate	5 – 3000 μ L/min	200 μ L/min
Nebulizer gas	0 – 90 psi	40 – 60 psi
Heater gas	0 – 90 psi	50 psi
Curtain gas	20 – 50 psi	30 psi
Temperature	0 – 750 $^{\circ}$ C	425 – 650 $^{\circ}$ C
DP	ESI ⁺ : 70 V ESI ⁻ : -70 V	ESI ⁺ : 0 – 400 V ESI ⁻ : -400 – 0 V
CE	ESI ⁺ : 0 – 150 V ESI ⁻ : -150 – 0 V	Compound dependent
Ion Spray Voltage	ESI ⁺ : 0 – 5500 V ESI ⁻ : -4500 – 0 V	ESI ⁺ : 5500 V ESI ⁻ : -4500 V
Time bins to sum	1 – 100	4

An additional parameter that should be noted are the time bins to sum. As discussed in Figure 9, the maximum time resolution of the TDC is 25 ps (\cong one time bin). The operator can select, how many time bins are merged to create a datapoint, e.g. if four time bins to sum are selected, the TDC will collect all signals throughout the several thousand TOF pulses and create the arrival time histogram in bins of 100 ps. This way, the spectral intensity will be apparently higher, as four single data points are now summed up to one datapoint, but also resolution is lost (see Figure 13). Since the total number of ion counts is not changed, the overall sensitivity is not increased and EICs show identical peak heights and peak areas. However, this option

can be useful to increase the apparent spectral intensity (by sacrificing resolution) in order to detect large, low abundant ions (e.g. for intact protein measurements) that would have otherwise been lost.

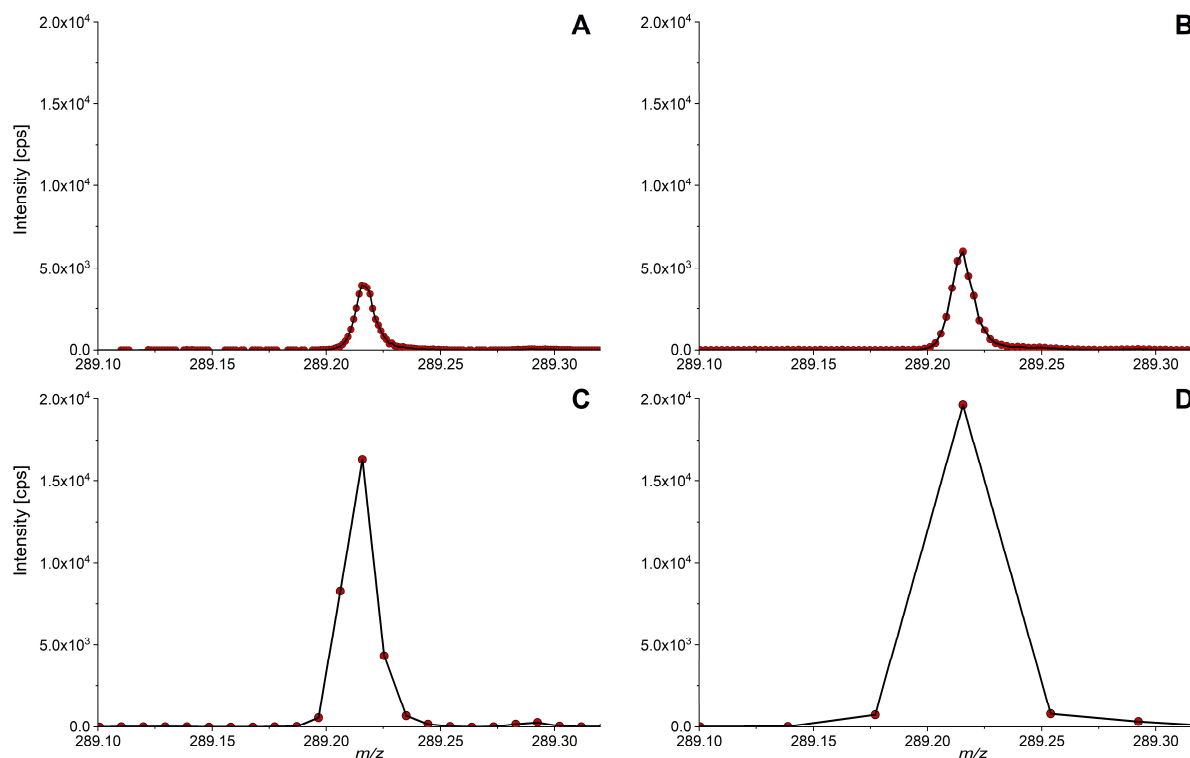


Figure 13. Influence of time bins to sum parameter on spectral resolution. Data shows repeated injections of 5 ng/mL testosterone with differing time bins to sum settings. A, 2 time bins to sum; B, 4 time bins to sum (standard setting); C, 16 time bins to sum; D, 64 time bins to sum. The choice of the bin size has a drastic influence on the spectral resolution. For each graph, the total sum of cps is identical. The amount of time bins to sum should be chosen to yield at least 10 data points per spectral peak.

1.2.6. Data Processing in Untargeted Lipidomics

Untargeted lipidomics assays, in particular when acquired with DIA techniques, yield vast raw data files with millions of data points. To be able to transform this information into interpretable results, multiple (pre-)processing steps via sophisticated bioinformatics are required.^{137,138} As an intermediate result, a data matrix, covering m/z values and retention times of unknowns and identified/annotated compounds together with their corresponding intensities in the respective samples, is obtained. Only after final statistical evaluation, results are then considered for hypothesis generation. A broad variety of commercialized, open access, or script-based tools is available and under constant development to support the exhaustive extraction of qualitative and quantitative information and its automation. As an integral part of untargeted analysis, data processing and statistical evaluation are briefly discussed in this subchapter.

The main goal of initial processing procedures is to strip down the raw data in order to cover only the relevant information. Accordingly, it is necessary to execute noise filtering (i.e. background subtraction), either by setting dynamic or absolute intensity thresholds under which signals are not further considered for subsequent steps.¹³⁸ In addition, a blank subtraction should be conducted. Here, unspecific signals acquired from a processed blank sample are excluded for the interpretation of the study outcome or are only respected if they are detected with a significantly higher intensity in the study samples. Afterwards, spectra are subjected to deisotoping and deadducting to reduce multiple signals, which can be derived from the same analyte, e.g. when several charge states or adducts of the analyte are observed, into a (monoisotopic) peak.^{139,140}

In chromatographic assays, data filtering is followed by the detection of actual features via peak finding algorithms, which assign series of data points to individual features, without considering remaining noise.^{138,141} Optionally, found peaks can be smoothed to improve its appearance or its S/N ratios.¹⁴² The data then has to be aligned across all study samples due to inter-sample variation of retention time and mass accuracy.¹⁴³ Depending on the chromatographic precision and the stability of the mass calibration, tolerance thresholds can be set to modulate alignment parameters. Moreover, alignment filters that determine the minimum number of samples in which the feature must be detected, can be applied.

Generally, MS acquisition is accompanied by the introduction of systematic or random errors, e.g. via fluctuations of the electrospray, which lead to variation of signal intensities. This issue can be addressed with normalization. In targeted assays, sufficient normalization is usually achieved by the addition of appropriate ISs. For untargeted data, a large variety of normalization strategies was developed, comprising scaling methods, QC-based, model-based, or IS-based approaches.¹⁴⁴ Suitable methods have to be chosen with care, as different normalization procedures regulate different types of variation. Due to the partially strong data manipulation, results should always be thoroughly checked for plausibility to avoid the interpretation of artificially induced results.

The fundamental part of any untargeted assay is compound annotation and identification.^{88,114,145} To increase the confidence level of identification,¹¹¹⁻¹¹⁴ as much information as possible should be gathered about the feature, with which several libraries can be screened for potential matches (see Table 3). Via the accurate mass and the isotope pattern of the precursor, sum formulas are initially derived. Thereafter, corresponding MS/MS spectra are matched to spectral fragment libraries, which were either iteratively created by experimental determination or consist of computer-generated fragment patterns (so-called *in silico* MS/MS libraries). MS/MS spectral matching drastically decreases the number of potential hits, yet, highly similar compounds with identical sum formulas can show overlaps in fragmentation. It is therefore advised to acquire orthogonal information like retention time or

collision cross section (CCS) to further enhance the confidence level. In case no library hit is found for an unknown peak of interest, the collected information should be used for structure elucidation, and, if necessary, additional experiments need to be conducted. The identity of a compound can be verified with high probability, when an equivalent standard reference matches the previously acquired findings.

Table 3. Selection of compound and mass spectral databases.

Database	Entries	Description
Pubchem ¹⁴⁶	235,035,188 structures	Largest compound database of small molecules
Chempider ¹⁴⁷	110,527,546 structures	Collection of chemical structures for small molecules
KEGG ¹⁴⁸	647,201 structures	Pathway database for metabolites of multiple species
MetaCyc ¹⁴⁹	15,655 structures	Pathway database for metabolites of multiple species
HMDB ¹⁵⁰	114,100 structures	Compound database of human metabolites including drugs, toxins, pollutants and nutritional products
CHEBI ¹⁵¹	55,878 structures	Compound database for small molecules of biological interest
LMSD ¹⁵²	43,413 structures	Compound database for lipids
METLIN ¹⁵³	> 431,000 spectra	MS database for metabolites with experimental and <i>in silico</i> spectra
mzCloud	24,303 spectra	MS database for metabolites and drugs with experimental spectra
MoNA	622,520 spectra	MS database for metabolites with experimental and <i>in silico</i> spectra

Access date July 2019.

Modern software is capable to automatically screen the available data and match the findings to embedded or connected libraries. Results are usually reported as a score that implies a probability for the reported identification. Furthermore, in particular when DIA techniques are used, implemented deconvolution algorithms are able to purify composite spectra by assigning only highly associated fragments to their precursors. Misannotations, however, cannot be ruled out and high priority findings must be manually checked for consistency to verify the presence of adequate data for reliable identification.¹⁵⁴ Here, additionally to the previous measures, extrapolation of lipid elution patterns in RPLC can be utilized to obtain accessory information.¹⁵⁵

1.2.7. Statistical Evaluation

With the completion of data processing, an aligned and, if necessary, normalized data matrix of several hundreds or thousands of features is obtained. To evaluate such complex results, various statistical calculation or visualization procedures are employed. Many statistical tests require normal distribution and are not able to deal with zero or empty values in the data array. Thus, logarithmic transformation and missing value imputation are often applied prior to statistical evaluation.¹⁴⁴ Furthermore, the data must be systematically screened for true outliers, as they will have a detrimental effect on the study outcome, if present.

A regular untargeted lipidomics study for biomarker discovery is designed to elaborate discrete features that have been significantly altered between properly chosen experimental groups. For this purpose, means/medians of these groups are compared by statistical hypothesis testing. In general, parametric tests, for which preliminary assumptions about the structure of the data are made, are distinguished from nonparametric tests. For example, *t*-tests are parametric as they are only valid for normally distributed data. Moreover, for Student's *t*-test equal variances between groups are assumed, whereas Welch's *t*-test is designed for unequal variances.¹⁵⁶ Accordingly, to correctly apply these tests, the proclaimed assumptions must be verified by testing normality (e.g. via the Shapiro-Wilk or Kolmogorov-Smirnov tests) and variance equality (Bartlett's test).¹⁵⁷ If these assumptions are not fulfilled or unknown, nonparametric tests like the Wilcoxon-Mann-Whitney-*U*-test are performed. The results of statistical hypothesis testing are usually reported as *p*-values, which are indicating probabilities for the trueness of the null hypothesis.

By design, *p*-values below a preset significance level α will occur by chance when the number of hypothesis tests is increased. This is a well-known issue of multiple testing in untargeted assays, which will lead to false positives (type I errors).¹⁵⁸ The *p*-values must therefore be adjusted to control for type I errors, which can be achieved by various *p*-value correction procedures like the false discovery rate evaluation or the Bonferroni approach.^{159,160}

Besides univariate significance tests, numerous multivariate statistical methods are available. The most prominently used methodology for untargeted assays is principal component analysis (PCA). It is highly suitable to transform multidimensional data into lower dimensional projections via linear combination of variables, that show maximum variation.¹⁵⁷ Via the obtained scores plot, relations between samples can be derived from grouping or separating trends. Also, when QCs are embedded in the analytical sequence, the stability of the instrument performance can be estimated with the scores plot, as technical replicates of the QCs must ideally be tightly clustered due to the expected low variation. In contrast, potential outliers can be spotted, since they show different variation and are not clustered with the other

experimental samples. Furthermore, the loadings plot allows to investigate the contribution of the individual variables to the observed variation.¹⁵⁷

Many more tools for extensive statistical analysis are available, e.g discriminant analysis, receiver operating characteristics (ROC), hierarchical clustering, volcano plots, heatmaps, and partial least squares methods. Together, they aid to visualize classification or differences between experimental groups and are consulted to evaluate highly significant features, which can then be postulated as potential biomarkers.

1.3. Bioanalytical Method Validation

Bioanalysis is specified as the ability to determine drug and/or metabolite concentrations in biological matrices. It is the basis of many pharmaceutical and clinical studies concerning pharmacokinetics, bioequivalence, toxicology, and therapeutic drug monitoring and plays a key role in the interpretation of pharmacodynamics and the regulation of drug safety and drug efficacy. In 1990, with the emergence of novel analytical advancements, in particular the introduction of commercialized LC-MS platforms, the demand for the harmonization of validation procedures was first recognized by the American Association of Pharmaceutical Scientists (AAPS) and the United States Food and Drug Administration (FDA).¹⁶¹ Several additional workshops throughout the last decades^{162,163} have resulted in the formulation of broadly accepted guidelines for bioanalytical method validation of therapeutics and biomarkers. The central aspects of these guidelines, which were recently updated in 2018,¹⁶⁴ and the requirements for chromatographic assays are discussed in the following section.

To prove that a method can reliably quantify the target analyte in study samples, several experiments have to be conducted and preset criteria must be fulfilled. At first, the selectivity and specificity of target analytes and ISs need to be verified by analysis of blank matrix samples from at least six sources. If the absence of significant interferences has been shown, a calibration curve for every analyte must be assessed in the sample matrix. This is achieved by spiking target analyte reference standards (ideally of authenticated analytical purity with the certificates of analyses provided) into an analyte-free matrix. Good practices require the addition of an appropriate IS for every analyte in the first steps of sample preparation or after sample collection, to correct for analyte loss or matrix effects.¹⁶⁵ ISs do not require certification, but their suitability must be demonstrated by the absence of interferences with the analytes. The used matrix should be identical to the matrix of the study samples and the calibration range should comprise all expected concentration levels. Furthermore, the simplest model to accurately correlate instrument response ratios (analyte response / corresponding IS response) and analyte concentration should be chosen. In total, a calibration curve should consist of a blank sample (without analytes and ISs), a zero calibrant (without analytes but with

ISs) and at least six non-zero calibrant levels. The lowest non-zero calibrant level is assigned as the lower limit of quantification (LLOQ), whereas the highest calibrant level indicates the upper limit of quantification (ULOQ). If a background noise or interference of the analyte is found in the blank sample, the analyte response at the LLOQ should exceed the blank response by at least 5-fold. After regression, calculated concentrations for calibrants should show a maximum bias of $\pm 15\%$ to their theoretical concentrations, except for LLOQ where $\pm 20\%$ bias are accepted. For successful calibration, this criteria must apply to at least 75 % of all calibrant samples.

Besides calibrants, quality control (QC) samples are prepared by spiking blank matrix, but different stock solutions of the reference standards should be used for this purpose. The QCs are employed to evaluate accuracy and precision at four concentration levels: (i) at the LLOQ; (ii) at the lower range (defined as 3-fold LLOQ); (iii) at the mid range (not precisely defined); (iv) at the high range (not precisely defined) of the calibration curve. The performance of the QCs is evaluated via preceding calibrant measurements for at least three independent runs. In each run, the respective QC levels must be analyzed with a minimum of five replicates. The same acceptance criteria that were previously defined for calibrant samples also apply to accuracy and precision of the QCs.

Moreover, QCs are utilized to assess analyte stability for different conditions, e.g. stability of the processed samples, freeze-thaw stability, or long-term stability. These measurements only need to be conducted at low and high range QC levels in at least triplicate. For analyte recovery, no strict acceptance criteria are defined. However, its extent should be determined via post-extraction experiments¹⁶⁶ at low, mid, and high range QC levels.

An overview about the required validation runs and the respective acceptance criteria is provided in Table 4. After successful validation, quantitative performance has to be monitored throughout the study samples. Here, slightly different criteria are recommended (see Table 4). The described guidelines were postulated to standardize validation procedures for targeted analysis and quantitative purposes. For untargeted assays, comparable guidelines do not yet exist and, although the same principles apply, compliance to FDA guidelines is not always fully applicable owing to the enhanced complexity and the lack of standards and blank matrices. Nevertheless, the metabolomics and lipidomics community is gathering good practices and iteratively enforces improved workflows by striving for the harmonization of extraction protocols, QC design, normalization, data evaluation, identification, and quantitative approaches.^{68,167-170} Together with the growing availability of databases and bioinformatic tools¹⁷¹ the validity of untargeted metabolomics and lipidomics is constantly improving to comply with fit-for-purpose recommendations.

Table 4. Validation requirements and acceptance criteria according to FDA guidelines. ¹⁶⁴

Parameter	Validation Recommendations	In-Study Analysis Recommendation
Calibration Curve	<p>Elements:</p> <ul style="list-style-type: none"> • 1 blank matrix sample (no analytes, no ISs) • 1 zero calibrant (no analytes, with IS) • 6 non-zero calibrant levels • use same matrix as study samples • use simplest regression model <p>Acceptance criteria:</p> <ul style="list-style-type: none"> • Non-zero calibrators should be $\pm 15\%$ of theoretical concentrations, at LLOQ $\pm 20\%$ are acceptable • 75 % and a minimum of 6 non-zero calibrant levels should meet the criteria in each validation run 	<p>Elements:</p> <ul style="list-style-type: none"> • 1 blank matrix sample (no analytes, no ISs) • 1 zero calibrant (no analytes, with IS) • 6 non-zero calibrant levels • use same matrix as study samples • use same regression model as in validation <p>Acceptance criteria:</p> <ul style="list-style-type: none"> • Non-zero calibrators should be $\pm 15\%$ of theoretical concentrations, at LLOQ $\pm 20\%$ are acceptable • 75 % and a minimum of 6 non-zero calibrant levels should meet the criteria in each validation run
Quality Control	<p>Elements:</p> <ul style="list-style-type: none"> • for accuracy and precision: 4 QCs (LLOQ, 3xLLOQ, mid, high) • for other validation runs: 3xLLOQ, mid, high <p>Acceptance criteria:</p> <ul style="list-style-type: none"> • accuracy: QCs should be $\pm 15\%$ of theoretical concentrations, at LLOQ $\pm 20\%$ are acceptable • precision: QCs should show a relative standard deviation of $\pm 15\%$, at LLOQ $\pm 20\%$ are acceptable • accuracy and precision are established with ≥ 3 independent runs and ≥ 5 replicates per level 	<p>Elements:</p> <ul style="list-style-type: none"> • ≥ 3 QCs (3xLLOQ, mid, high) and ≥ 2 replicates in each analytical run • total QCs should be 5 % of study samples (excluding calibrants and QCs) or ≥ 6 <p>Acceptance criteria:</p> <ul style="list-style-type: none"> • $\geq 67\%$ of QCs should be $\pm 15\%$ of their theoretical concentrations • $\geq 50\%$ of QCs per level should be $\pm 15\%$ of their theoretical concentrations

	<p>Elements:</p> <ul style="list-style-type: none"> 6 individual sources must be interference-free <p>Acceptance criteria:</p> <ul style="list-style-type: none"> blank matrix sample and zero calibrant sample should be free of interference IS response in the blank should not exceed 5 % of the average IS responses of the calibrants and QCs 	-
Selectivity	<p>Elements:</p> <ul style="list-style-type: none"> check cross-reacting molecules, concomitant medications, bio-transformed species, etc. for interference <p>Acceptance criteria:</p> <ul style="list-style-type: none"> blank matrix sample and zero calibrant sample should be free of interference 	<p>Acceptance criteria:</p> <ul style="list-style-type: none"> blank matrix sample and zero calibrant sample should be free of interference
Specificity	<p>Elements:</p> <ul style="list-style-type: none"> check as needed 	<p>Elements:</p> <ul style="list-style-type: none"> check as needed
Carryover	<p>Acceptance criteria:</p> <ul style="list-style-type: none"> carryover should not exceed 20 % of the LLOQ <p>Elements:</p> <ul style="list-style-type: none"> the lowest non-zero calibrant defines the LLOQ 	<p>Acceptance criteria:</p> <ul style="list-style-type: none"> carryover should not exceed 20 % of the LLOQ <p>-</p>
Sensitivity	<p>Acceptance criteria:</p> <ul style="list-style-type: none"> the analyte response at the LLOQ should be ≥ 5 times the analyte response of the zero calibrant <p>Elements:</p> <ul style="list-style-type: none"> the lowest non-zero calibrant defines the LLOQ 	<p>Acceptance criteria:</p> <ul style="list-style-type: none"> the analyte response at the LLOQ should be ≥ 5 times the analyte response of the zero calibrant if LLOQ criteria are not met, the next higher calibrant is selected as new LLOQ
Stability	<p>Elements:</p> <ul style="list-style-type: none"> ≥ 3 replicates at 2 QC levels (3xLLOQ, high) <p>Acceptance criteria:</p> <ul style="list-style-type: none"> accuracy of QCs should be ± 15 % of theoretical concentrations 	<p>Elements:</p> <ul style="list-style-type: none"> check as needed
Recovery	<p>Elements:</p> <ul style="list-style-type: none"> extracted samples at 3xLLOQ, mid, high versus extracts of post-spiked extraction blanks at 3xLLOQ, mid, high 	-

1.3.1. Calibration of Endogenous Compounds

For the preparation of calibrants and QCs, according to FDA guidelines, the use of blank matrix is instructed. However, for many endogenous analytes no true blank matrix is available. In some cases, this problem can be resolved by selecting matrix donors from specific patient groups that naturally show analyte levels below the anticipated LLOQ, e.g. young girls or post-menopausal women with typically low plasma testosterone levels. Still, challenging analytes require alternative options for calibration and accurate quantification. In general, four approaches have been proposed: (i) standard addition;¹⁷² (ii) background addition; (iii) surrogate matrices;^{173,174} (iv) surrogate calibrants (i.e. surrogate analytes).¹⁷⁵

Standard addition is conducted by spiking small, increasing increments of the analyte into aliquots of the intended sample. By extrapolation of the calibration line, the original analyte concentration can be calculated from the x-intercept (see Figure 14). Determination of its uncertainty can also be executed via the error of the y-intercept. This method is effective as it is independent from variations of matrix effects (see Figure 14). However, a high sample volume is needed, and many extra measurements are necessary. It is also not feasible when the original analyte concentration is close to the upper limit of the linear dynamic range as additional spiking of the analyte will result in signal saturation.

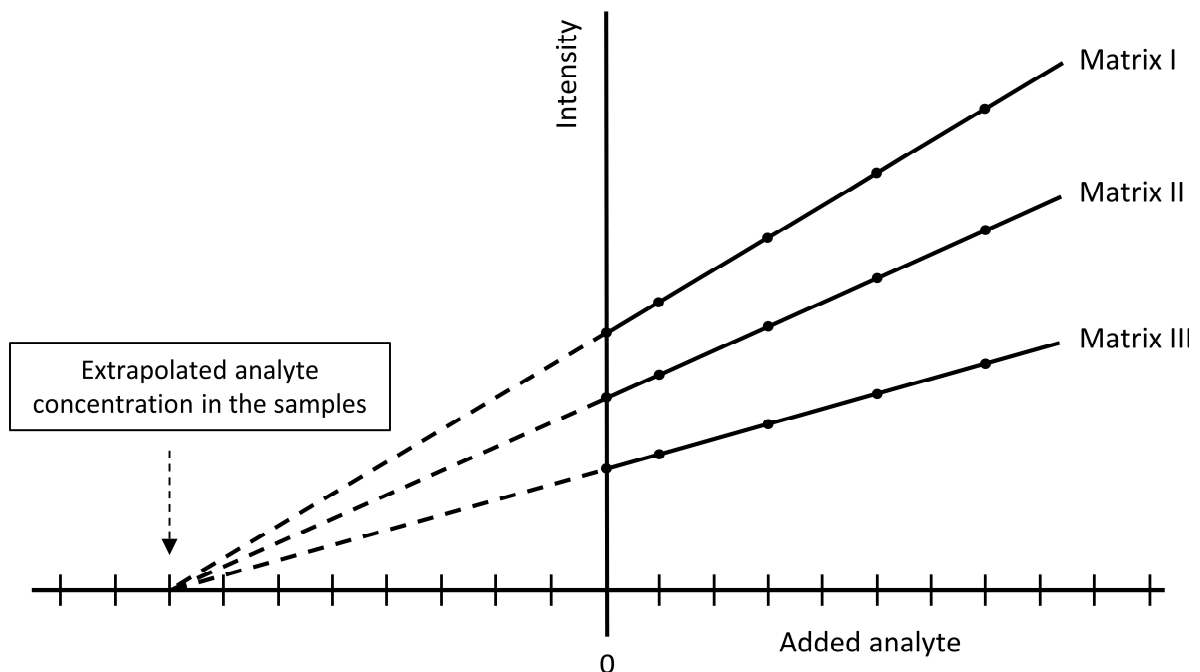


Figure 14. Standard addition. Samples with identical analyte concentration but differing matrix are shown. The differing matrix conditions do not affect the results.

When using background addition, a matrix pool with a known analyte concentration (determined via e.g. standard addition) is used to prepare calibrants and QCs. The analyte

content of the pool automatically represents the LLOQ and thus limits the achievable sensitivity that can be validated. The sum of the spiked analyte and the background concentration yields the final calibrant concentration, respectively. This method allows matrix-matched calibration but is not suitable if background concentrations of analytes are rather high. Moreover, the method has to be readjusted if a new batch of pooled matrix is used. Additional complications may arise if multiple analytes are quantified simultaneously and unacceptably high background concentrations are present. In the literature, also background subtraction is described.¹⁷⁴ Here, the total signal of a spiked sample is subtracted by the background signal of the pooled matrix. The spiked amount is then declared as the LLOQ. However, since the obtained signal does not reflect the conditions at the authentic LLOQ (decreased accuracy and precision due to a signal close to the noise level),¹⁶⁴ this method is not advised.

To be able to calibrate and validate the instrumental LLOQ, surrogate matrix methods can be used. Here, the analyte is spiked into a similar, modified or artificial matrix. The validity of the chosen matrix must be shown via parallelism of the calibration curve in surrogate matrix and the standard addition curve in true matrix.¹⁷⁶ If no matrix effects are present and if the recovery is 100 %, also neat solutions with sufficient analyte solubility can be chosen as surrogate matrices. Although the use of ISs is highly recommended, a simple approach to calibrate in neat solution with added ISs to control signal losses due to recovery and matrix effects (without evaluation of these parameters) is not applicable, as the LLOQ and the linear dynamic range cannot be accurately assessed.

Another option is to remove the target analyte from the true matrix via activated charcoal, specific antibodies, enzymes or chemical reactions/heat. Here, it has to be assured that prior to analyte spiking, all previously added material for analyte stripping is removed or deactivated. Yet, these methods bear the risk to alter the matrix effect or recovery and still parallelism has to be verified.

Moreover, artificial matrices can be utilized to simulate complex or scarce matrices. Most commonly phosphate-buffered saline and bovine serum albumin are used to mimic plasma and to enhance solubility for hydrophobic compounds compared to simple aqueous solutions.¹⁷⁷ Also more complex mixtures for diverse biological matrices are described.¹⁷³ A major challenge is the determination of a suitable surrogate matrix when multiple analytes must comply and exert parallelism.

Ultimately, an approach that utilizes authentic matrix and surrogate analytes for calibration are available. The surrogate calibrant must be a compound that shows the same analytical behavior as the targeted analyte. Amongst the best candidates are usually stable isotope labeled analogues of the target analyte, preferably marked with ¹³C-, ¹⁵N-, or ¹⁸O- atoms as these analogues were shown to be less prone to isotope exchange reactions and to more accurately reflect the physicochemical properties (including matrix effects) of the target analyte

than deuterated compounds.¹⁷³ As a matter of course, selectivity of target analytes and surrogate calibrants is required and any impurities of residual unlabeled analyte in the surrogate calibrant reference standard must be ruled out. Furthermore, it is often observed that the instrument response of surrogate calibrants is not identical to target analytes. This can have multiple causes like impurities, differences in ionization efficiencies due to kinetic isotope effects, or the compression of the isotope distribution pattern if ¹³C-labeled analogues are used (see Figure 15).¹⁷⁸ Therefore, instrument responses of surrogate calibrants have to be investigated and, if necessary, balanced to match the response of the corresponding target analytes. This can be achieved by modulation of CE or DP or by concentration adjustment of the surrogate calibrant. Alternatively, also a response factor can be applied but is not advised as the conditions at the LLOQ are not accurately reproduced when response factors are applied. After matching of response ratios (usually a tolerance of 5 % deviation is accepted),¹⁷⁶ parallelism has to be proved.

Although precise guidelines for validation of surrogate methods are not given, meeting the existing criteria that are valid for bioanalytical method validation should be anticipated as far as possible.

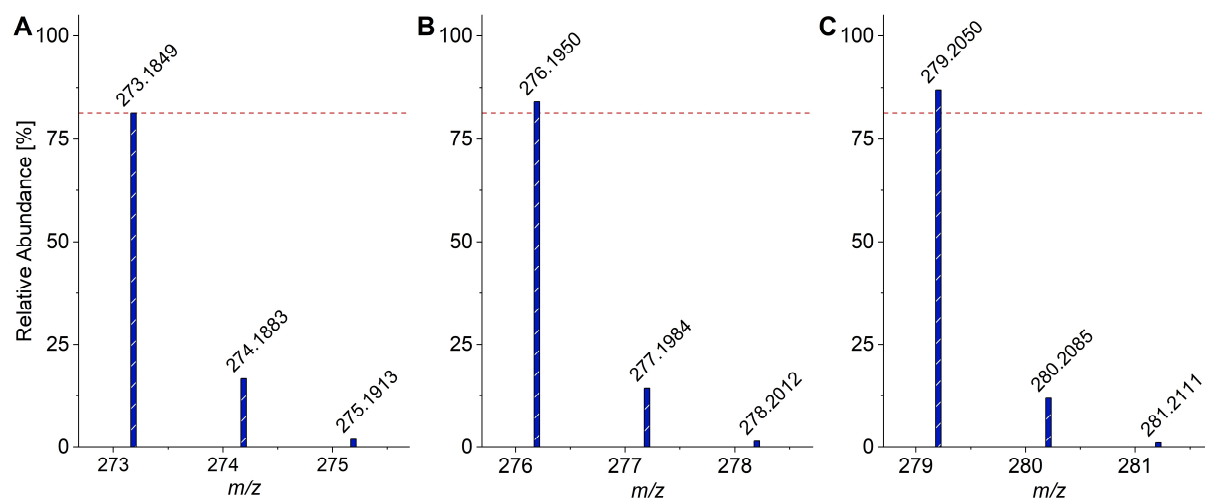


Figure 15. Compression of monoisotopic peaks. The isotopic distributions of respective [M+H]⁺ adducts of estradiol and its ¹³C-analogues are shown. A, Estradiol (C₁₈H₂₄O₂); B, ¹³C₃-estradiol; C, ¹³C₆-estradiol. With an increasing number of incorporated ¹³C-atoms the portion of the monoisotopic peak is increased.

1.4. References

- (1) Sanger, F.; Nicklen, S.; Coulson, A. R. *Proc Natl Acad Sci U S A* **1977**, *74*, 5463-5467.
- (2) Shendure, J.; Balasubramanian, S.; Church, G. M.; Gilbert, W.; Rogers, J.; Schloss, J. A.; Waterston, R. H. *Nature* **2017**, *550*, 345-353.
- (3) Lagarde, M.; Geloën, A.; Record, M.; Vance, D.; Spener, F. *Biochim Biophys Acta* **2003**, *1634*, 61.
- (4) Lee, S. H.; Lee, N.; Hong, Y.; Chung, B. C.; Choi, M. H. *Anal Chem* **2016**, *88*, 11624-11630.
- (5) Moss, G. P.; Smith, P. A. S.; Tavernier, D. In *Pure and Applied Chemistry*, 1995, p 1307.
- (6) Heawchaiyaphum, C.; Pientong, C.; Phusingha, P.; Vatanasapt, P.; Promthet, S.; Daduang, J.; Teeramatwanich, W.; Kongyingyoes, B.; Chuerduangphui, J.; Ekalaksananan, T. *J Proteomics* **2018**, *173*, 52-61.
- (7) Fahy, E.; Subramaniam, S.; Brown, H. A.; Glass, C. K.; Merrill, A. H., Jr.; Murphy, R. C.; Raetz, C. R.; Russell, D. W.; Seyama, Y.; Shaw, W.; Shimizu, T.; Spener, F.; van Meer, G.; VanNieuwenhze, M. S.; White, S. H.; Witztum, J. L.; Dennis, E. A. *J Lipid Res* **2005**, *46*, 839-861.
- (8) Fahy, E.; Subramaniam, S.; Murphy, R. C.; Nishijima, M.; Raetz, C. R.; Shimizu, T.; Spener, F.; van Meer, G.; Wakelam, M. J.; Dennis, E. A. *J Lipid Res* **2009**, *50* Suppl, S9-14.
- (9) Gorter, E.; Grendel, F. *J Exp Med* **1925**, *41*, 439-443.
- (10) Spector, A. A.; Yorek, M. A. *J Lipid Res* **1985**, *26*, 1015-1035.
- (11) Wymann, M. P.; Schneider, R. *Nat Rev Mol Cell Biol* **2008**, *9*, 162-176.
- (12) Stephenson, D. J.; Hoeflerlin, L. A.; Chalfant, C. E. *Transl Res* **2017**, *189*, 13-29.
- (13) Svensson, M.; Mossberg, A. K.; Pettersson, J.; Linse, S.; Svanborg, C. *Protein Sci* **2003**, *12*, 2805-2814.
- (14) Abdullah, L.; Evans, J. E.; Emmerich, T.; Crynen, G.; Shackleton, B.; Keegan, A. P.; Luis, C.; Tai, L.; LaDu, M. J.; Mullan, M.; Crawford, F.; Bachmeier, C. *Aging (Albany NY)* **2017**, *9*, 964-985.
- (15) Yamamoto, K.; Isogai, Y.; Sato, H.; Taketomi, Y.; Murakami, M. *Anal Bioanal Chem* **2011**, *400*, 1829-1842.
- (16) Kim, H. Y.; Lee, K. M.; Kim, S. H.; Kwon, Y. J.; Chun, Y. J.; Choi, H. K. *Oncotarget* **2016**, *7*, 67111-67128.
- (17) Pechlaner, R.; Kiechl, S.; Mayr, M. *Circulation* **2016**, *134*, 1651-1654.
- (18) Markgraf, D. F.; Al-Hasani, H.; Lehr, S. *Int J Mol Sci* **2016**, *17*.
- (19) Passos-Castilho, A. M.; Carvalho, V. M.; Cardozo, K. H.; Kikuchi, L.; Chagas, A. L.; Gomes-Gouvea, M. S.; Malta, F.; de Seixas-Santos Nastro, A. C.; Pinho, J. R.; Carrilho, F. J.; Granato, C. F. *BMC Cancer* **2015**, *15*, 985.
- (20) Vatarescu, M.; Bechor, S.; Haim, Y.; Pecht, T.; Tarnovscki, T.; Slutsky, N.; Nov, O.; Shapiro, H.; Shemesh, A.; Porgador, A.; Bashan, N.; Rudich, A. *J Endocrinol* **2017**, *233*, 293-305.
- (21) Li, J.; Ren, S.; Piao, H. L.; Wang, F.; Yin, P.; Xu, C.; Lu, X.; Ye, G.; Shao, Y.; Yan, M.; Zhao, X.; Sun, Y.; Xu, G. *Sci Rep* **2016**, *6*, 20984.
- (22) Meeusen, J. W.; Donato, L. J.; Bryant, S. C.; Baudhuin, L. M.; Berger, P. B.; Jaffe, A. S. *Arterioscler Thromb Vasc Biol* **2018**, *38*, 1933-1939.
- (23) Puri, R.; Duong, M.; Uno, K.; Kataoka, Y.; Nicholls, S. J. *Expert Opin Drug Discov* **2012**, *7*, 63-72.
- (24) Hammerling, J. A. *Laboratory Medicine* **2012**, *43*, 41-44.
- (25) Wu, Y.; Li, L. *J Chromatogr A* **2016**, *1430*, 80-95.
- (26) Rustam, Y. H.; Reid, G. E. *Analytical Chemistry* **2018**, *90*, 374-397.
- (27) Wang, J.; Wang, C.; Han, X. *Anal Chim Acta* **2019**, *1061*, 28-41.
- (28) Jurowski, K.; Kochan, K.; Walczak, J.; Barańska, M.; Piekoszewski, W.; Buszewski, B. *TrAC Trends in Analytical Chemistry* **2017**, *86*, 276-289.
- (29) Bligh, E. G.; Dyer, W. J. *Can J Biochem Physiol* **1959**, *37*, 911-917.
- (30) Folch, J.; Lees, M.; Sloane Stanley, G. H. *J Biol Chem* **1957**, *226*, 497-509.
- (31) Hara, A.; Radin, N. S. *Anal Biochem* **1979**, *100*, 364-370.
- (32) Löfgren, L.; Forsberg, G.-B.; Ståhlman, M. *Scientific Reports* **2016**, *6*, 27688.
- (33) Pellegrino, R. M.; Di Veroli, A.; Valeri, A.; Goracci, L.; Cruciani, G. *Anal Bioanal Chem* **2014**, *406*, 7937-7948.
- (34) Matyash, V.; Liebisch, G.; Kurzchalia, T. V.; Shevchenko, A.; Schwudke, D. *J Lipid Res* **2008**, *49*, 1137-1146.

- (35) Calderon, C.; Sanwald, C.; Schlotterbeck, J.; Drotleff, B.; Lammerhofer, M. *Anal Chim Acta* **2019**, 1048, 66-74.
- (36) Sarafian, M. H.; Gaudin, M.; Lewis, M. R.; Martin, F.-P.; Holmes, E.; Nicholson, J. K.; Dumas, M.-E. *Anal Chem* **2014**, 86, 5766-5774.
- (37) Lee, J. Y.; Yoo, C.; Jun, S. Y.; Ahn, C. Y.; Oh, H. M. *Bioresour Technol* **2010**, 101 Suppl 1, S75-77.
- (38) Schmid, P.; Hunter, E.; Calvert, J. *Physiol Chem Phys Med NMR* **1973**, 151-155.
- (39) Nagano, K.; Kano, H.; Arito, H.; Yamamoto, S.; Matsushima, T. *J Toxicol Environ Health A* **2006**, 69, 1827-1842.
- (40) Celltreat.
- (41) Hyotylainen, T.; Ahonen, L.; Poho, P.; Oresic, M. *Biochim Biophys Acta Mol Cell Biol Lipids* **2017**, 1862, 800-803.
- (42) Polson, C.; Sarkar, P.; Incledon, B.; Raguvaran, V.; Grant, R. *Journal of Chromatography B* **2003**, 785, 263-275.
- (43) Alshehry, Z. H.; Barlow, C. K.; Weir, J. M.; Zhou, Y.; McConville, M. J.; Meikle, P. J. *Metabolites* **2015**, 5, 389-403.
- (44) Cho, M.; Chen, Q.; Okpodu, C.; Boss, W. LC-GC(MAG. SEPAR. SCI.). **1992**, 10, 464-468.
- (45) Scherer, M.; Schmitz, G.; Liebisch, G. *Clin Chem* **2009**, 55, 1218-1222.
- (46) Pati, S.; Nie, B.; Arnold, R. D.; Cummings, B. S. *Biomed Chromatogr* **2016**, 30, 695-709.
- (47) Drotleff, B.; Hallschmid, M.; Lämmerhofer, M. *Analytica Chimica Acta* **2018**, 1022, 70-80.
- (48) Danne-Rasche, N.; Coman, C.; Ahrends, R. *Anal Chem* **2018**, 90, 8093-8101.
- (49) Yang, K.; Han, X. *Trends Biochem Sci* **2016**, 41, 954-969.
- (50) Jurowski, K.; Kochan, K.; Walczak, J.; Baranska, M.; Piekoszewski, W.; Buszewski, B. *Crit Rev Anal Chem* **2017**, 47, 418-437.
- (51) Li, J.; Vosegaard, T.; Guo, Z. *Progress in Lipid Research* **2017**, 68, 37-56.
- (52) Wang, M.; Wang, C.; Han, X. *Mass Spectrom Rev* **2017**, 36, 693-714.
- (53) Trufelli, H.; Palma, P.; Famiglini, G.; Cappiello, A. *Mass Spectrometry Reviews* **2011**, 30, 491-509.
- (54) Gathungu, R. M.; Larrea, P.; Sniatynski, M. J.; Marur, V. R.; Bowden, J. A.; Koelmel, J. P.; Starke-Reed, P.; Hubbard, V. S.; Kristal, B. S. *Analytical Chemistry* **2018**, 90, 13523-13532.
- (55) Yang, K.; Cheng, H.; Gross, R. W.; Han, X. *Anal Chem* **2009**, 81, 4356-4368.
- (56) Han, X.; Yang, K.; Yang, J.; Fikes, K. N.; Cheng, H.; Gross, R. W. *J Am Soc Mass Spectrom* **2006**, 17, 264-274.
- (57) Holčapek, M.; Liebisch, G.; Ekroos, K. *Analytical Chemistry* **2018**, 90, 4249-4257.
- (58) Paglia, G.; Kliman, M.; Claude, E.; Geromanos, S.; Astarita, G. *Anal Bioanal Chem* **2015**, 407, 4995-5007.
- (59) Fauland, A.; Kofeler, H.; Trotschmuller, M.; Knopf, A.; Hartler, J.; Eberl, A.; Chitraju, C.; Lankmayr, E.; Spener, F. *J Lipid Res* **2011**, 52, 2314-2322.
- (60) Schlotterbeck, J.; Kolb, A.; Lämmerhofer, M. *Journal of Separation Science* **2018**, 41, 4286-4295.
- (61) Schlotterbeck, J.; Cebo, M.; Kolb, A.; Lämmerhofer, M. *Anal Bioanal Chem* **2019**, 411, 479-491.
- (62) Anesi, A.; Guella, G. *J Chromatogr A* **2015**, 1384, 44-52.
- (63) Lísá, M.; Holčapek, M. *Analytical Chemistry* **2015**, 87, 7187-7195.
- (64) Yang, L.; Cui, X.; Zhang, N.; Li, M.; Bai, Y.; Han, X.; Shi, Y.; Liu, H. *Anal Bioanal Chem* **2015**, 407, 5065-5077.
- (65) Zhang, Q.; Xu, H.; Liu, R.; Gao, P.; Yang, X.; Jin, W.; Zhang, Y.; Bi, K.; Li, Q. *Analytical Chemistry* **2019**, 91, 3389-3396.
- (66) Contrepois, K.; Mahmoudi, S.; Ubhi, B. K.; Papsdorf, K.; Hornburg, D.; Brunet, A.; Snyder, M. *Scientific Reports* **2018**, 8, 17747.
- (67) Zhou, J.; Liu, C.; Si, D.; Jia, B.; Zhong, L.; Yin, Y. *Analytica Chimica Acta* **2017**, 972, 62-72.
- (68) Forest, A.; Ruiz, M.; Bouchard, B.; Boucher, G.; Gingras, O.; Daneault, C.; Robillard Frayne, I.; Rhainds, D.; Tardif, J.-C.; Rioux, J. D.; Des Rosiers, C. *Journal of Proteome Research* **2018**, 17, 3657-3670.
- (69) Yin, L.; Zhang, Z.; Liu, Y.; Gao, Y.; Gu, J. *Analyst* **2019**, 144, 824-845.
- (70) Passarelli, M. K.; Ewing, A. G.; Winograd, N. *Analytical Chemistry* **2013**, 85, 2231-2238.
- (71) Phelps, M. S.; Verbeck, G. F. *Analytical Methods* **2015**, 7, 3668-3670.

- (72) Buchberger, A. R.; DeLaney, K.; Johnson, J.; Li, L. *Analytical Chemistry* **2018**, 90, 240-265.
- (73) Merrill, C. B.; Basit, A.; Armirotti, A.; Jia, Y.; Gall, C. M.; Lynch, G.; Piomelli, D. *Scientific Reports* **2017**, 7, 5318.
- (74) Aston, F. W. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **1920**, 39, 449-455.
- (75) Griffiths, J. *Anal Chem* **2008**, 80, 5678-5683.
- (76) Yates lii, J. R. *Nature Methods* **2011**, 8, 633.
- (77) de Hoffmann, E.; Stroobant, V. *Mass Spectrometry - Principles and Application*, 3rd ed.; John Wiley & Sons, Ltd, 2007.
- (78) Siuzdak, G. *JALA: Journal of the Association for Laboratory Automation* **2004**, 9, 50-63.
- (79) Kebarle, P.; Tang, L. *Analytical Chemistry* **1993**, 65, 972A-986A.
- (80) Tang, L.; Kebarle, P. *Analytical Chemistry* **1993**, 65, 3654-3668.
- (81) Hopfgartner, G.; Bean, K.; Henion, J.; Henry, R. *J Chromatogr A* **1993**, 647, 51-61.
- (82) Van Berkel, G. J.; Kertesz, V. *Analytical Chemistry* **2007**, 79, 5510-5520.
- (83) Van Berkel, G. J.; Zhou, F. *Analytical Chemistry* **1995**, 67, 3958-3964.
- (84) Cappiello, A.; Famiglini, G.; Palma, P.; Pierini, E.; Termopoli, V.; Trufelli, H. *Analytical Chemistry* **2008**, 80, 9343-9348.
- (85) Trufelli, H.; Palma, P.; Famiglini, G.; Cappiello, A. *Mass Spectrom Rev* **2011**, 30, 491-509.
- (86) Dreisewerd, K. *Chemical Reviews* **2003**, 103, 395-426.
- (87) Haag, A. M. In *Modern Proteomics – Sample Preparation, Analysis and Practical Applications*, Mirzaei, H.; Carrasco, M., Eds.; Springer International Publishing: Cham, 2016, pp 157-169.
- (88) Kind, T.; Tsugawa, H.; Cajka, T.; Ma, Y.; Lai, Z.; Mehta, S. S.; Wohlgemuth, G.; Barupal, D. K.; Showalter, M. R.; Arita, M.; Fiehn, O. *Mass Spectrometry Reviews* **2018**, 37, 513-532.
- (89) Paul, W.; Steinwedel, H. In *Zeitschrift für Naturforschung A*, 1953, p 448.
- (90) Miller, P. E.; Denton, M. B. *Journal of Chemical Education* **1986**, 63, 617.
- (91) Somogyi, Á. In *Medical Applications of Mass Spectrometry*, Vékey, K.; Telekes, A.; Vertes, A., Eds.; Elsevier: Amsterdam, 2008, pp 93-140.
- (92) Henchman, M.; Steel, C. *Journal of Chemical Education* **1998**, 75, 1049.
- (93) Glish, G. L.; Burinsky, D. J. *Journal of the American Society for Mass Spectrometry* **2008**, 19, 161-172.
- (94) Griffiths, W. J.; Wang, Y. *Chemical Society Reviews* **2009**, 38, 1882-1896.
- (95) Shinholt, D. L.; Anthony, S. N.; Alexander, A. W.; Draper, B. E.; Jarrold, M. F. *Review of Scientific Instruments* **2014**, 85, 113109.
- (96) Cotter, R. J. *Time-of-flight mass spectrometry : instrumentation and applications in biological research; Washington (D.C.) : American chemical society*, 1997.
- (97) Chernushevich, I. V.; Loboda, A. V.; Thomson, B. A. *Journal of Mass Spectrometry* **2001**, 36, 849-865.
- (98) Juhasz, P.; Vestal, M. L.; Martin, S. A. *Journal of the American Society for Mass Spectrometry* **1997**, 8, 209-217.
- (99) Mamyrin, B. A.; Karataev, V. I.; Shmikk, D. V.; Zagulin, V. A. *Zhurnal Eksperimental'noj i Teoreticheskoy Fiziki* **1973**, 64, 82-89.
- (100) Vestal, M. L.; Juhasz, P.; Martin, S. A. *Rapid Communications in Mass Spectrometry* **1995**, 9, 1044-1050.
- (101) Wiley, W. C.; McLaren, I. H. *Review of Scientific Instruments* **1955**, 26, 1150-1157.
- (102) Andrews, G. L.; Simons, B. L.; Young, J. B.; Hawkridge, A. M.; Muddiman, D. C. *Analytical Chemistry* **2011**, 83, 5442-5446.
- (103) Dawson, J. H. J.; Guilhaus, M. *Rapid Communications in Mass Spectrometry* **1989**, 3, 155-159.
- (104) Guilhaus, M.; Selby, D.; Mlynski, V. *Mass Spectrometry Reviews* **2000**, 19, 65-107.
- (105) Roemmelt, A. T.; Steuer, A. E.; Kraemer, T. *Analytical Chemistry* **2015**, 87, 9294-9301.
- (106) Siegel, D.; Meinema, A. C.; Permentier, H.; Hopfgartner, G.; Bischoff, R. *Analytical Chemistry* **2014**, 86, 5089-5100.
- (107) Koppenaar, D. W.; Barinaga, C. J.; Denton, M. B.; Sperline, R. P.; Hieftje, G. M.; Schilling, G. D.; Andrade, F. J.; Barnes, J. H.; Iv, I. V. *Analytical Chemistry* **2005**, 77, 418 A-427 A.
- (108) Collins, R. D. *Vacuum* **1969**, 19, 105-111.

- (109) Cui, K.; Li, X.; Liu, Z.; Zhu, R. *IEEE Transactions on Radiation and Plasma Medical Sciences* **2017**, *1*, 391-399.
- (110) Zhong, F.; Liu, S.; Jin, W.; Simmons, D.; Bloomfield, N.: *Poster Contribution ASMS 2018*, 2018.
- (111) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W. M.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi, R.; Kopka, J.; Lane, A. N.; Lindon, J. C.; Marriott, P.; Nicholls, A. W.; Reilly, M. D., et al. *Metabolomics* **2007**, *3*, 211-221.
- (112) Viant, M. R.; Kurland, I. J.; Jones, M. R.; Dunn, W. B. *Current Opinion in Chemical Biology* **2017**, *36*, 64-69.
- (113) Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J. *Environmental Science & Technology* **2014**, *48*, 2097-2098.
- (114) Blazenovic, I.; Kind, T.; Ji, J.; Fiehn, O. *Metabolites* **2018**, *8*.
- (115) Cajka, T.; Fiehn, O. *Analytical Chemistry* **2016**, *88*, 524-545.
- (116) Hermes, N.; Jewell, K. S.; Wick, A.; Ternes, T. A. *J Chromatogr A* **2018**, *1531*, 64-73.
- (117) Laiko, V. V.; Dodonov, A. F.; Cotter, R. J. *Rapid Communications in Mass Spectrometry* **1994**, *8*, 720-726.
- (118) Chernushevich, I. V. *European Journal of Mass Spectrometry* **2000**, *6*, 471-479.
- (119) Fitzgerald, R. L.; Rivera, J. D.; Herold, D. A. *Clinical Chemistry* **1999**, *45*, 1224.
- (120) Decaestecker, T. N.; Clauwaert, K. M.; Van Bocxlaer, J. F.; Lambert, W. E.; Van den Eeckhout, E. G.; Van Peteghem, C. H.; De Leenheer, A. P. *Rapid Communications in Mass Spectrometry* **2000**, *14*, 1787-1792.
- (121) Koopmans, F.; Ho, J. T. C.; Smit, A. B.; Li, K. W. *PROTEOMICS* **2018**, *18*, 1700304.
- (122) Hopfgartner, G.; Tonoli, D.; Varesio, E. *Anal Bioanal Chem* **2012**, *402*, 2587-2596.
- (123) Plumb, R. S.; Johnson, K. A.; Rainville, P.; Smith, B. W.; Wilson, I. D.; Castro-Perez, J. M.; Nicholson, J. K. *Rapid Communications in Mass Spectrometry* **2006**, *20*, 1989-1994.
- (124) Xu, G.; Stupak, J.; Yang, L.; Hu, L.; Guo, B.; Li, J. *Rapid Communications in Mass Spectrometry* **2018**, *32*, 763-774.
- (125) Simons, B.; Kauhanen, D.; Sylvanne, T.; Tarasov, K.; Duchoslav, E.; Ekroos, K. *Metabolites* **2012**, *2*, 195-213.
- (126) Venable, J. D.; Dong, M.-Q.; Wohlschlegel, J.; Dillin, A.; Yates, J. R. *Nature Methods* **2004**, *1*, 39-45.
- (127) Gillet, L. C.; Navarro, P.; Tate, S.; Röst, H.; Selevsek, N.; Reiter, L.; Bonner, R.; Aebersold, R. *Molecular & Cellular Proteomics* **2012**, *11*, O111.016717.
- (128) Álvarez-Ruiz, R.; Picó, Y. *J Chromatogr A* **2019**, *1595*, 81-90.
- (129) Whitman, J. D.; Lynch, K. L. *Clinical Chemistry* **2019**, *65*, 862.
- (130) Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; VanderGheynst, J.; Fiehn, O.; Arita, M. *Nature Methods* **2015**, *12*, 523.
- (131) Zhang, Y.; Bilbao, A.; Bruderer, T.; Luban, J.; Strambio-De-Castillia, C.; Lisacek, F.; Hopfgartner, G.; Varesio, E. *Journal of Proteome Research* **2015**, *14*, 4359-4371.
- (132) Schlotterbeck, J.; Chatterjee, M.; Gawaz, M.; Lämmerhofer, M. *Analytica Chimica Acta* **2019**, *1046*, 1-15.
- (133) Raetz, M.; Duchoslav, E.; Bonner, R.; Hopfgartner, G. *Anal Bioanal Chem* **2019**.
- (134) Ivosev, G.; Cox, D. M.; Bloomfield, N.; Wasim, F.; LeBlanc, Y.: *Poster Contribution ASMS 2018*, 2018.
- (135) Collins, B. C.; Hunter, C. L.; Liu, Y.; Schilling, B.; Rosenberger, G.; Bader, S. L.; Chan, D. W.; Gibson, B. W.; Gingras, A.-C.; Held, J. M.; Hirayama-Kurogi, M.; Hou, G.; Krisp, C.; Larsen, B.; Lin, L.; Liu, S.; Molloy, M. P.; Moritz, R. L.; Ohtsuki, S.; Schlapbach, R., et al. *Nature Communications* **2017**, *8*, 291.
- (136) Cech, N. B.; Enke, C. G. *Mass Spectrometry Reviews* **2001**, *20*, 362-387.
- (137) Han, X. In *Lipidomics*, 2016, pp 121-150.
- (138) Smith, R.; Mathis, A. D.; Ventura, D.; Prince, J. T. *BMC Bioinformatics* **2014**, *15*, S9.
- (139) Mahieu, N. G.; Patti, G. J. *Analytical Chemistry* **2017**, *89*, 10397-10406.
- (140) Wang, L.; Xing, X.; Chen, L.; Yang, L.; Su, X.; Rabitz, H.; Lu, W.; Rabinowitz, J. D. *Analytical Chemistry* **2019**, *91*, 1838-1846.
- (141) Rafiei, A.; Sleno, L. *Rapid Communications in Mass Spectrometry* **2015**, *29*, 119-127.

- (142) Felinger, A.; Pap, T. L.; Inczédy, J. *Analytica Chimica Acta* **1991**, *248*, 441-446.
- (143) Lange, E.; Tautenhahn, R.; Neumann, S.; Gröpl, C. *BMC Bioinformatics* **2008**, *9*, 375.
- (144) Li, B.; Tang, J.; Yang, Q.; Li, S.; Cui, X.; Li, Y.; Chen, Y.; Xue, W.; Li, X.; Zhu, F. *Nucleic Acids Research* **2017**, *45*, W162-W170.
- (145) Vinaixa, M.; Schymanski, E. L.; Neumann, S.; Navarro, M.; Salek, R. M.; Yanes, O. *TrAC Trends in Analytical Chemistry* **2016**, *78*, 23-35.
- (146) Bolton, E. E.; Wang, Y.; Thiessen, P. A.; Bryant, S. H. In *Annual Reports in Computational Chemistry*, Wheeler, R. A.; Spellmeyer, D. C., Eds.; Elsevier, 2008, pp 217-241.
- (147) Pence, H. E.; Williams, A. *Journal of Chemical Education* **2010**, *87*, 1123-1124.
- (148) Kanehisa, M.; Goto, S.; Hattori, M.; Aoki-Kinoshita, K. F.; Itoh, M.; Kawashima, S.; Katayama, T.; Araki, M.; Hirakawa, M. *Nucleic Acids Research* **2006**, *34*, D354-D357.
- (149) Caspi, R.; Foerster, H.; Fulcher, C. A.; Kaipa, P.; Krummenacker, M.; Latendresse, M.; Paley, S.; Rhee, S. Y.; Shearer, A. G.; Tissier, C.; Walk, T. C.; Zhang, P.; Karp, P. D. *Nucleic Acids Research* **2007**, *36*, D623-D631.
- (150) Wishart, D. S.; Jewison, T.; Guo, A. C.; Wilson, M.; Knox, C.; Liu, Y.; Djoumbou, Y.; Mandal, R.; Aziat, F.; Dong, E.; Bouatra, S.; Sinelnikov, I.; Arndt, D.; Xia, J.; Liu, P.; Yallou, F.; Bjorn Dahl, T.; Perez-Pineiro, R.; Eisner, R.; Allen, F., et al. *Nucleic Acids Research* **2012**, *41*, D801-D807.
- (151) Degtyarenko, K.; de Matos, P.; Ennis, M.; Hastings, J.; Zbinden, M.; McNaught, A.; Alcántara, R.; Darsow, M.; Guedj, M.; Ashburner, M. *Nucleic Acids Research* **2007**, *36*, D344-D350.
- (152)
- (153) Guijas, C.; Montenegro-Burke, J. R.; Domingo-Almenara, X.; Palermo, A.; Warth, B.; Hermann, G.; Koellensperger, G.; Huan, T.; Uritboonthai, W.; Aisporna, A. E.; Wolan, D. W.; Spilker, M. E.; Benton, H. P.; Siuzdak, G. *Analytical Chemistry* **2018**, *90*, 3156-3164.
- (154) Koelmel, J. P.; Ulmer, C. Z.; Jones, C. M.; Yost, R. A.; Bowden, J. A. *Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids* **2017**, *1862*, 766-770.
- (155) Aicheler, F.; Li, J.; Hoene, M.; Lehmann, R.; Xu, G.; Kohlbacher, O. *Analytical Chemistry* **2015**, *87*, 7698-7704.
- (156) Miller, J.; Miller, J. C. *Statistics and chemometrics for analytical chemistry*; Pearson education, 2018.
- (157) Checa, A.; Bedia, C.; Jaumot, J. *Analytica Chimica Acta* **2015**, *885*, 1-16.
- (158) Bland, J. M.; Altman, D. G. *BMJ* **1995**, *310*, 170.
- (159) Benjamini, Y.; Hochberg, Y. *Journal of the Royal Statistical Society. Series B (Methodological)* **1995**, *57*, 289-300.
- (160) Storey, J. D. *Ann. Statist.* **2003**, *31*, 2013-2035.
- (161) Shah, V. P.; Midha, K. K.; Dighe, S.; McGilveray, I. J.; Skelly, J. P.; Yacobi, A.; Layloff, T.; Viswanathan, C. T.; Cook, C. E.; McDowall, R. D.; Pittman, K. A.; Spector, S.; Albert, K. S.; Bolton, S.; Cook, C. E.; Dighe, S.; Dobrinska, M.; Doub, W.; Eichelbaum, M.; Findlay, J. W. A., et al. *International Journal of Pharmaceutics* **1992**, *82*, 1-7.
- (162) Shah, V. P.; Midha, K. K.; Findlay, J. W. A.; Hill, H. M.; Hulse, J. D.; McGilveray, I. J.; McKay, G.; Miller, K. J.; Patnaik, R. N.; Powell, M. L.; Tonelli, A.; Viswanathan, C. T.; Yacobi, A. *Pharmaceutical Research* **2000**, *17*, 1551-1557.
- (163) Booth, B.; Arnold, M. E.; DeSilva, B.; Amaravadi, L.; Dudal, S.; Fluhler, E.; Gorovits, B.; Haidar, S. H.; Kadavil, J.; Lowes, S.; Nicholson, R.; Rock, M.; Skelly, M.; Stevenson, L.; Subramaniam, S.; Weiner, R.; Woolf, E. *The AAPS Journal* **2015**, *17*, 277-288.
- (164) *Bioanalytical Method Validation - Guidance for Industry*, US Food and Drug Administration, 2018.
- (165) Lowes, S.; Jersey, J.; Shoup, R.; Garofolo, F.; Savoie, N.; Mortz, E.; Needham, S.; Caturla, M. C.; Steffen, R.; Sheldon, C.; Hayes, R.; Samuels, T.; Di Donato, L.; Kamerud, J.; Michael, S.; Lin, Z.; Hillier, J.; Moussallie, M.; de Souza Teixeira, L.; Rocci, M., et al. *Bioanalysis* **2011**, *3*, 1323-1332.
- (166) Matuszewski, B. K.; Constanzer, M. L.; Chavez-Eng, C. M. *Analytical Chemistry* **2003**, *75*, 3019-3030.
- (167) Broadhurst, D.; Goodacre, R.; Reinke, S. N.; Kuligowski, J.; Wilson, I. D.; Lewis, M. R.; Dunn, W. B. *Metabolomics* **2018**, *14*, 72.
- (168) Dudzik, D.; Barbas-Bernardos, C.; Garcia, A.; Barbas, C. *Journal of Pharmaceutical and Biomedical Analysis* **2018**, *147*, 149-173.
- (169) Drotleff, B.; Lämmerhofer, M. *Analytical Chemistry* **2019**.
- (170) Cajka, T.; Smilowitz, J. T.; Fiehn, O. *Analytical Chemistry* **2017**, *89*, 12360-12368.

- (171) Cui, L.; Lu, H.; Lee, Y. H. *Mass Spectrometry Reviews* **2018**, *37*, 772-792.
- (172) Bader, M. *Journal of Chemical Education* **1980**, *57*, 703.
- (173) van de Merbel, N. C. *TrAC Trends in Analytical Chemistry* **2008**, *27*, 924-933.
- (174) Thakare, R.; Chhonker, Y. S.; Gautam, N.; Alamoudi, J. A.; Alnouti, Y. *Journal of Pharmaceutical and Biomedical Analysis* **2016**, *128*, 426-437.
- (175) Li, W.; Cohen, L. H. *Analytical Chemistry* **2003**, *75*, 5854-5859.
- (176) Jones, B. R.; Schultz, G. A.; Eckstein, J. A.; Ackermann, B. L. *Bioanalysis* **2012**, *4*, 2343-2356.
- (177) Wakamatsu, A.; Ochiai, S.; Suzuki, E.; Yokota, Y.; Ochiai, M.; Kotani, Y.; Sasahara, S.; Nakanaga, K.; Hashimoto, Y.; Ueno, S.; Kato, N.; Kawada, S.; Hayakawa, J.; Shimada, E.; Horita, S.; Sakai, K. *Bioanalysis* **2018**, *10*, 1349-1360.
- (178) MacNeill, R.; Sangster, T.; Moussallie, M.; Trinh, V.; Stromeyer, R.; Daley, E. *Bioanalysis* **2009**, *2*, 69-80.

2. List of Figures

Figure 1. Published research in metabolomics and lipidomics.....	1
Figure 2. Distribution and exemplary structures of biological lipid categories.....	2
Figure 3. Comparison of the lipid extraction profiles of typically used organic solvent systems.	4
Figure 4. Summary of common workflows.	7
Figure 5. Formation of the Taylor cone.	9
Figure 6. Voltage profile of the rods in a quadrupole analyzer.	11
Figure 7. Stability diagram of a quadrupole mass analyzer.....	12
Figure 8. Design of a QTOF instrument (TripleTOF 5600).....	14
Figure 9. Scheme of the spectrum acquisition with a TDC.	16
Figure 10. Principle of the high sensitivity and high resolution mode.	19
Figure 11. MS/MS ^{ALL} acquisition.	20
Figure 12. Exemplary scheme of a SWATH enabled acquisition cycle.	21
Figure 13. Influence of time bins to sum parameter on spectral resolution.....	25
Figure 14. Standard addition.	33
Figure 15. Compression of monoisotopic peaks.	35

3. List of Tables

Table 1. Performance parameters of exemplary HR hybrid MS instruments.....	10
Table 2. Operational ranges for optimization parameters.	24
Table 3. Selection of compound and mass spectral databases.	27
Table 4. Validation requirements and acceptance criteria according to FDA guidelines.....	31

4. Objectives of the Thesis

The aim of this thesis was to develop new methods and optimize existing workflows for lipidomics to support clinical research. On the one hand, the focus was to improve untargeted assays in general, which allow broad lipid profiling of hundreds to thousands of lipids per sample. On the other hand, a progressive analysis of specific, demanding lipid classes such as steroids was approached.

Steroid hormones are one of the most frequently monitored analytes during clinical studies. They are often by default quantified via immunoassays, which were reported to yield discrepant results, presumably due to their susceptibility to matrix effects and cross-reactivities. In consequence, researchers head for MS determination of steroids, which represent a challenging analyte group due to their partial low abundance in biological matrices. Accordingly, the first aim of this thesis was to develop a sensitive and robust LC-MS/MS method for absolute quantification of estradiol and testosterone in clinical plasma samples. However, instead of a standard LC-QqQ method, a merged targeted/untargeted method by UHPLC-QTOF analysis with SWATH acquisition was pursued. Nevertheless, a QqQ-like performance in terms of sensitivity, accuracy and precision had to be demonstrated to pass the strict validation requirements in bioanalysis. To reach the set goals, an advanced sample preparation protocol for maximizing recovery, purification and analyte enrichment was needed. As alternative calibration was required due to the lack of steroid-free plasma, the eligibility of matrix-matched surrogate calibration was evaluated. Eventually, in order to enhance the extractable information for the valuable study samples, simultaneously acquired TOF-MS data was screened for additional steroids to exhibit the full potential of QTOF analysis.

In targeted assays, normalization of analyte signals, which typically incorporate various sources of variation, is achieved by addition of suitable ISs and the calculation of response ratios. For untargeted data, which captures an exceeding number of detected features and comprises many unknown compounds, normalization is consequentially aggravated. Nevertheless, critical datasets that reveal only few significant differences of minor extent, strongly depend on accurate normalization, which must be able to reduce unwanted variance in order to detect robust, true positive findings. Various normalization strategies already exist, yet, no harmonized guidelines on how to select the best performing method without sacrificing data integrity are available. It was therefore aspired to gather defined requirements that enable a rational decision-making model for the respective selection of optimum normalization methodologies. The validity of these novel recommendations had to be presented via an appropriate, exemplary dataset.

Although an effective normalization is capable to drastically improve the data quality, the ultimate goal of untargeted acquisition should be absolute quantification of all or of as many compounds as possible during one analytical sequence. This achievement would enable an optimum comparability of results between different studies, matrices, disease states and so forth, which is currently not given with relative quantification and foldchanges. However, absolute quantification of unknown or uncalibrated targets is not trivial and requires additional efforts in study design and data processing. A potential strategy towards class-specific quantification via surrogate calibration with lipid class representatives was developed and investigated. Moreover, although no guidelines for the validation of such strategies exist, a fit-for-purpose verification of the methodology had to be performed.

5. Results and Discussion

5.1. Quantification of Steroid Hormones in Plasma Using a Surrogate Calibrant Approach and UHPLC-ESI-QTOF-MS/MS with SWATH-Acquisition Combined with Untargeted Profiling

Bernhard Drotleff[†], Manfred Hallschmid^{§,⊥,#}, Michael Lämmerhofer[†]

[†]Institute of Pharmaceutical Sciences, Pharmaceutical (Bio-)Analysis, University of Tübingen, Tübingen, Germany

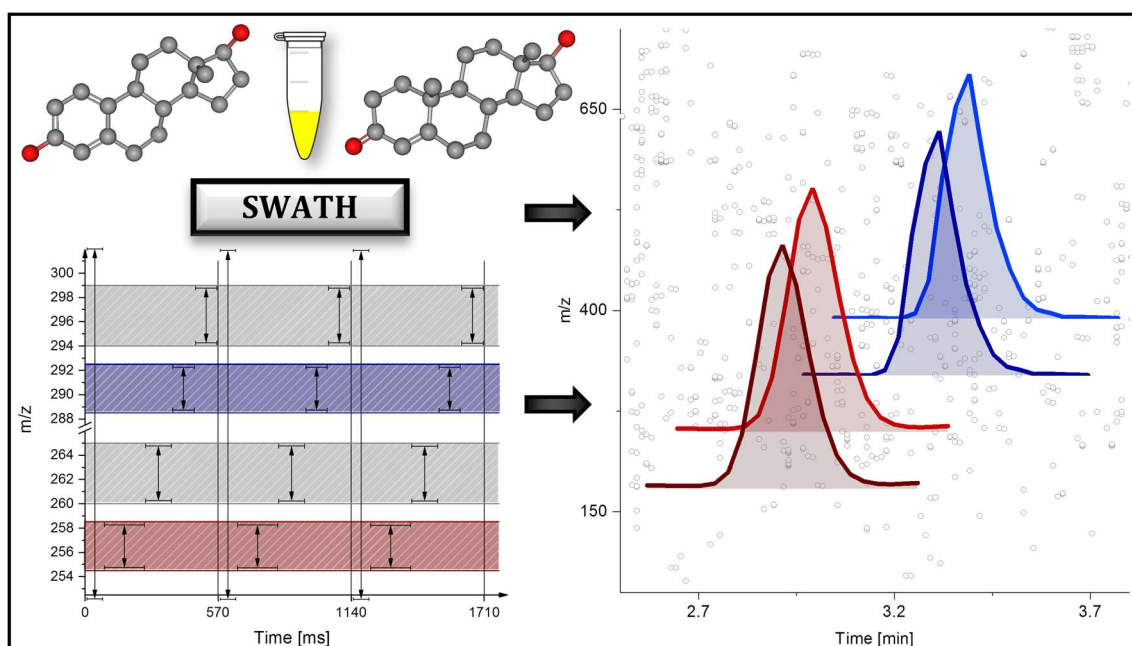
[§]Department of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Tübingen, Germany

[⊥]German Center for Diabetes Research (DZD), Munich-Neuherberg, Germany

[#]Institute for Diabetes Research and Metabolic Diseases of the Helmholtz Center Munich at the University of Tübingen (IDM), Tübingen, Germany

Reprinted with permission from *Analytica Chimica Acta*, Volume 1022, Pages 70-80, DOI: 10.1016/j.aca.2018.03.040

Copyright (2018) Elsevier B.V.



Graphical Abstract

5.1.1. Abstract

In spite of demonstrated lack of accuracy and consistency, quantification of steroid hormones is still most commonly executed via immunoassays. Mass spectrometric methods with triple quadrupole instruments are well established and, because of their proven robustness and sensitivity, best suited for targeted analysis. However, recent studies have shown that high-resolution mass spectrometers, like quadrupole time-of-flight instruments (QTOF), show comparable performance in terms of quantification and can generate additional sample information via untargeted profiling workflows. We demonstrate that adequate accuracy and selectivity for estradiol and testosterone can be achieved with a QTOF by data-independent acquisition with sequential window acquisition of all theoretical fragment-ion mass spectra (SWATH). Besides potential combination of targeted quantification and untargeted profiling, SWATH offers advantages with respect to sensitivity because the reduced total number of MS/MS experiments could be used to increase accumulation time without increasing cycle time. By applying a surrogate calibrant method leading to successful validation, a reliable method for absolute steroid quantification and high potential for steroid profiling has been developed. Linear calibration was achieved in the range from 10 - 1,000 pg mL⁻¹ for ¹³C₃-estradiol and from 20 - 15,000 pg mL⁻¹ for ¹³C₃-testosterone. Results for inter-day precision (¹³C₃-estradiol: 4.5 - 10.2 %; ¹³C₃-testosterone: 5.1 - 7.8 %) and inter-day accuracy (¹³C₃-estradiol: 94.6 - 112.8 %; ¹³C₃-testosterone: 98.2 - 107.7 %) were found to be well acceptable. Eventually, the method has been utilized to measure clinical samples of a study in which male volunteers obtained transdermal estradiol patches and sex hormone levels were quantified in plasma.

5.1.2. Introduction

17 β -estradiol (E) and 17 β -testosterone (T), the main steroid sex hormones in women and men, play crucial roles in human physiology and are frequently monitored analytes in routine diagnostics and clinical studies.¹ Despite the well-known disadvantages, like impact of matrix effects and cross-reactivities,² the majority of steroid analytics is still performed via immunoassays. Numerous studies have already shown inconsistency between assay results, especially in critical patient groups with low steroid levels.³⁻⁹ Accurate results, however, are mandatory for effective therapy and study interpretation. Consequently, the demand for reliable techniques, in particular liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS), is emerging in clinical analysis and clinical studies.¹⁰

Another challenge of steroid quantification in plasma is the absence of true blank matrix for calibration and assessment of assay selectivity. To overcome this problem, various alternative methods are described.¹¹ In order to obtain an authentic analytical environment, a surrogate calibrant approach¹² was selected for this method. Herein, calibration is done via an analyte-related substance, preferably a stable-isotope-labeled analogue (SIL), which is spiked into the true matrix. After initial matching of SIL response to target analyte response and verification of parallelism,¹³ the surrogate calibration is used for sample quantification.

The goal of this study was to develop and validate a sensitive LC-MS/MS method for the quantification of E and T in human plasma to verify and complement results previously gathered by a competitive chemiluminescent enzyme immunoassay. A large number of quantitative assays using LC hyphenated to triple quadrupole (QqQ) instruments were already published for these steroid hormones.^{10, 14-26} To reach low concentration levels of target analytes in various matrices, pre-column derivatization is often carried out, using e.g. Girard-P,^{27,28} dansyl chloride,²⁷ aminoxypropyl trimethylammonium bromide²⁹ (Amplifex Keto) for ketolic steroids such as T, and dansyl chloride²⁷ or 1,2-dimethylimidazole-5-sulfonyl chloride²¹ for phenolic steroids such as E. Due to robustness, high sensitivity and wide linear range, LC-ESI-QqQ is the method of first choice for targeted quantitative analysis of steroid hormones. Recently, however, quantification by LC coupled to high-resolution (HR) MS raised some interest due to good performance.³⁰⁻³² Usually, quantitative data with such HR-MS instruments (quadrupole/time-of-flight or quadrupole/orbitrap) are acquired in MRM^{HR} (also called parallel reaction monitoring, PRM) or data-dependent acquisition (DDA).^{33,34} In former acquisition mode, after a full scan MS experiment (survey scan) MS/MS experiments are programmed for the selected targets whereby precursor selection occurs by a quadrupole mass analyzer with unit mass followed by fragmentation and analysis of the product ions in the HR-mass analyzer. Highly selective MS/MS chromatograms can be extracted for the programmed targets (i.e. EICs for fragment ions of the selected precursors), while untargeted profiling is still possible at

the MS level.^{35,36} In DDA, subsequent to the full scan MS experiment, a series of MS/MS experiments, in which the most intensive precursor ions detected in the survey scan are fragmented, is carried out. Thus, MS/MS data are not collected comprehensively across the entire chromatogram and all study samples. The consequence is that quantitative analysis can be only performed with the precursor ion from the MS experiment. This restriction can be overcome by untargeted profiling with data-independent acquisition (DIA). In DIA, MS/MS fragmentation occurs without dependence on information from the survey scan. All precursors of the entire m/z range co-isolated by the quadrupole are co-fragmented simultaneously (termed MS^E, all ion fragmentation).³⁷ This yields complex composite spectra, which is the reason why this acquisition mode has not become very popular. However, precursor selection can also be performed in a stepped manner with sequential, intermediate-sized Q1 windows (e.g. 20 - 50 Da), thus covering the entire m/z range of interest. This acquisition mode has been developed for proteomics³⁸ but has been recently tested for small molecules as well, including metabolomics and lipidomics.³⁹⁻⁴⁵ Better performance than with DDA has been documented for this DIA called SWATH (sequential window acquisition of all theoretical fragment-ion mass spectra) due to better analyte (metabolite) coverage, better reproducibility, and less complex composite spectra.⁴³ Moreover, comprehensive MS/MS data are available and can be used for quantitative analysis. The application of a QTOF with SWATH acquisition for quantitative purposes has recently shown promising results.^{41,46} Here, we wanted to utilize the advantageous properties SWATH offers in terms of sensitivity, especially when surrogate calibration is used. In contrary to previous works, fully optimized SWATH experiments for generating specific and sensitive MS/MS fragment ion signals for quantification of target analytes without derivatization was established.

Concluding, we demonstrate the performance of UHPLC-ESI-QTOF-MS/MS analysis by DIA with SWATH for the simultaneous targeted quantitative analysis of E and T in human plasma samples from a clinical study in which male subjects were treated with transdermal E patches. Extension of the method to a combined targeted/untargeted profiling method is illustrated as well. Furthermore, reliable quantification based on peak areas of extracted MS/MS chromatograms of characteristic fragment ions in SWATH experiments is demonstrated.

5.1.3. Experimental Section

Materials

T, 17 β -testosterone-2,3,4-¹³C₃ (¹³C₃T, 100 μ g mL⁻¹ in methanol), E, 17 β -estradiol-2,3,4-¹³C₃ (¹³C₃E), 17 α -estradiol (epiestradiol, epiE) and phosphoric acid (85 %, w/v, ACS grade) were purchased from Sigma-Aldrich (Saint Louis, MO, USA). 17 β -testosterone-2,2,4,6,6-²H₅ (*d*₅T, 106.7 μ g mL⁻¹ in methanol) was purchased from IsoSciences (King of Prussia, PA, USA). 17 β -estradiol-2,4,16,16,17-²H₅ (*d*₅E, 100 μ g mL⁻¹ in acetonitrile) and 17 α -testosterone (epiT; 1.0 mg mL⁻¹ in acetonitrile) were purchased from Cerilliant (Round Rock, TX, USA). Details on standard solutions, (surrogate-) calibrants and quality controls can be found in supplementary data (Appendix A.). Cortisone and cortisol were purchased from Cayman Chemical (Ann Arbor, MI, USA). Type I purity water was obtained from a Purelab Ultra purification system (ELGA LabWater, Celle, Germany). Immunoassay measurements of study samples were done with an Immulite 2000 system (Siemens Diagnostics, Erlangen, Germany) using complying E and T kits for total quantification.

Immunoassay

In this fully automated, competitive chemiluminescent enzyme immunoassay the solid phase consist of beads coated with rabbit polyclonal antibodies specific for the respective target analyte. After introduction of the sample (T: 20 μ L; E: 25 μ L) and alkaline-phosphate conjugated with E or T, respectively, the target analytes compete with the analyte-enzyme complexes for the limited binding sites during an incubation period of 60 minutes. After washing to remove excess material and reagents, a chemiluminescent substrate (adamantly dioxetane phosphate ester) is added. Hydrolyzation of the substrate by alkaline phosphatase yields unstable anions, which, as a result of decomposition, generate constant emission of photons. Accordingly, light intensity is inversely proportional to target analyte concentration in the sample. Lyophilized serum quality controls (MassCheck Steroid Panel 2, tri-level) were purchased from Chromsystems (Graefelfing, Germany). Subjects providing blood samples gave written informed consent to the study that conformed to the Declaration of Helsinki as revised in 2008 and was approved by the local Ethics Committee on Research Involving Humans.

Sample Preparation

500 μ L of EDTA plasma were diluted with 500 μ L of 5 % H₃PO₄ (w/v) that contained 1.0 ng mL⁻¹ of *d*₅E and 0.4 ng mL⁻¹ of *d*₅T as internal standards (IS). After vortexing, the sample was loaded onto a dry Oasis PRiME HLB SPE cartridge (1 cc / 30 mg, Waters, Milford, MA, USA). Samples were processed applying negative pressure with a Vacmaster 20 manifold (Biotage,

Uppsala, Sweden). After the first loading step, the cartridges were washed with 1 mL of 50 % MeOH in H₂O (v/v). Analytes were then eluted with 2 × 500 µL MeOH and the eluate was dried using a Savant ISS110 SpeedVac concentrator (Thermo Fisher Scientific, Waltham, MA, USA). After reconstitution in 100 µL MeOH, samples were centrifuged for 5 min at 15,000 × g and 4 °C with a 5415R microcentrifuge (Eppendorf, Hamburg, Germany). The supernatant was transferred into a vial, which was crimped and stored at 4 °C in the autosampler. Samples were analyzed as soon as possible after preparation.

LC-Method

The chromatographic system consisted of a 1290 Infinity UHPLC system (Agilent Technologies, Waldbronn, Germany) and a PAL HTC-xt autosampler (CTC Analytics, Zwingen, Switzerland). Separation was performed on a Kinetex C18 column (50 mm × 2.1 mm, 2.6 µm, 100 Å pore size) with a KrudKatcher Ultra in-line filter (Phenomenex, Aschaffenburg, Germany) for column protection. Mobile phase A consisted of H₂O + 0.1 % formic acid (v/v) and mobile phase B of MeCN + 0.1 % formic acid (v/v). The flow rate was 0.3 mL min⁻¹ with a constant oven temperature of 30 °C. Injection volume was set to 10 µL. The following gradient was applied: 5 - 30 % B from 0.0 – 0.5 min, 30 – 45 % B from 0.5 – 3.2 min, 45 – 95 % B from 3.2 – 3.5 min, holding 95 % B from 3.5 – 4.0 min, 95 – 5 % B from 4.0 – 4.2 min, equilibration with 5 % B from 4.2 – 5.0 min.

MS-Method

Mass spectrometric detection was performed on a TripleTOF 5600+ mass spectrometer with a DuoSpray source (Sciex, Concord, Ontario, Canada). Optimized ion source parameters were as follows: curtain gas (N₂) 35 psi; nebulizer gas (N₂) 50 psi; heater gas (N₂) 80 psi, ion source voltage floating 4,000 V, source temperature 600 °C. Samples were measured in positive electrospray ionization (ESI) mode, running one TOF-MS experiment in the mass range of *m/z* 30 – 1,000 (survey scan; resolution ≥30,000, FWHM @ 829.5393 Da) and four SWATH-MS/MS experiments (resolution ≥15,000, FWHM @ 397.2122 Da) per cycle (method 1, see Table 1). Accumulation time (*t*_{Acc}) was set to the following values: TOF-MS scan: 20 ms; SWATH of T/¹³C₃T: 50 ms; SWATH of *d*₅T: 50 ms; SWATH of E/¹³C₃E: 300 ms; SWATH of *d*₅E: 100 ms. Total cycle time (*t*_{Cyc}) was delimited to 570 ms to attain at least ten data points per peak in regard to average peak widths of about 6 s. Enhanced product ion mode was enabled. For SWATH experiments of *d*₅-internal standards, enhancement was set to the monoisotopic mass of the used fragment, respectively. For SWATH experiments that covered two compounds, target analytes and surrogate calibrants, the enhancement mass was set to the calculated mean mass of both corresponding fragments.

Table 1. Overview of method-parameters for targeted analysis of E and T (method 1).^a

Analyte	Ion Species	Precursor [m/z]	Q1 window [m/z]	Fragment [m/z]	DP [V]	CE [V]	t _R [min]
E	[M-H ₂ O+H] ⁺	255.1743	254.5 - 258.5	159.0804	90	25	3.076 ± 0.011
¹³ C ₃ E	[M-H ₂ O+H] ⁺	258.1844		162.0905			3.076 ± 0.012
d ₆ -E	[M-H ₂ O+H] ⁺	260.2057	260.0 - 265.0	161.0930	90	25	3.044 ± 0.012
T	[M+H] ⁺	289.2162		109.0648			3.159 ± 0.009
¹³ C ₃ T	[M+H] ⁺	292.2263	288.5 - 292.5	112.0749	200	33	3.158 ± 0.009
d ₆ T	[M+H] ⁺	294.2476	294.0 - 299.0	113.0899	120	33	3.136 ± 0.010

^aDeclustering potential (DP), collision energy (CE), retention time (t_R). SWATH (Q1) windows covered both, target analyte and corresponding surrogate calibrant. Their specific fragments have a mass difference of 3 Da and do not show interferences. To create a SWATH-MS/MS experiment, also a TOF-MS experiment has to be performed in each cycle. This precededent TOF-MS run (30 - 1,000 Da) had a t_{acc} of 20 ms, CE of 10 V and a DP of 100 V.

Mass calibration was done via infusion of sodium acetate (0.1 mg mL⁻¹ in MeCN:H₂O, 1:1, v/v) every 25th injection. The whole analytical system was controlled by the Analyst 1.7 TF software (Sciex).

Data Analysis and Quantification.

Calibration curves were constructed using weighted least-square linear regression (weighting factor: 1/x) of six different calibrant levels by plotting peak area ratios of ¹³C₃E/d₅E and ¹³C₃T/d₅T against respective surrogate calibrant concentrations. The resulting equations were used to determine target analyte concentrations in real samples via E/d₅E and T/d₅T ratios, respectively. Two QCs, QC_{low} (¹³C₃T: 60 pg mL⁻¹; ¹³C₃E: 30 pg mL⁻¹) and QC_{high} (¹³C₃T: 12,000 pg mL⁻¹; ¹³C₃E: 800 pg mL⁻¹) were embedded after every 20th sample in the sequence to verify stable method performance. To control for accuracy and linearity of calibration, five determinations of the calibration were equally distributed across the whole sequence. Quantification was based on fragment ions (Table 1). Fragment peak areas were extracted using a ± 10 mDa mass window in the associated SWATH experiments. Automated integration with the MultiQuant 3.0 software (Sciex) was done using a MQIII algorithm, Gaussian smoothing (width: 2 data points), noise percentage of 90 %, baseline subtraction window of 0.1 min and a peak splitting factor of 2. Excel 2007 (Microsoft, Redmond, WA, USA), SPSS Statistics 23 (IBM, Armonk, NY, USA) and Origin 2017 (OriginLab, Northampton, MA, USA) were used for further data evaluation.

5.1.4. Results and Discussion

Sample Preparation

E and T are bound to plasma proteins like SHBG (sex hormone-binding globulin).⁴⁷ Their release by organic solvents used for protein precipitation would demand an evaporation step prior to SPE which is needed for E/T enrichment. Hence, 5 % H₃PO₄ was selected for protein precipitation⁴⁸⁻⁵¹ because the resultant supernatant could be directly loaded onto the Oasis PRiME HLB material, which does not require pre-conditioning and equilibration prior to the loading step. 50 % MeOH in H₂O (v/v) was selected as optimum washing eluent and complete analyte elution with good recoveries of E, ¹³C₃E, T and ¹³C₃T could be achieved with 2 × 500 μL MeOH. By drying and reconstitution in 100 μL MeOH, a total sample pre-concentration factor of 5 was achieved to reach sufficient levels of sensitivity (for details see Appendix A.).

LC-MS Method

A fast UHPLC method with gradient elution (5 min including re-equilibration) was developed using a core-shell C18 column (Kinetex® C18, 2.6 μm). Faster elution by higher flow rates was not considered because the detection sensitivity significantly dropped at flow rates higher than 0.3 mL min^{-1} .⁵² Close to baseline separation of E and T was achieved ($R_s = 0.98$) (Fig. A.3C) and in spite of a fast gradient sufficient assay specificity was ensured by selective mass spectrometric detection.

The low concentrations of E expected in male plasma samples required dedicated optimization of MS parameters to reach maximal sensitivity for E. For assessment of most sensitive conditions, ionization efficiencies of analytes were tested with APCI and ESI in positive and negative mode. Best sensitivity for E was achieved in negative APCI mode, but ionization of T was unacceptable in negative APCI and negative ESI. Accordingly, analysis in positive mode was mandatory since polarity switching in ms time scale is not possible for the TripleTOF 5600+. Whereas the $[\text{M}+\text{H}]^+$ -precursor ion could be detected for T, E only showed an in-source fragmentation product $[\text{M}-\text{H}_2\text{O}+\text{H}]^+$, which was selected as the precursor. For acquisition, data-independent acquisition mode using SWATH, a sequential window-based MS/MS acquisition methodology with intermediate Q1 precursor window sizes, was executed. It allows flexible adjustment and thus optimization of MS parameters for each SWATH window separately and leads to a comprehensive set of MS/MS data in the selected Q1 precursor windows. Since SWATH acquisition used parameters, which secured ≥ 10 spectra available across each peak, enough data points were available to enable generation of MS/MS chromatograms, i.e. EICs of fragments, with some advantages as described below (see also Fig. A.4). Activation of the enhanced product ion mode showed >3 times increase in signal intensities. This feature optimizes the ion pulsing process for a specific fragment and improves the duty cycle.⁵³ However, only a narrow m/z -region around the targeted fragment is enhanced by this process and ions outside this region are lost for detection and excluded. Because of this effect, precursor ions of analytes were not observed in the SWATH-MS/MS experiments in the present case (Fig. A.5).

Comparison of SWATH and MRM^{HR} Sensitivity by their Instrumental LODs

Instead of individual product ion MS/MS experiments with unit mass Q1 precursor selection (MRM^{HR}) for each analyte, SWATH-MS/MS experiments were created (Table 1). By selection of appropriate window sizes (4 Da for E, T and their corresponding $^{13}\text{C}_3$ analogues; 5 Da for the deuterated internal standards), fragments of target analytes and corresponding surrogate calibrants could be detected in the same SWATH window. Because of fragmentation interferences, separate SWATH windows had to be created for d_5 -analogues. Optimized window sizes assured assay specificity for the fragment ions used for quantification.

Sensitivity, on the other hand, generally increases with increasing accumulation time t_{Acc} (see Fig. A.6 and Fig. A.7).

SWATH acquisition allowed to reduce the total number of MS/MS experiments and allowed to distribute the maximally available t_{Acc} between fewer experiments. This enabled to increase t_{Acc} for each analyte as compared to MRM^{HR}. In order to compare the sensitivity of MRM^{HR} and SWATH, the instrumental limits of detection (LODs) were determined for three different methods: The SWATH method with the parameters described in section 2. and Table 1, an MRM^{HR} method with t_{Acc} equal to the SWATH experiment (i.e. 300 ms for E and 50 ms for T) (MRM_{eq}) and an MRM method with half the t_{Acc} (i.e. 150 ms for E and 25 ms for T) (MRM_{1/2}). The MRM_{1/2} method was designed since it represents the most realistic equivalent to the SWATH method as only half the t_{Acc} is available due to the double number of experiments if each of the analyte and ¹³C₃-calibrant is acquired by separate product ion MS/MS experiments. Besides t_{Acc} , all other mass spectrometric parameters (see LC-method subchapter and Table 1) were kept identical for each method to ensure best achievable comparability. Furthermore, all methods were run with enabled “enhanced product ion mode”, had identical cycle times t_{Cyc} and a uniform t_{Acc} (20 ms) for the mandatory TOF-MS experiment. To assess instrumental LODs in the low concentration range, an 8-point calibration of both target analytes in MeOH was analyzed in triplicate. Instrumental LODs were lowest for the SWATH method (5.8 and 8.1 pg mL⁻¹ for E and T, respectively; about factor 2-3 lower than with MRM^{HR} even at equal t_{Acc} ; see also Table A.7).

Assay Specificity

While SWATH was shown to increase sensitivity, specificity is lost owing to the broader Q1 isolation window. Validation therefore ultimately requires verification of sufficient assay specificity. First of all, possible interferences deriving from SILs have to be ruled out. The attempt to cover target analyte and corresponding ¹³C₃- and *d*₅-analogues in one single 8 Da-wide SWATH window, respectively, failed since interferences were observed both for E and T. Investigation showed that fragmentation of *d*₅-standards caused significant interference due to overlapping isotope patterns of *d*₅-fragments and ¹³C₃-fragments. Accordingly, a separate 5 Da SWATH window was created for analysis of each *d*₅-standard. Further optimization showed that two additional SWATH experiments of 4 Da width are adequate to cover corresponding pairs of target analytes and surrogate ¹³C₃-calibrants, respectively. Fragmentations in these windows were free of interference and showed sufficient specificity (see Fig. A.11 – A.16). In untargeted SWATH methods, windows are usually overlapping by 1 Da. In our targeted approach a gap of at least 1.5 Da had to remain between the SWATH windows to avoid interferences. This is owed to the fact that the Q1 is not capable of doing an exact cutout of *m/z* ranges. Also ions with an *m/z* slightly (~ 1 Da) below or above SWATH

window limits will pass through the Q1, which can lead to unwanted interference. Cross-validation via commercial quality controls has finally been utilized to verify assay specificity. Also, epiT and epiE, epimers of T and E with identical fragmentation, were analyzed and showed chromatographic baseline separation (epiT to T, Δt_R : 0.42 min; epiE to E, Δt_R : 0.27 min) (see Fig. A.10). Assay specificity (i.e. lack of interferences) of $^{13}\text{C}_3$ - and d_5 -standards was determined by analyzing six different blank plasma samples. No interfering peaks in a retention time window of ± 0.1 min of the respective analyte were detected.

Calibration and Limits of Quantification

With optimized conditions, both E and T could be detected with high sensitivity. Unfortunately, for T the signal leveled off at concentrations above $1,000 \text{ pg mL}^{-1}$ due to detector saturation. De-optimization, by raising DP from 120 to 200 V, led to a shift of the linearity range which then covered the relevant concentration range between 20 pg mL^{-1} (instrumental LLOQ) to the upper limit of quantification (ULOQ) of $15,000 \text{ pg mL}^{-1}$ (see also Fig. A.8).

Due to absence of blank matrix for matrix-matched calibration, a surrogate calibrant approach was adopted. To ensure accuracy of quantification via $^{13}\text{C}_3$ -surrogate calibrants, parallelism of the calibration curves between surrogate calibrants and the corresponding standard addition curve of the target analyte has to be verified.¹³ In the present case, the maximum difference of the slopes of T and $^{13}\text{C}_3\text{T}$ during three inter-day measurements was 3.7 % (slope of $^{13}\text{C}_3\text{T}$ divided by slope of T) and 3.2 % for E and $^{13}\text{C}_3\text{E}$ (slope of $^{13}\text{C}_3\text{E}$ divided by slope of E) (see Fig. A.18). Therefore, $^{13}\text{C}_3\text{T}$ and $^{13}\text{C}_3\text{E}$ have been found to be adequate surrogate calibrants for quantitative analysis of T and E in human plasma.

LLOQs in real samples were determined adopting the criteria set forth by the FDA guideline for bioanalytical method validation (analyte response at least 5 times the response of the blank response, precision of 20 % and accuracy of 80 - 120 %). Thus, 10 pg mL^{-1} for E and $^{13}\text{C}_3\text{E}$, and 20 pg mL^{-1} for T and $^{13}\text{C}_3\text{T}$ were set as LLOQs in real samples (Fig. 1). During validation these values were shown to meet the acceptance criteria for LLOQs.

Method Validation

Method validation was performed on the basis of the FDA guideline on bioanalytical method validation with minor modifications (e.g. 5 replicates over 3 independent days instead of one replicate over 5 independent days for assessment of inter-day accuracy and precision). The detailed results can be found in Appendix A. Matrix effect (ME), extraction recovery (RE) and process efficiency (PE) were evaluated according to Matuszewski et. al⁵⁴ with $^{13}\text{C}_3$ -labelled analogues of analytes, which are expected to suffer equally from ME as the coeluted target analytes. The results are shown in Table 2. E shows an average ME (ion suppression) of 81.3 % and T a more significant average ME of 60.4 %, which made its correction by internal

standards (here d_5 -analogues) mandatory. Average recoveries for E and T were 88.0 and 84.4 %, respectively.

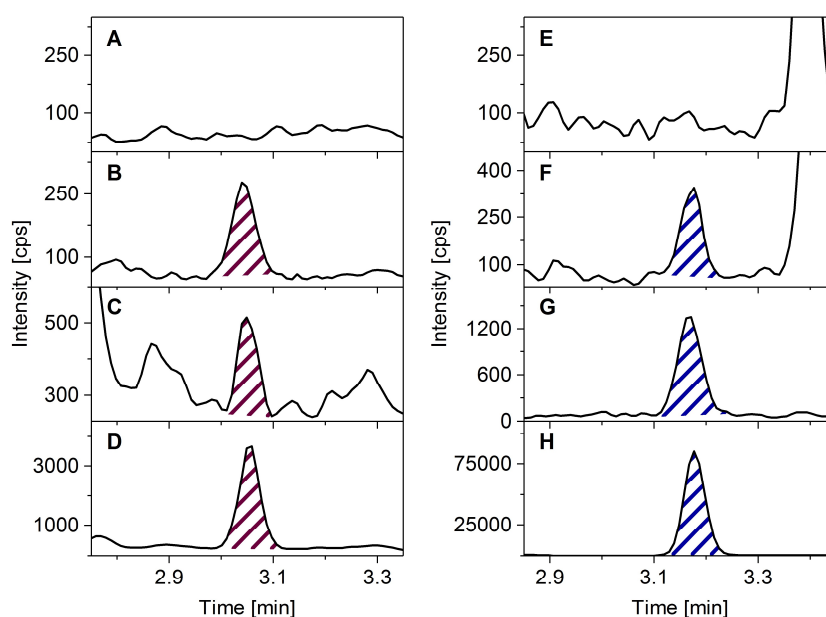


Fig. 1. Chromatograms (product ion EIC) in true plasma matrix. A: $^{13}\text{C}_3\text{E}$ in blank matrix; B: $^{13}\text{C}_3\text{E}$ spiked at LLOQ (10.0 pg mL^{-1}); C: E in real sample at LLOQ (10.7 pg mL^{-1}); D: E in real sample (242 pg mL^{-1}); E: $^{13}\text{C}_3\text{T}$ in blank matrix; F: $^{13}\text{C}_3\text{T}$ spiked at LLOQ (20.0 pg mL^{-1}); G: T in commercial control Level I (201 pg mL^{-1} , lowest concentration of all samples); H: T in real sample ($7,507 \text{ pg mL}^{-1}$).

Intra-assay and inter-day precisions and accuracies were determined in plasma using the surrogate calibrants. Four QCs were used to validate precision and accuracy: QC_{LLOQ} , $\text{QC}_{3 \times \text{LLOQ}}$, QC_{Mid} , QC_{ULOQ} . These QCs were measured on three days in quintuplicate ($n = 5$) (Table 3). Precisions were $<10 \%$ in the entire range and accuracies between 95 and 115 % recoveries clearly confirm that assay specificity of the current SWATH methodology is adequate. Adequate method performance was further confirmed by cross-validation with commercial QCs (lyophilized true plasma matrix controls with certified E and T concentrations). Results are shown in Table A.13. Precisions matched those of above validation study and bias remained within acceptable limits (6 - 15 %). Adequate analyte stability during sample storage, freeze-thaw cycles, autosampler stability and short-term stability at ambient temperature was verified as well (see Table A.11).

Intra-Assay Cross-Validation with Alternative Quantifiers

With the employed DIA using SWATH, comprehensive high-resolution MS/MS data are available across the steroid hormone peaks in all samples. Thus, it becomes possible to select post-acquisition the most appropriate ion from a peak group, viz. precursor ion from either TOF-MS or MS/MS experiments, or any fragment ion from MS/MS experiments, to generate

the EIC chromatograms for quantification. This opens up the possibility to use the most intensive ion as quantifier ion, provided it has sufficient specificity, and any of the other ion traces as qualifier ions, similar to QqQ-based quantification assays but with high mass resolution readout and no need of pre-acquisition decision on the selected ion transitions. It enables another level of validation via controlling the results by additional fragment or precursor ion EICs or ion ratios.⁵⁵ In other words, in order to control if the chosen fragment for quantification is selective, other fragments or precursors of the same analyte can be used for quantification and both sets of results can be compared. For example, for T a linear calibration from 500 to 15,000 pg mL⁻¹ could be achieved for the precursor from the MS experiment. Also, a second fragment with m/z 97.0648 (MS/MS fragment 2), with comparable sensitivity to the original quantifier fragment of m/z 109.0648, yielded a linear calibration function for the entire

Table 2. Matrix effect, extraction recovery, process efficiency.^b

Analyte		ME [%]	RE [%]	PE [%]
¹³ C ₃ E	QC _{LLOQ}	82.2 ± 12.7	94.2 ± 11.1	77.4 ± 11.5
	QC _{3x LLOQ}	79.6 ± 11.8	93.3 ± 9.7	74.3 ± 14.1
	QC _{Mid}	83.3 ± 5.7	79.4 ± 9.2	66.1 ± 8.4
	QC _{ULOQ}	80.2 ± 6.7	85.2 ± 3.9	68.4 ± 3.7
¹³ C ₃ T	QC _{LLOQ}	55.2 ± 13.3	86.9 ± 12.2	48.0 ± 6.5
	QC _{3x LLOQ}	58.7 ± 14.5	89.2 ± 10.9	52.4 ± 9.8
	QC _{Mid}	63.2 ± 4.5	78.7 ± 5.8	49.8 ± 6.8
	QC _{ULOQ}	64.4 ± 7.1	82.6 ± 5.5	53.2 ± 4.8

^bSingle determinations of 5 different lots were used to create QCs in neat standard solution, post-extraction spiked plasma and pre-extraction spiked plasma. Error was calculated by addition of relative errors of mean values. T shows a relatively ineffective PE. Since the LLOQ of 20 pg mL⁻¹ is below normal reference levels in patients, a PE of around 50 % can be accepted. Concentrations, see Table 3 footnote c.

range. All patient sample concentrations were also calculated for these alternative signals. Using MS/MS fragment 2 as alternative quantifier the results were in good agreement to the original results (scatter plot linear regression: $y = 1.000 (\pm 0.003) x + 158.2 (\pm 16.7)$, $R^2 = 0.99684$) (Fig. A.21A). Using the precursor ion trace of the TOF-MS experiment, the agreement was still acceptable yet significantly worse (scatter plot linear regression: $y = 0.9541 (\pm 0.006) x - 126.0 (\pm 31.6)$, $R^2 = 0.9875$) (Fig. A.21B) indicating the potential problem in terms of specificity of single stage MS data. Consequently, also these results consolidate adequate assay specificity and method performance.

Combined Targeted/Untargeted Profiling (Towards Steroidomics)

Contrary to classical targeted assays with triple quadrupole instruments the current method provides simultaneously lipid profiles in an untargeted manner. Additional information can be derived from TOF-MS experiments (survey scans) or SWATH-MS/MS experiments. In many cases, signals in TOF-MS lack of specificity or show insufficient sensitivity (see Fig. 4). Comprehensive data of superior quality can be acquired by additional MS/MS experiments. To demonstrate the potential of SWATH for steroidomic analysis, an exemplary MS-method with six extra SWATH experiments was created to cover the relevant mass range of interest (Table A.14; method 2).

Table 3. Validation results of precision and accuracy.^c

Analyte	QC _{LLOQ}		QC _{3xLLOQ}		QC _{Mid}		QC _{ULOQ}		
	Accuracy [%]	Precision [%]	Accuracy [%]	Precision [%]	Accuracy [%]	Precision [%]	Accuracy [%]	Precision [%]	
¹³ C ₃ E	Intra-day (n = 5)	114.7	3.9	99.2	2.5	96.9	4.4	93.9	8.5
	Inter-day (n = 15)	112.8	5.5	99.2	5.0	94.6	4.5	96.9	10.2
¹³ C ₃ T	Intra-day (n = 5)	111.7	4.7	102.8	2.7	108.2	4.5	95.7	6.6
	Inter-day (n = 15)	107.7	7.8	102.1	6.5	104.1	7.6	98.2	5.1

^cConcentrations were as following: QC_{LLOQ}, 10 pg mL⁻¹ for ¹³C₃E and 20 pg mL⁻¹ for ¹³C₃T; QC_{3xLLOQ}, 30 pg mL⁻¹ for ¹³C₃E and 60 pg mL⁻¹ for ¹³C₃T; QC_{Mid}, 250 pg mL⁻¹ for ¹³C₃E and 2,500 pg mL⁻¹ for ¹³C₃T; QC_{ULOQ} 1,000 pg mL⁻¹ for ¹³C₃E and 15,000 pg mL⁻¹ for ¹³C₃T.

The four previously optimized SWATH windows for E and T quantification remained unaltered, so that the capability of combined untargeted profiling and targeted quantification of E and T can be documented. A mass range from m/z 250 - 370 was additionally covered by SWATH MS/MS experiments, which mostly comprises unconjugated steroids. To use this narrower range for MS/MS experiments allows to design smaller precursor selection windows which is favorable for assay specificity in steroid analysis. The peak spotting plot in Fig. 2 and Fig. A.22 reveals a total of 1,613 molecular features in the TOF-MS survey scan.

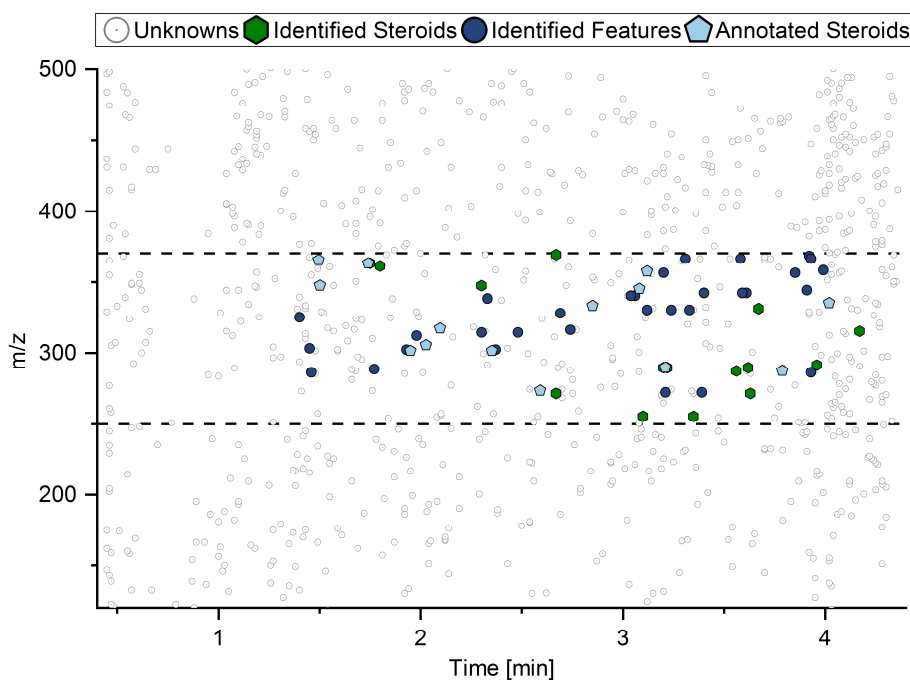


Fig. 2. Aligned peak spotting in 9 repeated measurements of a pooled plasma QC sample (m/z range from 120 to 500 is shown, for extended overview see Fig. A.22) applying method 2 (see Table A.14). 1,613 molecular features with a peak intensity over 2,000 cps were found in the survey scan after blank subtraction, de-isotoping and de-adducting. Dashed lines represent the mass range covered comprehensively by SWATH MS/MS experiments. Identified steroids were verified by injection of authentic standards and matching of t_R and mass spectra. Identified features showed matching precursor m/z and high level agreement of mass spectra (LipidBlast⁵⁶, MassBank⁶⁰) identified by MS-DIAL⁶¹ software. Annotated steroids were found by matching m/z of precursors from steroids covered in the LipidMaps⁶² database.

For verification of the utility of this expanded steroidomics profiling method, the trilevel commercial controls were analyzed. These commercial QCs specify concentrations for a variety of other steroids, besides E and T, dehydroepiandrosterone, dehydroepiandrosterone-sulfate, androstenedione, hydroxyprogesterone, dihydrotestosterone (DHT) and progesterone which could be identified by matching precursor mass, isotope pattern and MS/MS fragmentation (see Fig. 3). Furthermore, verification of identity was achieved by controlling for linearity of the obtained three-point calibration of the trilevel controls (see Table A.15). Cortisol, cortisone, epiE and epiT could also be specifically identified in samples by comparison with available standards (see Fig. 3). Other steroids only annotated by exact mass and coherent fragmentation were aldosterone, corticosterone, deoxycortisol, deoxycorticosterone, estrone and pregnenolone. Furthermore, several acylcarnitines could be identified via spectral matching to the LipidBlast⁵⁶ database, concluding that also other non-steroidal, lipophilic compounds are captured by sample preparation.

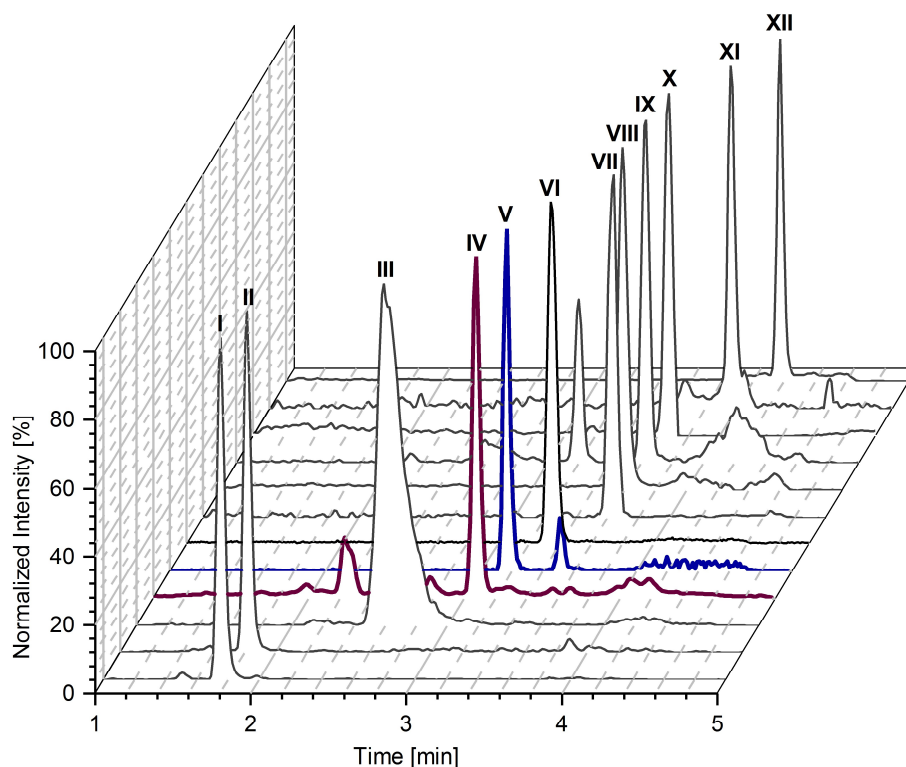


Fig. 3. Overlay of normalized chromatograms of identified steroids (targets IV and V; non-targeted steroids I-III and VI-XII) in commercial control. I: Cortisol (fragment, m/z 327.1955 \pm 0.02); II: Cortisone (fragment, m/z 343.1904 \pm 0.02); III: Dehydroepiandrosterone-Sulfate (fragment, m/z 213.1638 \pm 0.02); IV: Estradiol (fragment, m/z 159.0804 \pm 0.02); V: Testosterone (fragment, m/z 109.0648 \pm 0.02); VI: epiE (fragment, m/z 159.0804 \pm 0.02); VII: epiT (fragment, m/z 109.0648 \pm 0.02); VIII: Androstenedione (fragment, m/z 97.0648 \pm 0.02); IX: Dehydroepiandrosterone (precursor, m/z 271.2062 \pm 0.02); X: Hydroxyprogesterone (fragment, m/z 97.0648 \pm 0.02); XI: Dihydrotestosterone (fragment, m/z 255.2113 \pm 0.02); XII: Progesterone (fragment, m/z 97.0648 \pm 0.02). Method 2 (see Table A.14).

The currently employed combined targeted/untargeted profiling by data-independent acquisition with SWATH provides other benefits. Availability of comprehensive MS/MS data within the dedicated m/z range across the chromatogram and all samples allows to extract both MS chromatograms of precursors but also MS/MS chromatograms of fragments which is not possible with common data-dependent acquisition. This enables uncompromised retrospective data processing post-acquisition. Quantitative analysis can be either performed on precursors or fragments, which ever gives better assay specificity and/or higher sensitivity. This is documented in Fig. 4 exemplarily for non-targeted dihydrotestosterone (DHT). DHT is a bioactive metabolite of T formed by the enzyme 5α -reductase and is the biologically most active form of T. In a targeted assay with a triple quadrupole and SRM acquisition for E and T, no information on DHT could be obtained. In the combined targeted/untargeted screening approach, presented in Table A.14, DHT is detected in the different samples as well. This allows deriving information, at least for relative quantification (e.g. for differential steroidomics between sample groups). However, the signal is very poor in the TOF-MS chromatogram of

the precursor (S/N (PeakView estimate) = 2.3) due to many interferences and a high noise level (Fig. 4A). Although the signal is reduced in the MS/MS chromatogram of the precursor (Fig. 4B), the S/N ratio was significantly improved due to a lower noise level. Upon extraction of the MS/MS chromatogram from the precursor with m/z 255.2113 the S/N ratio increased by a factor of about 4 because the majority of interferences were eliminated and the noise level further reduced (Fig. 4C). Data processing on this signal is certainly advantageous for retrospective relative quantification of samples. The fact that in DIA with SWATH all signals are acquired and comprehensive MS as well as MS/MS data are available, provides researchers the flexibility to use the optimal MS or MS/MS signal for data processing. If taken into account that MS parameters were not optimized for the untargeted SWATH experiments, even higher sensitivity might be possible. Also, ion ratios can be further processed for confirmation underpinning the advantage of DIA.⁵⁵

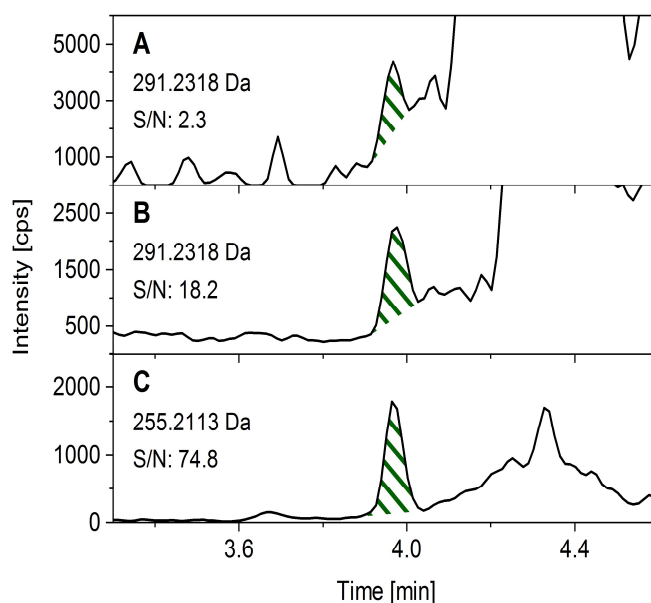


Fig. 4. Comparison of signal quality for non-targeted dihydrotestosterone. Signals were obtained from commercial QC Lvl. III (1,050 $\mu\text{g ml}^{-1}$). A: TOF-MS of precursor ion; B: SWATH-MS/MS of precursor ion; C: SWATH-MS/MS of fragment ion. Method 2 (see Table A.14). (S/N values are estimates calculated with PeakView).

Clinical Study and comparison with immunoassay results

In a clinical study, the effect of E and insulin on food intake in men was investigated. Here, two groups of healthy young men (each $n = 16$) received transdermal E (100 $\mu\text{g}/24\text{h}$) or placebo via transdermal patches for three days. According to a 2×2 design, the experiment comprised two individual sessions in each subject with intranasal insulin (160 IU) and, respectively, placebo administration. In each session, plasma samples were collected at five different time points, totaling 320 samples. These samples were measured by method 1 (Table

1) and also quantified by a competitive chemiluminescent enzyme immunoassay (IA, Immulite 2000). Whereas E levels of 22.2 % of samples were below the LLOQ of the immunoassay (20 pg mL⁻¹), only one sample (0.3 %) could not be quantified by mass spectrometry (LLOQ: 10 pg mL⁻¹). IA results were compared to mass spectrometric (MS) results by correlation scatter plots (Fig. 5) and Bland-Altman plots (Fig. 6). At first sight, the scatter plot for E presumes acceptable agreement between methods. However, the Pearson correlation coefficient r of 0.8913 expresses the high variability in the lower region between 40 - 100 pg mL⁻¹. This gets more clearly visible in the Bland-Altman plot, where we can see that differences increase with lower E levels and reach over ± 60 %. The scatter plot for T shows disagreement, especially in the upper region above 5,000 pg mL⁻¹. A further look into the Bland-Altman plot shows that there is strong variability over the whole range. Although 2s limits are narrower than for E, differences of over ± 50 % can be observed, which is unacceptable for clinical measurements. The reason for the partially strong disagreement could be the known disadvantages of immunoassays, as they are prone to cross reactivity, general sample condition like lipemia or hemolysis⁵⁷ or other interferences. Several groups already investigated agreement between different methodologies for steroid quantification and found large discrepancy exceeding clinical acceptance limits.³⁻⁹ Vesper et. al⁸ reported the high variabilities of estradiol assays in general and found mean bias of up to 22.5 % for MS methods compared to up to 235 % for immunoassays. Wang et. al⁹ found that the Immulite 2000 is likely to systematically underestimate T concentration and showed discrepancy of over 60 % compared to LC-MS/MS, which correlates well to our observations. Overall, variability of quantitative results was found to be substantially smaller for MS methods than for immunoassays.⁵⁸ Consequently, institutions like the Joint Committee for Traceability in Laboratory Medicine (JCTLM) only accept MS assays as reference methods⁵⁹ and the National Institute of Standards and Technology (NIST) is working on establishing reliable LC-MS/MS methods for steroid quantification.²³

To control for trueness of the mass spectrometric method, commercial QCs with known concentration levels were purchased and quantified (see Table A.13). By reaching the clinical acceptance limit of 85 - 115 % accuracy and <15 % precision, the MS method was proven to yield reasonable results.

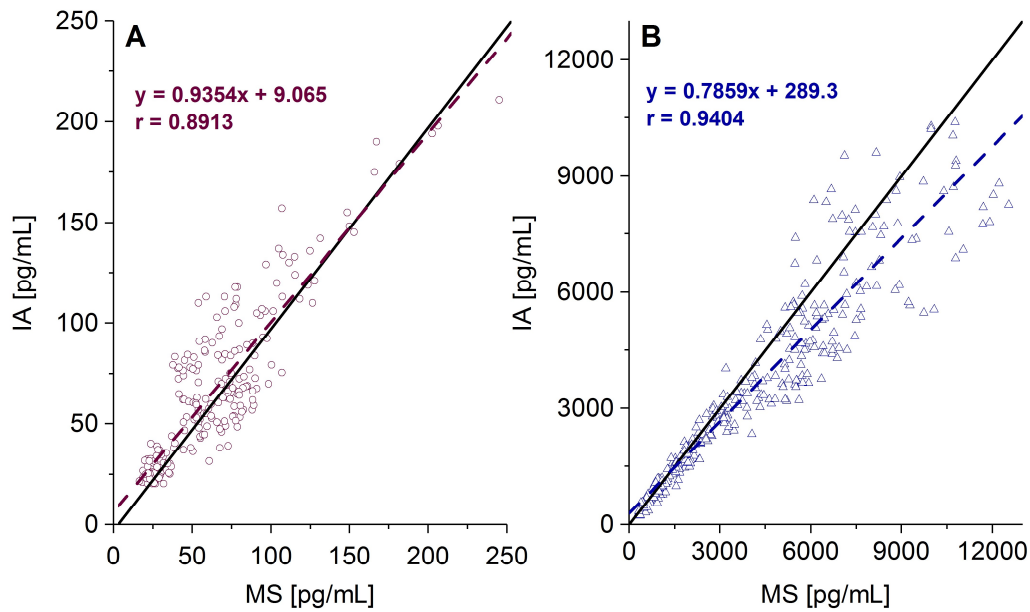


Fig. 5. Scatter plots for comparison of results from immunoassay (IA) and mass spectrometry (MS). Plot (A) shows results for E and plot (B) for T. Solid lines resemble the optimum line of parity. Dashed lines are results of linear regression analysis of results obtained with the two methods.

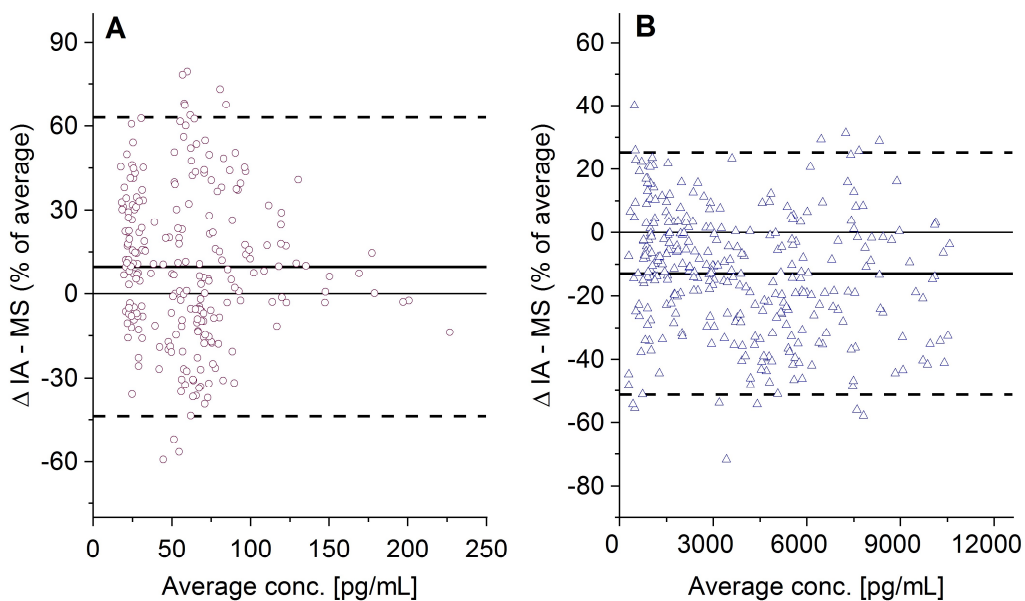


Fig. 6. Bland-Altman plots for comparison of results from immunoassay (IA) and mass spectrometry (MS). Plot (a) for E with mean difference (9.7 %, solid line) and 2s limits (95 % limits of agreement; +2s = 63.1 %, -2s = - 43.7 %, dashed lines). Plot (b) for T with mean difference (-13.0 %, solid line) and 2s limits (+2s = 25.3 %, -2s = - 51.2 %, dashed lines).

Moreover, processing of survey scan data revealed additional information on study participants. For instance, a 3.2-fold increase of hydroxyprogesterone (d_5T -normalized) in placebo patch groups compared to E patch groups could be displayed (Fig. 7). Application of transdermal E therefore seems to interact in hydroxyprogesterone metabolism. To support this hypothesis, we retrospectively analyzed hydroxyprogesterone/ d_5T response ratios in QC samples (QC_{low} and QC_{high}; n = 36), which were run equally distributed across the entire

sample sequence and were derived from an identical plasma pool. Precision, calculated as relative standard deviation, was 23.6 %. Moreover, hydroxyprogesterone ratios in commercial QCs (n = 9 per level) showed following precision: Level I (0.30 ng mL⁻¹): 29.7 %; Level II (1.54 ng mL⁻¹): 16.6 %; Level III (8.96 ng mL⁻¹): 8.3 %. These values are well below the biological variance observed in the study samples and below the common acceptance limit for assay precision of 30% RSD for biomarker studies (usually applied as criteria in untargeted profiling methods). Other examples of significantly regulated steroids were found as well (Fig. A.23, Fig A.24). In general, it is shown that such an assay combined favorably hypothesis-driven targeted quantification and untargeted profiling which allowed to generate new hypotheses without extra measurements, without additional samples, and without additional human/animal experiments. Consequently, such a combined targeted/untargeted assay can be regarded in line with the 3R-principle for human and animal studies (3R principle means to avoid animal experiments altogether (Replacement), to limit the number of animals (Reduction) and their suffering (Refinement) in tests to an absolute minimum), because it collects more information per sample.

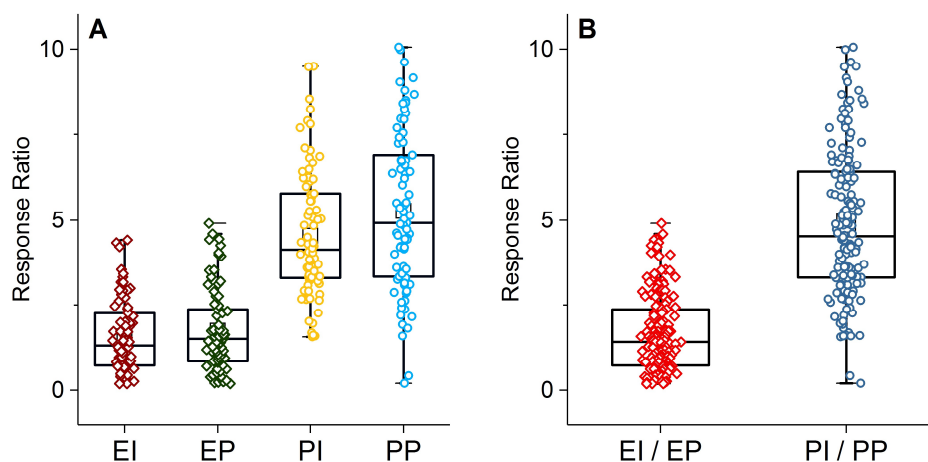


Fig. 7. Relative quantification of hydroxyprogesterone. EI: E patch & insulin treatment; EP: E patch & placebo treatment; PI: placebo patch & insulin treatment; PP: placebo patch and placebo treatment. Boxplots for each of the four groups (A) and for grouped E patch and grouped placebo patch samples (B). For B, a 3.2-fold increase (median values) in hydroxyprogesterone was found in placebo patch groups (*U*-test, *p*-value: 3.3×10^{-47}). Signals were obtained from TOF-MS scan (precursor signals).

5.1.5. Conclusions

The DIA technique SWATH, primarily designed for untargeted analysis of peptides in proteomics, was shown capable of accurate and reliable quantification via HR-MS/MS data. While controlling for specificity, advantageous analysis in terms of analyte coverage and sensitivity compared to regular MRM^{HR} was demonstrated. This way simultaneous low-level quantification of E and T was achieved without derivatization nor polarity switching. Especially for endogenous compounds that require alternative quantification via surrogate calibrants, SWATH enables beneficial experiment design by reduction of the total number of MS and MS/MS experiments favorable for keeping cycle times short. Owing to the feature of combined targeted/untargeted analysis, valuable secondary information is recorded and accessible post-acquisition. High quality untargeted MS/MS data, e.g. for steroid profiling, can be collected by optional, user-modulated SWATH experiments. Validation according to international guidelines (with some minor modifications) and accurate quantification of certified, commercial quality controls underline the value of this acquisition technique. By exploiting the potential of SWATH for sensitive and quantitative steroid analysis, most likely in conjunction with extended chromatography, the avenue towards steroidomics has been paved herein.

Acknowledgments

We acknowledge the financial support by the “Struktur- und Innovationsfonds Baden-Württemberg (SI-BW)”, the German Science Funds (DFG no. INST 37/821-1 FUGG) and by grants from the German Federal Ministry of Education and Research (BMBF) to the German Center for Diabetes Research (DZD e.V.; 01GI0925).

5.1.6. References

- (1) Holst, J. P.; Soldin, O. P.; Guo, T.; Soldin, S. J. *Clin Lab Med* **2004**, *24*, 105-118.
- (2) Tate, J.; Ward, G. *Clin Biochem Rev* **2004**, *25*, 105-120.
- (3) Chen, Y.; Yazdanpanah, M.; Hoffman, B. R.; Diamandis, E. P.; Wong, P. Y. *Clin Biochem* **2009**, *42*, 484-490.
- (4) Taieb, J.; Mathian, B.; Millot, F.; Patricot, M. C.; Mathieu, E.; Queyrel, N.; Lacroix, I.; Somma-Delpero, C.; Boudou, P. *Clin Chem* **2003**, *49*, 1381-1395.
- (5) Fanelli, F.; Belluomo, I.; Di Lallo, V. D.; Cuomo, G.; De lasio, R.; Baccini, M.; Casadio, E.; Casetta, B.; Vicennati, V.; Gambineri, A.; Grossi, G.; Pasquali, R.; Pagotto, U. *Steroids* **2011**, *76*, 244-253.
- (6) Handelsman, D. J.; Newman, J. D.; Jimenez, M.; McLachlan, R.; Sartorius, G.; Jones, G. R. *Clin Chem* **2014**, *60*, 510-517.
- (7) Stanczyk, F. Z.; Cho, M. M.; Endres, D. B.; Morrison, J. L.; Patel, S.; Paulson, R. J. *Steroids* **2003**, *68*, 1173-1178.
- (8) Vesper, H. W.; Botelho, J. C.; Vidal, M. L.; Rahmani, Y.; Thienpont, L. M.; Caudill, S. P. *Steroids* **2014**, *82*, 7-13.
- (9) Wang, C.; Catlin, D. H.; Demers, L. M.; Starcevic, B.; Swerdloff, R. S. *J Clin Endocrinol Metab* **2004**, *89*, 534-543.
- (10) Rauh, M. *J Steroid Biochem Mol Biol* **2010**, *121*, 520-527.
- (11) Thakare, R.; Chhonker, Y. S.; Gautam, N.; Alamoudi, J. A.; Alnouti, Y. *J Pharm Biomed Anal* **2016**, *128*, 426-437.
- (12) Li, W. L.; Cohen, L. H. *Anal Chem* **2003**, *75*, 5854-5859.
- (13) Jones, B. R.; Schultz, G. A.; Eckstein, J. A.; Ackermann, B. L. *Bioanalysis* **2012**, *4*, 2343-2356.
- (14) Ceglarek, U.; Kortz, L.; Leichtle, A.; Fiedler, G. M.; Kratzsch, J.; Thiery, J. *Clin Chim Acta* **2009**, *401*, 114-118.
- (15) Guo, T.; Taylor, R. L.; Singh, R. J.; Soldin, S. J. *Clin Chim Acta* **2006**, *372*, 76-82.
- (16) Harwood, D. T.; Handelsman, D. J. *Clin Chim Acta* **2009**, *409*, 78-84.
- (17) Jeanneret, F.; Tonoli, D.; Rossier, M. F.; Saugy, M.; Boccard, J.; Rudaz, S. *J Chromatogr A* **2016**, *1430*, 97-112.
- (18) Wooding, K. M.; Hankin, J. A.; Johnson, C. A.; Chosich, J. D.; Baek, S. W.; Bradford, A. P.; Murphy, R. C.; Santoro, N. *Steroids* **2015**, *96*, 89-94.
- (19) Dai, W.; Huang, Q.; Yin, P.; Li, J.; Zhou, J.; Kong, H.; Zhao, C.; Lu, X.; Xu, G. *Anal Chem* **2012**, *84*, 10245-10251.
- (20) Lee, S. H.; Lee, N.; Hong, Y.; Chung, B. C.; Choi, M. H. *Anal Chem* **2016**, *88*, 11624-11630.
- (21) Keski-Rahkonen, P.; Desai, R.; Jimenez, M.; Harwood, D. T.; Handelsman, D. J. *Anal Chem* **2015**, *87*, 7180-7186.
- (22) McCulloch, R. D.; Robb, D. B. *Anal Chem* **2017**, *89*, 4169-4176.
- (23) Boggs, A. S.; Bowden, J. A.; Galligan, T. M.; Guillette, L. J., Jr.; Kucklick, J. R. *Anal Bioanal Chem* **2016**, *408*, 4179-4190.
- (24) Zhou, H.; Wang, Y.; Gatcombe, M.; Farris, J.; Botelho, J. C.; Caudill, S. P.; Vesper, H. W. *Anal Bioanal Chem* **2017**.
- (25) Johanning, J.; Heinkele, G.; Precht, J. C.; Brauch, H.; Eichelbaum, M.; Schwab, M.; Schroth, W.; Murdter, T. E. *Anal Bioanal Chem* **2015**, *407*, 7497-7502.
- (26) Polet, M.; De Wilde, L.; Van Renterghem, P.; Van Gansbeke, W.; Van Eenoo, P. *Anal Chim Acta* **2017**.
- (27) Wang, C.; Wu, C.; Zhang, L.; Zhang, J. *Anal Chem* **2016**, *88*, 7878-7884.
- (28) Guo, N.; Liu, P.; Ding, J.; Zheng, S. J.; Yuan, B. F.; Feng, Y. Q. *Anal Chim Acta* **2016**, *905*, 106-114.
- (29) Star-Weinstock, M.; Williamson, B. L.; Dey, S.; Pillai, S.; Purkayastha, S. *Anal Chem* **2012**, *84*, 9310-9317.
- (30) Pataj, Z.; Liebisch, G.; Schmitz, G.; Matysik, S. *J Chromatogr A* **2016**, *1439*, 82-88.
- (31) Tonoli, D.; Furstenberger, C.; Boccard, J.; Hochstrasser, D.; Jeanneret, F.; Odermatt, A.; Rudaz, S. *Chem Res Toxicol* **2015**, *28*, 955-966.
- (32) Zhou, J. T.; Liu, C. L.; Si, D. D.; Jia, B.; Zhong, L. J.; Yin, Y. X. *Anal Chim Acta* **2017**, *972*, 62-72.

- (33) Triebel, A.; Trotzmüller, M.; Hartler, J.; Stojakovic, T.; Kofeler, H. C. *J Chromatogr B Analyt Technol Biomed Life Sci* **2017**, 1053, 72-80.
- (34) Cajka, T.; Fiehn, O. *Anal Chem* **2016**, 88, 524-545.
- (35) Palermo, A.; Botre, F.; de la Torre, X.; Zamboni, N. *Anal Chim Acta* **2017**, 964, 112-122.
- (36) Wang, S.; Zhou, L.; Wang, Z.; Shi, X.; Xu, G. *Anal Chim Acta* **2017**, 966, 34-40.
- (37) Zhou, J.; Li, Y.; Chen, X.; Zhong, L.; Yin, Y. *Talanta* **2017**, 164, 128-136.
- (38) Gillet, L. C.; Navarro, P.; Tate, S.; Rost, H.; Selevsek, N.; Reiter, L.; Bonner, R.; Aebersold, R. *Mol Cell Proteomics* **2012**, 11, O111 016717.
- (39) Hopfgartner, G.; Tonoli, D.; Varesio, E. *Anal Bioanal Chem* **2012**, 402, 2587-2596.
- (40) Arnhard, K.; Gottschall, A.; Pitterl, F.; Oberacher, H. *Anal Bioanal Chem* **2015**, 407, 405-414.
- (41) Roemmelt, A. T.; Steuer, A. E.; Kraemer, T. *Anal Chem* **2015**, 87, 9294-9301.
- (42) Bonner, R.; Hopfgartner, G. *Bioanalysis* **2016**, 8, 1735-1750.
- (43) Zhu, X.; Chen, Y.; Subramanian, R. *Anal Chem* **2014**, 86, 1202-1209.
- (44) Cajka, T.; Fiehn, O. *Trends Analyt Chem* **2014**, 61, 192-206.
- (45) Chatterjee, M.; Rath, D.; Schlotterbeck, J.; Rheinlaender, J.; Walker-Allgaier, B.; Alnaggar, N.; Zdanyte, M.; Müller, I.; Borst, O.; Geisler, T.; Schaffer, T. E.; Lammerhofer, M.; Gawaz, M. *Eur Heart J* **2017**, 38, 1993-2005.
- (46) Siegel, D.; Meinema, A. C.; Permentier, H.; Hopfgartner, G.; Bischoff, R. *Anal Chem* **2014**, 86, 5089-5100.
- (47) Hammond, G. L. *J Endocrinol* **2016**, 230, R13-25.
- (48) Kulle, A. E.; Riepe, F. G.; Melchior, D.; Hiort, O.; Holterhus, P. M. *J Clin Endocrinol Metab* **2010**, 95, 2399-2409.
- (49) Polson, C.; Sarkar, P.; Incledon, B.; Raguvaran, V.; Grant, R. *J Chromatogr B Analyt Technol Biomed Life Sci* **2003**, 785, 263-275.
- (50) Ding, J.; Neue, U. D. *Rapid Commun Mass Spectrom* **1999**, 13, 2151-2159.
- (51) Samtani, M. N.; Jusko, W. J. *Biomed Chromatogr* **2007**, 21, 585-597.
- (52) Kebarle, P.; Tang, L. *Analytical Chemistry* **1993**, 65, A972-A986.
- (53) Chernushevich, I. V.; Loboda, A. V.; Thomson, B. A. *J Mass Spectrom* **2001**, 36, 849-865.
- (54) Matuszewski, B. K.; Constanzer, M. L.; Chavez-Eng, C. M. *Anal Chem* **2003**, 75, 3019-3030.
- (55) Naz, S.; Gallart-Ayala, H.; Reinke, S. N.; Mathon, C.; Blankley, R.; Chaleckis, R.; Wheelock, C. E. *Anal Chem* **2017**, 89, 7933-7942.
- (56) Kind, T.; Liu, K. H.; Lee, D. Y.; DeFelice, B.; Meissen, J. K.; Fiehn, O. *Nat Methods* **2013**, 10, 755-758.
- (57) Stanczyk, F. Z.; Lee, J. S.; Santen, R. J. *Cancer Epidemiol Biomarkers Prev* **2007**, 16, 1713-1719.
- (58) Vesper, H. W.; Bhasin, S.; Wang, C.; Tai, S. S.; Dodge, L. A.; Singh, R. J.; Nelson, J.; Ohorodnik, S.; Clarke, N. J.; Salameh, W. A.; Parker, C. R., Jr.; Razdan, R.; Monsell, E. A.; Myers, G. L. *Steroids* **2009**, 74, 498-503.
- (59) Vesper, H. W.; Botelho, J. C.; Shacklady, C.; Smith, A.; Myers, G. L. *Steroids* **2008**, 73, 1286-1292.
- (60) Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K.; Oda, Y.; Kakazu, Y.; Kusano, M.; Tohge, T.; Matsuda, F.; Sawada, Y.; Hirai, M. Y.; Nakanishi, H.; Ikeda, K.; Akimoto, N., et al. *J Mass Spectrom* **2010**, 45, 703-714.
- (61) Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; VanderGheynst, J.; Fiehn, O.; Arita, M. *Nat Methods* **2015**, 12, 523-526.
- (62) Sud, M.; Fahy, E.; Cotter, D.; Brown, A.; Dennis, E. A.; Glass, C. K.; Merrill, A. H., Jr.; Murphy, R. C.; Raetz, C. R.; Russell, D. W.; Subramaniam, S. *Nucleic Acids Res* **2007**, 35, D527-532.

5.1.7. Supplementary Data

Materials and Methods

Fig. A.1 shows structures of the analytes, surrogate calibrants and internal standards. Suppliers are specified in the main document.

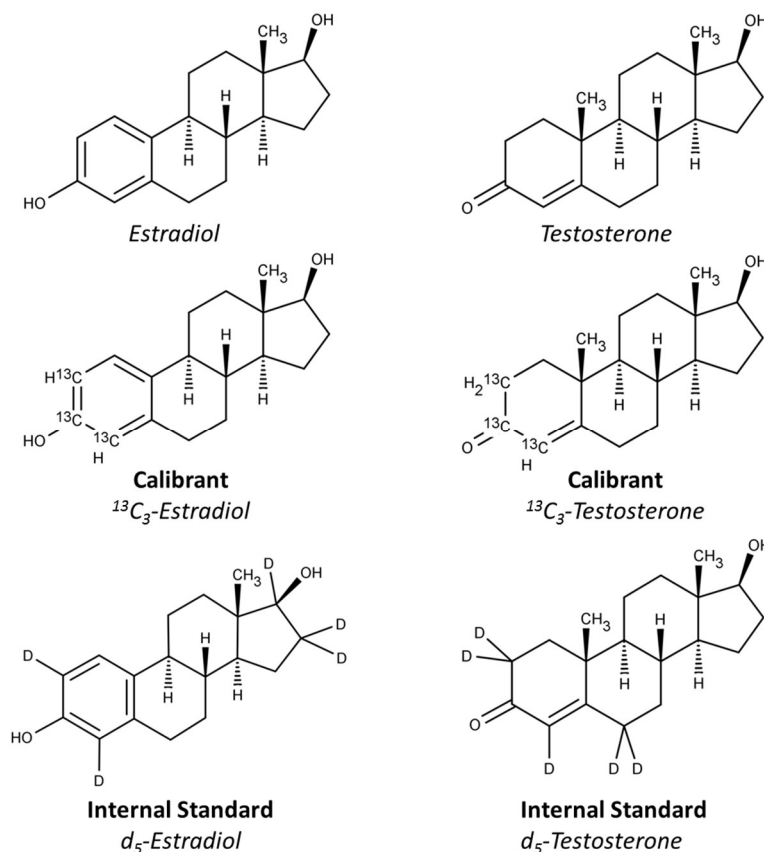


Fig. A.1 Structure of steroid hormone analytes. Estradiol (E) and testosterone (T) as well as corresponding $^{13}\text{C}_3$ -labelled surrogate calibrants ($^{13}\text{C}_3\text{E}$ and $^{13}\text{C}_3\text{T}$) and deuterated internal standards ($d_5\text{E}$ and $d_5\text{T}$).

Solvents for mobile phase preparation were MS grade. Acetonitrile (MeCN, Ultra LC-MS grade), methanol (MeOH, Ultra LC-MS grade), 2-propanol (Ultra LC-MS grade) and formic acid (98 %, w/v, ACS grade) were purchased from Carl Roth (Karlsruhe, Germany). Stock solutions of each standard in MeOH (1.0 mg mL^{-1}) were further diluted in multiple steps, using a 10 mL volumetric flask and MeOH, to working solutions of following concentrations: $1,000 \text{ ng mL}^{-1}$, 100 ng mL^{-1} , 10 ng mL^{-1} and 1 ng mL^{-1} . Prior to preparation of spiked plasma samples, responses of $^{13}\text{C}_3$ -standards (surrogate calibrants) were matched with corresponding target analyte responses by altering $^{13}\text{C}_3$ -standard concentrations. A response factor ratio (RF; $^{13}\text{C}_3$ -standard divided by target analyte standard) of 1.00 ± 0.05 was deemed acceptable (Jones et. al) and adjusted (see Table A.1 and Table A.2). The RF-fitted $^{13}\text{C}_3$ -standard solutions were

used for spiking pooled EDTA plasma, yielding matrix-matched calibrants and quality controls (QCs). Six non-zero surrogate calibrants with concentrations from 10 - 1,000 pg mL⁻¹ for ¹³C₃E and 20 - 15,000 pg mL⁻¹ for ¹³C₃T were prepared. For validation, four QCs: QC_{LLOQ}, QC_{3xLLOQ}, QC_{Mid}, QC_{ULOQ} (¹³C₃E: 10, 30, 250, 1,000 pg mL⁻¹; ¹³C₃T: 20, 60, 2,500, 15,000 pg mL⁻¹) were spiked separately. Two QCs, QC_{low} (¹³C₃T: 60 pg mL⁻¹; ¹³C₃E: 30 pg mL⁻¹) and QC_{high} (¹³C₃T: 12,000 pg mL⁻¹; ¹³C₃E: 800 pg mL⁻¹) were embedded in the sample sequence to verify stable method performance.

Table A.1. Evaluation of initial response factor of target analytes and corresponding surrogate calibrants.^a

T Area	¹³ C ₃ T Area	Response factor	E Area	¹³ C ₃ E Area	Response factor
55058	52818	0.959	11442	13662	1.194
53584	52932	0.988	11753	14352	1.221
61523	58915	0.958	13395	16353	1.221
87084	82572	0.948	14838	17622	1.188
	Average	0.963		Average	1.206
	StDev	0.017		StDev	0.018
	RSD [%]	1.777		RSD [%]	1.458

^aResults were obtained by multiple analysis of 1.0 ng mL⁻¹ of each compound in MeOH. Response factor = Area surrogate calibrant / Area target analyte.

Table A.2. Final response factor of target analytes and corresponding surrogate calibrants after adjusting concentrations of surrogate calibrants.^b

T Area	¹³ C ₃ T Area	Response factor	E Area	¹³ C ₃ E Area	Response factor
45014	46481	1.033	20200	21020	1.041
60660	60419	0.996	17600	18280	1.039
52848	51930	0.983	28780	28390	0.986
76736	78098	1.018	18129	17973	0.991
	Average	1.007		Average	1.022
	StDev	0.022		StDev	0.029
	RSD [%]	2.209		RSD [%]	2.872

^bResults were obtained by multiple analysis after adjusting surrogate analyte concentration to match corresponding target analyte response. Final concentrations in MeOH: T: 1.00 ng mL⁻¹; ¹³C₃T: 1.04 ng mL⁻¹; E: 1.00 ng mL⁻¹; ¹³C₃E: 0.83 ng mL⁻¹. A response factor ratio of 1.00 ± 0.05 with an RSD ≤ 3 % was deemed acceptable.

Method Development

Sample Preparation

Several low level steroid hormone plasma samples were pooled and spiked with 100 pg mL⁻¹ of analytes and corresponding ¹³C₃-labelled analogues (¹³C₃E and ¹³C₃T). Two distinct protocols were used for protein precipitation and a variety of distinct SPE cartridges were tested for sample clean up.

I. Protein precipitation using 1.0 mL plasma with ice-cold MeOH (2.5:1, v:v)

- 2 mL of supernatant into speed vac
- resuspend in 1.0 mL 25 % MeOH
- load on SPE cartridge & wash with 1.0 mL 5 % MeOH
- elute with MeOH (2 x 200 µL)
- dry in speed vac & resuspend in 100 µL MeOH
- centrifuge and measure supernatant

II. Incubate 1.0 mL plasma with 1.0 mL 5 % H₃PO₄ for 30 min

- load on SPE cartridge & wash with 1.0 mL 5 % MeOH
- elute with MeOH (2 x 200 µL)
- dry in speed vac & resuspend in 100 µL MeOH
- centrifuge and measure supernatant

A 1:20 dilution of the resultant extracts was injected into UHPLC-ESI-QTOF-MS/MS (for method description see main document). The following results were obtained (Table A.3 and Table A.4).

Table A.3. Optimization of sample preparation comprising the steps protein precipitation and SPE. Results for T and ¹³T₃E.

Cartridge (Supplier)	Protocol	T [S/N]
Supel-Select HLB, 1 mL (Sigma-Aldrich)	I	443.6
Sep-Pak C18, 1 mL (Waters)	I	484.7
Oasis PRiME HLB, 1 mL (Waters)	I	485.1
Chromabond C18ec, 1 mL (Macherey-Nagel)	I	679.6
Oasis HLB, 1 mL (Waters)	I	695.5
Supra-Clean C18-S, 1 mL (PerkinElmer)	I	835.2
Sep-Pak C18, 1 mL (Waters)	II	955.1
Oasis PRiME HLB, 1 mL (Waters)	II	1269.1
Supra-Clean C18-S, 1 mL (PerkinElmer)	II	1299.7

Table A.4. Optimization of sample preparation comprising the steps protein precipitation and SPE. Results for E and ¹³C₃E.

Cartridge (Supplier)	Protocol	E [S/N]
Oasis PRiME HLB, 1 mL (Waters)	I	36.3
Supel-Select HLB, 1 mL (Sigma-Aldrich)	I	42.8
Supra-Clean C18-S, 1 mL (PerkinElmer)	I	49.3
Sep-Pak C18, 1 mL (Waters)	I	53.6
Oasis HLB, 1 mL (Waters)	I	57.2
Chromabond C18ec, 1 mL (Macherey-Nagel)	I	68.2
Supra-Clean C18-S, 1 mL (PerkinElmer)	II	88.1
Sep-Pak C18, 1 mL (Waters)	II	104.0
Oasis PRiME HLB, 1 mL (Waters)	II	131.9

As can be seen from Table A.3 and Table A.4, the following SPE cartridges showed the best results:

Supra-Clean C18-S, 1 mL (PerkinElmer)	50 mg / 1mL
Oasis PRiME HLB, 1 mL (Waters)	30 mg / 1 mL
Chromabond C18ec, 1 mL (Macherey-Nagel)	100 mg / 1mL
Sep-Pak C18, 1 mL (Waters)	100 mg / 1mL

From these experiments, the following optimized sample preparation protocol was derived (Fig. A.2) which is described in detail in the main document.

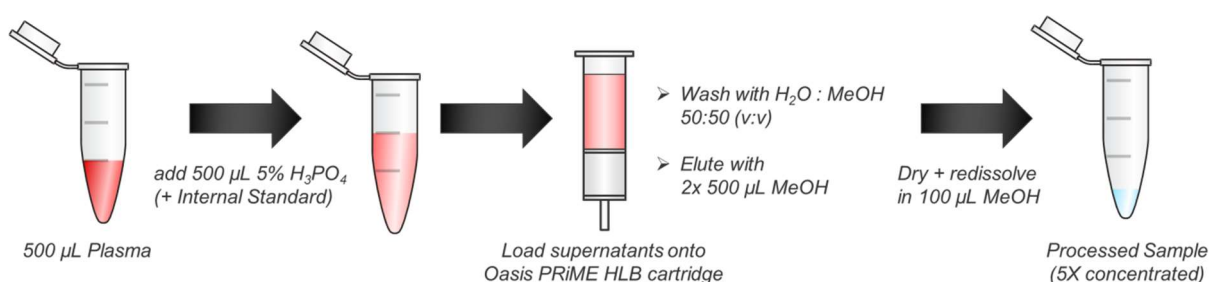


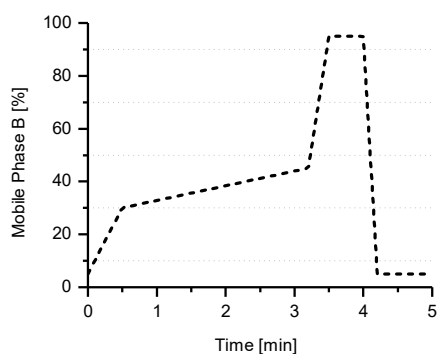
Fig. A.2. Optimized sample preparation procedure.

To conclude, time-consuming drying or further aqueous dilution was avoided by selecting 5 % H_3PO_4 for initial precipitation. Efficient removal of residual proteins from the resultant supernatant was achieved by subsequent purification steps. As the Oasis PRiME HLB material does not require conditioning and equilibration, samples could be directly loaded onto the cartridges. 50 % MeOH in H_2O (v/v) was selected as optimum washing eluent. Lower percentages of MeOH led to higher noise and decreased sensitivity due to ion suppression of leftover matrix compounds. Complete analyte elution of E and $^{13}\text{C}_3\text{E}$ could be achieved with 1 \times 500 μL MeOH. However, a second 500 μL MeOH elution was necessary to improve recovery of T and $^{13}\text{C}_3\text{T}$. MeOH + 0.1 % formic acid (v/v), MeCN and 2-propanol were also tested as elution solvents but only 2-propanol led to comparable results. Accordingly, MeOH was chosen as it shows a favorable evaporation rate. By drying and reconstitution in 100 μL MeOH, a total sample concentration factor of 5 was achieved to reach sufficient levels of sensitivity.

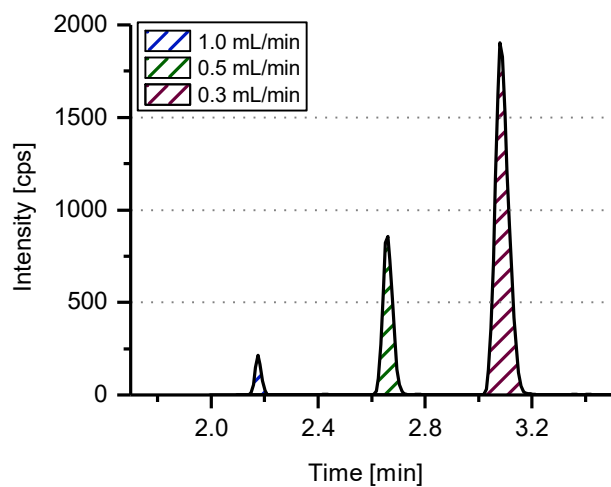
Liquid Chromatography

A fast UHPLC method with gradient elution (5 min including reequilibration) was developed using a core-shell C18 column (Kinetex® C18, 2.6 μm) (Fig. A.3A). Faster elution by higher flow rates was not considered because the detection sensitivity significantly dropped at flow rates higher than 0.3 mL min⁻¹ (Fig. A.3B). Decreasing the flow rate, starting from 0.5 mL min⁻¹, led to higher sensitivity. However, poor peak shape at very low flow rates was limiting this optimization parameter so that 0.3 mL min⁻¹ was set as optimum. Higher percentages than 0.1 % (v/v) of formic acid in mobile phases did not show any improvement in chromatography or sensitivity. The chosen gradient was optimized for sensitivity, selectivity and run time. Resolution between the two target analytes was not further optimized in favor of short analysis times. Yet, although only partial, close to baseline resolution was achieved (Fig. A.3C), there was no assay specificity problem arising from the two steroids due to selective detection by the specific SWATH acquisition.

(A)



(B)



(C)

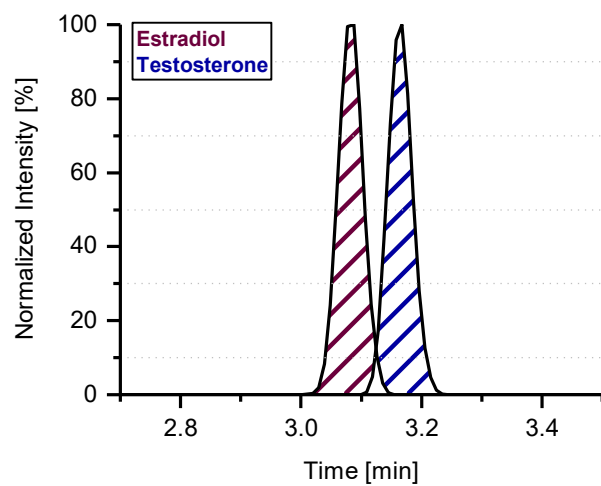


Fig. A.3. UHPLC separation on Kinetex[®] C18, 2.6 μ m (50 x 2.1 mm) column. (A) Gradient profile, (B) effect of flow rate on sensitivity, and (C) separation of E and T under final conditions ($R_s = 0.98$). Mobile phases, A: H₂O + 0.1 % FA, B: ACN + 0.1 % FA.

MS Detection

Concentration levels of E in human plasma are extremely low, in particular in male, in the low pg mL^{-1} range (see Table A.5). On the other hand, T concentrations in male plasma are in the ng mL^{-1} range and can be readily detected. Second, E shows higher ionization efficiencies in negative ion mode due to its phenolic group and weakly acidic character, while T cannot be detected in negative ion mode (Table A.6). Unfortunately, polarity switching is not possible on the TripleTOF®5600+ system. Derivatization of E to yield derivatives which give good ionization yields in positive mode has sometimes been pursued to overcome this problem, yet was not considered in our study. Thus, another solution had to be found.

Table A.5. Reference values for T and E in adult men.

Testosterone in adult men, total	2.4 – 9.5 ng mL^{-1} *
Testosterone in study subjects (Immulin®)	0.5 – 12.0 ng mL^{-1}
Estradiol in adult men, total	8.0 – 35.0 pg mL^{-1} **
Estradiol in study subjects (Immulin®)	20 – 500 pg mL^{-1}

*Mayo Medical Laboratories:

<http://www.mayomedicallaboratories.com/test-updates/attachment.php?id=33420>

**Mayo Medical Laboratories:

<http://www.mayomedicallaboratories.com/test-updates/attachment.php?id=31374>

Table A.6. Direct infusion of analyte solution ($1,000 \text{ ng mL}^{-1}$) to find suitable electrospray ionization mode for $[\text{M}+\text{H}]^+$ or $[\text{M}-\text{H}]^-$ ions (ESI).

Analyte	Positive Mode	Negative Mode
Testosterone	✓ ✓ ✓	X
Estradiol	✓	✓ ✓

In ESI, an in-source fragment for E is generated by water cleavage in the positive ion mode (Fig. A.4). It can be seen, that the S/N is significantly improved for this in-source fragment (middle panel) as compared to the protonated precursor of E (top panel). The sensitivity (S/N) could be further improved by selecting the fragment with m/z 159.0804 from the in-source fragment $[\text{M}-\text{H}_2\text{O}+\text{H}]^+$ as precursor ion (bottom panel) for detection in positive ion mode (Fig. A.4). Furthermore, all source and gas parameters were optimized for maximal sensitivity for E (not shown; optimized settings reported in main document).

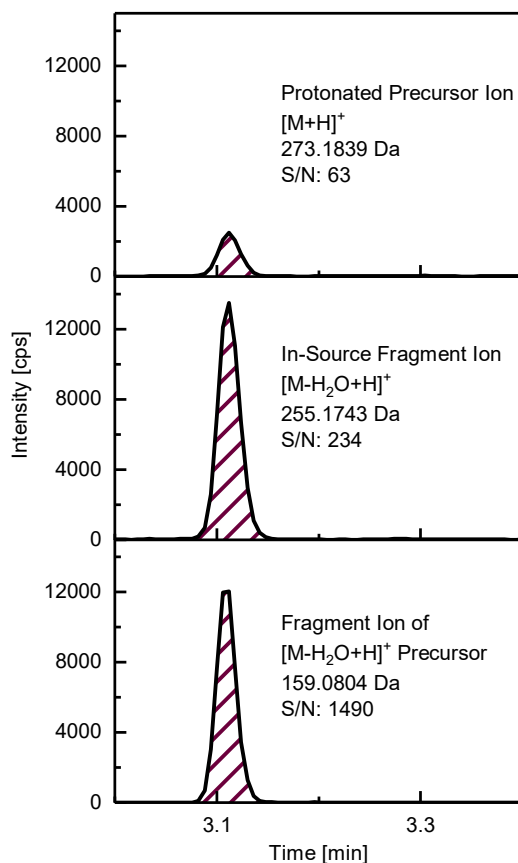


Fig. A.4. Selection of E ion for quantitative analysis and comparison of sensitivities. S/N ratios of precursor ion $[M+H]^+$, in-source fragment $[M-H_2O+H]^+$ and fragment of in-source fragment.

A further sensitivity gain for E was furnished by using the enhanced product ion mode (high sensitivity mode of MS/MS) with fragment enhancement for the fragment with m/z 159.0804 of the in-source fragment with m/z 255.1743 as precursor ion. The result is illustrated in Fig. A5A and Fig. A.5B. The intensity of the fragment m/z 159.0804 is increased by factor of about 3. For T even a higher increase of about 10-fold was achieved (Fig. A.5C and Fig. A.5D). The information about the precursor ions is unfortunately lost in the MS/MS spectra of the enhanced mode (Fig. A.5B and Fig. A.5D).

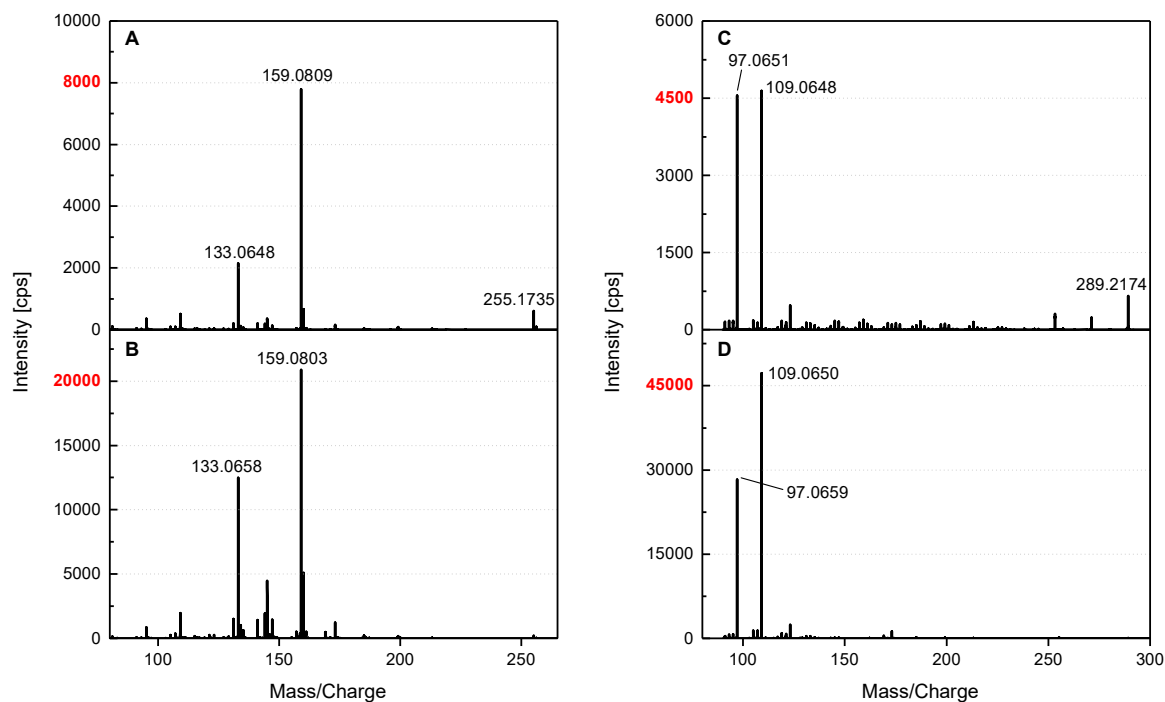


Fig. A.5. MS/MS spectra of E (precursor, in-source fragment with m/z 255.1743) without fragment enhancement (A) and with fragment enhancement on m/z 159.0804 (B). Fragment with m/z 133.0648 was not used as it showed insufficient specificity. MS/MS spectra of T (precursor, m/z 289.2162) without fragment enhancement (C) and with fragment enhancement on m/z 109.0648 (D).

A convenient strategy to improve sensitivity is to increase accumulation time (Fig. A.6). Unfortunately, the increase of the accumulation time comes at expense of the cycle time. When cycle times are too large, a sufficient number of data points across the peak is not possible anymore. As shown in Fig. A.7, when the accumulation time was increased from 50 to 100 ms the S/N ratio could be improved roughly by factor of 2. Further increase to 250 ms provided an S/N of only 102 (ca. 70 % related to 100 ms accumulation time). While both experiments, 50 and 100 ms accumulation time, yielded more than 10 data points per peak, the latter experiment with 250 ms did not provide a sufficient number of data points and thus the peak was not correctly described (signal was cut off).

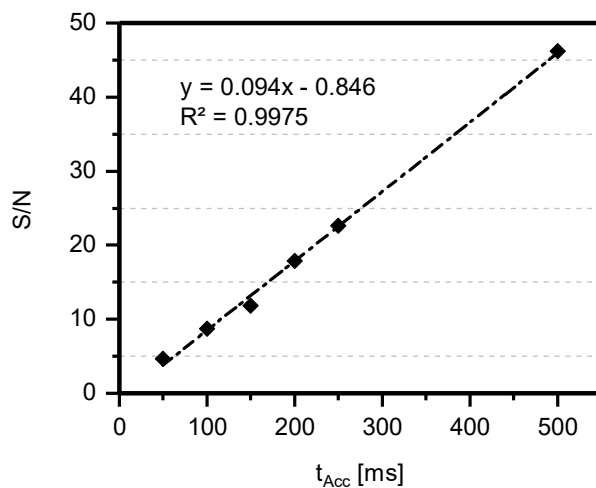


Fig. A.6. Effect of accumulation times on sensitivity (S/N) of E (10 pg mL^{-1}) in SWATH acquisition mode.

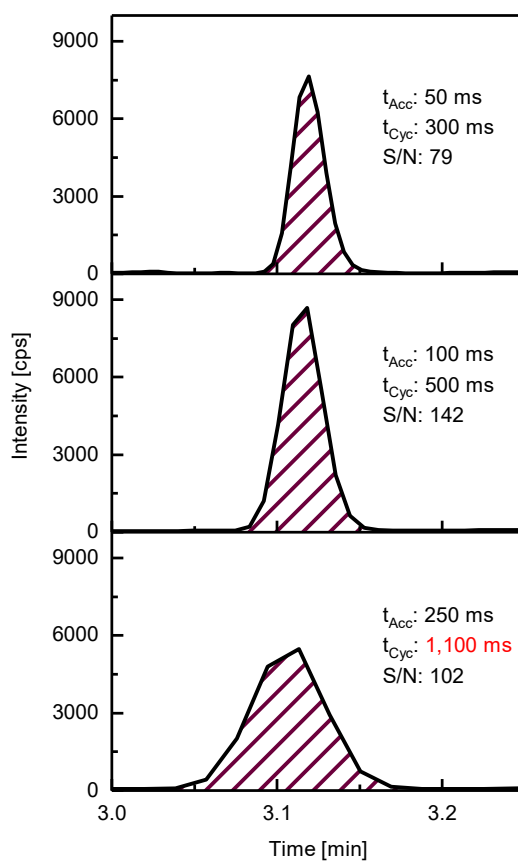


Fig. A.7. Effect of accumulation time (t_{Acc}) and cycle time (t_{Cyc}) on sensitivity of E (500 pg mL^{-1}) and data points across the peak (product ion mode). S/N values are estimates calculated with PeakView.

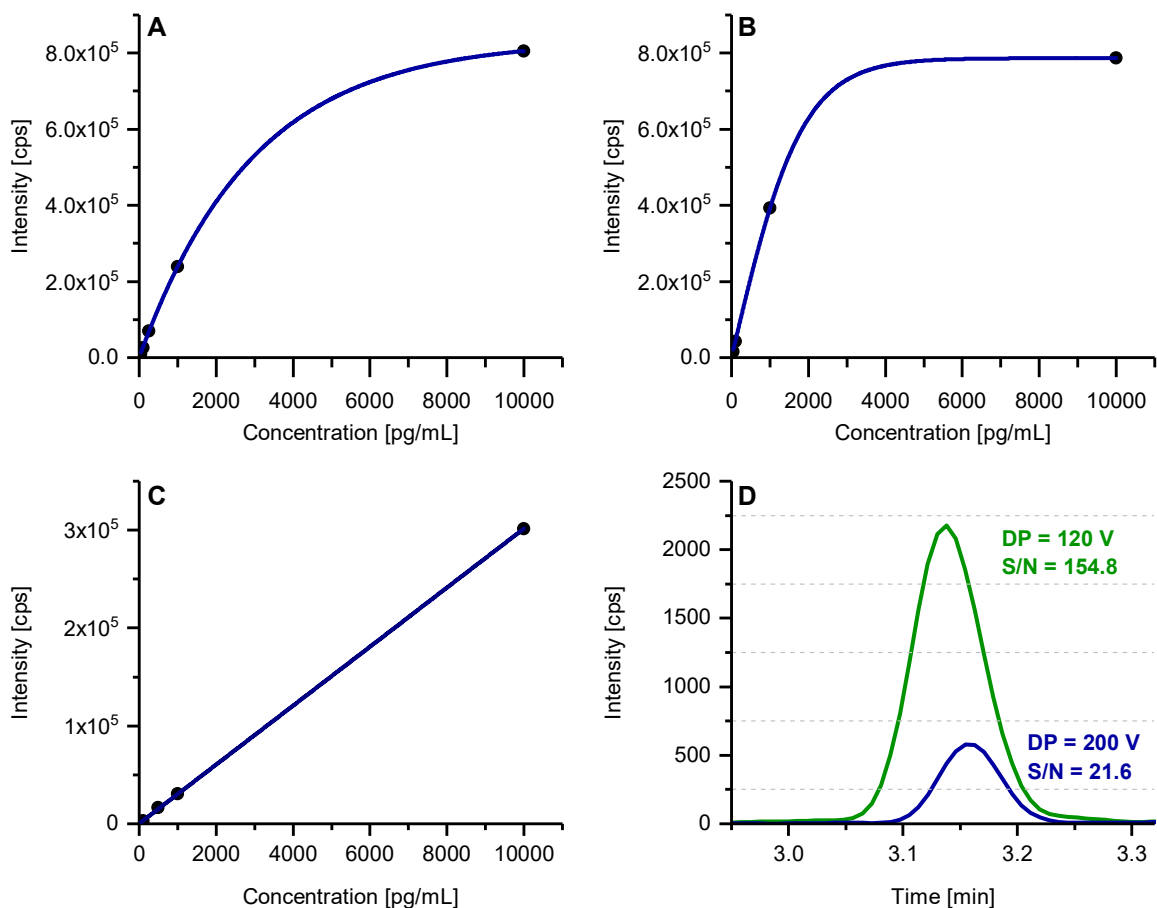


Fig. A.8. De-optimization of ion-source parameters (declustering potential DP) for T in order to have linearity in the relevant concentration range. (A) DP optimized for maximal sensitivity (120 V), (B) calibration function at DP 160 V slightly de-optimized, (C) DP de-optimized (200 V) so that linearity is observed in the relevant concentration range, and (D) overlay of XICs of T (50 pg mL^{-1}) at two distinct DPs, 200 V (blue) and 120 V (green).

With such resultant highly optimized conditions, both E and T can be detected at very low levels (ca. 10 pg mL^{-1}). Unfortunately, for T the linear range does not cover the relevant concentration range of the real samples under investigation. At concentrations above $1,000 \text{ pg mL}^{-1}$ the signal levels off due to detector saturation (Fig. A.8A). For the range of about $10 - 15,000 \text{ pg mL}^{-1}$ only a quadratic relationship between concentration and response ratio could be achieved. Consequently, for T the MS parameters needed to be de-optimized. While a number of options were available, the DP was selected for this purpose. Slightly increasing DP from 120 V to 160 V has only a minor effect (Fig. A.8B). When the DP was increased to 200 V, linearity was observed for T over the concentration range of 50 pg mL^{-1} to 10 ng mL^{-1} (Fig. A.8C). A comparison of XICs from MS/MS data at two distinct DP of 120 and 200 V is shown in Fig. A.8D. LOD and LOQ were significantly higher, yet sufficient for the present purpose.

Fig. A.9 shows the comparison of signal intensities obtained for different acquisition modes, MRM^{HR} and SWATH. MRM_{eq} represents the product-ion mode equivalent to equal accumulation time as in SWATH acquisition (300 ms for E and 50 ms for T). $\text{MRM}_{1/2}$ on the other hand is the product-ion mode with half of the accumulation time which is more realistic because it is the equivalent with same cycle time. The corresponding peak areas and S/N ratios can be found in Fig. A.9. Corresponding LODs of each method are listed in Table A.7.

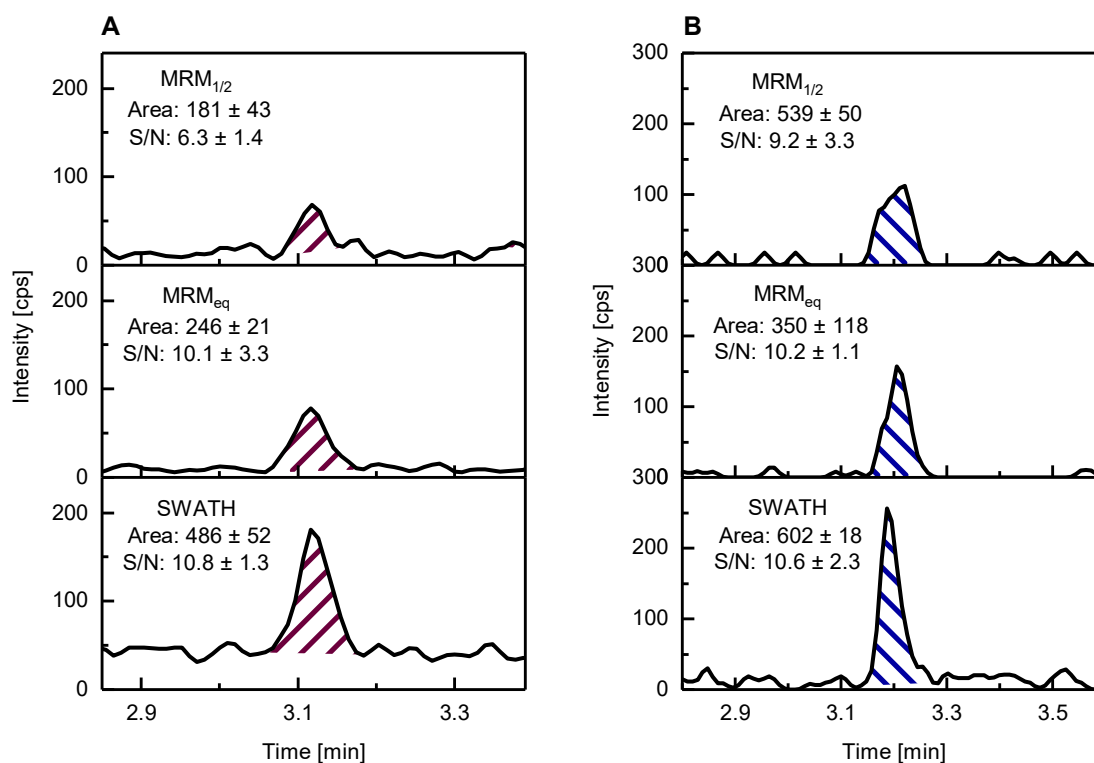


Fig. A.9. Comparison of signals obtained from MRM^{HR} and SWATH acquisition. A: E (20 pg mL⁻¹ in MeOH), B: T (40 pg mL⁻¹ in MeOH). (S/N values are estimates calculated with PeakView).

For determination of instrumental LODs, an 8-point calibration of both target analytes in MeOH, ranging from 20 pg mL⁻¹ to 200 pg mL⁻¹, was analysed in triplicate. Via the standard error of the intercept (se_b), instrumental LODs could be calculated from the calibration curve (see Table A.7). It is shown that the SWATH method achieved the lowest instrumental LODs even at equal t_{Acc} .

Table A.7. Comparison of instrumental LODs (in pg mL⁻¹) of MRM^{HR} and SWATH.^c

Method	Analyte	t _{Acc} [ms]	LOD
MRM _{1/2}	E	150	13.6
	T	25	15.3
MRM _{eq}	E	300	9.6
	T	50	14.1
SWATH	E	300	5.8
	T	50	8.1

^cInstrumental LODs were calculated by $3x se_b$ divided by the slope of the calibration curve. The mandatory TOF-MS experiment had a t_{Acc} of 20 ms in all experiments. To reach uniform t_{Cyc}, additional experiments were added in the MRM_{1/2} method. For associated chromatograms see Fig. A.9.

Method Validation and Calibration

Assay Specificity

Widening the Q1 windows during precursor selection in SWATH acquisition (as compared to the product ion scan mode with its unit mass precursor selection for fragmentation) raises the risk for interferences in the TOF readout. Stringent control and validation of assay specificity is therefore necessary. A number of potential interferences with the same sum formula as T can be found in the LipidMaps database (Table A.8). Amongst them, epitestosterone (epiT) is a likely interference due to the fact that it shows the same fragments as well. It is an important steroid from an analytical perspective, in particular in doping control, as it is monitored as well and used to derive the T/epiT ratio. At T/epiT levels larger than 4/1 further investigation for potential T abuse is conducted. For this reason, a mixture of T and epiT was injected. As can be seen in Fig. A.10, epiT elutes at different retention time and therefore does not cause any problems. Similarly, epiestradiol (epiE) is a potential interference but was also eluted at different retention time compared to E. Other isomers of Table A.9 are expected to show different fragmentations and/or retention time.

Table A.8. Potential interferences for T (isomeric structures from LipidMaps database search).

Common Name	Systematic name	Formula	Mass
Testosterone	17beta-hydroxyandrost-4-en-3-one	C ₁₉ H ₂₈ O ₂	288.2089
Dehydroepiandrosterone	3beta-hydroxyandrost-5-en-17-one	C ₁₉ H ₂₈ O ₂	288.2089
17 α -Testosterone (Epitestosterone)	17alpha-hydroxyandrost-4-en-3-one	C ₁₉ H ₂₈ O ₂	288.2089
-	5beta-androstane-3,17-dione	C ₁₉ H ₂₈ O ₂	288.2089
-	5alpha-androstane-3,17-dione	C ₁₉ H ₂₈ O ₂	288.2089
Dehydroandrosterone	-	C ₁₉ H ₂₈ O ₂	288.2089
1-Testosterone	17beta-hydroxy-5alpha-androst-1-en-3-one	C ₁₉ H ₂₈ O ₂	288.2089

Table A.9. Potential interferences for E (isomeric structures from LipidMaps database search).^d

Common Name	Systematic name	Formula	Mass
-	estra-5,7,9-triene-3beta,17beta-diol	C ₁₈ H ₂₄ O ₂	272.1776
-	8alpha-estra-1,3,5(10)-trien-3,17beta-diol	C ₁₈ H ₂₄ O ₂	272.1776
-	estra-5,7,9-triene-3alpha,17alpha-diol	C ₁₈ H ₂₄ O ₂	272.1776
17 α -Estradiol (Epiestradiol)	estra-1,3,5(10)-triene-3,17alpha-diol	C ₁₈ H ₂₄ O ₂	272.1776

^dA database search (LipidMaps, HMDB) for potential interference for the [M-H₂O+H]⁺ precursor (C₁₈H₂₂O) did not lead to any results.

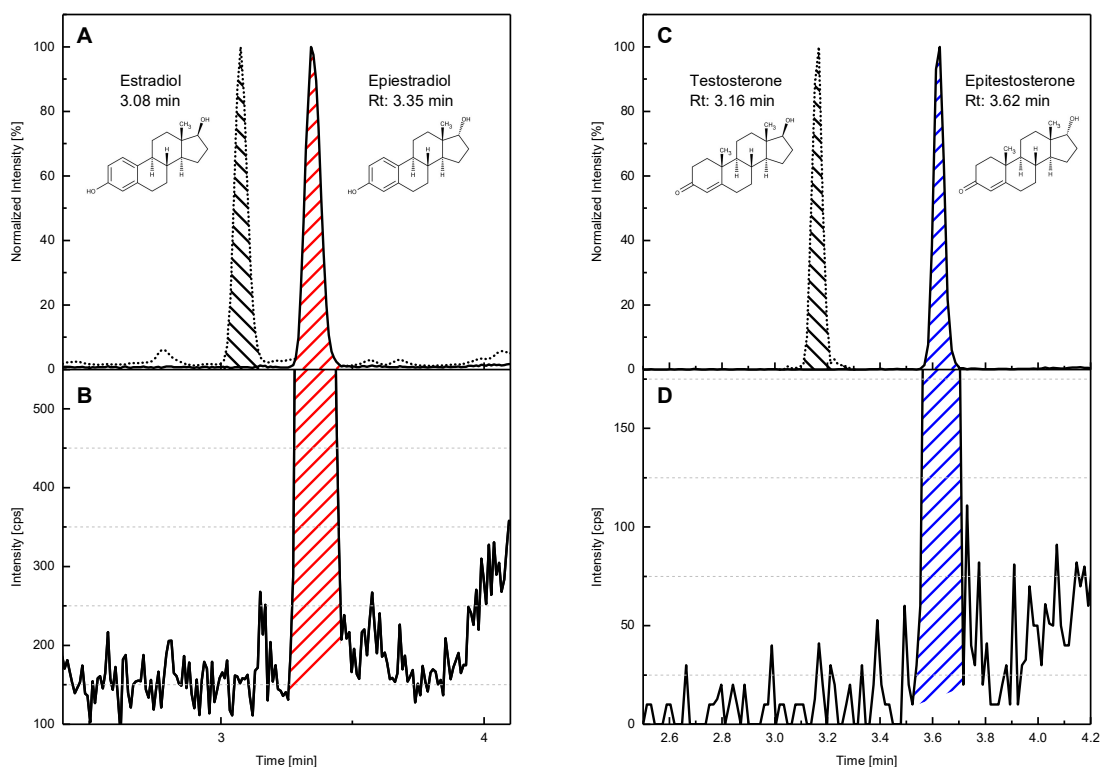


Fig. A.10. Assay specificity testing for steroid epimers. Injection of E & epiE (A) and T & epiT (C) shows baseline separation of peaks in both cases. Injection of only epiE (B) or only epiT (D) shows no interaction for target analytes.

Assay specificity of SWATH was compared to MRM^{HR} (product ion mode with same fragment as used for SWATH, other mass spectrometric parameters remained identical to SWATH method). It can be seen in Fig. A.11 that no interference is observed in the XIC traces of T, ¹³C₃-T and d₅T upon injection of the respective other targets in MRM^{HR} in which Q1 precursor selection occurs with unit mass. The situation is different for SWATH. When precursor selection occurred with 8 Da windows, a signal for d₅T was monitored in the ¹³C₃-T XIC trace, and likewise for T and ¹³C₃-T in the XIC of d₅T (Fig. A.12) demonstrating insufficient assay specificity when the Q1 window for precursor selection was too wide in SWATH. However, when the Q1 window was narrowed to 4 Da (5 Da for d₅T), no interferences were observed (see Fig. A.13). The situation was the same for E (see Fig. A.14, Fig. A.15 and Fig. A.16).

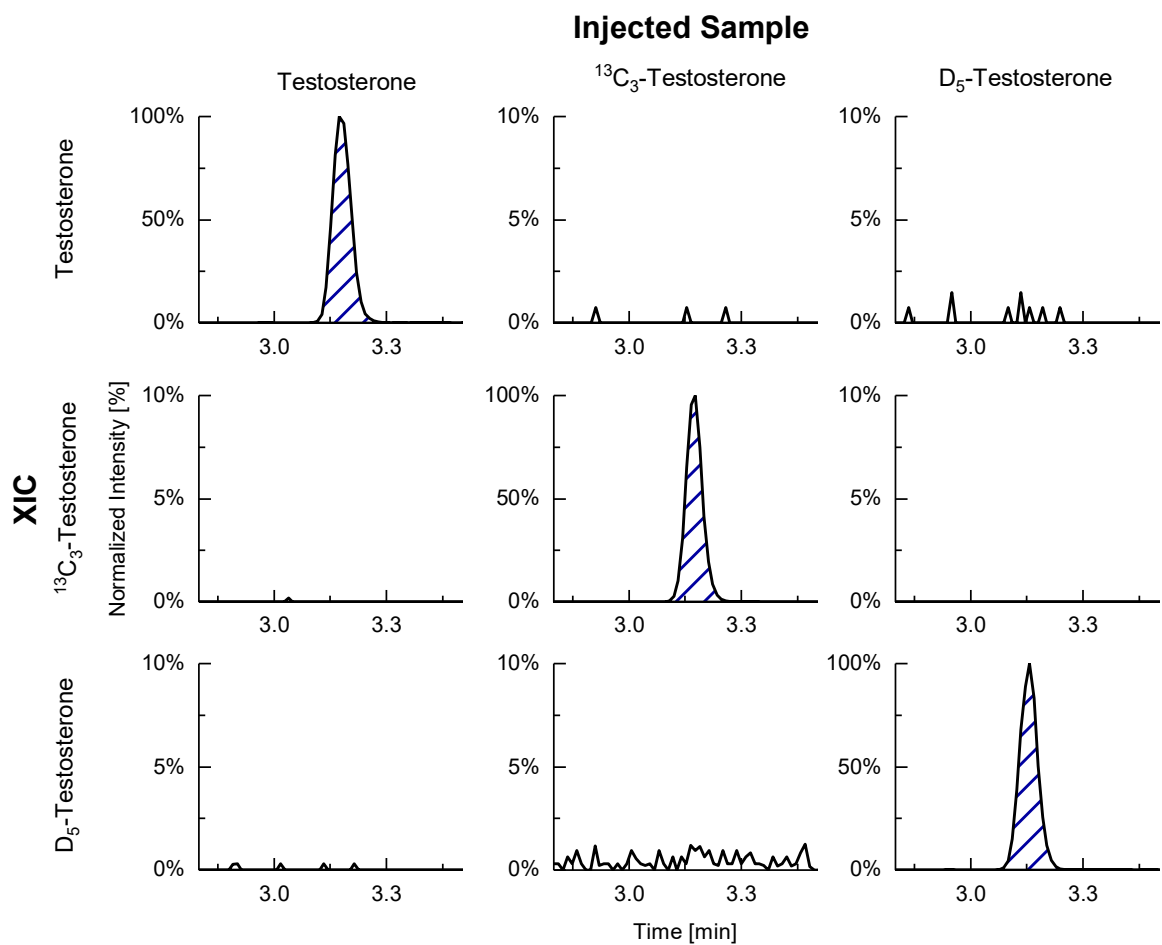


Fig. A.11. T: MRM^{HR} for every transition (5 ng mL⁻¹ per analyte).

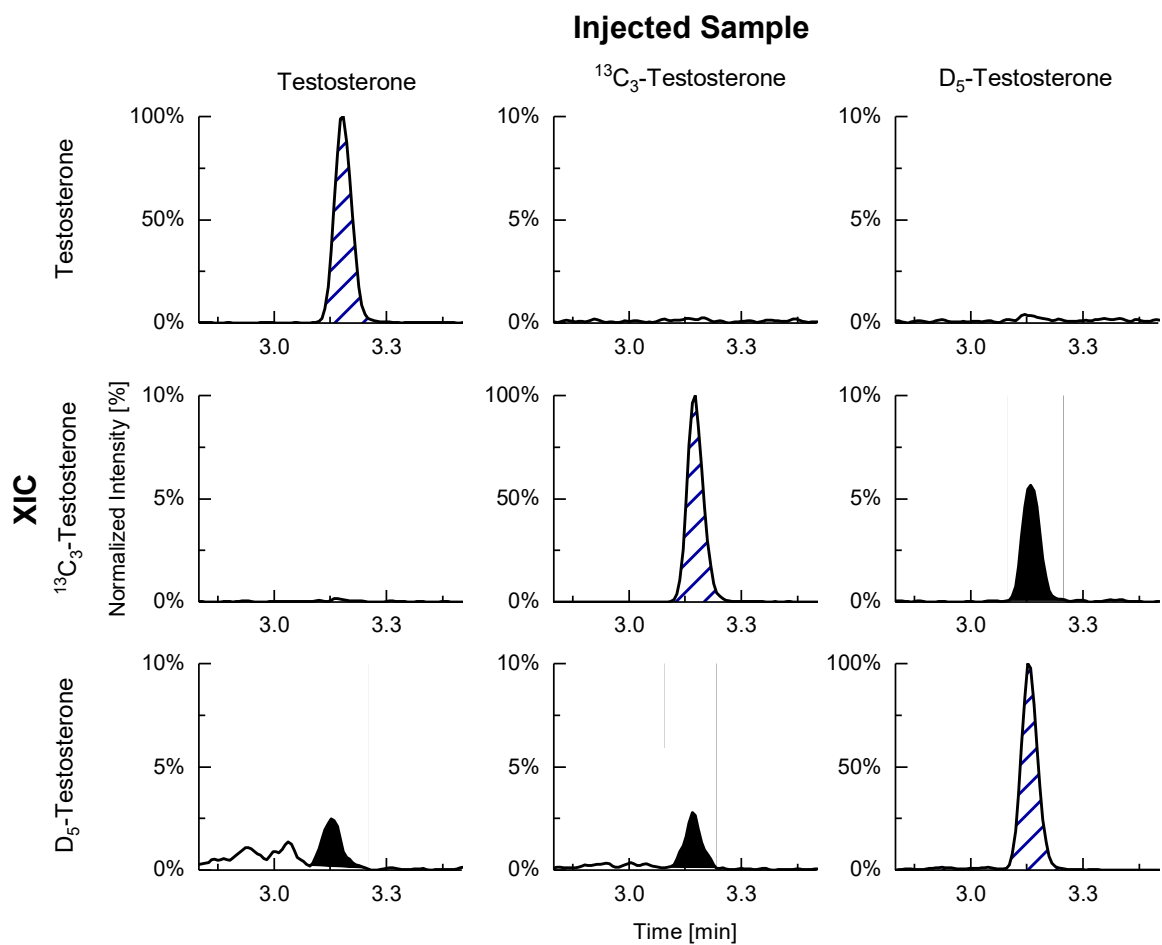


Fig. A.12. T: SWATH with 8 Da window for T, ¹³C₃T & d₅T (5 ng mL⁻¹ per analyte). Interfering signals are marked with a black filling.

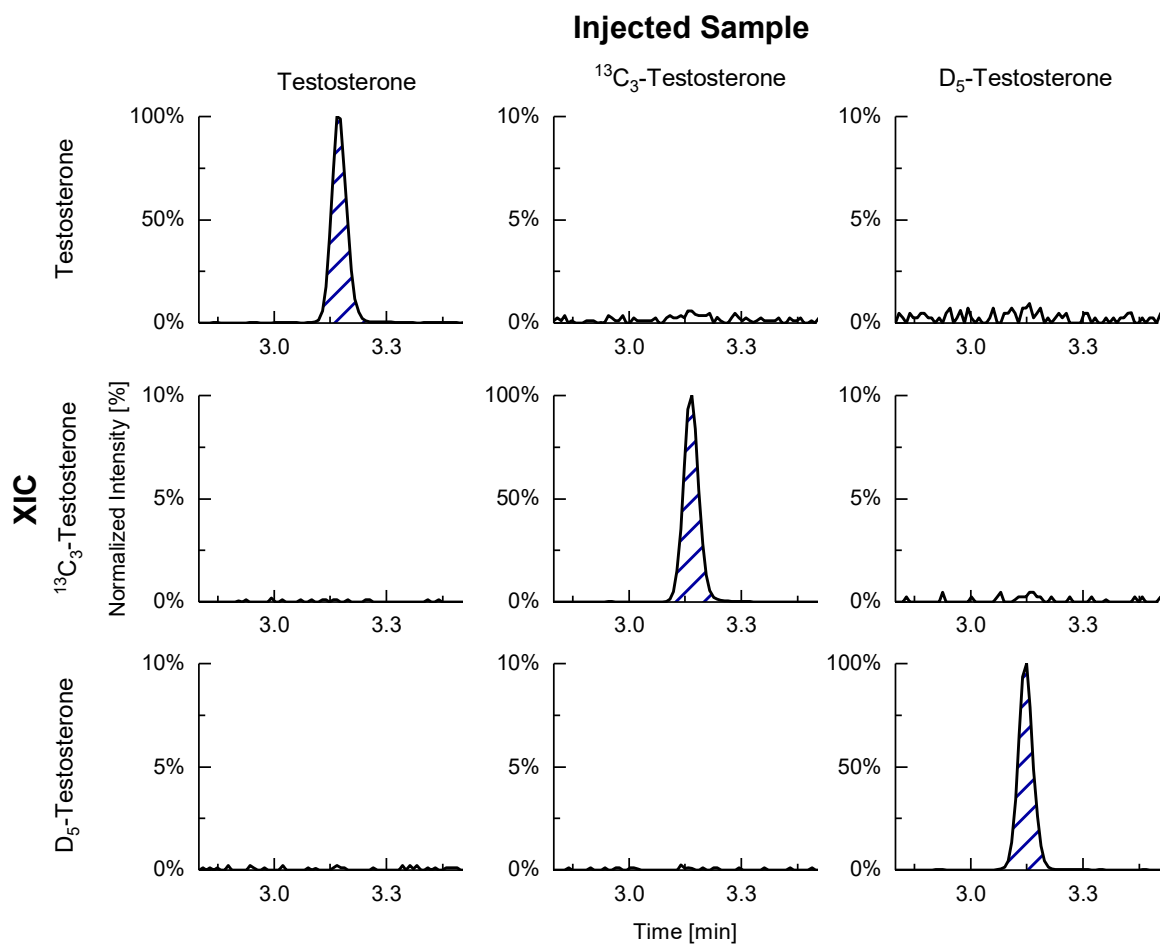


Fig. A.13. T: SWATH with 4 Da window for T & ¹³C₃T, 5 Da window for d₅T (5 ng mL⁻¹ per analyte).

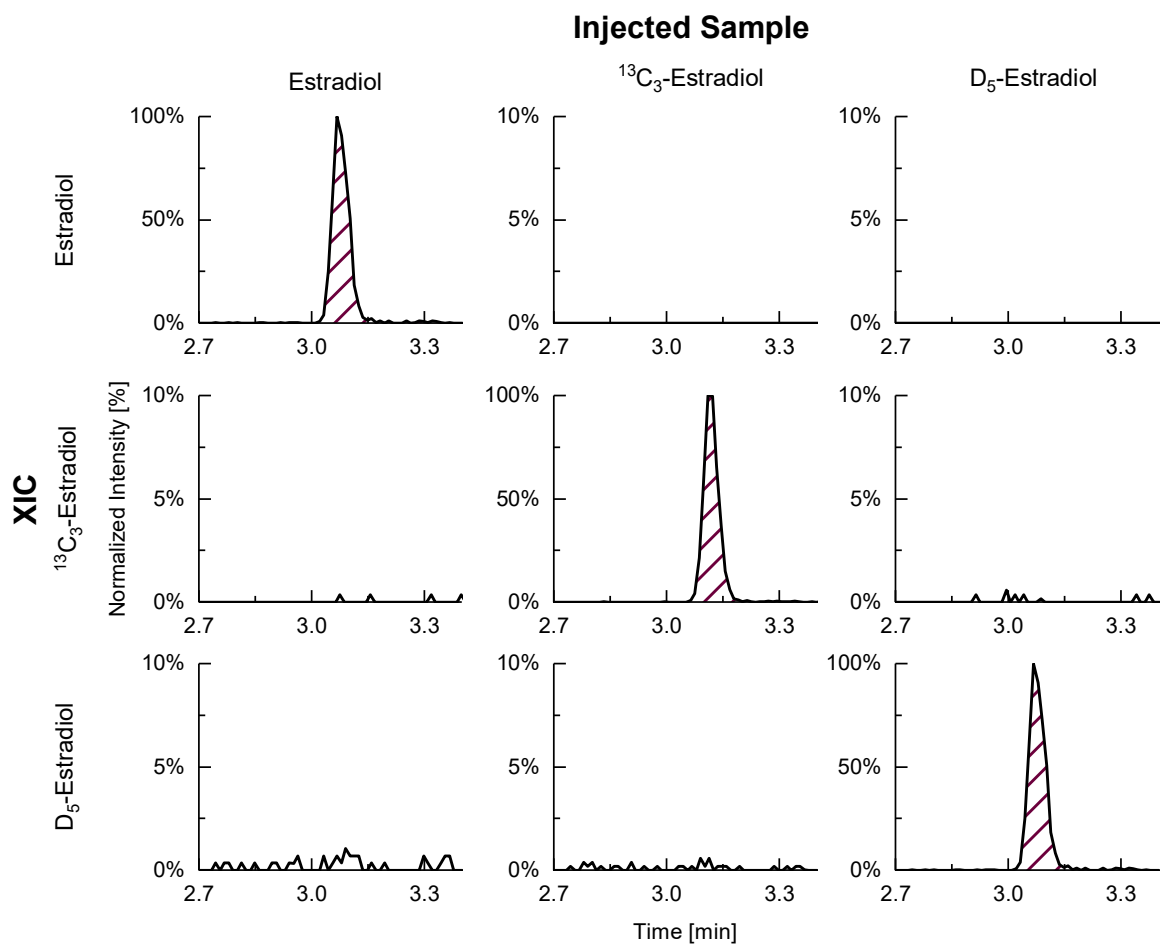


Fig. A.14. E: MRM^{HR} for every transition (5 ng mL⁻¹ per analyte).

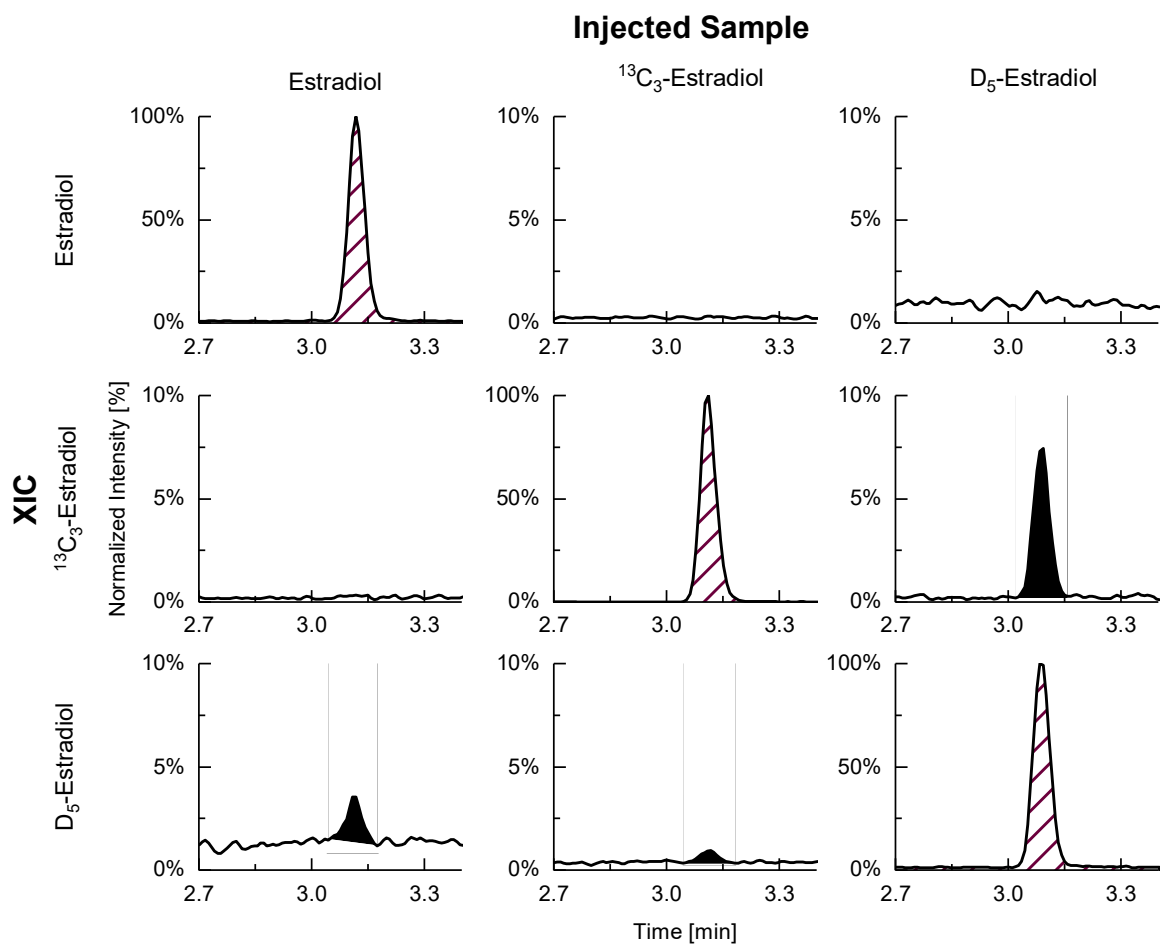


Fig. A.15. E: SWATH with 8 Da window for E, ¹³C₃E & d₅E (5 ng mL⁻¹ per analyte). Interfering signals are marked with a black filling.

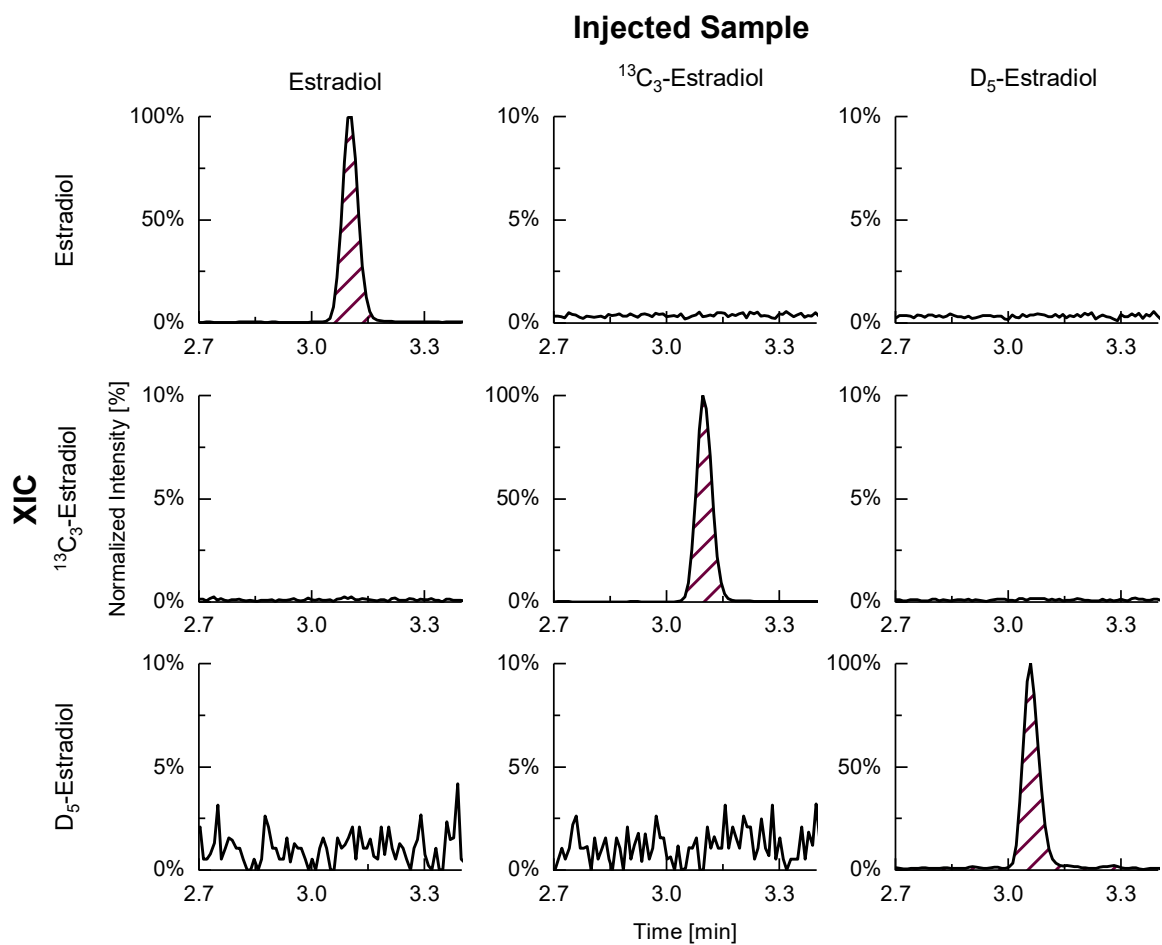


Fig. A.16. E: SWATH with 4 Da window for E & ¹³C₃E, 5 Da window for d₅E (5 ng mL⁻¹ per analyte).

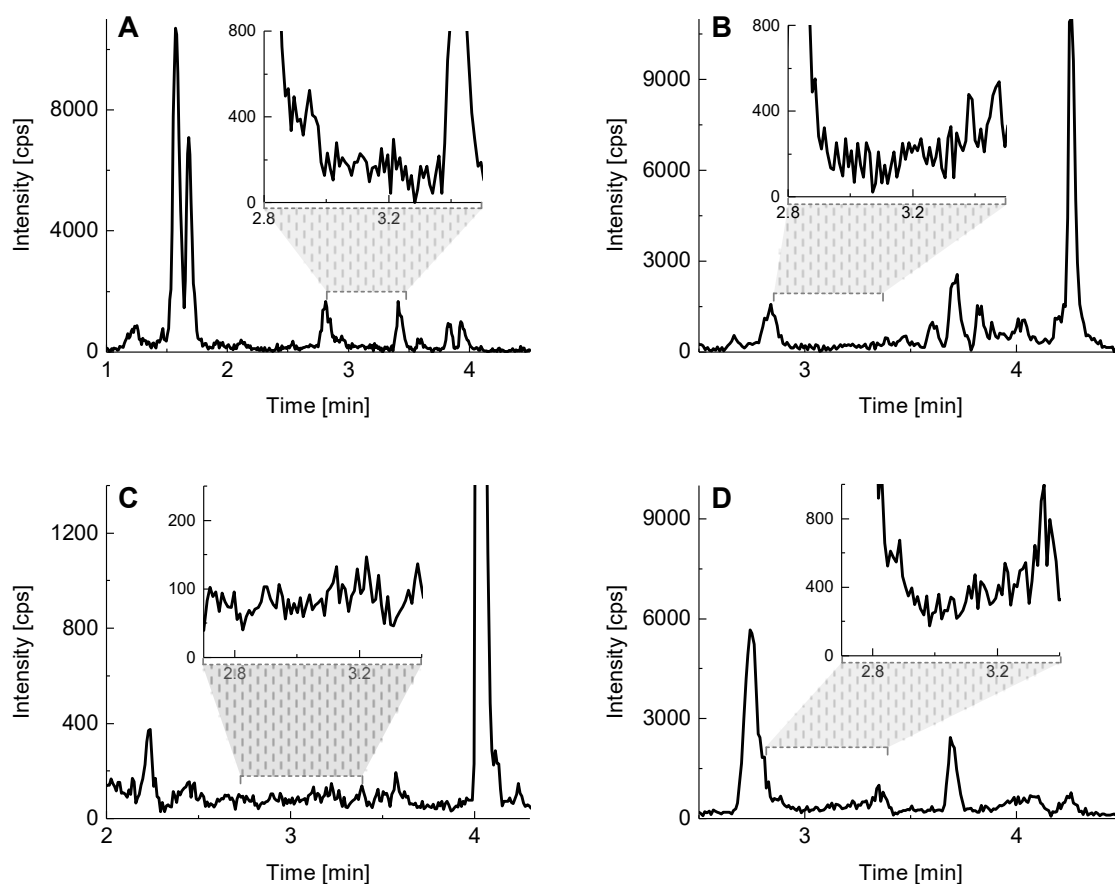


Fig. A.17. Specificity in blank matrix. XICs of fragments of $^{13}\text{C}_3\text{T}$ (a), $d_5\text{T}$ (b), $^{13}\text{C}_3\text{E}$ (c) and $d_5\text{E}$ (d) are shown in a blank matrix sample. Relevant regions of retention time are shown in a zoomed-in window, respectively.

Calibration by the Surrogate Calibrant Approach

By spiking equal amounts of target analyte and $^{13}\text{C}_3$ -standard into plasma, two calibration curves are generated: one shows the standard addition of the target analyte and the other shows the calibration of the $^{13}\text{C}_3$ -standard (Fig. A.18; Table A.10). In order to use the surrogate calibration for quantification, both curves must be parallel after initial matching of responses, which was done by adjusting concentrations of surrogate calibrants (Table A.1 and Table A.2). Parallelism for both substance pairs was controlled for during the whole study, including validation and sample measurements.

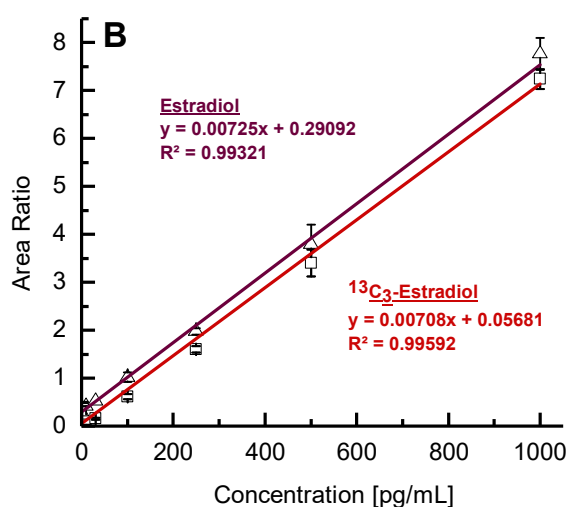
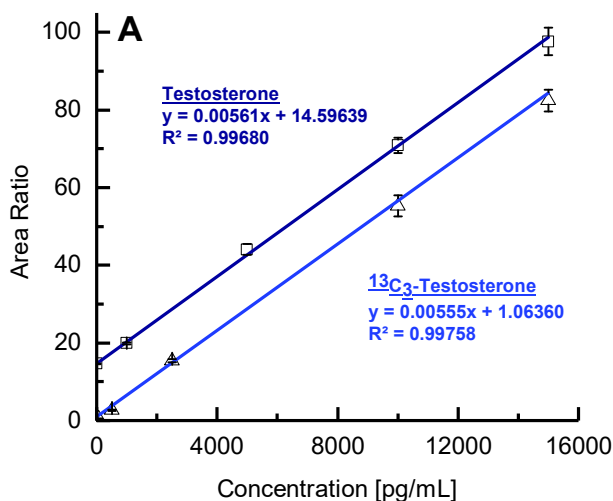


Fig. A.18. Parallelism of standard addition curve and corresponding surrogate calibrant curve for (A) T & $^{13}\text{C}_3\text{T}$, (B) E & $^{13}\text{C}_3\text{E}$. Following concentrations were spiked into plasma calibrants, respectively: T: 0/20/60/1,000/5,000/10,000/15,000 pg mL⁻¹; $^{13}\text{C}_3\text{T}$: 20/60/500/2,500/10,000/15,000 pg mL⁻¹; E: 0/10/30/100/250/500/1000 pg mL⁻¹; $^{13}\text{C}_3\text{E}$: 10/30/100/250/500/1000 pg mL⁻¹.

Table A.10. Control for parallelism during validation.^e

Analyte	Day	Slope surrogate calibrant	Slope target analyte	Slope ratio	Slope deviation [%]
Estradiol	1	0.00708	0.00725	0.977	2.3
	2	0.00895	0.00867	1.032	3.2
	3	0.00775	0.00759	1.021	2.1
Testosterone	1	0.00555	0.00561	0.989	1.1
	2	0.00600	0.00623	0.963	3.7
	3	0.00505	0.00505	1.000	0.0

^eA slope ratio ($^{13}\text{C}_3$ -standard/target analyte standard) of 1.00 ± 0.05 was deemed acceptable.

Matrix Effects, Extraction Recoveries and Process Efficiencies

Matrix Effect by Post-Column Infusion

Besides validation of matrix effects, extraction recoveries and process efficiencies by the protocol proposed by Matuszewski discussed in the main document, matrix effects were also investigated by post-column infusion experiments. $^{13}\text{C}_3$ -labelled surrogate calibrants and d_5 -labelled internal standard solutions were infused via a T-piece to the column effluent and ionized in the ESI source. The XICs of the corresponding solutes were monitored (Fig. A.19 and Fig. A.20). It can be seen that at the relevant retention time of the analyte signals are not suppressed significantly indicating negligible or minor matrix effects.

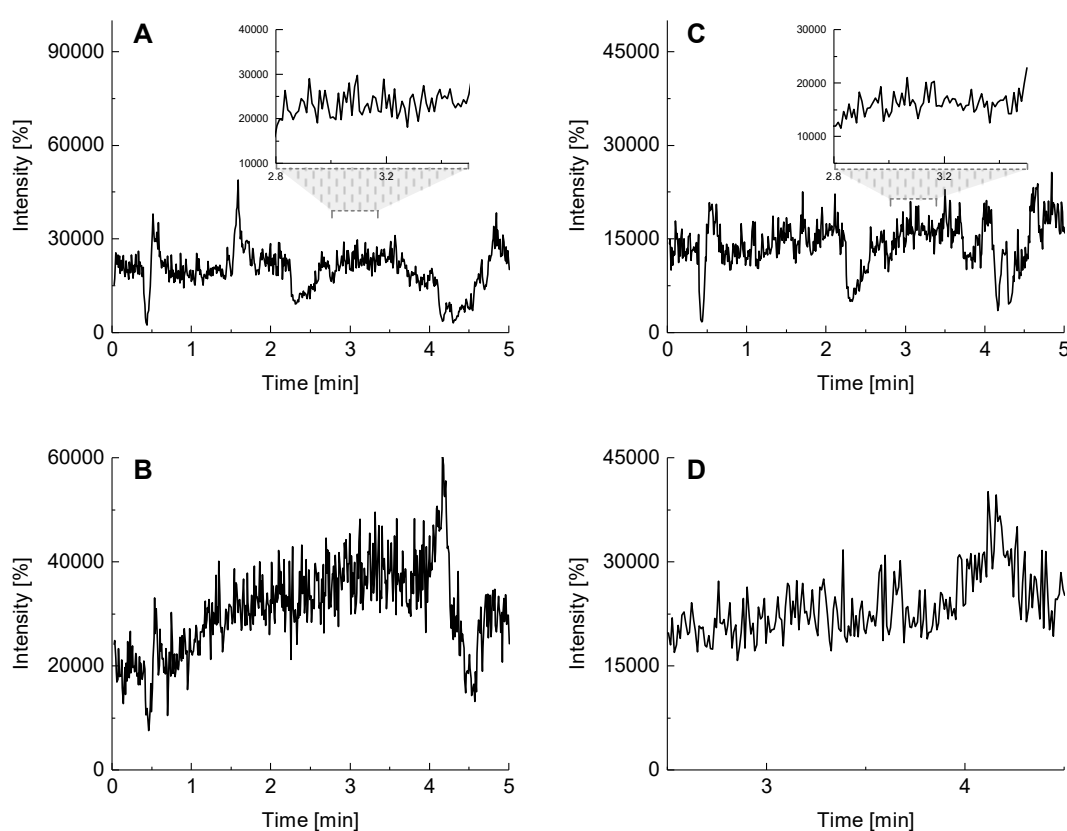


Fig. A.19. Post-column infusion of $^{13}\text{C}_3\text{T}$ for plasma (A) and MeOH (B) and post-column infusion of $d_5\text{T}$ for plasma (C) and MeOH (D). Relevant regions of retention time are shown in zoomed-in window. XICs of infused substance for relevant fragment mass are shown, respectively.

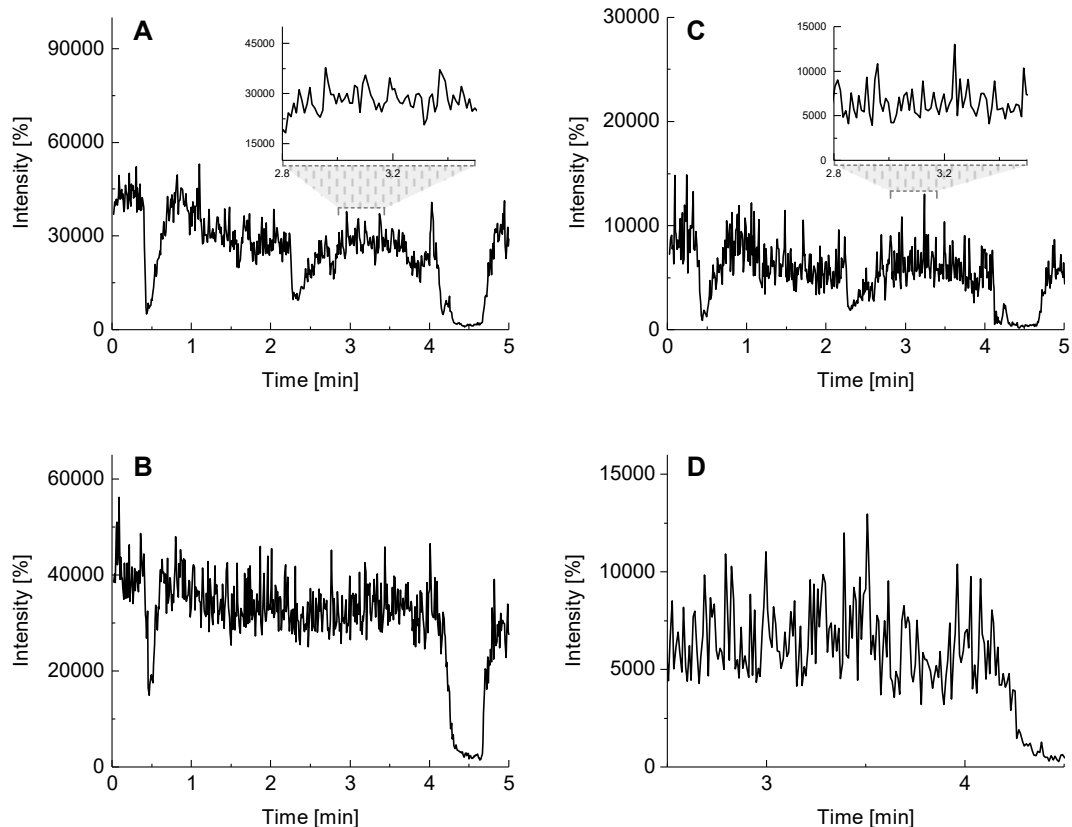


Fig. A.20. Post-column infusion of $^{13}\text{C}_3\text{E}$ for plasma (a) and MeOH (b) and post-column infusion of $d_5\text{E}$ for plasma (c) and MeOH (d). Relevant regions of retention time are shown in zoomed-in window. XICs of infused substance for relevant fragment mass are shown, respectively.

Stability

Different types of stabilities were validated. Freeze-thaw stability was assessed by freezing freshly prepared QC_{low} and QC_{high} samples ($n = 3$) for 24 h. Three cycles of thawing and freezing for 4 h, respectively, were done and samples were analyzed after final thawing. Furthermore, long term stability of samples ($n = 3$) that were stored for 6 months at -20°C was tested. Short term stability was validated by analyzing freshly prepared samples that were frozen for 24 h, thawed and kept at room temperature (25°C) for 6 h. The respective response ratios were compared to the response ratios of freshly prepared QC_{low} and QC_{high} samples. By leaving these freshly prepared samples in the autosampler at 4°C for 10 h, also post-preparative stability could be assessed. It can be seen that all stability assessments are within acceptable limits (Table A.11) Since freeze-thaw stability of $^{13}\text{C}_3\text{E}$ showed an accuracy of only 80.9 % for the high level QC, multiple freeze-thaw cycles were avoided.

Table A.11. Validation of stability.^f

Sample	QC	¹³ C ₃ -Estradiol		¹³ C ₃ -Testosterone	
		Accuracy [%]	RSD [%]	Accuracy [%]	RSD [%]
Fresh QC	low	100.0	7.0	100.0	3.5
	high	100.0	5.6	100.0	8.3
Freeze-thaw QC	low	93.5	13.0	87.8	8.1
	high	80.9	1.6	88.8	15.6
Long term QC	low	105.1	8.8	98.4	10.8
	high	89.6	8.4	100.0	0.2
Short term QC	low	103.4	21.9	90.9	17.3
	high	90.3	15.4	86.5	4.8
Post-prep. QC	low	98.8	4.2	91.0	6.6
	high	102.0	2.5	89.1	1.8

^fResponse ratios of freshly prepared QCs were set to 100 % accuracy for comparison.

Application

Results from Study Samples

The method was applied to measure plasma concentrations of T and E in a clinical study in which healthy young men received a 3-day treatment with transdermal E (Estradot 50 patches, Novartis Pharma, Nuremberg, Germany, 100 µg / 24 h) or placebo and the respective effect on calorie intake was monitored. Sixty-four sets of samples obtained at 5 different time points were quantified, so that the total sample number was 320.

Table A.12. Analysis results.^g

Sample ID	Estradiol [pg mL ⁻¹]	Testosterone [pg mL ⁻¹]	Estradiol Patch	Insulin treatment
10A10	47.6 ± 4.2	7097.9 ± 380.8	X	X
10A2	48.4 ± 4.2	8411.7 ± 451.2	X	X
10A4	52.1 ± 4.5	9490.5 ± 508.9	X	X
10A6	52.4 ± 4.5	9721.8 ± 521.3	X	X
10A8	52.1 ± 4.5	8319.4 ± 446.2	X	X
10B10	35.5 ± 3.3	8947.3 ± 479.8	X	✓
10B2	30.4 ± 3.0	8161.9 ± 437.8	X	✓
10B4	32.9 ± 3.1	9974.5 ± 534.9	X	✓
10B6	42.1 ± 3.8	10770.2 ± 577.5	X	✓
10B8	505.6 ± 40.8	9966.3 ± 534.4	X	✓
11A10	65.3 ± 5.5	1671.6 ± 90.2	✓	✓
11A2	54.7 ± 4.7	2244.9 ± 120.8	✓	✓
11A4	87 ± 7.2	2314.5 ± 124.6	✓	✓
11A6	69.3 ± 5.8	2145.2 ± 115.5	✓	✓
11A8	69 ± 5.8	2033.9 ± 109.5	✓	✓
11B10	57.9 ± 4.9	2030.3 ± 109.4	✓	X
11B2	50.9 ± 4.4	1561.4 ± 84.3	✓	X
11B4	65.7 ± 5.5	1617.6 ± 87.3	✓	X
11B6	66.3 ± 5.6	1437.8 ± 77.7	✓	X
11B8	64.7 ± 5.4	1555.2 ± 83.9	✓	X
12A10	19.6 ± 2.3	4271.4 ± 229.4	X	✓
12A2	22.1 ± 2.4	3061.8 ± 164.6	X	✓
12A4	23.8 ± 2.5	3517.3 ± 189.0	X	✓
12A6	19.8 ± 2.3	4174.8 ± 224.2	X	✓

12A8	27.1 ± 2.7	5277.5 ± 283.2	X	√
12B10	23.4 ± 2.5	4139.4 ± 222.3	X	X
12B2	23.3 ± 2.5	5733.4 ± 307.7	X	X
12B4	24.9 ± 2.6	6577.0 ± 352.9	X	X
12B6	20.3 ± 2.3	7191.8 ± 385.8	X	X
12B8	19.8 ± 2.3	6225.7 ± 334.0	X	X
13A10	92.1 ± 7.6	3226.4 ± 173.4	√	X
13A2	115.3 ± 9.4	2741.0 ± 147.4	√	X
13A4	104.2 ± 8.5	3264.3 ± 175.4	√	X
13A6	120.6 ± 9.8	2922.8 ± 157.1	√	X
13A8	149.7 ± 12.2	3592.0 ± 193.0	√	X
13B10	53.9 ± 4.6	1516.1 ± 81.8	√	√
13B2	69.5 ± 5.8	2027.4 ± 109.2	√	√
13B4	71.3 ± 6.0	2308.7 ± 124.3	√	√
13B6	68.8 ± 5.8	2193.8 ± 118.1	√	√
13B8	67.3 ± 5.6	2682.1 ± 144.2	√	√
14A10	16.7 ± 2.2	10070.5 ± 540	X	X
14A2	12.0 ± 2.0	7464.4 ± 400.4	X	X
14A4	19.1 ± 2.3	10787.2 ± 578.4	X	X
14A6	17.0 ± 2.2	11680.3 ± 626.3	X	X
14A8	18.0 ± 2.2	10797 ± 578.9	X	X
14B10	22.4 ± 2.5	11041.5 ± 592	X	√
14B2	29.6 ± 2.9	12224.4 ± 655.4	X	√
14B4	25.9 ± 2.7	12542 ± 672.4	X	√
14B6	21.8 ± 2.4	11925.7 ± 639.4	X	√
14B8	21.6 ± 2.4	12028.7 ± 644.9	X	√
16A10	16.2 ± 2.1	6354.0 ± 340.9	X	X
16A2	15.7 ± 2.1	6403.3 ± 343.6	X	X
16A4	15.2 ± 2.1	6714.9 ± 360.2	X	X
16A6	15.5 ± 2.1	6120.4 ± 328.4	X	X
16A8	17.0 ± 2.2	6390.3 ± 342.9	X	X
16B10	23.1 ± 2.5	5855 ± 314.2	X	√
16B2	21.1 ± 2.4	7207.0 ± 386.6	X	√
16B4	19.8 ± 2.3	5940.9 ± 318.8	X	√
16B6	26.4 ± 2.7	6755.7 ± 362.4	X	√
16B8	26.8 ± 2.7	7047.3 ± 378.1	X	√
18A10	95.4 ± 7.8	1840.7 ± 99.2	√	√
18A2	104.6 ± 8.6	3467.5 ± 186.3	√	√
18A4	102.2 ± 8.4	2749.7 ± 147.9	√	√
18A6	112.0 ± 9.2	2933.7 ± 157.7	√	√
18A8	112.3 ± 9.2	3171.1 ± 170.4	√	√
18B10	128.6 ± 10.5	636.1 ± 34.9	√	X
18B2	242.1 ± 19.6	1169.3 ± 63.3	√	X
18B4	203 ± 16.4	974.1 ± 52.9	√	X
18B6	178.7 ± 14.5	1137.6 ± 61.6	√	X
18B8	199.7 ± 16.2	1110.8 ± 60.2	√	X
19A10	49.6 ± 4.3	828.7 ± 45.1	√	X
19A2	73.8 ± 6.1	1825.5 ± 98.4	√	X
19A4	50.9 ± 4.4	2035.7 ± 109.6	√	X
19A6	71.1 ± 5.9	1892.9 ± 102	√	X
19A8	59.4 ± 5.0	1652.2 ± 89.1	√	X
19B10	38.7 ± 3.5	1062 ± 57.6	√	√
19B2	45.8 ± 4.0	1443.2 ± 77.9	√	√
19B4	49.8 ± 4.3	1615.8 ± 87.2	√	√
19B6	42.1 ± 3.8	1433.8 ± 77.4	√	√
19B8	76.7 ± 6.4	1612.3 ± 87	√	√
1A10	77.2 ± 6.4	1275.7 ± 69	√	√
1A2	79.7 ± 6.6	2192.5 ± 118	√	√
1A4	85.0 ± 7.0	1857.7 ± 100.1	√	√
1A6	86.2 ± 7.1	1987.7 ± 107.1	√	√
1A8	83.7 ± 6.9	1712.4 ± 92.3	√	√
1B10	72.3 ± 6	1663.6 ± 89.7	√	X
1B2	68.3 ± 5.7	2783.8 ± 149.7	√	X
1B4	76.9 ± 6.4	3103 ± 166.8	√	X
1B6	70.8 ± 5.9	2774.9 ± 149.2	√	X
1B8	74.9 ± 6.2	2109.5 ± 113.6	√	X
20A10	20 ± 2.3	4568.1 ± 245.3	X	√
20A2	20.1 ± 2.3	5476.4 ± 293.9	X	√

20A4	17.5 ± 2.2	5310.5 ± 285	X	✓
20A6	19 ± 2.3	4349.3 ± 233.5	X	✓
20A8	20.1 ± 2.3	4573.5 ± 245.5	X	✓
20B10	27.6 ± 2.8	5931.7 ± 318.3	X	X
20B2	28.7 ± 2.9	6753.7 ± 362.3	X	X
20B4	26.8 ± 2.7	6711.8 ± 360.1	X	X
20B6	29.4 ± 2.9	7447.0 ± 399.5	X	X
20B8	24.1 ± 2.6	6726.5 ± 360.9	X	X
23A10	15.5 ± 2.1	6372.3 ± 341.9	X	✓
23A2	20.5 ± 2.4	6503 ± 348.9	X	✓
23A4	24.3 ± 2.6	7076.5 ± 379.6	X	✓
23A6	21.6 ± 2.4	7674.1 ± 411.6	X	✓
23A8	19.8 ± 2.3	8009.3 ± 429.6	X	✓
23B10	23.8 ± 2.5	7128.9 ± 382.4	X	X
23B2	28.7 ± 2.9	8140.3 ± 436.6	X	X
23B4	30.7 ± 3.0	8882.0 ± 476.3	X	X
23B6	33.7 ± 3.2	10389.0 ± 557.1	X	X
23B8	41.5 ± 3.7	10778.7 ± 578	X	X
24A10	66.1 ± 5.6	540.7 ± 29.9	✓	X
24A2	90.3 ± 7.4	300.4 ± 17.3	✓	X
24A4	73.9 ± 6.2	368.5 ± 20.8	✓	X
24A6	75.8 ± 6.3	297.8 ± 17.2	✓	X
24A8	75.1 ± 6.3	368.7 ± 20.8	✓	X
24B10	41.3 ± 3.7	911.2 ± 49.5	✓	✓
24B2	36.0 ± 3.3	824.2 ± 44.9	✓	✓
24B4	47.6 ± 4.2	824.2 ± 44.9	✓	✓
24B6	38.2 ± 3.5	718.7 ± 39.3	✓	✓
24B8	38.3 ± 3.5	701.7 ± 38.4	✓	✓
25A10	16.8 ± 2.2	6887.1 ± 369.5	X	X
25A2	19.1 ± 2.3	8258.7 ± 443	X	X
25A4	21.5 ± 2.4	8929.5 ± 478.9	X	X
25A6	22.3 ± 2.5	8175.1 ± 438.5	X	X
25A8	21.5 ± 2.4	9287.3 ± 498.1	X	X
25B10	16.0 ± 2.1	8794.3 ± 471.6	X	✓
25B2	18.1 ± 2.2	10696.4 ± 573.5	X	✓
25B4	16.5 ± 2.2	11694.1 ± 627	X	✓
25B6	15.8 ± 2.1	10722.2 ± 574.9	X	✓
25B8	18.0 ± 2.2	10557.4 ± 566.1	X	✓
26A10	75.8 ± 6.3	1014.2 ± 55	✓	✓
26A2	82.2 ± 6.8	1274.3 ± 68.9	✓	✓
26A4	103.9 ± 8.5	989.3 ± 53.7	✓	✓
26A6	94.6 ± 7.8	840.7 ± 45.8	✓	✓
26A8	89 ± 7.3	1003.5 ± 54.5	✓	✓
26B10	86.7 ± 7.2	942.2 ± 51.2	✓	X
26B2	108.7 ± 8.9	883.3 ± 48.1	✓	X
26B4	121.9 ± 9.9	822 ± 44.8	✓	X
26B6	93.8 ± 7.7	962.0 ± 52.2	✓	X
26B8	100.9 ± 8.3	916.8 ± 49.8	✓	X
27A10	41.5 ± 3.7	1530.2 ± 82.6	✓	X
27A2	44.1 ± 3.9	3289.7 ± 176.8	✓	X
27A4	62.8 ± 5.3	2299.3 ± 123.7	✓	X
27A6	66.1 ± 5.6	2598.7 ± 139.8	✓	X
27A8	55.4 ± 4.7	2361.1 ± 127.1	✓	X
27B10	41.2 ± 3.7	1217.2 ± 65.9	✓	✓
27B2	48.1 ± 4.2	2211.0 ± 119.0	✓	✓
27B4	38.7 ± 3.5	1878.6 ± 101.2	✓	✓
27B6	59.9 ± 5.1	1925.2 ± 103.7	✓	✓
27B8	51.1 ± 4.4	1530.9 ± 82.6	✓	✓
29A10	23.4 ± 2.5	8054.2 ± 432.0	X	X
29A2	21.0 ± 2.4	7288.2 ± 391.0	X	X
29A4	30.6 ± 3.0	9356.4 ± 501.8	X	X
29A6	36.2 ± 3.3	8498.3 ± 455.8	X	X
29A8	34.0 ± 3.2	8825.1 ± 473.3	X	X
29B10	18.5 ± 2.3	5501.4 ± 295.2	X	✓
29B2	19.1 ± 2.3	7026.3 ± 376.9	X	✓
29B4	22.9 ± 2.5	6681.9 ± 358.5	X	✓
29B6	24.8 ± 2.6	7120.4 ± 382.0	X	✓

29B8	27.1 ± 2.7	6105.9 ± 327.6	X	√
2A10	70.0 ± 5.9	492.7 ± 27.3	√	X
2A2	81.4 ± 6.7	1032.5 ± 56.0	√	X
2A4	73.8 ± 6.1	878.8 ± 47.8	√	X
2A6	85.4 ± 7.1	1002.9 ± 54.4	√	X
2A8	77.6 ± 6.4	783.4 ± 42.7	√	X
2B10	68.1 ± 5.7	581.9 ± 32.0	√	√
2B2	56.4 ± 4.8	1000.0 ± 54.3	√	√
2B4	73.6 ± 6.1	849.2 ± 46.2	√	√
2B6	67.6 ± 5.7	916.5 ± 49.8	√	√
2B8	88.2 ± 7.3	958.7 ± 52.1	√	√
30A10	63.8 ± 5.4	929.7 ± 50.5	√	√
30A2	70 ± 5.9	1278.4 ± 69.1	√	√
30A4	72.9 ± 6.1	1208.5 ± 65.4	√	√
30A6	79.2 ± 6.6	1108.4 ± 60.1	√	√
30A8	71.1 ± 5.9	1029.9 ± 55.9	√	√
30B10	57.5 ± 4.9	465.1 ± 25.9	√	X
30B2	69.5 ± 5.8	778.0 ± 42.4	√	X
30B4	71.3 ± 6.0	631.9 ± 34.7	√	X
30B6	72.8 ± 6.1	621.2 ± 34.1	√	X
30B8	66.6 ± 5.6	581.9 ± 32.0	√	X
31A10	25.4 ± 2.6	5705.5 ± 306.2	X	X
31A2	25.6 ± 2.7	6728.9 ± 361	X	X
31A4	26.1 ± 2.7	7250.7 ± 389	X	X
31A6	24.3 ± 2.6	5818.6 ± 312.2	X	X
31A8	21.5 ± 2.4	7468.8 ± 400.6	X	X
31B10	32.7 ± 3.1	4148.1 ± 222.8	X	√
31B2	20.1 ± 2.3	5424.7 ± 291.1	X	√
31B4	27.2 ± 2.8	5445.2 ± 292.2	X	√
31B6	18.5 ± 2.3	5210.1 ± 279.6	X	√
31B8	24.6 ± 2.6	5666.9 ± 304.1	X	√
32A10	581.1 ± 46.9	1101.7 ± 59.7	√	X
32A2	55.2 ± 4.7	1109.9 ± 60.1	√	X
32A4	55.9 ± 4.8	838.2 ± 45.7	√	X
32A6	304.2 ± 24.6	879.5 ± 47.8	√	X
32A8	741.7 ± 59.9	793.9 ± 43.3	√	X
32B10	88.2 ± 7.3	1562.5 ± 84.3	√	√
32B2	97.4 ± 8	1921.2 ± 103.5	√	√
32B4	104.1 ± 8.5	1440.7 ± 77.8	√	√
32B6	98.4 ± 8.1	1338.6 ± 72.4	√	√
32B8	123.4 ± 10.1	1275.9 ± 69	√	√
33A10	19.1 ± 2.3	3102.6 ± 166.8	X	√
33A2	15.8 ± 2.1	5806.8 ± 311.6	X	√
33A4	17.2 ± 2.2	5701.3 ± 306.0	X	√
33A6	17.0 ± 2.2	5287.0 ± 283.8	X	√
33A8	19.8 ± 2.3	4846.3 ± 260.2	X	√
33B10	17.5 ± 2.2	5550.9 ± 297.9	X	X
33B2	15.2 ± 2.1	6066.2 ± 325.5	X	X
33B4	19.5 ± 2.3	5836.7 ± 313.2	X	X
33B6	16.0 ± 2.1	6453 ± 346.2	X	X
33B8	21.0 ± 2.4	6928.1 ± 371.7	X	X
34A10	34.7 ± 3.2	4643.7 ± 249.3	√	√
34A2	50.9 ± 4.4	2888.2 ± 155.3	√	√
34A4	67.8 ± 5.7	2658.9 ± 143	√	√
34A6	76.9 ± 6.4	2475.3 ± 133.2	√	√
34A8	67.8 ± 5.7	2536 ± 136.4	√	√
34B10	85.5 ± 7.1	1949.6 ± 105	√	X
34B2	67.5 ± 5.7	3809.2 ± 204.6	√	X
34B4	93.5 ± 7.7	3129.8 ± 168.2	√	X
34B6	91.5 ± 7.5	3317.2 ± 178.2	√	X
34B8	74.6 ± 6.2	3132.9 ± 168.4	√	X
35A10	18.3 ± 2.2	4234 ± 227.4	X	X
35A2	14.0 ± 2.0	3695.3 ± 198.5	X	X
35A4	18.3 ± 2.2	4052.4 ± 217.6	X	X
35A6	17.6 ± 2.2	3892 ± 209	X	X
35A8	18.5 ± 2.3	4841.8 ± 259.9	X	X
35B10	23.4 ± 2.5	2356.4 ± 126.8	X	√
35B2	22.1 ± 2.4	3225 ± 173.3	X	√

35B4	23.4 ± 2.5	3709.3 ± 199.3	X	√
35B6	25.9 ± 2.7	3245.8 ± 174.4	X	√
35B8	25.3 ± 2.6	3729.6 ± 200.3	X	√
37A10	13.7 ± 2.0	4167.9 ± 223.8	X	X
37A2	17.0 ± 2.2	5478.9 ± 294	X	X
37A4	20.5 ± 2.4	5162.1 ± 277.1	X	X
37A6	21.1 ± 2.4	6187.8 ± 332	X	X
37A8	22.4 ± 2.5	6172.2 ± 331.2	X	X
37B10	22.1 ± 2.4	4360.7 ± 234.1	X	√
37B2	29.1 ± 2.9	4552.9 ± 244.4	X	√
37B4	25.4 ± 2.6	5584.2 ± 299.7	X	√
37B6	26.6 ± 2.7	5001.3 ± 268.5	X	√
37B8	31.6 ± 3.0	5802.1 ± 311.4	X	√
38A10	41.8 ± 3.7	1954.7 ± 105.3	√	X
38A2	45.0 ± 4.0	3153.2 ± 169.5	√	X
38A4	43.6 ± 3.9	3186.7 ± 171.3	√	X
38A6	49.9 ± 4.3	3043.5 ± 163.6	√	X
38A8	47.8 ± 4.2	2555 ± 137.4	√	X
38B10	63.5 ± 5.4	1381.6 ± 74.7	√	√
38B2	71.4 ± 6.0	2090.1 ± 112.6	√	√
38B4	55.1 ± 4.7	2372.7 ± 127.7	√	√
38B6	51.7 ± 4.5	2093.2 ± 112.7	√	√
38B8	81.9 ± 6.8	1877.1 ± 101.2	√	√
39A10	19.1 ± 2.3	5344.6 ± 286.8	X	√
39A2	18.8 ± 2.3	5307.8 ± 284.9	X	√
39A4	22.9 ± 2.5	6116.2 ± 328.2	X	√
39A6	21 ± 2.4	6198.3 ± 332.6	X	√
39A8	20.1 ± 2.3	5476.9 ± 293.9	X	√
39B10	21.0 ± 2.4	7663.8 ± 411.1	X	X
39B2	32.2 ± 3.1	9237.5 ± 495.4	X	X
39B4	23.1 ± 2.5	9739.4 ± 522.3	X	X
39B6	20.0 ± 2.3	7843.2 ± 420.7	X	X
39B8	24.4 ± 2.6	7665.6 ± 411.2	X	X
3A10	28.2 ± 2.8	5209.6 ± 279.6	X	√
3A2	32.5 ± 3.1	5677.4 ± 304.7	X	√
3A4	29.1 ± 2.9	5717.3 ± 306.8	X	√
3A6	27.7 ± 2.8	5723.4 ± 307.1	X	√
3A8	28.2 ± 2.8	5477.3 ± 294	X	√
3B10	21.3 ± 2.4	4327.4 ± 232.4	X	X
3B2	18.5 ± 2.3	5700.8 ± 305.9	X	X
3B4	23.9 ± 2.6	6462.6 ± 346.7	X	X
3B6	22.8 ± 2.5	5971.2 ± 320.4	X	X
3B8	24.6 ± 2.6	5880.2 ± 315.5	X	X
4A10	55.1 ± 4.7	836.9 ± 45.6	√	X
4A2	59.2 ± 5	1222.1 ± 66.1	√	X
4A4	55.6 ± 4.8	1063.3 ± 57.6	√	X
4A6	63.8 ± 5.4	959.6 ± 52.1	√	X
4A8	59.2 ± 5	935.3 ± 50.8	√	X
4B10	75.6 ± 6.3	811.0 ± 44.2	√	√
4B2	64.7 ± 5.4	1082.7 ± 58.7	√	√
4B4	74.9 ± 6.2	834.0 ± 45.4	√	√
4B6	76.6 ± 6.4	854.1 ± 46.5	√	√
4B8	78.6 ± 6.5	978.1 ± 53.1	√	√
5A10	11.4 ± 1.9	3481.1 ± 187	X	√
5A2	12.8 ± 2	3899.8 ± 209.5	X	√
5A4	13.5 ± 2	5159.5 ± 276.9	X	√
5A6	14.3 ± 2.1	4440.7 ± 238.4	X	√
5A8	14.7 ± 2.1	5059.5 ± 271.6	X	√
5B10	<10.0	4062.2 ± 218.2	X	X
5B2	10.7 ± 1.9	4705.7 ± 252.6	X	X
5B4	13.8 ± 2	5604.5 ± 300.8	X	X
5B6	12.7 ± 2	4638.6 ± 249	X	X
5B8	10.0 ± 1.9	5530.2 ± 296.8	X	X
6A10	16.5 ± 2.2	3136.7 ± 168.6	X	X
6A2	19.6 ± 2.3	5134.7 ± 275.6	X	X
6A4	18.6 ± 2.3	5085.6 ± 273	X	X

6A6	20.1 ± 2.3	5511.0 ± 295.8	X	X
6A8	21.3 ± 2.4	4816.6 ± 258.6	X	X
6B10	19.3 ± 2.3	5150.5 ± 276.4	X	√
6B2	21.5 ± 2.4	6432.9 ± 345.1	X	√
6B4	24.1 ± 2.6	7507.7 ± 402.7	X	√
6B6	22 ± 2.4	7639.3 ± 409.8	X	√
6B8	22.6 ± 2.5	7507.0 ± 402.7	X	√
7A10	124.9 ± 10.2	325.9 ± 18.6	√	√
7A2	145.6 ± 11.8	458.6 ± 25.5	√	√
7A4	164.3 ± 13.3	491.6 ± 27.3	√	√
7A6	162.8 ± 13.2	444.1 ± 24.8	√	√
7A8	147.1 ± 11.9	380.7 ± 21.5	√	√
7B10	46.3 ± 4.1	778.2 ± 42.5	√	X
7B2	41.7 ± 3.7	1548.5 ± 83.6	√	X
7B4	62.7 ± 5.3	923.5 ± 50.2	√	X
7B6	56.5 ± 4.8	1087.2 ± 58.9	√	X
7B8	51.4 ± 4.4	877.1 ± 47.7	√	X
9A10	49.3 ± 4.3	3123.3 ± 167.9	√	X
9A2	71.9 ± 6	3560.1 ± 191.3	√	X
9A4	79.9 ± 6.6	3913 ± 210.2	√	X
9A6	76.9 ± 6.4	3856.1 ± 207.1	√	X
9A8	85.5 ± 7.1	4405.5 ± 236.5	√	X
9B10	49.8 ± 4.3	1623.9 ± 87.6	√	√
9B2	74.1 ± 6.2	2536.4 ± 136.4	√	√
9B4	64.5 ± 5.4	2062.9 ± 111.1	√	√
9B6	51.1 ± 4.4	1833.1 ± 98.8	√	√
9B8	65.7 ± 5.5	2337.5 ± 125.8	√	√

Errors were calculated using the formula $\frac{\Delta x}{|x|} = \sqrt{\left(\frac{\sqrt{(\Delta y)^2 + (\Delta b)^2}}{y-b}\right)^2 + \left(\frac{\Delta a}{a}\right)^2}$ with a being the slope and b as the intercept of the respective calibration line $y = a * x + b$ ($^{13}\text{C}_3\text{T}$: $y = (4.483\text{e}^{-3} \pm 3.6\text{e}^{-5}) * x + (3.916\text{e}^{-2} \pm 2.278\text{e}^{-2})$; $^{13}\text{C}_3\text{E}$: $y = (6.041\text{e}^{-3} \pm 9.9\text{e}^{-5}) * x + (-4.61\text{e}^{-3} \pm 1.049\text{e}^{-2})$). For the relative error of y ($\frac{\Delta y}{y}$), the average relative standard deviation of each calibration point of the five in-sequence calibrations was used ($^{13}\text{C}_3\text{T}$: 5.30 %; $^{13}\text{C}_3\text{E}$: 7.91 %). The first number of the sample ID identifies the individual patient, the following capital letter describes the chronological sequence of experiments (B after A) and the last number stands for the 5 different time points of sampling (2: 8:30 am; 4: 9:08 am; 6: 9:26 am; 8: 9:56 am; 10: 10:55 am).

Cross-validation

Intra-assay cross-validation

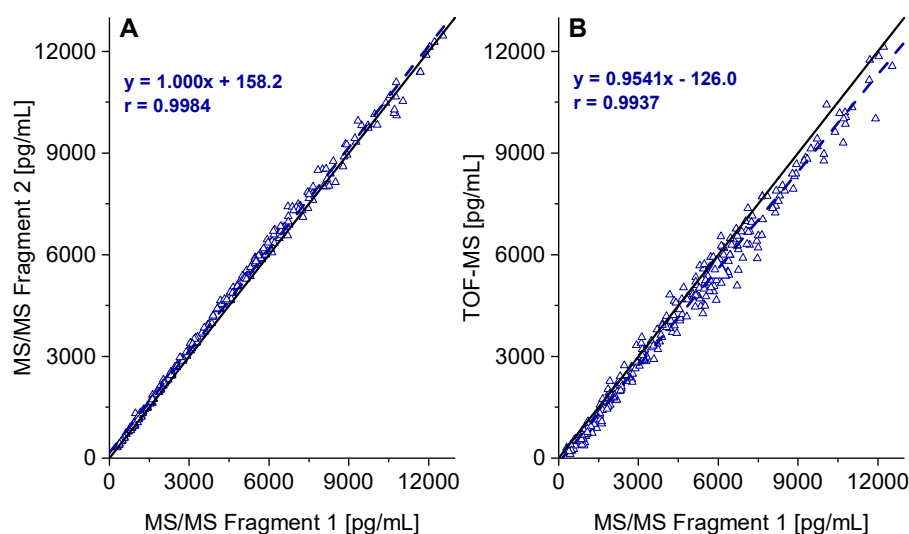


Fig. A.21. Scatter plots of results for intra-assay cross-validation via alternative quantifiers. A: Scatter plot for SWATH-MS/MS fragment 2 (m/z 97.0648); B: Scatter plot for TOF-MS precursor (m/z 289.2162). Solid lines resemble the optimum line of parity. Dashed lines are results of linear regression analysis.

Cross-validation with commercial controls

The lyophilized true plasma matrix controls (MassCheck Steroid Panel 2) with certified E and T concentrations were resuspended and treated like regular samples. For T, all three available levels could be employed. For E only the lower 2 levels were used since the upper level, with a concentration of $2,500 \text{ pg mL}^{-1}$, exceeded the calibration range of the current method. Three aliquots of each level were prepared and analyzed in triplicate. Normality tests and following Grubb's tests did not show any outliers so that the mean of all measurements of each level was calculated and compared to the target value of the commercial control. Results are shown in Table A.13. Precisions matched those of above validation study and bias remained within acceptable limits (6 - 15 %).

Table A.13. Validation by independent commercial quality controls.^h

Analyte	Level	Target conc. [pg mL ⁻¹]	Target range [pg mL ⁻¹]	Accuracy [%]	Precision [%]
E	I	82	57 - 107	104.3	11.0
	II	411	329 - 493	112.0	5.9
T	I	201	141 - 261	93.8	11.2
	II	1,520	1,210 - 1,820	105.0	6.3
	III	7,820	6,260 - 9,380	114.8	6.9

^hFor level III of the E controls the concentration exceeded the calibration range.

Combined Targeted/Untargeted Profiling (Towards Steroidomics)

Extended SWATH Method

Another method (method 2) with six additional SWATH experiments was created to demonstrate the ability of combined targeted quantitative analysis with untargeted profiling (Table A.14). In consequence, the total cycle time slightly increased to 710 ms. Experiments that covered target analytes were not altered in DP or CE. All other SWATH experiments were set to a DP of 100 V, a CE of 30 V and a collision energy spread (CES) of 10 V. For design of window widths, swathTuner (Zhang, Y.; Bilbao, A.; Bruderer, T.; Luban, J.; Strambio-De-Castillia, C.; Lisacek, F.; Hopfgartner, G.; Varesio, E. *J Proteome Res* 2015, 14, 4359-4371.) was used. Recommended window widths had to be adjusted to prevent interferences between analytes with overlapping retention periods and similar fragmentation. For example, the fragment with m/z 97.0648 (or 109.0648, respectively) is formed by androstenedione (t_R : 3.56 min) and epiT (t_R : 3.58 min). In order to be able to detect these signals distinctively, SWATH windows have to be designed in a way to cover each precursor in a separate window. Therefore, SWATH experiment 4 (Table A.14) was shortened to cover solely androstenedione (precursor m/z: 287.2006). epiT (precursor m/z: 289.2162) fragment signals could be detected separately in SWATH experiment 5, free of androstenedione interference.

Table A.14. SWATH design for combined targeted/untargeted profiling (method 2).ⁱ

Experiment	Range [m/z]	Acc [ms]	CE [V]	CES [V]
TOF	30 - 1,000	20	10	N/A
SWATH 1	254.5 - 258.5	300	25	0
SWATH 2	260.0 - 265.0	40	25	0
SWATH 3	250.0 - 276.6	40	30	10
SWATH 4	275.6 - 287.2	40	30	10
SWATH 5	288.5 - 292.5	30	33	0
SWATH 6	294.0 - 299.0	30	33	0
SWATH 7	286.2 - 310.3	40	30	10
SWATH 8	309.3 - 327.3	40	30	10
SWATH 9	326.3 - 344.8	40	30	10
SWATH 10	343.8 - 370.0	40	30	10

ⁱSWATH windows 1 and 2 were the same as for targeted analysis of E. SWATH windows 5 and 6 were the same as for targeted analysis of T. t_{Acc} of target windows were adjusted to reach a total t_{Cyc} of 710 ms (\approx 10 data points per peak). Collision energy spread (CES).

Towards Comprehensive Steroidomics

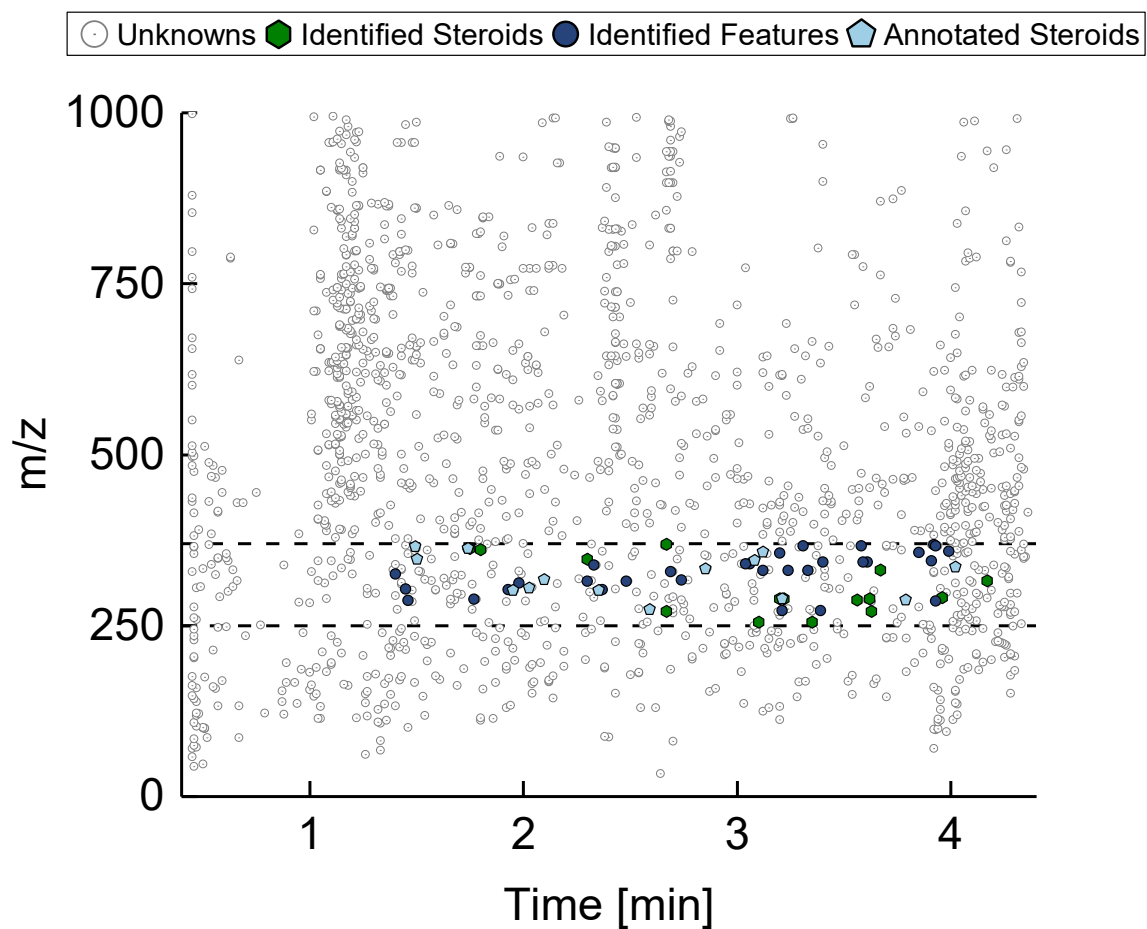


Fig. A.22. Aligned peak spotting in 9 repeated measurements of a pooled plasma QC sample applying method 2 (see Table A.14). Associated to Fig. 2 in main document, but with extended m/z range.

Table A.15. Identified steroids in commercial control samples.^j

Analyte	t _R [min]	TOF- Precursor [m/z]	MS/MS- Signal [m/z]	Calibration Range ^g [ng mL ⁻¹]	Calibration Function	R ²	S/N at Level I
Dehydroepi- androsterone- sulfate	2.67	271.2056		263.0 - 5,019	6.78 * 10 ⁻⁵ x - 6.82	0.9963	14.8 ± 2.4
		369.1730	213.1638	263.0 - 5,019	1.15 * 10 ⁻⁵ x - 2.58	0.9961	14.9 ± 2.4
Estradiol	3.16	255.1743		-	-	-	
		255.1743	159.0804	0.082 - 2.500	9.92 * 10 ⁻³ x - 0.38	0.9986	19.85 ± 2.1
Testosterone	3.22	289.2162		0.201 - 7.820	1.41 * 10 ⁻³ x - 0.07	0.9920	10.5 ± 1.0
			109.0648	0.201 - 7.820	2.26 * 10 ⁻³ x - 0.06	0.9976	41.2 ± 2.8
Androstene- dione	3.56	287.2006		-	-	-	
		287.2006	287.2006	0.277 - 9.58	5.957e-4x + 0.03607	0.9986	18.7 ± 2.8
		287.2006	97.0648	0.277 - 9.58	9.27e-4x - 0.07300	0.9972	19.1 ± 3.2
Dehydroepi- androsterone	3.63	271.2056					
Hydroxy- progesterone	3.67		271.2062	2.01 - 37.8	8.267e ⁻⁵ x - 0.08546	0.9951	10.8 ± 2.3
		331.2268	331.2268	0.300 - 8.96	6.69 * 10 ⁻⁴ x - 0.01	0.9980	32.4 ± 2.1
Dihydro- testosterone	3.96		97.0648	0.300 - 8.96	5.93 * 10 ⁻⁴ x - 0.05	0.9958	10.1 ± 1.8
		291.2318	255.2113	0.083 - 1.05	1.63 * 10 ⁻⁴ x - 0.01	0.9918	10.1 ± 0.2
Progesterone	4.17	315.2318	315.2318	0.310 - 15.1	2.77 * 10 ⁻⁴ x - 0.04	0.9975	11.7 ± 2.5
			97.0648	0.310 - 15.1	7.00 * 10 ⁻⁴ x - 0.15	0.9973	10.3 ± 0.7

^jRange of trilevel commercial control samples. Weighting was 1/x for all compounds. CE in untargeted windows was set to a general value of 30 V with a spread of ± 10 V, fragmentation was therefore not optimized. Calibration function is given for ion signal mentioned in corresponding row. *For these compounds d₅E was used as internal standard, for others d₅T was used (MS/MS-signals).

Untargeted findings

Further processing of TOF-MS or SWATH-MS/MS data besides E and T quantification can reveal secondary information about individual samples or study groups. By relative quantification and statistical comparison of features additional insights in sample properties, which would have been lost with common QqQ instruments, are gathered. For exemplary purposes, untargeted investigation of all study samples was executed via MarkerView 1.2.1 (Sciex). For normalization, d₅T and d₅E were assigned to find features according to their t_R. Grouping (EI = E patch & insulin treatment; EP = E patch & placebo treatment; PI = placebo patch & insulin treatment; PP = placebo patch and placebo treatment) and statistical analysis showed that hydroxyprogesterone levels were elevated in groups with placebo patches compared to groups that were administered E patches (see main document Fig. 6). Also epiT was found to be down-regulated in E patch groups (Fig. A.23), complying with the finding of down-regulation of T in E-treated groups. Additional findings are shown in Fig. A.24.

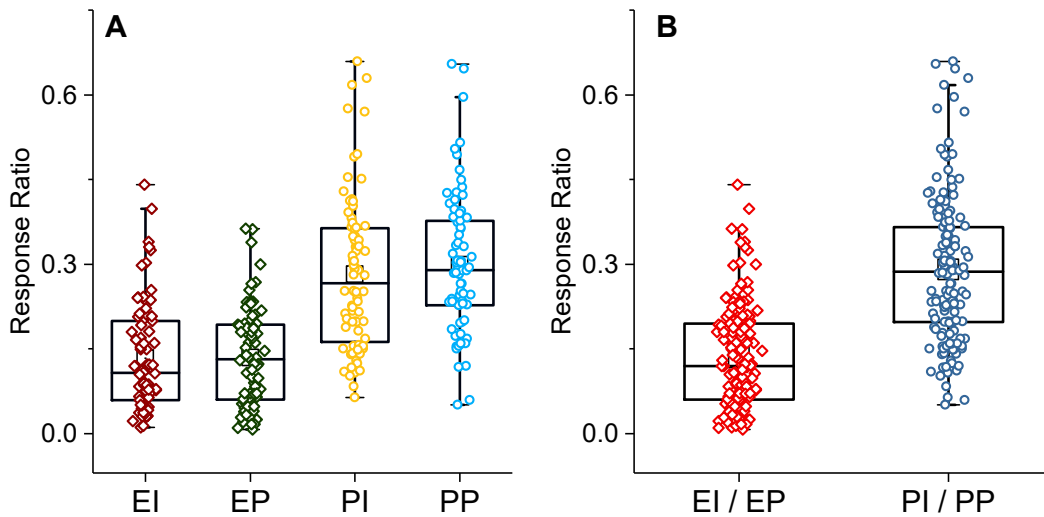


Fig. A.23. Relative quantification of epiT. Boxplots for each of the four groups (A) and for grouped E patch and grouped placebo patch samples (B). For B, a 2.3-fold increase (median values) in epiT was found in placebo patch groups (U-test, p -value: 8.8×10^{-29}). Signals were obtained from SWATH-MS/MS scan (fragment signals).

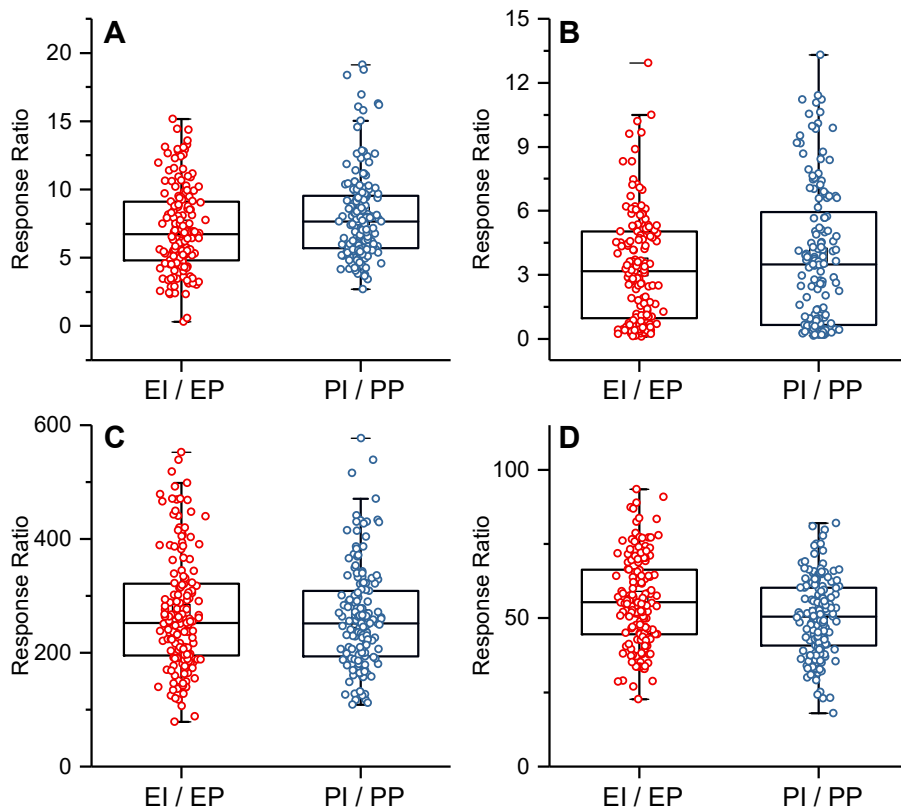


Fig. A.24 Additional Boxplots. (A) Androstenedione (U-test, p -value: 9.2×10^{-3}), (B) Progesterone (U-test, p -value: >0.05), (C) Cortisol (U-test, p -value: >0.05), (D) Cortisone (U-test, p -value: 1.1×10^{-3}).

5.2. Guidelines for Selection of Internal Standard-based Normalization Strategies in Untargeted Lipidomic Profiling by LC-HR-MS/MS

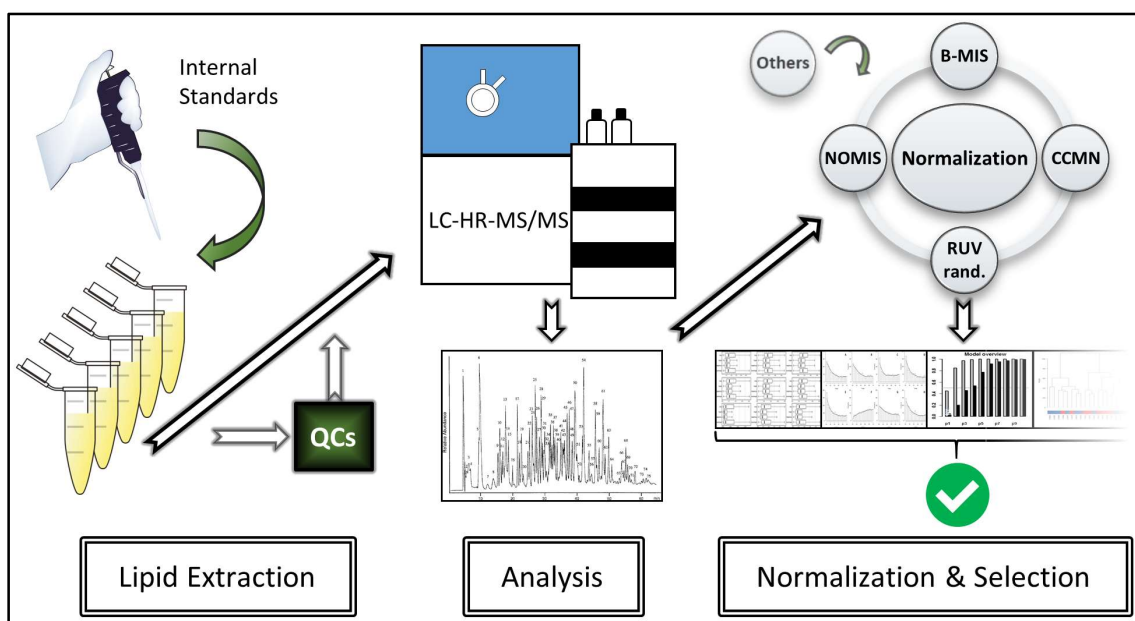
Bernhard Drotleff, Michael Lämmerhofer

Institute of Pharmaceutical Sciences, Pharmaceutical (Bio-)Analysis, University of Tübingen,
Tübingen, Germany

Reprinted with permission from *Analytical Chemistry*, in press,

DOI: 10.1021/acs.analchem.9b01505

Copyright (2019) American Chemical Society



Graphical Abstract

5.2.1. Abstract

Due to variation in instrument response caused by various sources of errors throughout an analytical assay, data normalization plays an indispensable role in untargeted LC-MS profiling, yet limited accepted guidelines on this topic exist. In this work, a systematic comparison of several normalization techniques, mainly focusing on internal standard-based approaches, has been performed to derive some general recommendations. For generation of untargeted lipidomic data, a comprehensive ultra-high performance liquid chromatography (UHPLC)-electrospray ionization (ESI)-quadrupole time of flight (QTOF)-MS/MS method was utilized. To monitor instrument stability and evaluate normalization performance, quality control (QC) samples, prepared from aliquots of all experimental samples, were embedded in the sequence. Stable isotope labeled standards, representing differing lipid classes, were spiked to each sample as internal standards for postacquisition normalization. Various metrics were used to compare distinct normalization strategies, with reduction of variation in QC samples being the critical requirement for acceptance of successful normalization. The comparison of intragroup coefficients of variation (CVs), median absolute deviations (MADs), and variance enables simple selection of the best performance of normalization with improved and coherent results. Furthermore, the importance for normalization in critical data sets, showing only minor effects between groups with high variation and outliers, is pointed out. Apart from normalization, also, influences of used raw data types are demonstrated. In addition, effects of various factors throughout the processing workflow were investigated and optimized. Eventually, implementation of quality control samples, even if not required for normalization, provided a useful basis for assessing data quality. Due to lack of consensus for selecting optimum normalization, suggestions for validating data integrity are given.

5.2.2. Introduction

Untargeted metabolomic or lipidomic screening methods, in particular by LC-MS hyphenation, have evolved to become widely used techniques in biomarker discovery.^{1,2} Although superior sensitivity, analyte coverage, and selectivity are shown,³ precision in LC-MS is limited by fluctuating ionization efficiencies of the standard electrospray ionization (ESI) source.⁴ Despite protocols and instrumentation being steadily improved and optimized, additional introduction of variation by random or systematic errors throughout the analytical process cannot be fully eliminated.

Excessive variation in the data produces outliers and significantly impairs statistical power,⁵ especially when effect size between experimental groups is moderate or weak. Furthermore, employment of animal models, typically limited by ethical constraints complying to the 3R principle or the investigation of rare clinical conditions, usually yields statistically underpowered results due to small sample sizes. Improvement of precision is therefore of utmost importance to minimize variation and achieve maximum significance. In targeted approaches, many sources of errors are usually controlled for by addition of stable isotope labeled (SIL) internal standards (ISs) and determination of analyte/IS ratios as the response. In untargeted omics, various strategies for normalization were developed, with most of them being model-driven approaches⁶ and only a minority relying on response ratios. One of the obstacles for utilizing ISs in untargeted profiling is limited suitability and availability of labeled compounds. Experiments applying global SIL samples (e.g., U-¹³C-labeled yeast or fungi) have shown promising results^{7,8} but are also accompanied by additional costs, higher potential for signal interference, and complex data processing efforts.

Many studies attempt normalization with a selection of multiple ISs; however, correct assignment of ISs to unlabeled features is still controversial. Satisfying results have been achieved by assignment via similar chemical properties like retention time (tR-IS),⁹ analyte class,¹⁰ or m/z ratio.¹¹ However, normalization based on chemical similarities might not be able to fully control the observed variation in analytical sequences.¹² Recently, Boysen et. al¹³ developed a strategy named B-MIS (best-matched internal standard normalization) for which assignment depends on best achievable precision in repeated measurements of a pooled quality control (QC) sample.¹⁴ Furthermore, other normalization methods relying on IS-derived models have been developed. Sysi-Aho et. al¹² described a model based on similarities in variation profiles of recorded features and ISs using multiple linear regression (NOMIS, normalization using optimal selection of multiple internal standards). Another model, focusing on systematic cross-contribution effects of analytes and ISs, termed CCMN¹⁵ (cross-contribution compensating multiple standard normalization), has also been successfully established. With RUV-random¹⁶ (remove unwanted variance-random), a statistical method

using linear mixed effects modeling to normalize data via ISs or quality control metabolites, a limited number of currently recognized IS-based normalization methodologies is available.

The described techniques are accompanied by various alternative normalization methods. Besides scaling and other IS-independent models, signal correction relying on QCs embedded in the sequence is commonly practiced in untargeted assays.¹⁷ These QC samples, preferably consisting of an equivalent mixture of all samples, were shown to be suitable to correct instrumental drifts and are used to monitor the stability of system performance.¹⁸ For normalization, a QC-based regression model, for example, LOWESS (locally weighted scatter plot smoothing),¹⁹ is deployed and executed to adjust real sample signals according to the analytical order. Although diverse normalization methods have been shown to yield improved results, suitability is always dependent on the nature of the samples and the occurring variation and therefore needs to be tested. Various projects like Normalyzer,²⁰ NOREVA,²¹ or MetaboGroupS²² provide platforms to perform and evaluate different normalization techniques on data, which simplifies investigation and selection of optimum normalization.

In this study, we focused on comparison of IS-based normalization and used different strategies to assess performance. Moreover, when a combined study design with QC samples and ISs is used, novel guidelines for selecting optimal normalization based on results for intragroup metrics of variation are suggested.

5.2.3. Methods

Materials

Ultra LC-MS grade acetonitrile, methanol, and 2-propanol as well as ACS grade formic acid (98%, w/v) were supplied by Carl Roth (Karlsruhe, Germany). Ammonium formate was purchased from Sigma–Aldrich (Saint-Louis, MO, USA). SPLASH LipidoMIX was obtained from Avanti Polar Lipids (Alabaster, AL, USA). Arachidonic acid-*d*₈, α -linolenic acid-*d*₁₄, and linoleic acid-*d*₄ were acquired from Cayman Chemical (Ann Arbor, MI, USA). Type I purity water was provided by a Purelab Ultra purification system (ELGA LabWater, Celle, Germany).

Sample Preparation

Plasma samples²³ were stored at $-80\text{ }^{\circ}\text{C}$ and thawed on ice on the day of sample preparation. Aliquots of 25 μL were used for untargeted lipid extraction^{24,25} via 2-propanol-based protein precipitation (55 μL of 2-propanol + 20 μL of methanol). For normalization, LipidoMIX and SIL fatty acids (arachidonic acid-*d*₈, α -linolenic acid-*d*₁₄, linoleic acid-*d*₄) were added to the methanolic portion prior to precipitation. After vortexing, samples were centrifuged for 10 min at 15 000*g* and 4 $^{\circ}\text{C}$ with a 5415R microcentrifuge (Eppendorf, Hamburg, Germany). The supernatant was transferred to a sealed glass vial and stored in a PAL HTC-xt autosampler

(CTC Analytics, Zwingen, Switzerland) at 4 °C under the absence of light. Samples were analyzed by ultra-high performance liquid chromatography (UHPLC)-ESI-quadrupole time of flight (QTOF)-MS/MS as soon as possible after preparation.

Selection of Internal Standards and Study Design

The number and selection of implemented ISs play key roles in untargeted omics as they can greatly affect the outcome of the study. A generally valid approach is using as many ISs as applicable to closely resemble the composition in the sample.²⁶ In this study, the main source of employed ISs was contained in the SPLASH LipidoMIX, which covers quantitative amounts of deuterated lipids designed to relatively reflect the ratios in human plasma. Due to differing concentrations and instrument response, suitable SIL lipids for IS normalization were selected after preset criteria for asymmetry factor (>0.9 and <1.4), minimum intensity (≥ 500 cps (counts per second)), and signal-to-noise (S/N) ratio (>30) were met (Tables S-1 and S-2). Furthermore, five plasma samples and five blank samples were processed without ISs and positively checked for the absence of signal interferences. QCs, prepared from pooled aliquots of all plasma samples, were embedded in the randomized sequence after every fifth injection. These samples were used for LOWESS normalization and for monitoring instrument performance. The analytical order is shown in Table S-3.

LC-MS/MS Method Parameters

Lipid separation was performed according to the method of Tsugawa et al.²⁷ Chromatography was carried out on a 1290 Infinity UHPLC system (Agilent Technologies, Santa Clara, CA, USA) with an Acquity UPLC CSH C18 column (100 mm \times 2.1 mm, 1.7 μ m, 130 Å) and a VanGuard Acquity UPLC CSH C18 precolumn (5 mm \times 2.1 mm, 1.7 μ m, 130 Å) (Waters, Milford, MA, USA). Mobile phase A was 60:40 MeCN/H₂O (v/v) with 0.1% formic acid (v/v) and 10 mM ammonium formate. Mobile phase B consisted of 90:9:1 IPA/MeCN/H₂O (v/v/v) with 0.1% formic acid (v/v) and 10 mM ammonium formate. The gradient (0.0 min, 15% B; 2.0 min, 30% B; 2.5 min, 48% B; 11.00 min, 82% B; 11.50 min, 99% B; 12.00 min, 99% B; 12.10 min, 15% B, 15.00 min, 15% B) was operated at a flow rate of 0.6 mL/min and a constant oven temperature of 65 °C. Injection volume was set to 3 μ L in positive and 5 μ L in negative ionization modes.

The chromatographic system was hyphenated to a TripleTOF 5600+ mass spectrometer with a DuoSpray source (Sciex, Framingham, MA, USA) operated with the ESI probe. Ion source parameters were as follows: curtain gas (N₂), 35 psi; nebulizer gas (N₂), 60 psi; heater gas (N₂), 60 psi; ion source voltage floating, +5500 V (positive mode) and -4500 V (negative mode); declustering potential, ± 80 V; source temperature, 350 °C. For comprehensive acquisition of MS/MS data, 20 SWATH (sequential window acquisition of all theoretical

fragment-ion mass spectra) experiments with a collision energy of ± 45 V (mode dependent) and a spread of ± 15 V were created. This method was applied as it allows rapid lipid identification via a well-established workflow for untargeted lipidomics.²⁸ Only TOF signals were used for normalization; therefore, SWATH experiment design is not further discussed. An accumulation time of 200 ms was assigned to the TOF-MS experiment for precursor detection in the mass range of m/z 50–2000. Total cycle time summed up to 750 ms, yielding at least 10 points per peak with an average peak width at base of 8 s. The resolution on the MS level was over 30 000 (fwhm @ m/z 829.5393), and it was over 15 000 (fwhm @ m/z 397.2122) on the MS/MS level in high sensitivity mode. Mass calibration was achieved via detection of sodium acetate clusters by infusion (0.1 mg/mL in MeCN/H₂O, 1:1, v/v) every 10th injection. Samples were first analyzed in positive and subsequently in negative ESI mode. The analytical system was controlled by Analyst 1.7 TF software (Sciex).

Data Processing

Data processing for untargeted lipidomic screening was done via MS-DIAL²⁷ (version 3.2026). The software enables processing of SWATH data, including peak finding, alignment, and lipid identification, relying on t_R , m/z values and isotope ratios as well as MS/MS similarity of an incorporated in silico library.²⁷ Processing parameters are listed in Table S-4. Raw area and raw height data sets from aligned TOF results were exported and normalized, respectively. LOWESS was performed via the LOWESS Normalization Tool (Riken, Saitama, Japan). mTIC normalization (based on sum of peak heights of identified metabolites)²⁹ was done via the embedded function in MS-DIAL. The IS-based methods CCMN, NOMIS, RUV-random, B-MIS, and t_R -IS were executed in R Studio 1.1.383 (R Foundation for Statistical Computing, Vienna, Austria) and Excel 2016 (Microsoft, Redmond, WA, USA). For RUV-random, only ISs and no additional quality control metabolites were selected. Factor k , required as the preliminary input for RUV-random, was set to 3 for the shown data. The influence of k is further discussed in the Supporting Information. In contrast to the proposed B-MIS method,¹³ no cutoffs for the coefficient of variation (CV) in raw data (>10%) or CV improvement compared to raw data ($\geq 40\%$) for acceptability of an IS were used. Instead, selection between IS normalization and raw data was always done according to the data type that produced the lowest CVs in the pooled QCs, respectively. Raw and normalized data matrices of LOWESS, mTIC, B-MIS, and t_R -IS were log-transformed prior to statistical testing to achieve comparability to the log-transformation-inheriting methods of NOMIS, CCMN, and RUV-random. Due to log-transformation, given CVs in this paper are calculated as geometric CVs.³⁰ In order to visually detect outliers and assess intragroup variation, relative log abundance (RLA) plots³¹ and principal component analysis (PCA) were considered. Furthermore, group-specific CV, median absolute deviation (MAD), and variance (Var) were calculated for each feature. For

detection of significant differences between groups, two-tailed Student's *t* tests were performed. In general, taking only the number of significant features with *p*-values <0.05 into account is not regarded as an accurate metric for comparing the performance of normalization.¹⁶ However, comparing *p*-value distribution is well accepted^{16,21} and can give better hints about the consistency of the data. Moreover, multiple testing errors (type I errors) have to be avoided, which is usually achieved by adjusting *p*-values according to an eligible procedure like the Bonferroni or Benjamini-Hochberg³² correction. The number of features with adjusted *p*-values under a certain significance level can be interpreted as an indirect measure of goodness of *p*-value distribution. The more the distribution deviates from the optimum, the less significantly different features will be found after *p*-value correction. Traditional correction methods, previously mentioned, can often be too conservative, leading to no significant findings, in particular when the difference from the null hypothesis is only weak or medium. Therefore, alternative procedures like SGoF (sequential goodness of fit metatest), designed to increase statistical power,^{33,34} were considered and reported. Other parameters for evaluation of data quality were the percentage of features with a CV <30% in QC samples, mean area under the curve (AUC) of receiver operating characteristics (ROC) of the most significant features, number of features with an AUC higher than 0.8, and cumulative goodness of fit estimate (R^2Y) as well as cumulative goodness of prediction estimate (Q^2Y) for the optimum number of predictive components after partial least-squares discriminant analysis (PLS-DA). Statistical processing was carried out in R using the following packages: *metabolomics*³⁵ (RLA plots), *NormalizeMets*³⁶ (NOMIS, CCMN, and RUV-random normalization), *sgof*³⁷ (SGoF *p*-value adjustment), *ropls*³⁸ (PLS-DA), *caTools*³⁹ (ROC and AUC calculation), and *qvalue*⁴⁰ (Benjamini-Hochberg correction with false discovery rate (FDR) estimation). For further data evaluation, PeakView 2.2 (Sciex), MarkerView 1.2.1 (Sciex), FunRich 3.1.3 (<http://www.funrich.org>), and Origin 2018 (OriginLab, Northampton, MA, USA) were used.

Data Sets

To evaluate normalization methods, data sets showing only weak to medium differences between experimental groups were chosen. Lipidomics data from plasma samples of mice were originated from a previous study, and no additional animal experiments were carried out for this work.²³ The biological results and interpretation will be reported elsewhere (manuscript in preparation). The used study consisted of 4 groups of mice, which had different genomic backgrounds and received differing diets following a 2 × 2 scheme. For simplification, we refer to the groups as knockout (KO) and wild-type control (WT) and differentiate diets as high-fat diet (HFD) and control diet (CD). Samples were acquired in negative (ESI⁻ data set) and positive (ESI⁺ data set) modes. The comparison of the main scientific interest was group KO-

HFD versus group WT-HFD (both $n = 10$). Presented data is based on this comparison unless stated otherwise. In the raw data, these groups showed high variation and outliers, resulting in failure to detect significant differences. By normalization, methods that were able to decrease variation, recover outliers, and map minor differences in the data were investigated.

5.2.4. Results and Discussion

Comparison of Raw Data Types

Raw data of untargeted assays can be extracted as peak areas or peak heights. In general, the use of peak heights is recommended due to enhanced robustness to signal interferences, less influence of integration errors, and therefore improved precision, especially in the case of low abundant features.²⁸ However, using peak heights requires sufficient stability and reproducibility of chromatographic performance. Hence, suitability of the different raw data types was first investigated for the utilized analytical method. In order to evaluate performance, raw height and raw area values were normalized. In addition, two mixed data matrices of response values, one composed of ratios of feature heights and IS areas (height/area) and the other of ratios of feature areas and IS heights (area/height), were created and normalized with IS-based methods. For comparison, the number of features with a p -value < 0.05 , the number of significant features after SGoF adjustment, and the percentage of features that showed a CV $< 30\%$ in the QCs were determined (Figure 1). The ESI⁻ data set contained outliers, which decreased the number of true positive findings by introduction of high intragroup variation. For their removal, outliers were visually selected by examination of within-group RLA plots (Figures 2 and S-2) and PCA (data not shown). After removal of outliers from the raw data, sets were again normalized and compared (Figures 1B and S-1B). For both polarities, with and without outliers, raw height data cover more features that show a CV $< 30\%$ in QCs (FDA precision criteria for biomarker studies) and produce a higher number of significant findings after SGoF correction than raw area data. This trend is also consistent throughout all normalization techniques, which underlines the beneficial quality of raw height data for untargeted lipidomics. For the mixed data matrices, area/height and height/area, no clear tendency can be identified for which one is better throughout all the normalization methods. However, height/area combined data outperform the respective raw area data type in most cases for the shown metrics. In B-MIS data, this mixed data set even yields the best results (Figures 1A and S-1B). A possible explanation could be the high quality of IS area signals, which were shown to be interference-free with high S/N ratios (Table S-2) and frequently of higher precision in QCs than height signals (Table S-5). Therefore, raw area data might represent the true response for ISs more accurately, which makes height/area-like data sets worth investigating. NOMIS displays favorable results for raw height and area/height data,

but performance is questionable as no true positives are obtained for outlier removed data (Figure 1B). Further investigation revealed that removal of SIL fatty acid ISs led to good performance of NOMIS normalization (data not shown), which assumes higher dependency on selected ISs than the other IS-based methods. Overall, CCMN and RUV-random show strongly varying results without a clear disposition to an optimum raw data type. In conclusion, these data sets suggest B-MIS for the maximum yield of significant findings and the most consistent results throughout raw data types. Further results are listed in Tables S-6–S-9.

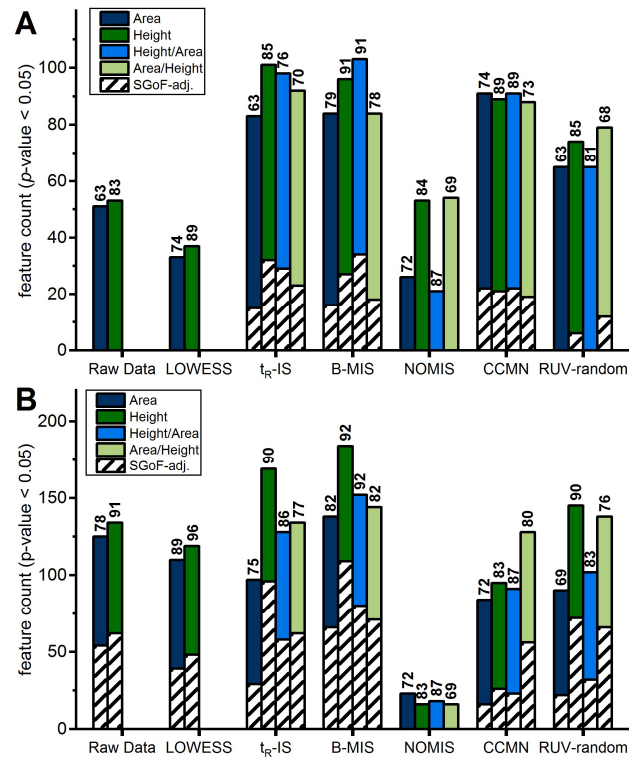


Figure 1. Comparison of raw data type and normalization (ESI⁻ data set). Number of features with a p -value < 0.05 are represented by colored columns. Overlay of the lined pattern columns show the number of SGoF-adjusted p -values < 0.05 . Values above the columns show the respective percentage of features in the QCs that showed a CV $< 30\%$. For RUV-random, k was set to 3. (A) ESI⁻ data set covering all samples. (B) ESI⁻ data set after removal of outliers (B9257, D0181, D3396, B9324, D0167, D3397, QC3, QC7).

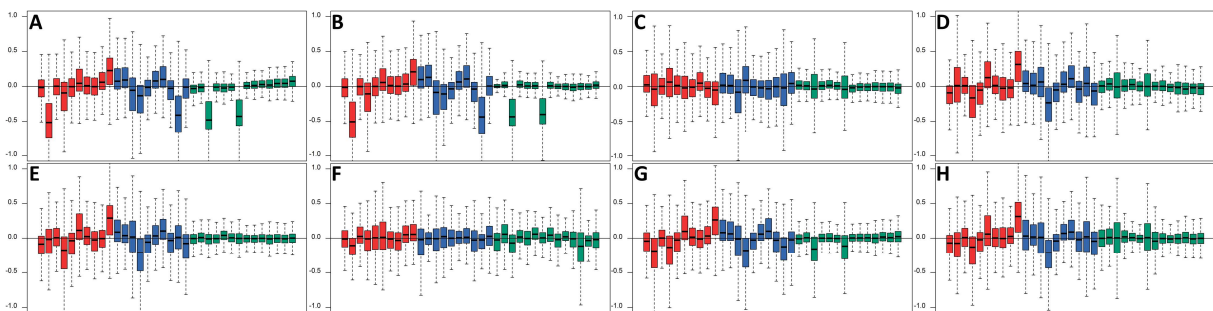


Figure 2. Within-group RLA plots for the ESI⁻ data set (comprised of box-whisker plots). Different experimental groups are represented by different colors: red (KO-HFD), blue (WT-HFD), and green (QC). RLA plots were sorted according to group. RLA plots sorted after analytical order were also investigated to detect systematic signal drifts, which was not the case. (A) Raw height data, (B) LOWESS, (C) mTIC, (D) t_R -IS, (E) B-MIS, (F) NOMIS, (G) CCMN, and (H) RUV-random ($k = 3$).

Evaluation of IS-Based Normalization Methods

As peak heights were proven to deliver better quality data throughout different normalization methods, only raw peak height-derived data were considered for the extensive comparison of the performance. For estimation of intragroup variation and detection of outliers, RLA plots are a simple yet powerful tool (Figures 2 and S-2). However, relying on visual judgment for the selection of optimum normalization can be challenging. Another approach via plotting group-dependent CVs, MADs, and variances was therefore chosen (Figure 3). By addition of a reference line, representing the median value of the respective metric in non-normalized data, the simple evaluation of normalization performance regarding reduction or an increase of intragroup variation is given. With embedded QC samples, a potent feature is enabled to the data to facilitate the choice of optimum normalization. Regardless of the underlying strategy, normalization should never increase the metrics of variation in technical replicates like QCs. Methods resulting in an increase of variation metrics compared to raw data (median of the boxplot on the right side of the red line; see Figure 3A,D,G) must therefore be dismissed for further evaluation. In this case, t_R -IS, NOMIS, and RUV-random have to be removed for further consideration due to the increased MAD in QCs (Figure 3D). In general, the reduction in variation in experimental groups after applying a particular normalization method should not come along at the expense of increasing variation in QC samples (Figure 3D–F: NOMIS). Moreover, normalization should decrease variation not only in QCs but also in experimental groups (KO-HFD: Figure 3B,E,H; WT-HFD: Figure 3C,F,I). Following these norms (Figure 4), mTIC, B-MIS, and CCMN represent complying methods for the ESI⁻ data set and are selected for further evaluation via parameters like p -value distribution (Figure S-3), number of true positives (Figure 1), mean AUC, or PLS-DA results (Table S-6). Results for metrics of variation for other data sets are shown in Figures S-4–S-6. According to this suggested workflow, B-MIS and CCMN, yielding acceptable results for PLS-DA and a comparable number of true positives, are the optimum candidates for the best-performing optimization in the ESI⁻ data set. If more than one normalization is deemed acceptable, overlapping true positive findings between methods can be determined (Figures S-8–S-10). This way, only the most robust results are considered for hypothesis generation.

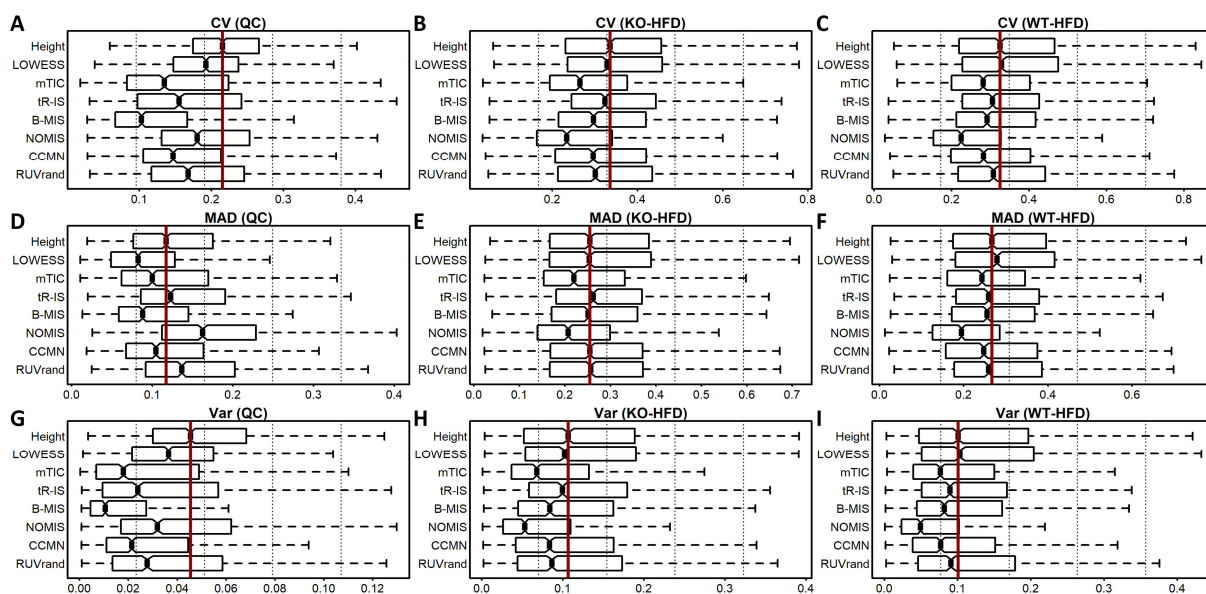


Figure 3. Box-whisker plots of intragroup metrics of variation (ESI- data set). The red line represents the median value of the respective metric in raw height data. (A) CV in QCs, (B) CV in KO-HFD, (C) CV in WT-HFD, (D) MAD for QCs, (E) MAD for KO-HFD, (F) MAD for WT-HFD, (G) Var for QCs, (H) Var for KO-HFD, and (I) Var for WT-HFD.

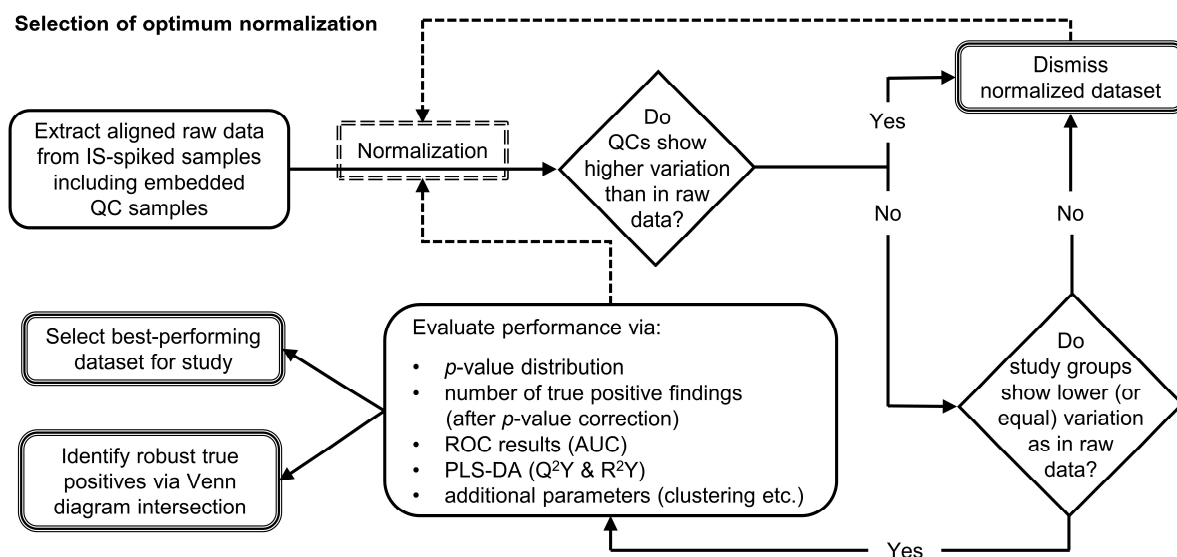


Figure 4. Flowchart for selecting the normalization strategy. Dashed lines represent optional repetition of the scheme until the anticipated number of normalization methods for comparison is reached. If no normalization method fulfills the criteria, raw data is used for further processing.

Investigation of B-MIS Normalization

IS assignment in B-MIS normalization can be dependent on the number of incorporated QC samples and can therefore greatly affect the outcome of the results. Enhancing the number of QCs for B-MIS seems likely to yield robust IS assignment that more accurately returns features and ISs of similar behavior. In order to examine this hypothesis, B-MIS was executed with increasing numbers of QC samples, starting with a minimum of three samples as originally proposed.¹³ In addition, three data matrices derived from randomly assigned ISs (random 1–3) were compared to B-MIS. Results for intragroup variation metrics are shown in Figure 5.

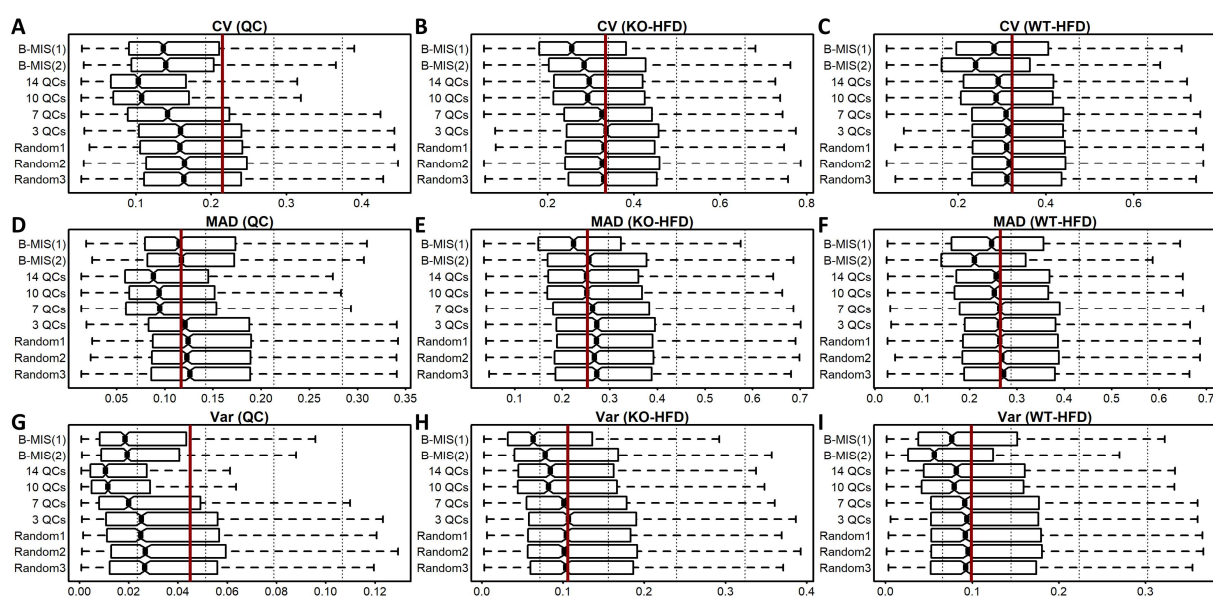


Figure 5. Box-whisker plots of intragroup metrics of variation to show the QC dependency of B-MIS normalization and the comparison with random IS assignment. Results for the comparison of KO-HFD versus WT-HFD in the ESI⁻ data set. The red line represents the median value of the respective metric in raw height data. In B-MIS (1), the IS assignment was optimized for the best precision in the KO-HFD group. In B-MIS (2), the IS assignment was optimized for best precision in the WT-HFD group. (A) CV in QCs, (B) CV in KO-HFD, (C) CV in WT-HFD, (D) MAD for QCs, (E) MAD for KO-HFD, (F) MAD for WT-HFD, (G) Var for QCs, (H) Var for KO-HFD, and (I) Var for WT-HFD.

As expected, variation in QCs and experimental groups is continuously decreased with the increasing number of employed QCs. Here, it should also be pointed out that usage of only three QCs led to results that were not significantly different from the results obtained from random IS assignment. Maximizing the number of QCs should therefore be anticipated but might certainly not always be applicable, for example, when sample volume is limited. We suggest incorporating IS-spiked QCs after a regular scheme for untargeted lipidomics¹⁸ in dense intervals if achievable. Thus, not only are instrument performance monitored and B-MIS results improved, but also other normalization methods like LOWESS are able to ideally address specific problems of sample variation postanalysis. To further explore the potential of B-MIS, IS assignment can also be executed to achieve the best precision in experimental

groups. However, care has to be taken as low intragroup variation cannot always be expected in study samples. As shown in Figure 5, IS assignment based on KO-HFD (B-MIS (1)) and on WT-HFD (B-MIS (2)) both strongly improve variation for the respective experimental group and additionally show lower variation in QC samples than the raw data. According to Figure 4, these B-MIS variants would therefore also be deemed acceptable.

Effect of Normalization on Data with Strong Intergroup Differences

To demonstrate the effects of normalization on highly diverse sample groups, mice that received differing diets were additionally compared (KO-HFD vs KO-CD, Figure 6). As expected, lipidomic profiling revealed numerous profound differences. Except for NOMIS, which underperformed for all negative mode data, and RUV-random (Supporting Information, influence of factor k), the number of true positive findings is almost identical for raw height, LOWESS, t_R -IS, B-MIS, and CCMN. In addition, true positive findings are also highly similar across the different normalization methods (Figure 6B), and even raw data yield excellent results. The observed differences must therefore be of greater amplitude than the effect of normalization. Accordingly, precise optimization of normalization is not implicitly demanded when the majority of the detected differences are of a strong extent. However, if the maximization of the performance is anticipated, most true positive findings are obtained with B-MIS, which also achieves the best results for reducing intragroup variation (Figure S-7).

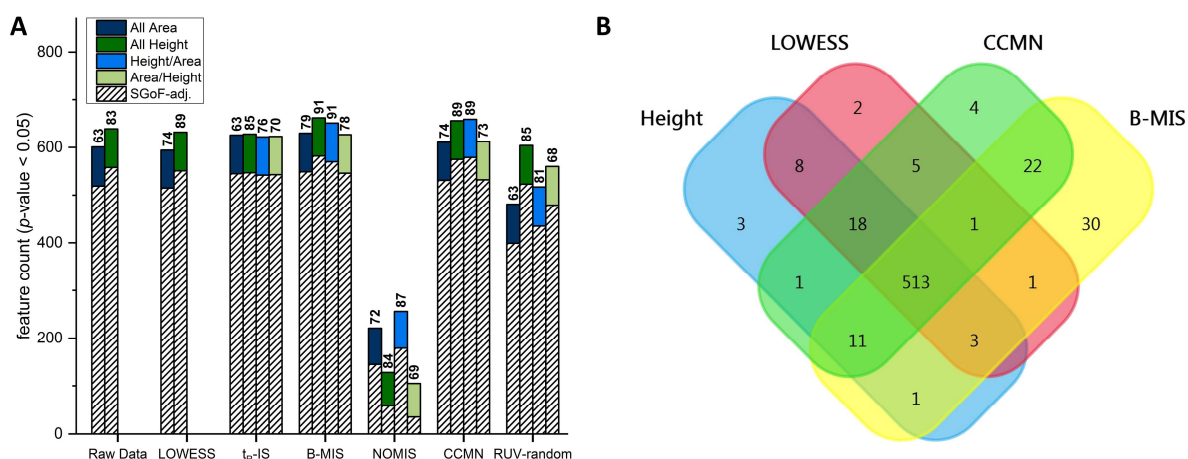


Figure 6. Normalization results for experimental groups with strong intergroup differences. Results for KO-HFD versus KO-CD in the ESI⁻ data set. (A) Column diagram of features with a p -value < 0.05 (colored columns). The overlay of the lined pattern columns shows the number of SGoF-adjusted p -values < 0.05 . Values above the columns show the respective percentage of features in the QCs that showed a CV $< 30\%$. (B) Venn diagram of true positive findings (SGoF-adjusted p -value < 0.05) across different normalization methods based on raw height.

5.2.5. Conclusion

Choosing the optimum normalization is challenging in untargeted omics assays, and obtained results can vary drastically between methods. Special care has to be taken, particularly in model-based approaches prone to overfitting.⁴¹ In this study, a comprehensive evaluation of IS-based normalization methods and highly influential factors like IS selection and raw data type was anticipated.

Overall, the selection of raw height data as response signals was shown to be the most robust choice for untargeted lipidomics data. Nevertheless, the investigation of alternative raw data types, especially mixed data matrices containing ratios of feature heights and IS areas, can yield optimized results due to enhanced precision. Complex normalization approaches like NOMIS and RUV-random showed higher susceptibility to raw data types, IS selection, and additional factors like k (factors of unwanted variance; see the Supporting Information). In contrast, B-MIS, relying on simple analyte/IS ratios, yielded comprehensible results between raw data types and had a high robustness to poorly performing ISs as they were automatically ignored during IS assignment. With an increased number of ISs truly reflecting the main sample composition, B-MIS results are also likely to improve, given that global B-MIS with global SILs is considered as ideal. However, cross-contributing effects of analytes and ISs are not addressed with B-MIS but can be better controlled with CCMN. Ultimately, various normalization methods should be executed and assessed for each data set, as the source of variation can vary and no generally valid normalization can be claimed.

To compare normalization performance between methods, several parameters for comparison like RLA plots, PCA plots, and p -value distribution are proposed, with most of them relying on visual interpretation and individual selection rather than distinct guidelines. Here, numerical characteristics like the amount of true positive findings or intragroup CVs, MADs, and variances could help to define and harmonize the selection process. In general, the inclusion of QC samples in untargeted omics workflows, following a reasonable sample/QC ratio, was shown to be beneficial. Thus, not only is instrument performance easily monitored and QC-based normalization applied, but also IS assignment in B-MIS is executed with increased precision (Figure 5). Furthermore, with the QC-based evaluation of differing normalization methods, obeying a simple scheme (Figure 4), the selection of optimum normalization results can be justified. The key principle is that normalization of data should never lead to increased variation in replicate measurements of identical QC samples. If the variation in experimental groups is additionally decreased, the examined normalization method qualifies to be considered for further evaluation.

Eventually, the general implementation of QCs in untargeted assays enables an effective control mechanism for data integrity postnormalization. Together with an exemplary data set,

the R code to automatically compare raw data, B-MIS, CCMN, NOMIS, and RUV-random via computed statistical parameters and plots of intragroup metrics of variation is provided at <https://github.com/LaemmerhoferLab/Selection-of-IS-Normalization>.

Acknowledgments

We are grateful for support by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation); Project number 374031971 – TRR 240. Prof. Robert Lukowski is gratefully acknowledged for providing the murine plasma samples.

5.2.6. References

- (1) Smith, R. D. *Clin. Chem.* 2012, 58, 528-530.
- (2) Stephenson, D. J.; Hoeflerlin, L. A.; Chalfant, C. E. *Transl. Res.* 2017, 189, 13-29.
- (3) Emwas, A. H. *Methods Mol. Biol.* 2015, 1277, 161-193.
- (4) Mizuno, H.; Ueda, K.; Kobayashi, Y.; Tsuyama, N.; Todoroki, K.; Min, J. Z.; Toyo'oka, T. *Biomed. Chromatogr.* 2017, 31, e3864.
- (5) Forshed, J. *J Proteome Res.* 2017, 16, 3954-3960.
- (6) Ejigu, B. A.; Valkenborg, D.; Baggerman, G.; Vanaerschot, M.; Witters, E.; Dujardin, J. C.; Burzykowski, T.; Berg, M. *OMICS* 2013, 17, 473-485.
- (7) Rampler, E.; Coman, C.; Hermann, G.; Sickmann, A.; Ahrends, R.; Koellensperger, G. *Analyst* 2017, 142, 1891-1899.
- (8) Bueschl, C.; Kluger, B.; Lemmens, M.; Adam, G.; Wiesenberger, G.; Maschietto, V.; Marocco, A.; Strauss, J.; Bodi, S.; Thallinger, G. G.; Krska, R.; Schuhmacher, R. *Metabolomics* 2014, 10, 754-769.
- (9) Bijlsma, S.; Bobeldijk, I.; Verheij, E. R.; Ramaker, R.; Kochhar, S.; Macdonald, I. A.; van Ommen, B.; Smilde, A. K. *Anal. Chem.* 2006, 78, 567-574.
- (10) Cajka, T.; Smilowitz, J. T.; Fiehn, O. *Anal. Chem.* 2017, 89, 12360-12368.
- (11) Katajamaa, M.; Oresic, M. *Bmc Bioinformatics* 2005, 6, 179.
- (12) Sysi-Aho, M.; Katajamaa, M.; Yetukuri, L.; Oresic, M. *Bmc Bioinformatics* 2007, 8, 93.
- (13) Boysen, A. K.; Heal, K. R.; Carlson, L. T.; Ingalls, A. E. *Anal. Chem.* 2018, 90, 1363-1369.
- (14) van der Kloet, F. M.; Bobeldijk, I.; Verheij, E. R.; Jellema, R. H. *J Proteome Res.* 2009, 8, 5132-5141.
- (15) Redestig, H.; Fukushima, A.; Stenlund, H.; Moritz, T.; Arita, M.; Saito, K.; Kusano, M. *Anal. Chem.* 2009, 81, 7974-7980.
- (16) De Livera, A. M.; Sysi-Aho, M.; Jacob, L.; Gagnon-Bartsch, J. A.; Castillo, S.; Simpson, J. A.; Speed, T. P. *Anal. Chem.* 2015, 87, 3606-3615.
- (17) Broadhurst, D.; Goodacre, R.; Reinke, S. N.; Kuligowski, J.; Wilson, I. D.; Lewis, M. R.; Dunn, W. B. *Metabolomics* 2018, 14, 72.
- (18) Dunn, W. B.; Wilson, I. D.; Nicholls, A. W.; Broadhurst, D. *Bioanalysis* 2012, 4, 2249-2264.
- (19) Cleveland, W. S. *Am. Stat.* 1981, 35, 54-54.
- (20) Chawade, A.; Alexandersson, E.; Levander, F. *J Proteome Res.* 2014, 13, 3114-3120.
- (21) Li, B.; Tang, J.; Yang, Q.; Li, S.; Cui, X.; Li, Y.; Chen, Y.; Xue, W.; Li, X.; Zhu, F. *Nucleic Acids Res.* 2017, 45, W162-W170.
- (22) Wang, S.; Chen, X.; Dan, D.; Zheng, W.; Hu, L.; Yang, H.; Cheng, J.; Gong, M. *Anal. Chem.* 2018, 90, 11124-11130.
- (23) Illison, J.; Tian, L.; McClafferty, H.; Werno, M.; Chamberlain, L. H.; Leiss, V.; Sassmann, A.; Offermanns, S.; Ruth, P.; Shipston, M. J.; Lukowski, R. *Diabetes* 2016, 65, 3621-3635.
- (24) Sarafian, M. H.; Gaudin, M.; Lewis, M. R.; Martin, F. P.; Holmes, E.; Nicholson, J. K.; Dumas, M. E. *Anal. Chem.* 2014, 86, 5766-5774.
- (25) Calderón, C.; Sanwald, C.; Schlotterbeck, J.; Drotleff, B.; Lämmerhofer, M. *Anal. Chim. Acta* 2019, 1048, 66-74.
- (26) Wang, M.; Wang, C.; Han, X. *Mass Spectrom. Rev.* 2017, 36, 693-714.
- (27) Tsugawa, H.; Cajka, T.; Kind, T.; Ma, Y.; Higgins, B.; Ikeda, K.; Kanazawa, M.; VanderGheynst, J.; Fiehn, O.; Arita, M. *Nat. Methods* 2015, 12, 523-526.
- (28) Schlotterbeck, J.; Chatterjee, M.; Gawaz, M.; Lämmerhofer, M. *Anal. Chim. Acta* 2019, 1046, 1-15.
- (29) Fiehn, O. *Curr. Protoc. Mol. Biol.* 2016, 114, 21.33.1-21.33.11.
- (30) Martinez, M. N.; Bartholomew, M. J. *Pharmaceutics* 2017, 9, 14.
- (31) De Livera, A. M.; Dias, D. A.; De Souza, D.; Rupasinghe, T.; Pyke, J.; Tull, D.; Roessner, U.; McConville, M.; Speed, T. P. *Anal. Chem.* 2012, 84, 10768-10776.
- (32) Benjamini, Y.; Hochberg, Y. *J R Stat. Soc. B* 1995, 57, 289-300.
- (33) Carvajal-Rodriguez, A.; de Una-Alvarez, J.; Rolan-Alvarez, E. *Bmc Bioinformatics* 2009, 10, 209.
- (34) Castro-Conde, I.; de Una-Alvarez, J. *Biometrical J.* 2015, 57, 108-122.

- (35) De Livera, A. M.; Bowne, J. B. *Package metabolomics for R*, 2014, <https://CRAN.R-project.org/src/contrib/Archive/metabolomics>.
- (36) De Livera, A. M.; Olshansky, G. *Package NormalizeMets for R*, 2018, <https://CRAN.R-project.org/package=NormalizeMets>.
- (37) Castro-Conde, I.; Una Alvarez, J. *Package sgof for R*, 2016, <https://CRAN.R-project.org/package=sgof>.
- (38) Thevenot, E. A.; Roux, A.; Xu, Y.; Egan, E.; Junot, C. J. *Proteome Res.* 2015, 14, 3322-3335.
- (39) Tuszynski, J. *Package caTools for R*, 2018. <https://CRAN.R-project.org/package=caTools>.
- (40) Storey, J. D.; Bass, A. J.; Dabney, A.; Robinson, D. *Package qvalue for R*, 2019, <http://github.com/jdstorey/qvalue>.
- (41) Fan, S.; Kind, T.; Cajka, T.; Hazen, S. L.; Tang, W. H. W.; Kaddurah-Daouk, R.; Irvin, M. R.; Arnett, D. K.; Barupal, D. K.; Fiehn, O. *Anal. Chem.* 2019, 91, 3590-3596.

5.2.7. Supporting Information

Table S-1. Internal standards in QCs and real samples.^a

Compound name	Conc. [ng/mL]	t _R [min]	Adduct type		IS	
			pos	neg	pos	neg
18:1(d7) LPC	340.0	1.61 ± 0.01	[M+H] ⁺	[M+FA-H] ⁻	√	√
18:1(d7) LPE	70.7	1.65 ± 0.00	[M+H] ⁺	[M-H] ⁻	√	√
15:0-18:1(d7) PI	121.3	4.53 ± 0.00	[M+H] ⁺	-	√	-
15:0-18:1(d7) PS	56.0	4.69 ± 0.03	-	[M-H] ⁻	-	√
15:0-18:1(d7) PG	388.0	4.78 ± 0.04	-	[M-H] ⁻	-	√
d18:1-18:1(d9) SM	412.0	4.87 ± 0.00	[M+H] ⁺	[M+FA-H] ⁻	√	√
15:0-18:1(d7) PC	2,143	5.24 ± 0.00	[M+H] ⁺	[M+FA-H] ⁻	√	√
15:0-18:1(d7) PE	76.0	5.42 ± 0.00	[M+H] ⁺	[M-H] ⁻	√	√
15:0-18:1(d7) DAG	125.3	6.67 ± 0.00	[M+NH ₄] ⁺	-	√	-
15:0-18:1(d7)-15:0 TAG	764.0	10.59 ± 0.00	[M+NH ₄] ⁺	-	√	-
Arachidonic acid(d8)	133.3	2.73 ± 0.01	-	[M-H] ⁻	-	√
α-Linolenic(d14)	133.3	2.28 ± 0.01	-	[M-H] ⁻	-	√
Linoleic acid(d4)	133.3	2.84 ± 0.01	-	[M-H] ⁻	-	√

^aValues are given for the final concentration in the supernatant after sample preparation. t_R values are based on results in QC samples. LPC: lysophosphatidylcholine; LPE: lysophosphatidylethanolamine, PI: phosphatidylinositol, PS: phosphatidylserine; PG: phosphatidylglycerol, SM: sphingomyelin, PC: phosphatidylcholine, PE: phosphatidylethanolamine; DAG: diacylglycerol; TAG: triacylglycerol.

Table S-2. Results for acceptance criteria of internal standards.^b

Compound name	positive mode		negative mode	
	S/N	Asymmetry factor	S/N	Asymmetry factor
18:1(d7) LPC	3,756 ± 255	1.07 ± 0.12	3,960 ± 707	1.11 ± 0.14
18:1(d7) LPE	761 ± 122	1.20 ± 0.08	1,843 ± 180	1.05 ± 0.12
15:0-18:1(d7) PI	190 ± 25	1.13 ± 0.24	-	-
15:0-18:1(d7) PS	-	-	385 ± 48	1.35 ± 0.16
15:0-18:1(d7) PG	-	-	3,087 ± 395	1.33 ± 0.19
d18:1-18:1(d9) SM	3,683 ± 453	0.90 ± 0.07	1,734 ± 217	1.36 ± 0.13
15:0-18:1(d7) PC	14,784 ± 2,023	1.32 ± 0.12	9,225 ± 1,410	1.24 ± 0.09
15:0-18:1(d7) PE	848 ± 109	1.17 ± 0.10	2,010 ± 321	1.18 ± 0.10
15:0-18:1(d7) DAG	1,234 ± 264	1.17 ± 0.07	-	-
15:0-18:1(d7)-15:0 TG	2,895 ± 632	1.09 ± 0.08	-	-
Arachidonic acid(d8)	-	-	3,292 ± 496	1.05 ± 0.14
α-Linolenic(d14)	-	-	2,484 ± 552	1.04 ± 0.13
Linoleic acid(d4)	-	-	35 ± 6	1.05 ± 0.11

^bS/N values are based on MarkerView estimates.

Table S-3. Sequence of analysis (ESI⁻ and ESI⁺ dataset).^c

Sample Name	Class	Analytical order	Sample Name	Class	Analytical order
QC_eq1_IDA	QC_eq	0	B9258	KO-HFD	29
QC_eq2_IDA	QC_eq	1	D0181	KO-HFD	30
QC_1	QC	2	D3371	KO-HFD	31
D0163	WT-CD	3	QC_6	QC	32
D0138	WT-CD	4	QC_7	QC	33
D3432	KO-CD	5	D0186	KO-CD	34
D0143	WT-CD	6	D3379	WT-CD	35
D3381	WT-CD	7	D3357	WT-HFD	36
QC_2	QC	8	D3383	WT-CD	37
D3377	KO-CD	9	D3396	KO-HFD	38
B9256	KO-HFD	10	QC_8	QC	39
D0137	WT-CD	11	D3394	KO-HFD	40
B9305	WT-HFD	12	D3380	WT-CD	41
B9343	KO-CD	13	D3305	WT-HFD	42
QC_3	QC	14	D3397	WT-HFD	43
B9257	KO-HFD	15	D3370	WT-HFD	44
D0191	KO-CD	16	QC_9	QC	45
D0183	KO-HFD	17	D0139	WT-CD	46
D0182	KO-HFD	18	D0168	KO-CD	47
D3392	KO-HFD	19	D0176	WT-HFD	48
QC_4	QC	20	B9345	WT-CD	49
D3417	WT-HFD	21	B9324	WT-HFD	50
D0185	WT-CD	22	QC_10	QC	51
D0167	WT-HFD	23	D0142	KO-CD	52
D3385	KO-CD	24	QC_11	QC	53
D3435	KO-CD	25	QC_12	QC	54
QC_5	QC	26	QC_13	QC	55
B9278	WT-HFD	27	QC_14	QC	56
D3353	KO-HFD	28			

^cQC_eq samples were injected for equilibration and conditioning of the column. Sample D3371 was removed from further data evaluation (including processing via MS-DIAL; see Table S-4) as a classification error was suspected (see Figure S-22 and S-23).

Table S-4. MS-DIAL parameters.^d

Parameter	Setting	
	ESI ⁺	ESI ⁻
Data collection range [min]	0.5 – 13.0	0.5 – 13.0
Mass range (as <i>m/z</i> ratio)	50 – 1250	50 – 1250
MS ¹ tolerance (as <i>m/z</i> ratio)	0.01	0.01
MS ² tolerance (as <i>m/z</i> ratio)	0.025	0.025
Smoothing level	1	1
Minimum points per peak	4	4
Minimum peak height [cps]	500	500
Identification tolerance for <i>t_R</i> [min]	1.0	1.0
ID score	80 %	80 %
Alignment reference sample	QC_3	QC_3
Alignment <i>t_R</i> tolerance [min]	0.2	0.2
Alignment MS ¹ tolerance (as <i>m/z</i> ratio)	0.02	0.02
Alignment filter: at least found in one group	70 %	70 %
Aligned features after blank subtraction	2,083	1,103
Identified aligned features	529	179
Annotated aligned features	29	52

^dBlank subtraction was done by removing signals that had a foldchange < 5 in the averaged samples compared to the averaged signals in three replicates of a processed blank sample. Missing values were replaced with a value that corresponded to 10% of the minimum in all samples (standard method embedded in MS-DIAL software). Although QC_3 was registered as an outlier in the ESI⁻ dataset, data was not essentially reprocessed. The iterative peak finding and alignment process of MS-DIAL (see ref. 27 Supplementary Figure 10), utilizing all samples after preliminary peak finding in the reference file, ensures maximum coverage. Requirements for reference files are, that they should not be blanks or external QCs and that no chromatographic and/or mass shifts are present. Comparative reprocessing using QC_5 resulted in 98.3 % identical aligned features with identical response values.

Table S-5. Precision of internal standards in QC samples (ESI⁻ and ESI⁺ dataset).^e

Compound name	positive mode		negative mode	
	CV raw area [%]	CV raw height [%]	CV raw area [%]	CV raw height [%]
18:1(d7) LPC	3.1	3.5	5.8	7.6
18:1(d7) LPE	4.8	6.3	7.3	8.4
15:0-18:1(d7) PI	6.9	5.1	-	-
15:0-18:1(d7) PS	-	-	6.6	5.3
15:0-18:1(d7) PG	-	-	6.1	6.6
d18:1-18:1(d9) SM	12.9	4.5	22.6	9.2
15:0-18:1(d7) PC	2.1	2.8	2.2	3.9
15:0-18:1(d7) PE	5.8	7.8	6.2	8.9
15:0-18:1(d7) DAG	5.1	5.1	-	-
15:0-18:1(d7)-15:0 TG	12.5	11.0	-	-
Arachidonic acid(d8)	-	-	7.1	8.7
α -Linolenic(d14)	-	-	11.1	12.7
Linoleic acid(d4)	-	-	19.7	10.1

^eFor calculation outlying QC samples were not considered (positive mode: QC10, QC11; negative mode: QC3, QC7). Although d18:1-18:1(d9) SM and Linoleic acid(d4) show great variation in raw area, B-MIS results were comparable even if both internal standards were not used and even if area values were replaced with height values in these cases.

Table S-6. Extracted parameters for comparison of normalization (ESI⁻ dataset).^f

Normalization	<i>p</i> -value < 0.05	<i>p</i> -value < 0.05 (SGoF)	Median CV (QCs)	Mean AUC	PLS-DA R ² Y / Q ² Y
Raw area	51	0	25.6	0.815 (14)	-
Raw height	53	0	21.5	0.816 (19)	-
LOWESS (area)	33	0	22.4	0.800 (9)	-
LOWESS (height)	37	0	19.2	0.801 (9)	-
mTIC	70	5	13.4	0.823 (19)	-
t _R -IS (area)	83	15	23.5	0.852 (44)	-
t _R -IS (height)	101	32	15.6	0.855 (48)	-
t _R -IS (height/area)	98	29	16.6	0.847 (43)	-
t _R -IS (IS height)	92	23	19.5	0.860 (47)	-
B-MIS (area)	84	16	14.1	0.847 (35)	0.998 / 0.522 (4)
B-MIS (height)	96	27	10.3	0.839 (39)	0.991 / 0.535 (4)
B-MIS (height/area)	103	34	10.4	0.842 (44)	0.993 / 0.451 (4)
B-MIS (area/height)	84	18	13.7	0.844 (36)	0.998 / 0.634 (4)
NOMIS (area)	26	0	21.3	0.797 (11)	-
NOMIS (height)	53	0	18.0	0.831 (21)	-
NOMIS (height/area)	21	0	16.8	0.786 (10)	-
NOMIS (IS height)	54	0	22.8	0.835 (22)	-
CCMN (area)	91	22	19.4	0.839 (28)	0.989 / 0.589 (3)
CCMN (height)	89	21	14.7	0.842 (32)	0.976 / 0.468 (3)
CCMN (height/area)	91	22	14.5	0.836 (31)	0.976 / 0.429 (3)
CCMN (IS height)	88	19	19.7	0.845 (29)	0.989 / 0.615 (3)
RUV-random (area)	65	0	23.5	0.841 (25)	-
RUV-random (height)	74	6	16.7	0.831 (35)	-
RUV-random (height/area)	65	0	19.3	0.827 (28)	-
RUV-random (area/height)	79	12	21.7	0.844 (38)	-

^fCVs are reported in %. Mean AUC values are derived from ROC of the 25 most significant features. The number in parentheses lists the amount of features with an AUC > 0.8. R²Y and Q²Y are reported for the optimum number of predictive components *p* (in parentheses). When the first predictive component was not significant, PLS-DA was rejected and no results are given.

Table S-7. Extracted parameters for comparison of normalization (ESI⁻ dataset, outliers removed).⁹

Normalization	<i>p</i> -value < 0.05	<i>p</i> -value < 0.05 (SGoF)	Median CV (QCs)	Mean AUC	PLS-DA R ² Y / Q ² Y
Raw area	125	54	13.9	0.919 (139)	0.997 / 0.554 (3)
Raw height	134	62	10.7	0.929 (143)	0.712 / 0.406 (1)
LOWESS (area)	110	39	10.0	0.912 (124)	0.997 / 0.483 (3)
LOWESS (height)	119	48	7.7	0.918 (119)	0.966 / 0.450 (2)
mTIC	158	84	10.0	0.921 (166)	0.998 / 0.709 (4)
t _R -IS (area)	97	29	19.3	0.927 (135)	0.960 / 0.337 (2)
t _R -IS (height)	169	96	12.8	0.936 (183)	0.684 / 0.380 (1)
t _R -IS (height/area)	128	58	13.0	0.921 (142)	0.622 / 0.266 (1)
t _R -IS (IS height)	134	62	16.0	0.939 (163)	0.965 / 0.454 (2)
B-MIS (area)	138	66	11.6	0.914 (148)	0.975 / 0.512 (2)
B-MIS (height)	184	109	8.6	0.937 (185)	0.953 / 0.499 (2)
B-MIS (height/area)	152	80	8.9	0.929 (160)	0.969 / 0.437 (2)
B-MIS (area/height)	144	71	11.5	0.936 (160)	0.963 / 0.503 (2)
NOMIS (area)	23	0	20.7	0.842 (19)	-
NOMIS (height)	16	0	17.8	0.833 (25)	-
NOMIS (height/area)	18	0	16.1	0.830 (19)	-
NOMIS (IS height)	16	0	22.4	0.826 (21)	-
CCMN (area)	84	16	20.2	0.904 (99)	0.739 / 0.225 (1)
CCMN (height)	95	26	17.1	0.911 (106)	0.729 / 0.294 (1)
CCMN (height/area)	91	23	15.6	0.869 (48)	0.706 / 0.226 (1)
CCMN (IS height)	128	56	14.7	0.919 (139)	0.998 / 0.508 (3)
RUV-random (area)	90	22	21.5	0.929 (143)	0.999 / 0.779 (4)
RUV-random (height)	145	72	13.8	0.912 (124)	0.737 / 0.394 (1)
RUV-random (height/area)	102	32	16.8	0.918 (119)	0.716 / 0.310 (1)
RUV-random (area/height)	138	66	17.6	0.921 (166)	0.995 / 0.619 (3)

⁹Following outliers were removed: B9257, D0181, D3396, B9324, D0167, D3397, QC3, QC7. CVs are reported in %. Mean AUC values are derived from ROC of the 25 most significant features. The number in parentheses lists the amount of features with an AUC > 0.8. R²Y and Q²Y are reported for the optimum number of predictive components *p* (in parentheses). When the first predictive component was not significant, PLS-DA was rejected and no results are given.

Table S-8. Extracted parameters for comparison of normalization (ESI⁺ dataset).^h

Normalization	<i>p</i> -value < 0.05	<i>p</i> -value < 0.05 (SGoF)	Median CV (QCs)	Mean AUC	PLS-DA R ² Y / Q ² Y
Raw area	67	0	28.0	0.845 (34)	-
Raw height	78	0	25.5	0.841 (35)	-
LOWESS (area)	87	0	24.6	0.849 (37)	-
LOWESS (height)	99	0	21.6	0.854 (56)	-
mTIC	119	0	18.7	0.846 (42)	-
t _R -IS (area)	82	0	22.2	0.855 (42)	-
t _R -IS (height)	97	0	18.8	0.855 (35)	-
t _R -IS (height/area)	94	0	19.5	0.844 (37)	-
t _R -IS (IS height)	82	0	22.0	0.842 (32)	-
B-MIS (area)	89	0	18.1	0.840 (34)	-
B-MIS (height)	94	0	14.2	0.846 (39)	-
B-MIS (height/area)	93	0	14.4	0.839 (31)	-
B-MIS (area/height)	70	0	17.9	0.836 (26)	-
NOMIS (area)	95	0	25.3	0.839 (26)	-
NOMIS (height)	202	78	21.6	0.862 (85)	0.980 / 0.417 (3)
NOMIS (height/area)	110	0	22.8	0.837 (31)	-
NOMIS (IS height)	199	74	24.4	0.866 (65)	0.984 / 0.434 (3)
CCMN (area)	133	14	21.1	0.861 (64)	-
CCMN (height)	167	44	17.4	0.862 (57)	0.995 / 0.472 (3)
CCMN (height/area)	150	28	17.9	0.870 (70)	-
CCMN (IS height)	146	24	20.8	0.859 (62)	-
RUV-random (area)	181	57	21.3	0.865 (74)	0.708 / 0.078 (1)
RUV-random (height)	168	44	17.1	0.871 (58)	-
RUV-random (height/area)	182	60	18.0	0.845 (34)	0.692 / 0.081 (1)
RUV-random (area/height)	147	24	20.6	0.841 (35)	-

^hCVs are reported in %. Mean AUC values are derived from ROC of the 25 most significant features. The number in parentheses lists the amount of features with an AUC > 0.8. R²Y and Q²Y are reported for the optimum number of predictive components *p* (in parentheses). When the first predictive component was not significant, PLS-DA was rejected and no results are given.

Table S-9. Extracted parameters for comparison of normalization (ESI⁺ dataset, outliers removed).ⁱ

Normalization	<i>p</i> -value < 0.05	<i>p</i> -value < 0.05 (SGoF)	Median CV (QCs)	Mean AUC	PLS-DA R ² Y / Q ² Y
Raw area	209	83	18.6	0.943 (290)	0.765 / 0.338 (1)
Raw height	228	101	14.7	0.958 (302)	0.767 / 0.363 (1)
LOWESS (area)	203	78	15.2	0.941 (279)	0.782 / 0.351 (1)
LOWESS (height)	230	103	12.3	0.960 (288)	0.783 / 0.375 (1)
mTIC	277	148	14.7	0.951 (378)	0.698 / 0.358 (1)
t _R -IS (area)	195	71	19.5	0.950 (261)	0.995 / 0.705 (3)
t _R -IS (height)	183	59	15.9	0.939 (246)	0.993 / 0.600 (3)
t _R -IS (height/area)	210	86	16.2	0.952 (279)	0.994 / 0.723 (3)
t _R -IS (IS height)	158	35	19.3	0.938 (212)	0.994 / 0.556 (3)
B-MIS (area)	230	104	16.8	0.960 (278)	0.748 / 0.319 (1)
B-MIS (height)	234	107	13.0	0.959 (285)	0.745 / 0.309 (1)
B-MIS (height/area)	257	129	13.1	0.961 (309)	0.752 / 0.338 (1)
B-MIS (area/height)	202	77	16.5	0.945 (265)	0.995 / 0.666 (3)
NOMIS (area)	108	0	24.3	0.888 (133)	0.996 / 0.217 (3)
NOMIS (height)	279	151	20.3	0.933 (305)	0.947 / 0.442 (2)
NOMIS (height/area)	108	0	21.9	0.892 (141)	0.994 / 0.221 (3)
NOMIS (IS height)	269	141	23.6	0.935 (284)	0.948 / 0.439 (2)
CCMN (area)	195	70	20.2	0.939 (265)	0.775 / 0.279 (1)
CCMN (height)	145	23	14.9	0.932 (215)	0.794 / 0.264 (1)
CCMN (height/area)	206	81	17.2	0.947 (269)	0.767 / 0.290 (1)
CCMN (IS height)	141	20	18.9	0.932 (221)	0.998 / 0.728 (3)
RUV-random (area)	237	110	20.4	0.928 (284)	0.705 / 0.294 (1)
RUV-random (height)	148	25	15.2	0.918 (189)	0.997 / 0.695 (3)
RUV-random (height/area)	242	114	17.6	0.932 (278)	0.693 / 0.286 (1)
RUV-random (area/height)	149	27	18.9	0.931 (188)	0.998 / 0.696 (3)

ⁱFollowing outliers were removed: D0181, D0182, D3392, B9324, D0167, D3396, QC10, QC11. CVs are reported in %. Mean AUC values are derived from ROC of the 25 most significant features. The number in parentheses lists the amount of features with an AUC > 0.8. R²Y and Q²Y are reported for the optimum number of predictive components *p* (in parentheses). When the first predictive component was not significant, PLS-DA was rejected and no results are given.

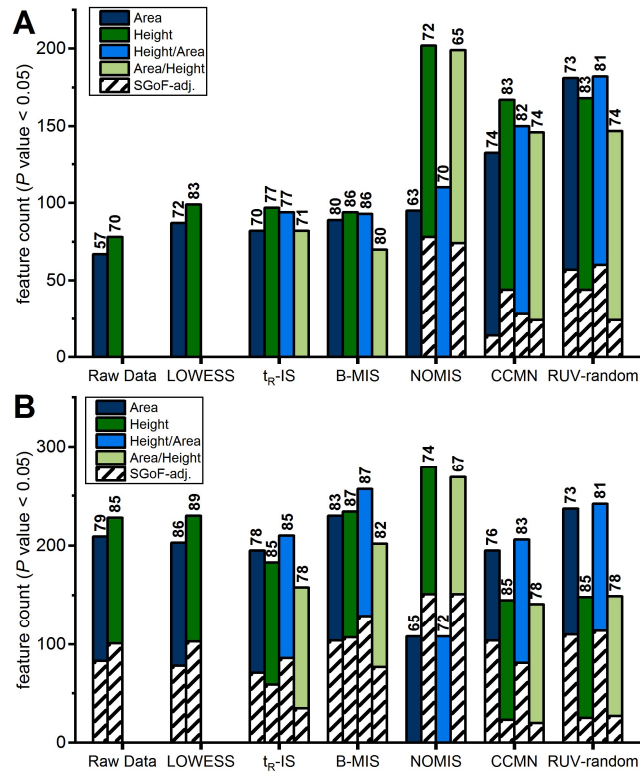


Figure S-1. Comparison of raw data type and normalization (ESI⁺ dataset). Number of features with a p -value < 0.05 are represented by colored columns. Overlay of the lined pattern columns show the number of SGoF-adjusted p -values < 0.05. Values above the columns show the respective percentage of features in the QCs that showed a CV < 30 %. For RUV-random k was set to 3. A: ESI⁺ dataset covering all samples. B: ESI⁺ dataset after removal of outliers (D0181, D0182, D3392, B9324, D0167, D3396, QC10, QC11).

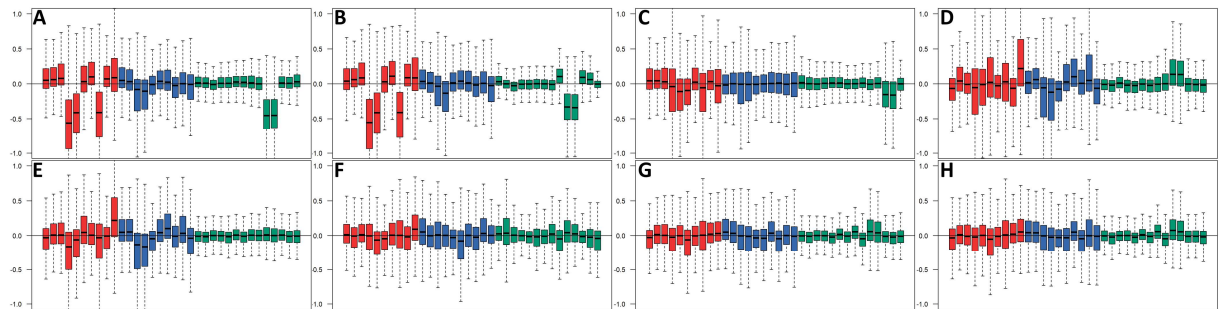


Figure S-2. Within-group RLA plots (ESI⁺ dataset). Different experimental groups are represented by different colors: red (KO-HFD), blue (WT-HFD), green (QC). RLA plots (comprised of Box-Whisker plots) were sorted according to group. RLA plots sorted after analytical order were also investigated to detect systematic signal drifts, which was not the case. A: raw height data, B: LOWESS, C: mTIC, D: t_R-IS, E: B-MIS, F: NOMIS, G: CCMN, H: RUV-random ($k = 3$).

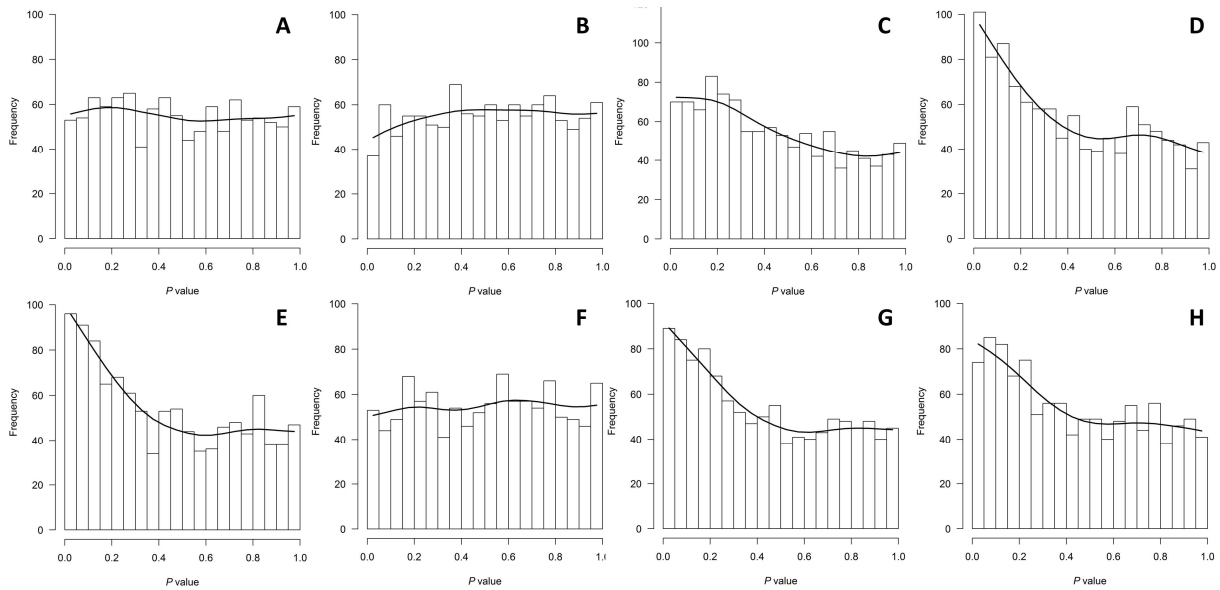


Figure S-3. p-value distribution for different normalization methods (ESI⁻ dataset). Results for comparison KO-HFD versus WT-HFD. Histogram has 20 breaks so that each column represents a 0.05 interval. A: raw height, B: LOWESS, C: mTIC, D: t_R-IS, E: B-MIS, F: NOMIS, G: CCMN, H: RUV-random ($k = 3$).

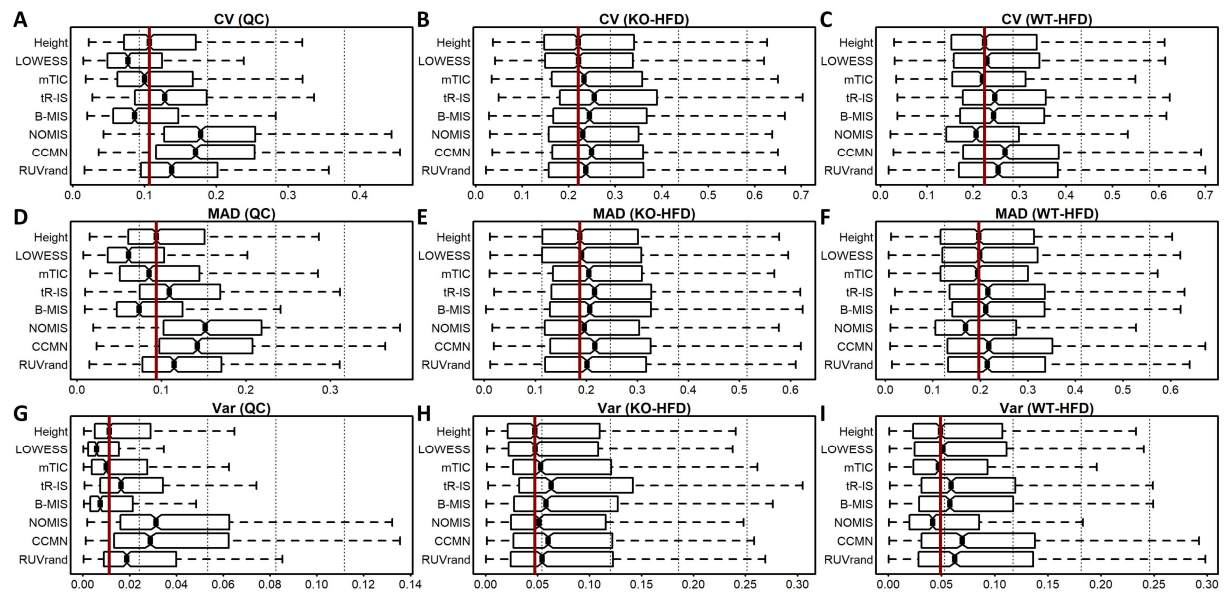


Figure S-4. Comparison of intra-group metrics of variation (ESI⁻ dataset, outliers removed). Following outliers were removed: B9257, D0181, D3396, B9324, D0167, D3397, QC3, QC7. Box-Whisker plots are shown. The red line represents the median value of the respective metric in raw height data. A: CV in QCs, B: CV in KO-HFD, C: CV in WT-HFD, D: MAD for QCs, E: MAD for KO-HFD, F: MAD for WT-HFD, G: Var for QCs, H: Var for KO-HFD, I: Var for WT-HFD. According to the proposed guidelines (Figure 4) LOWESS is accepted as a valid normalization.

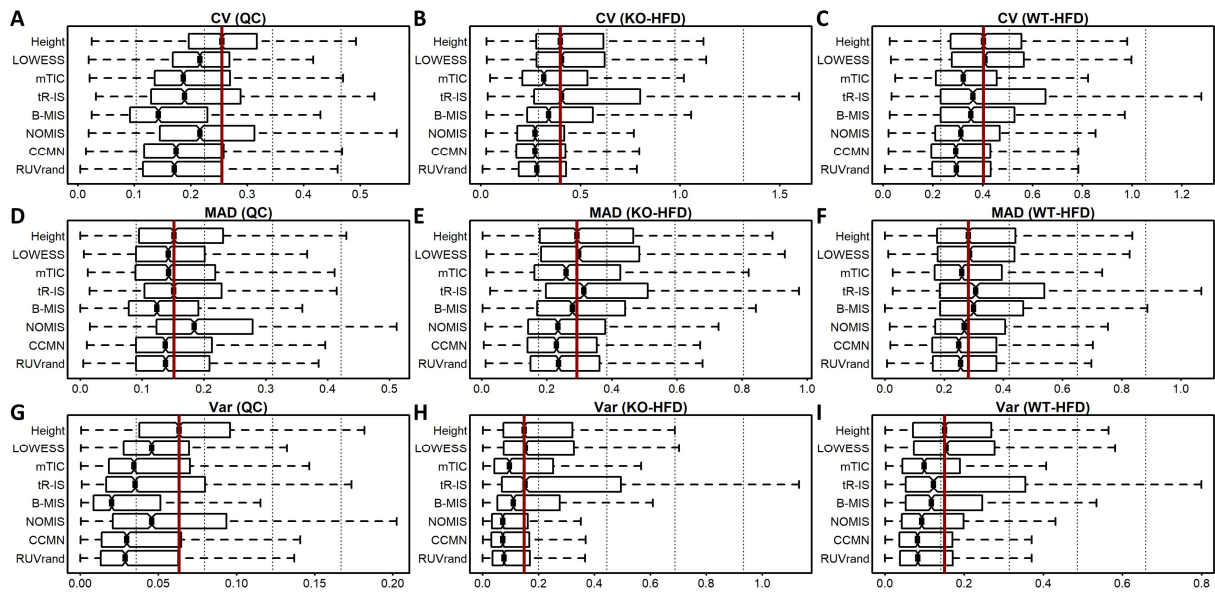


Figure S-5. Comparison of intra-group metrics of variation (ESI⁺ dataset). Box-Whisker plots are shown. The red line represents the median value of the respective metric in raw height data. A: CV in QCs, B: CV in KO-HFD, C: CV in WT-HFD, D: MAD for QCs, E: MAD for KO-HFD, F: MAD for WT-HFD, G: Var for QCs, H: Var for KO-HFD, I: Var for WT-HFD. According to the proposed guidelines (Figure 4) LOWESS, mTIC, CCMN and RUV-random are accepted as valid normalizations.

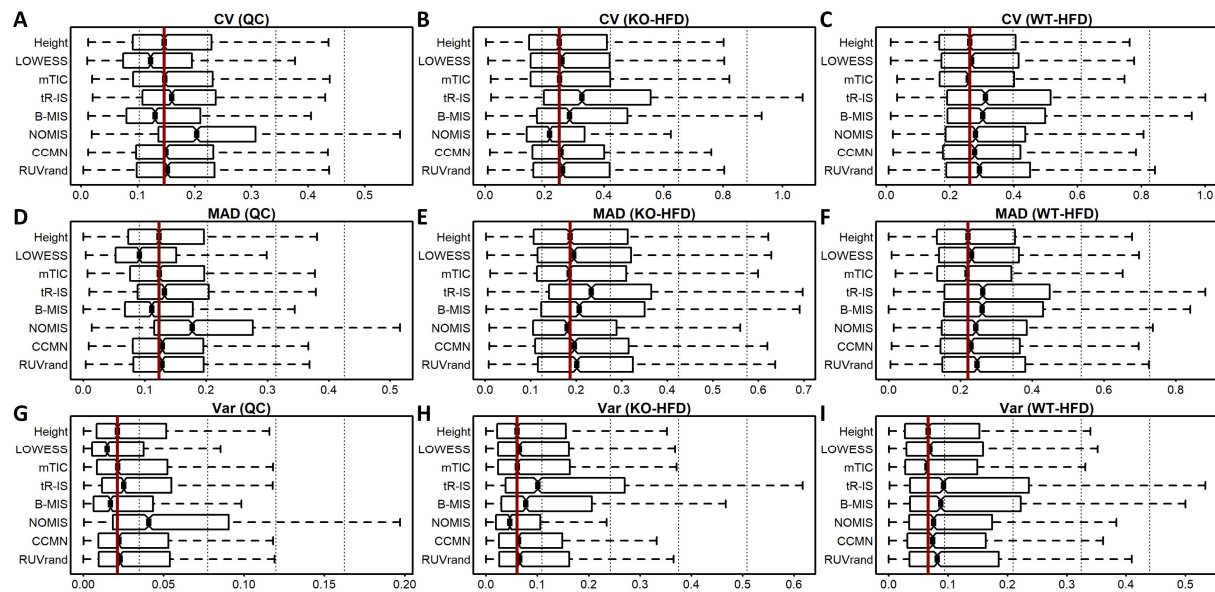


Figure S-6. Comparison of intra-group metrics of variation (ESI⁺ dataset, outliers removed). Following outliers were removed: D0181, D0182, D3392, B9324, D0167, D3396, QC10, QC11. Box-Whisker plots are shown. The red line represents the median value of the respective metric in raw height data. A: CV in QCs, B: CV in KO-HFD, C: CV in WT-HFD, D: MAD for QCs, E: MAD for KO-HFD, F: MAD for WT-HFD, G: Var for QCs, H: Var for KO-HFD, I: Var for WT-HFD. According to the proposed guidelines (Figure 4) LOWESS and mTIC are accepted as valid normalizations.

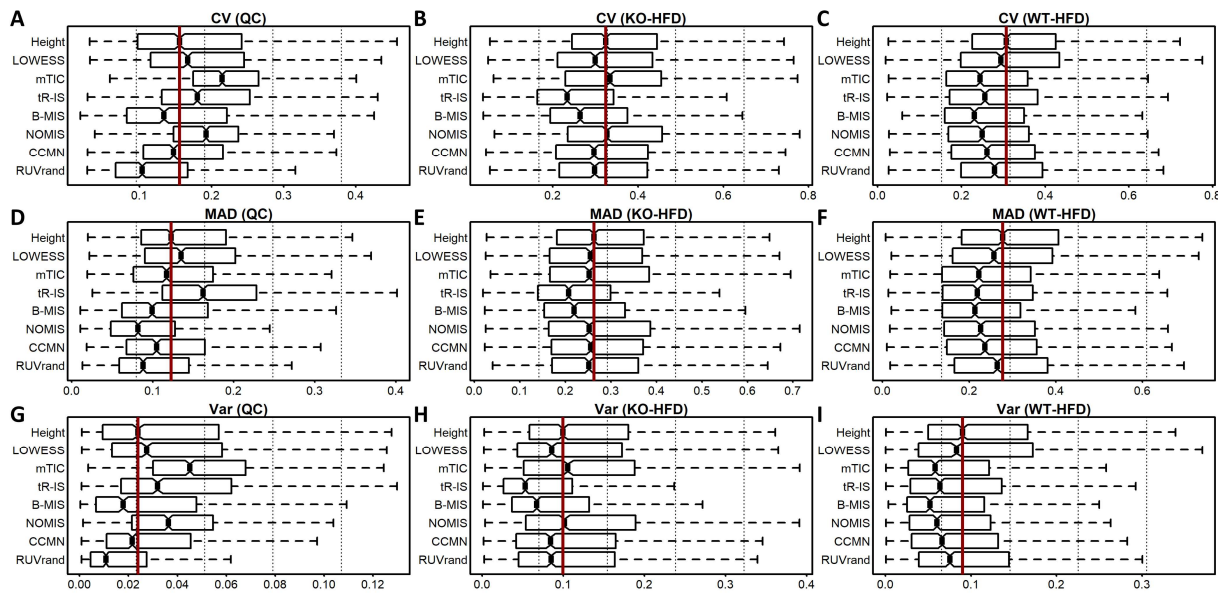


Figure S-7. Comparison of intra-group metrics of variation (ESI⁻ dataset, KO-HFD vs. KO-CD). Box-Whisker plots are shown. The red line represents the median value of the respective metric in raw height data. A: CV in QCs, B: CV in KO-HFD, C: CV in KO-CD, D: MAD for QCs, E: MAD for KO-HFD, F: MAD for KO-CD, G: Var for QCs, H: Var for KO-HFD, I: Var for KO-CD. According to the proposed guidelines (Figure 4) B-MIS, CCMN and RUV-random are accepted as valid normalizations.

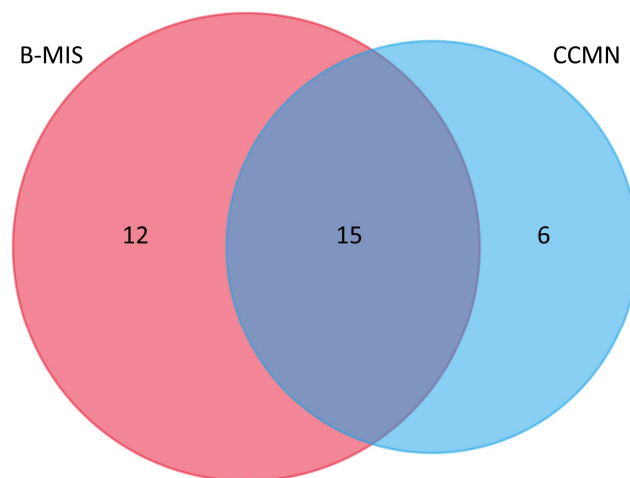


Figure S-8. True positive findings for accepted normalization methods, (ESI⁻ dataset). After selection of normalization methods according to Figure 4 and results in Supplemental Table 6, B-MIS and CCMN were deemed acceptable. Robust true positive findings are represented by the intersection area in the Venn diagram.

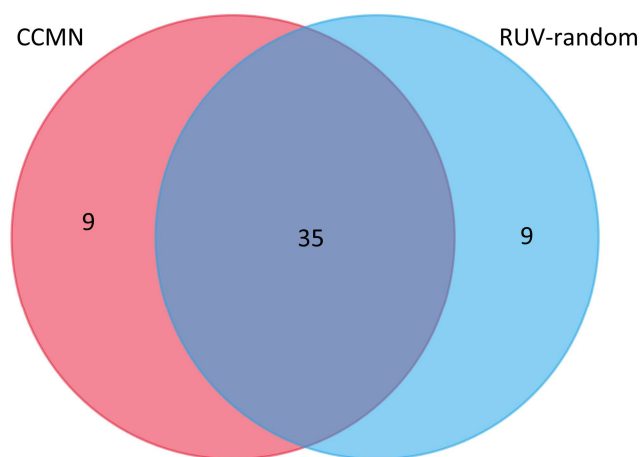


Figure S-9. True positive findings for accepted normalization methods (ESI⁺ dataset). After selection of normalization methods according to Figure 4 and results in Supplemental Table 8, CCMN and RUV-random were deemed acceptable. Robust true positive findings are represented by the intersection area in the Venn diagram.

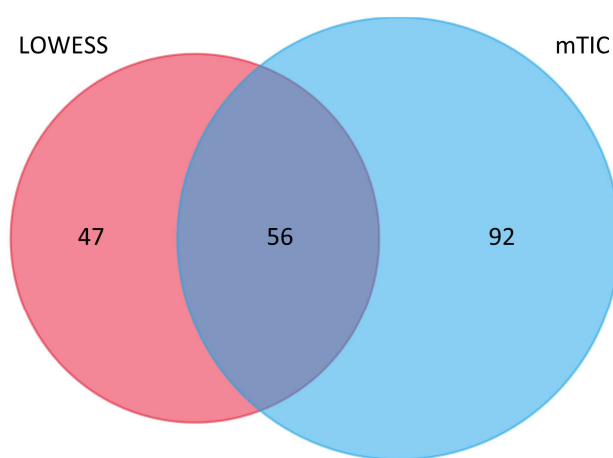


Figure S-10. True positive findings for accepted normalization methods (ESI⁺ dataset, outliers removed). After selection of normalization methods according to Figure 4 and results in Supplemental Table 9, CCMN and RUV-random were deemed acceptable. Robust true positive findings are represented by the intersection area in the Venn diagram.

Text T-1. Influence of factor k for RUV-random normalization.

RUV-random normalization requires input of additional parameters like k , which is describing the number of factors of unwanted variance. As only ISs were selected as control metabolites, values for k ranged between 1 and the total number of ISs. De Livera et. al¹⁶ recommend utilizing a plot that shows the variance explained by the principal components of control metabolite abundances against k factors (see Figure S-11). Additionally RLA plots and p -value histograms can be taken into account.³¹ In order to investigate the influence of k , RUV-random was executed with all variants of k and results were compared using intra-group metrics of variation (see Figures S-13 – S-16). A clear tendency of increasing variation in QC samples with increasing k is observed for all datasets. Although RUV-random shows great potential for reducing variation in experimental groups this effect regularly goes along with an increase of variation in QCs. Especially for outlier removed datasets (Figure S-14 & S-16), RUV-random is not able to outperform raw data and only small values for k should be considered. However, RUV-random shows great potential to improve variation in the ESI⁺ dataset (see Figure S-15). Enhanced investigation of RUV-random and factor k by comparison of intra-group metrics of variation is therefore mandatory.

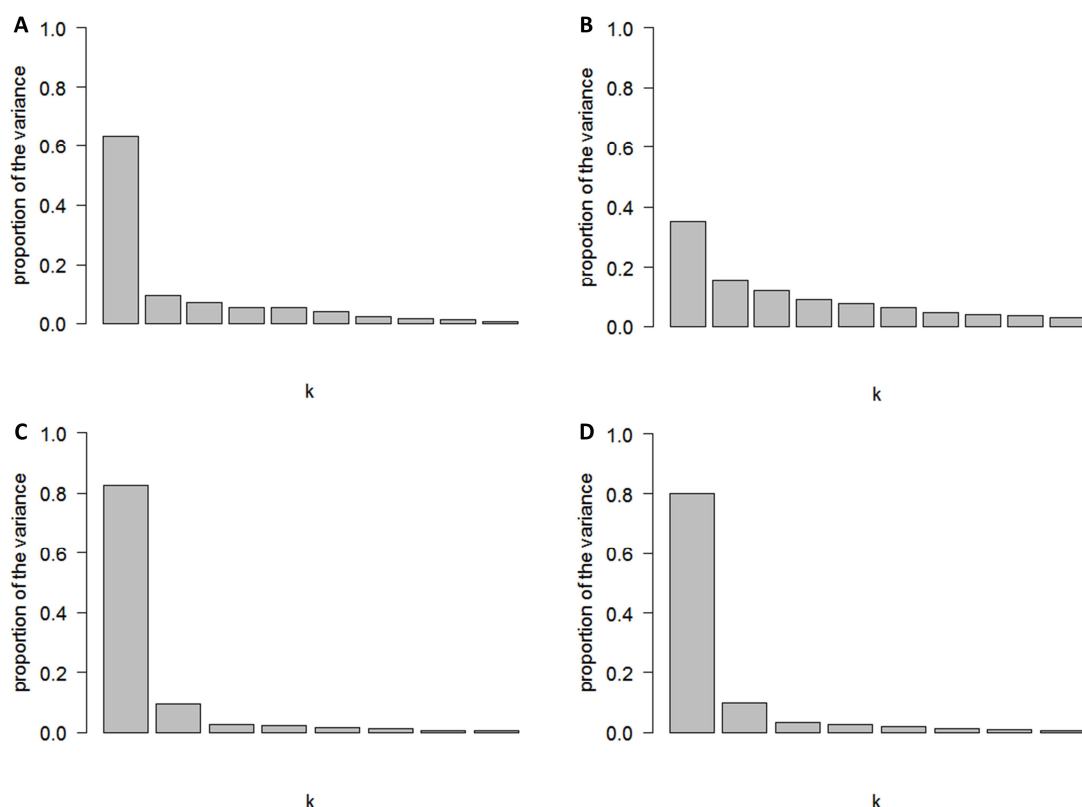


Figure S-11. Explained variance by k factors of variance for RUV-random. A: ESI⁻ dataset, B: ESI⁻ dataset, outliers removed, C: ESI⁺ dataset, D: ESI⁺ dataset, outliers removed.

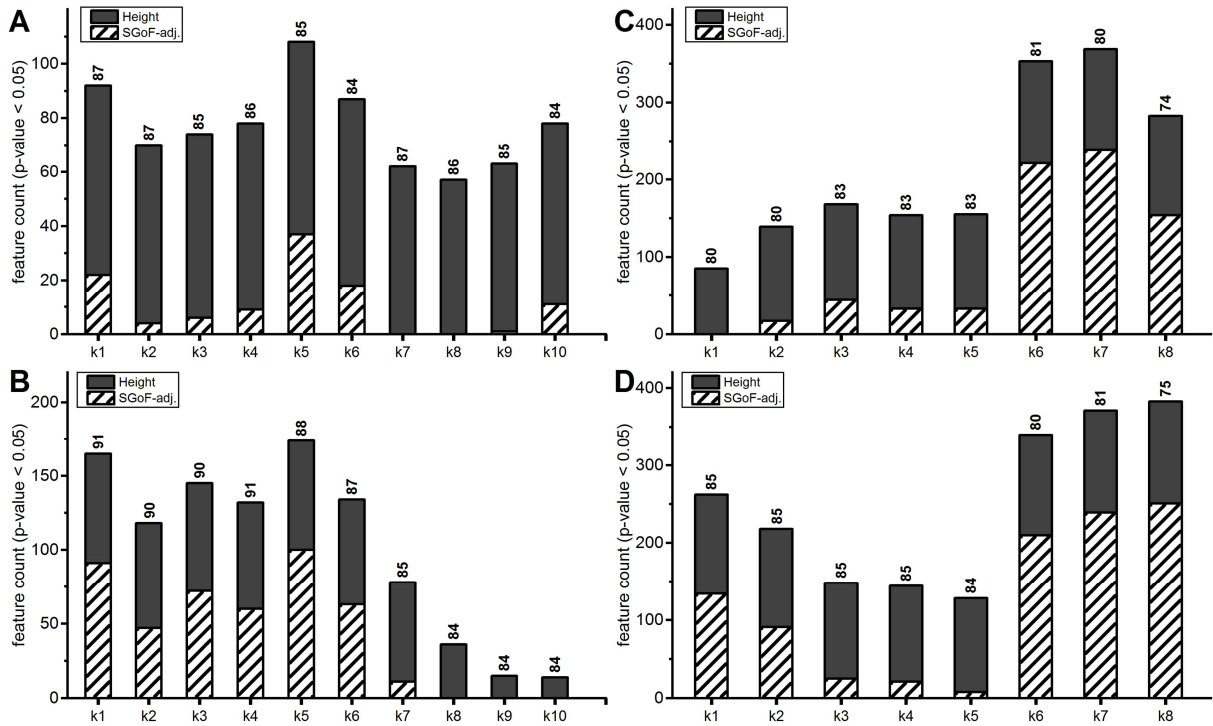


Figure S-12. Influence of factor k on RUV-random. Number of features with a p -value < 0.05 are represented by dark columns. Overlay of the lined pattern columns show the number of SGoF-adjusted p -values < 0.05 . Values above the columns show the respective percentage of features in the QCs that showed a CV $< 30\%$. A: ESI⁻ dataset, B: ESI⁻ dataset, outliers removed, C: ESI⁺ dataset, D: ESI⁺ dataset, outliers removed.

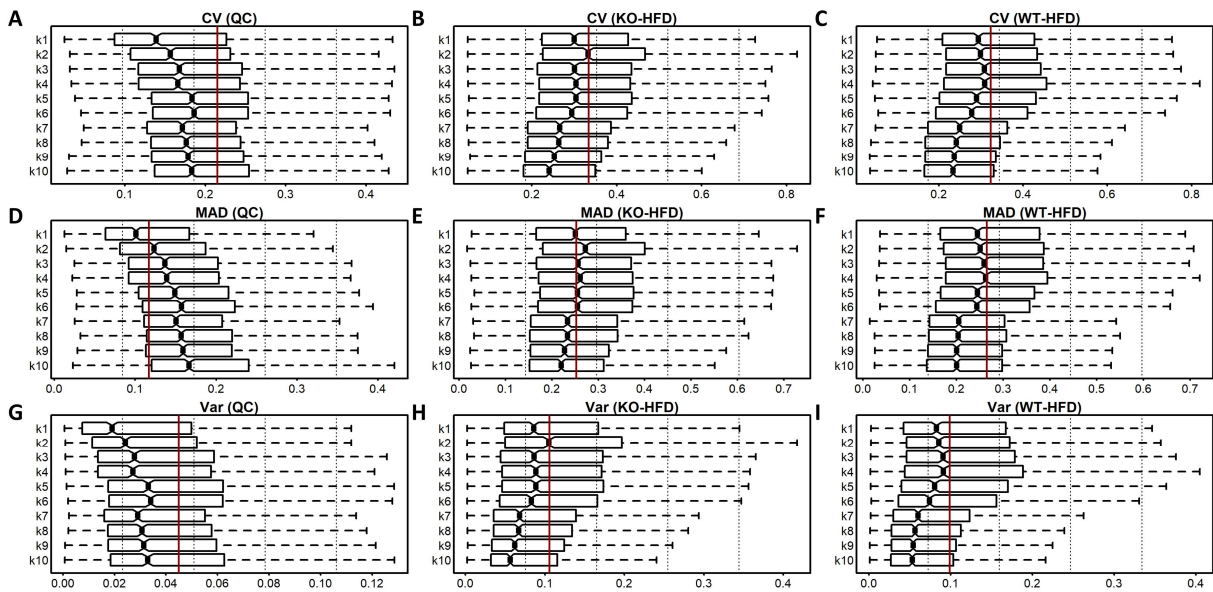


Figure S-13. Comparison of intra-group metrics of variation for RUV-random normalization depending on factor k (ESI⁻ dataset). Box-Whisker plots are shown. The red line represents the median value of the respective metric in raw height data. A: CV in QCs, B: CV in KO-HFD, C: CV in KO-CD, D: MAD for QCs, E: MAD for KO-HFD, F: MAD for KO-CD, G: Var for QCs, H: Var for KO-HFD, I: Var for KO-CD.

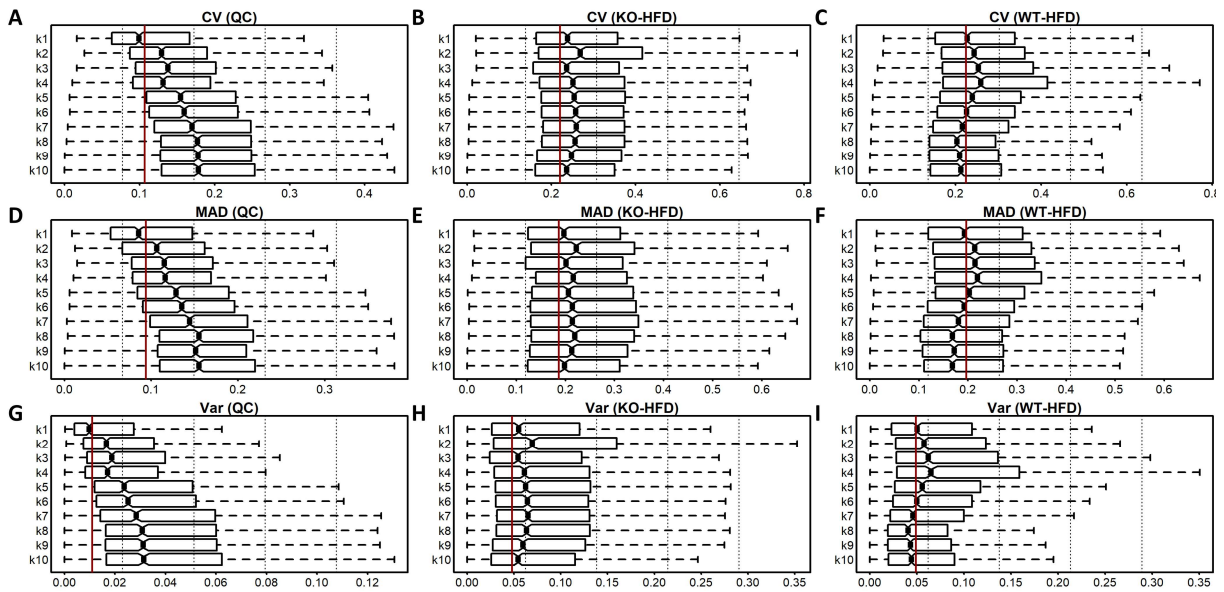


Figure S-14. Comparison of intra-group metrics of variation for RUV-random normalization depending on factor k (ESI⁻ dataset, outliers removed). Box-Whisker plots are shown. Following outliers were removed: B9257, D0181, D3396, B9324, D0167, D3397, QC3, QC7. The red line represents the median value of the respective metric in raw height data. A: CV in QCs, B: CV in KO-HFD, C: CV in WT-HFD, D: MAD for QCs, E: MAD for KO-HFD, F: MAD for WT-HFD, G: Var for QCs, H: Var for KO-HFD, I: Var for WT-HFD.

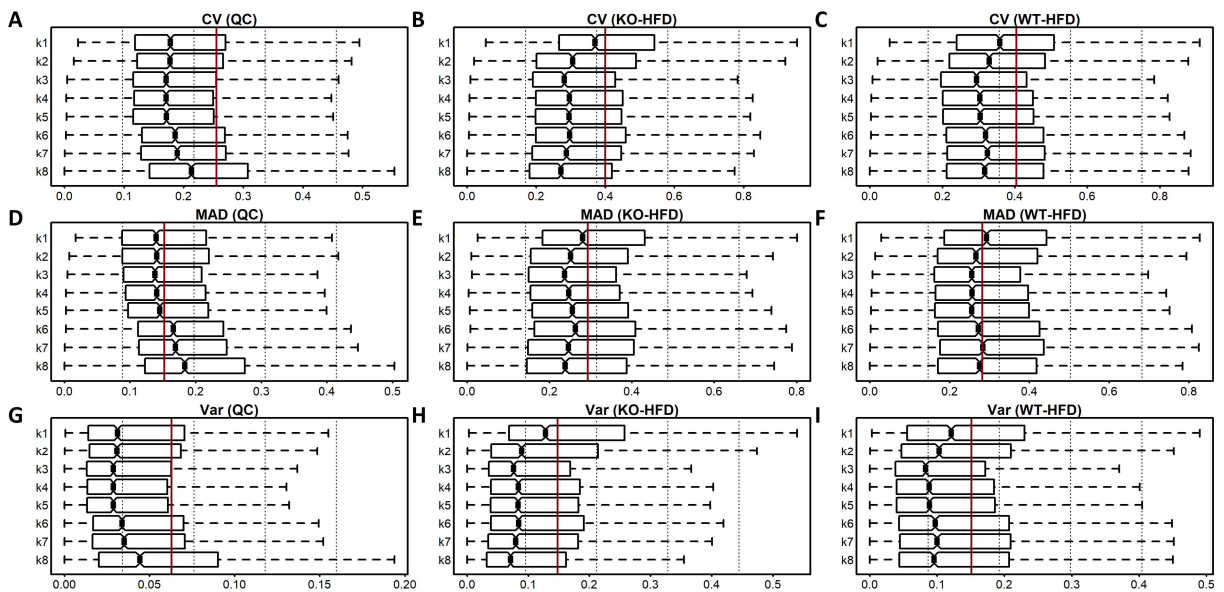


Figure S-15. Comparison of intra-group metrics of variation for RUV-random normalization depending on factor k (ESI⁺ dataset). Box-Whisker plots are shown. The red line represents the median value of the respective metric in raw height data. A: CV in QCs, B: CV in KO-HFD, C: CV in KO-CD, D: MAD for QCs, E: MAD for KO-HFD, F: MAD for KO-CD, G: Var for QCs, H: Var for KO-HFD, I: Var for KO-CD.

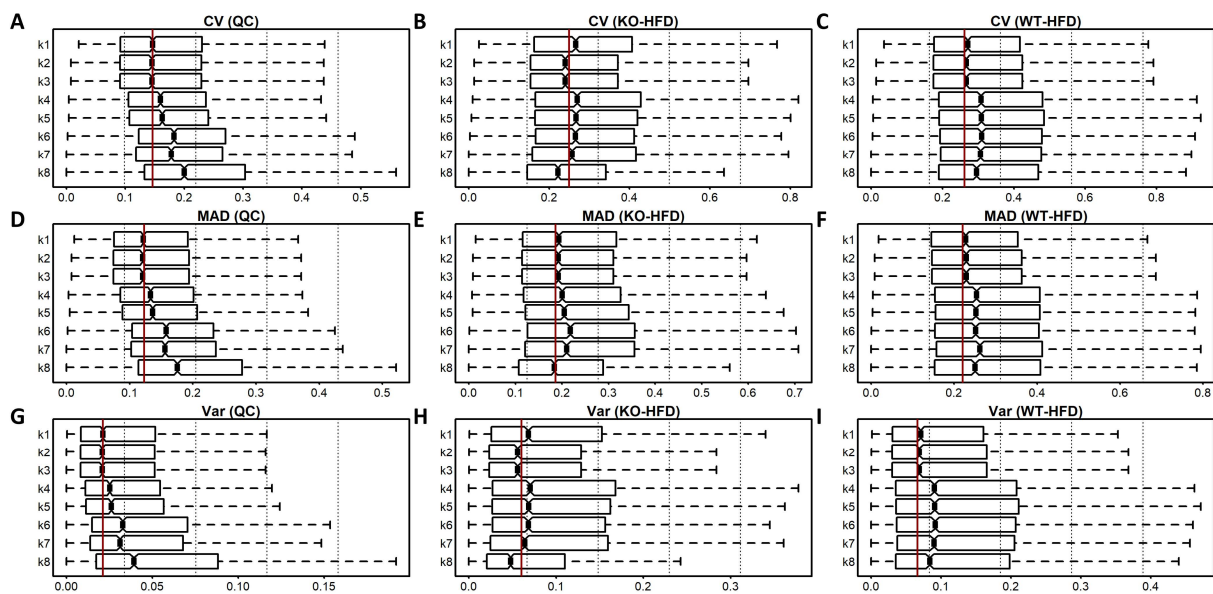


Figure S-16. Comparison of intra-group metrics of variation for RUV-random normalization depending on factor k (ESI⁺ dataset, outliers removed). Box-Whisker plots are shown. Following outliers were removed: D0181, D0182, D3392, B9324, D0167, D3396, QC10, QC11. The red line represents the median value of the respective metric in raw height data. A: CV in QCs, B: CV in KO-HFD, C: CV in WT-HFD, D: MAD for QCs, E: MAD for KO-HFD, F: MAD for WT-HFD, G: Var for QCs, H: Var for KO-HFD, I: Var for WT-HFD.

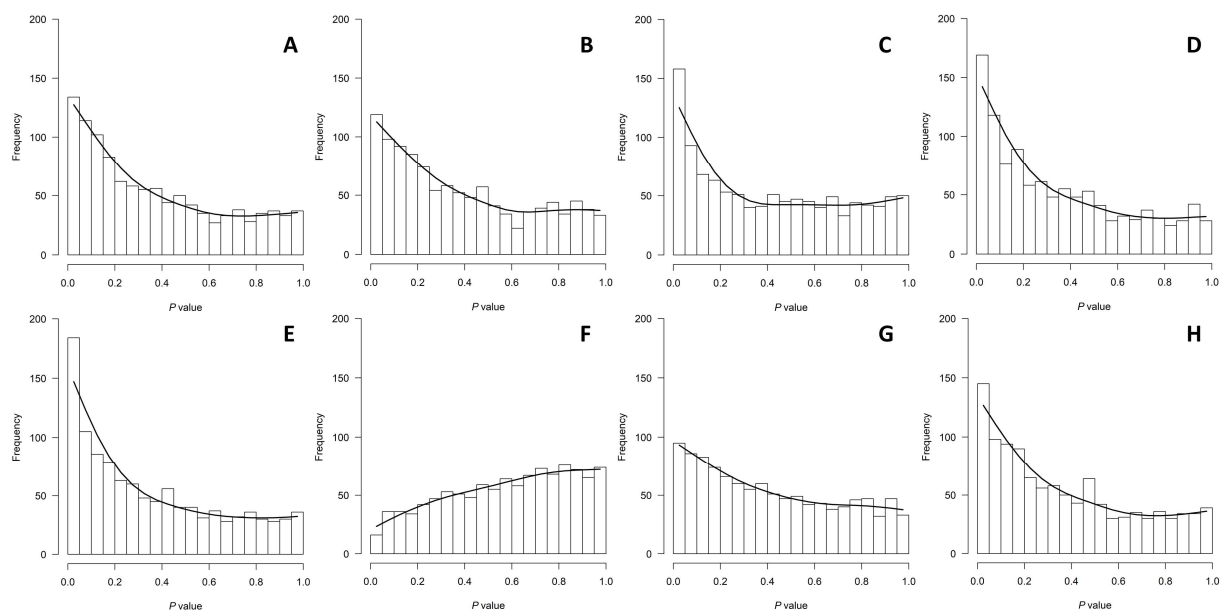


Figure S-17. p-value distribution for different normalization methods (ESI⁻ dataset, outliers removed). Results for comparison KO-HFD versus WT-HFD. Histogram has 20 breaks so that each columns represents a 0.05 interval. A: raw height, B: LOWESS, C: mTIC, D: t_R -IS, E: B-MIS, F: NOMIS, G: CCMN, H: RUV-random ($k = 3$).

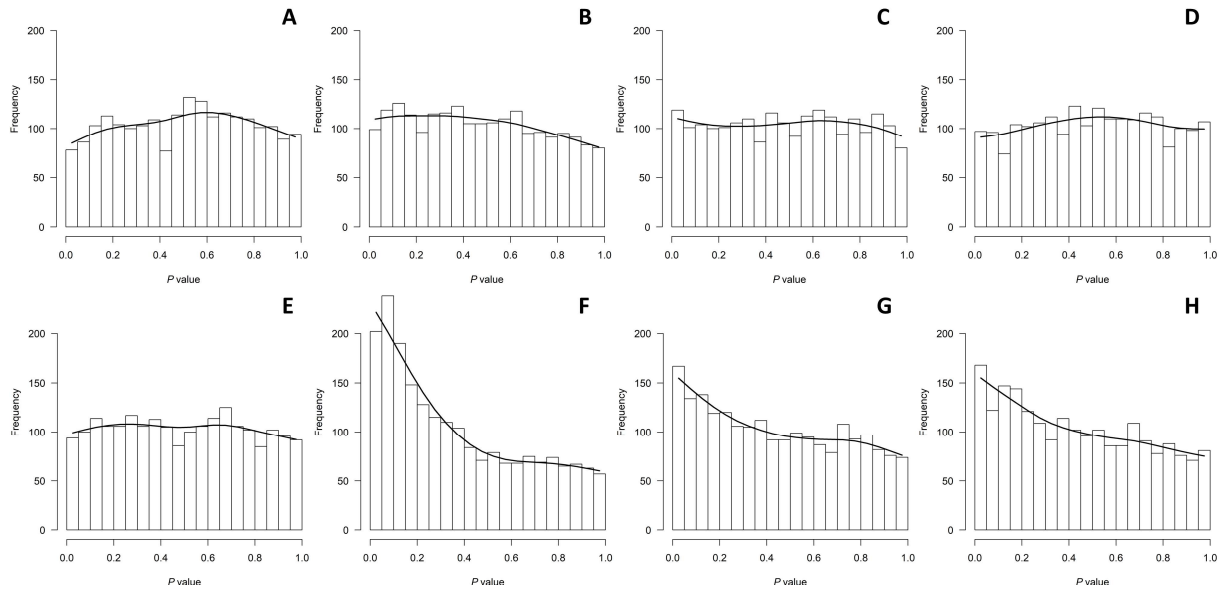


Figure S-18. p-value distribution for different normalization methods (ESI⁺ dataset). Results for comparison KO-HFD versus WT-HFD. Histogram has 20 breaks so that each columns represents a 0.05 interval. A: raw height, B: LOWESS, C: mTIC, D: t_R -IS, E: B-MIS, F: NOMIS, G: CCMN, H: RUV-random ($k = 3$).

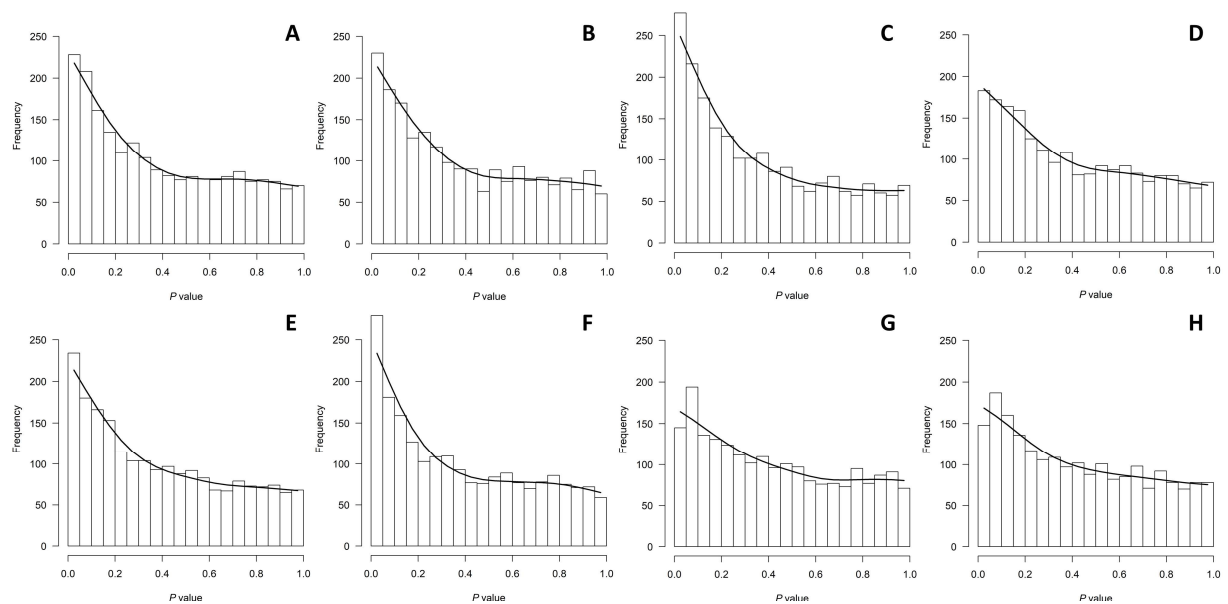


Figure S-19. p-value distribution for different normalization methods (ESI⁺ dataset, outliers removed). Results for comparison KO-HFD versus WT-HFD. Histogram has 20 breaks so that each columns represents a 0.05 interval. A: raw height, B: LOWESS, C: mTIC, D: t_R -IS, E: B-MIS, F: NOMIS, G: CCMN, H: RUV-random ($k = 3$).

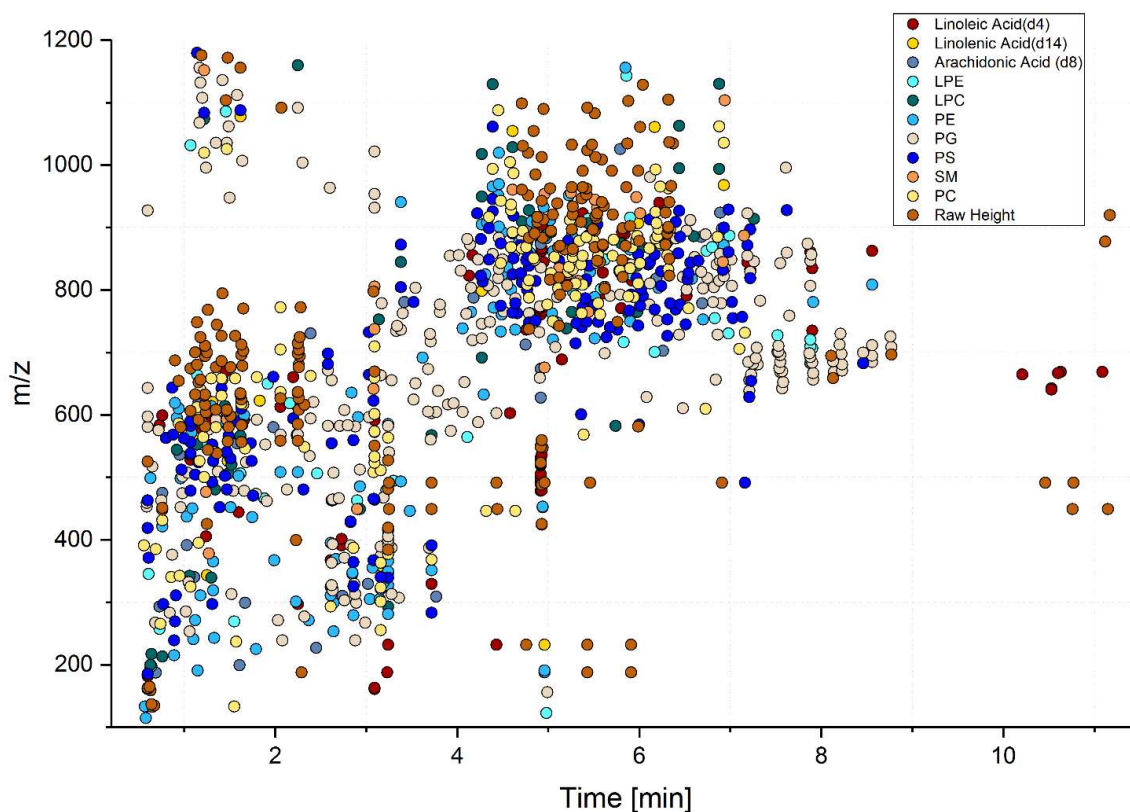


Figure S-20. IS assignment of aligned features using B-MIS (ESI⁻ dataset). Each feature is plotted with the color of the corresponding IS after B-MIS normalization. In the ESI⁻ dataset, 19.7 % of identified lipids, for which a class-specific IS was available, were normalized via their corresponding, class-specific IS. Furthermore, < 1 % of identified lipids was not normalized as no improvement was obtained with internal standardization (overall, raw height data for 18.9 % of all features was maintained). In the ESI⁺ dataset, 25.7 % of all identified lipids with a potential class-specific IS were normalized via this particular IS. Amongst these lipid candidates, 6.3 % remained un-normalized (in total 18.1 % of all features were not normalized)

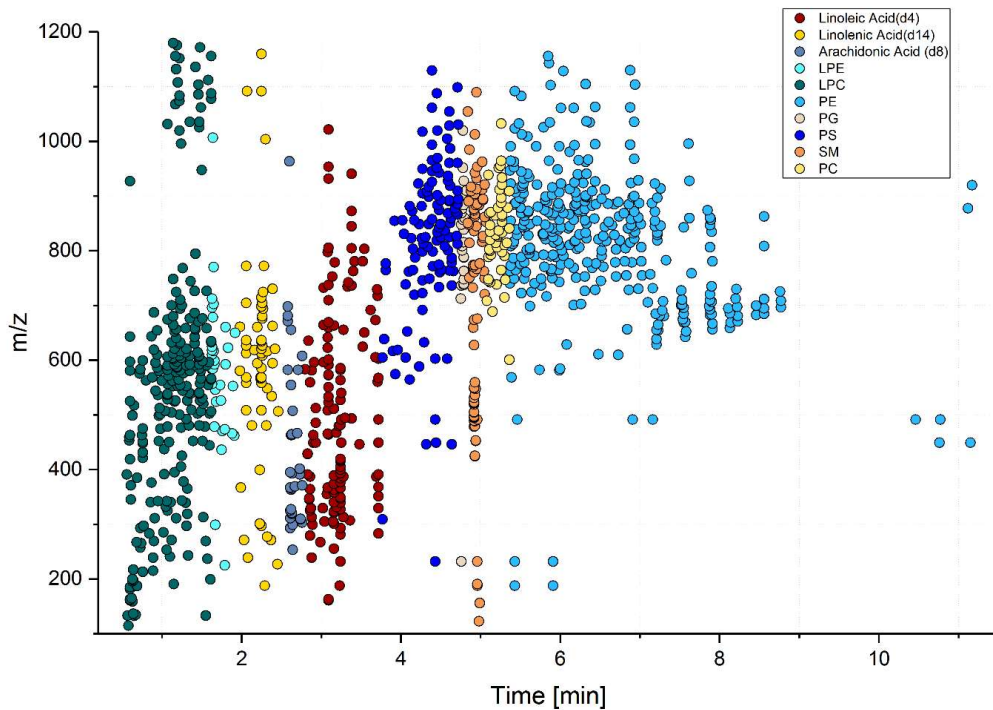


Figure S-21. IS assignment of aligned features using t_R -IS (ESI⁻ dataset). Each feature is plotted with the color of the corresponding IS after t_R -IS normalization.

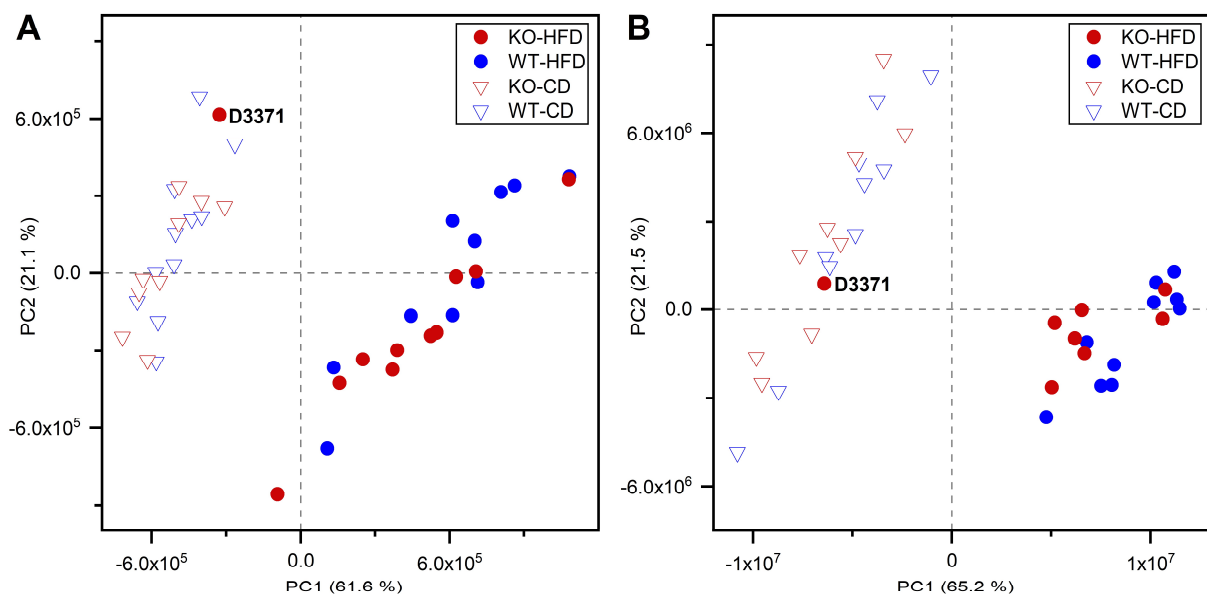


Figure S-22. PCA plots of raw height data (no weighting, mean center scaling). PCA results show that groups can be clearly separated according to the received diet. Sample D3371, classified as KO-HFD, does not group with other HFD samples but rather groups with CD samples. A misclassification error throughout the study is therefore indicated. A: ESI⁻ dataset, B: ESI⁺ dataset (to obtain a better overview, 3 samples (outliers D0181, D0182, D3392; see Figure S-24 and S-25) are not shown as they are outside the chosen scale).

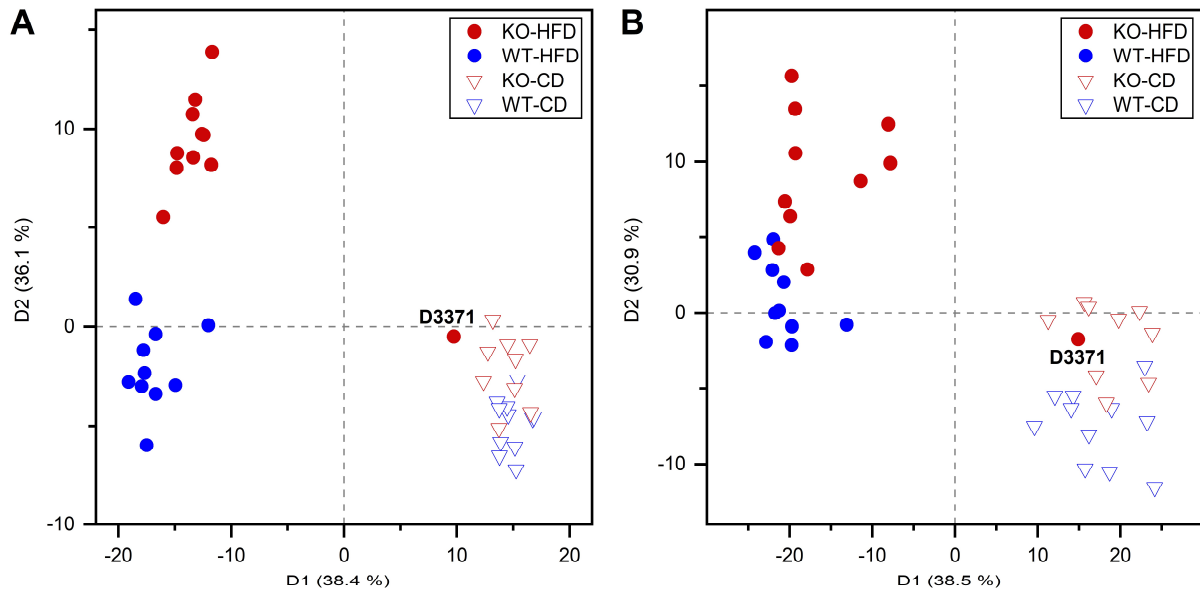


Figure S-23. PCA-DA plots of raw height data (logarithmic weighting, mean center scaling). PCA-DA (principal component analysis and discriminant analysis) results show that groups can be clearly separated according to the received diet. Sample D3371, classified as KO-HFD, does not group with other HFD samples but rather groups with CD samples. A misclassification error throughout the study is therefore indicated. A: ESI⁻ dataset, B: ESI⁺ dataset.

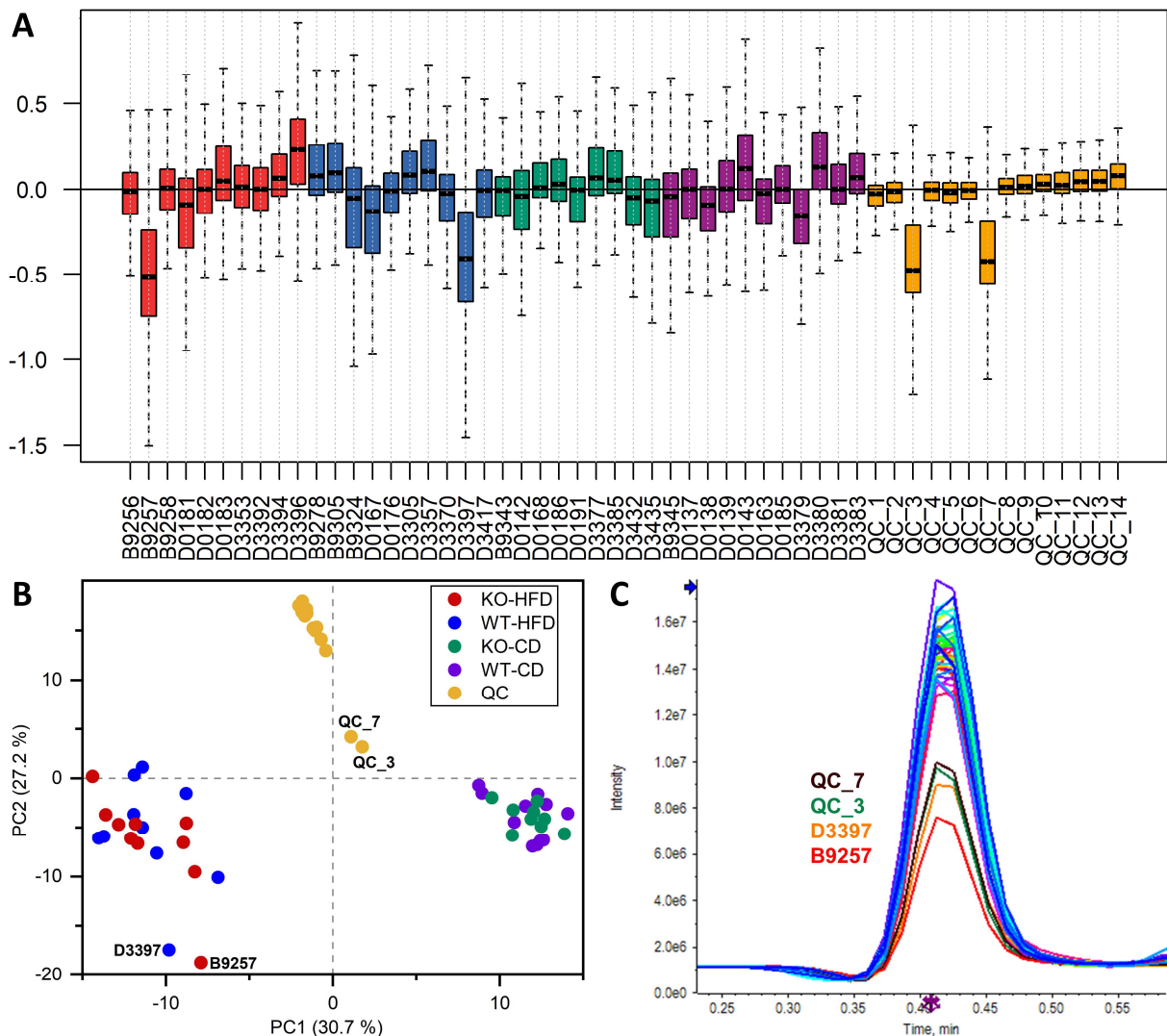


Figure S-24. Outlier evaluation for the ESI⁻ dataset. A: Within-group RLA plot (comprised of Box-Whisker plots), showing intra-group variation of single samples. Samples with a negative shift of the median indicate overall lower abundance of features. Furthermore, intra-group variation can be assessed via the size of the box and the extension of Whiskers. KO-HFD (red), WT-HFD (blue), KO-CD (green), WT-CD (purple), QCs (yellow). B: PCA plot prior to outlier removal (logarithmic weighting, mean center scaling). Samples showing a negative shift in A are also not grouped in PCA. C: Overlay of all TICs (only exemplary peak interval is shown for better overview). Whereas most TIC traces show normal distribution, QC_7, QC_3, D3397 and B9257 have a significantly lower intensity. An injector malfunction is therefore suspected. Besides these samples, D0181, D3396, B9324 and D0167 were additionally removed as outliers as they showed a light hemolytic trend and had above-average intra-group variation (see size of boxes and Whiskers in A).

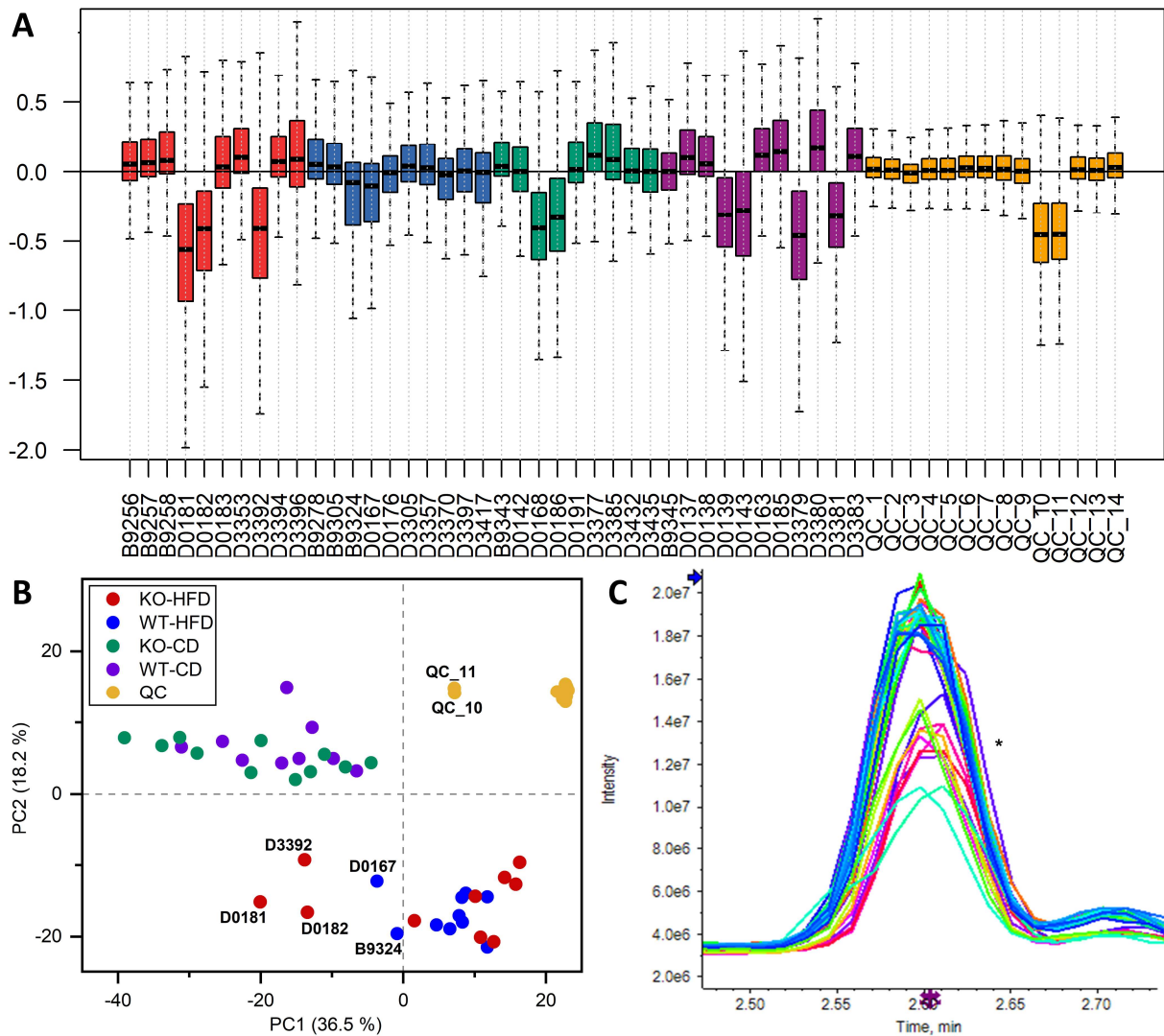


Figure S-25. Outlier evaluation for the ESI⁺ dataset. A: Within-group RLA plot (comprised of Box-Whisker plots), showing intra-group variation of single samples. Samples with a negative shift of the median indicate overall lower abundance of features. Furthermore, intra-group variation can be assessed via the size of the box and the extension of Whiskers. KO-HFD (red), WT-HFD (blue), KO-CD (green), WT-CD (purple), QCs (yellow). B: PCA plot prior to outlier removal (logarithmic weighting, mean center scaling). C: Overlay of all TICs (only exemplary peak interval is shown for better overview). Whereas most TIC traces show normal distribution, D3392, D3381, D3379, D0186, D0182, D0181, D0168, D0139, D0143, QC_10 and QC_11 (marked with a *) have a significantly lower intensity. Further investigation showed, that 96 % of all features in these samples had a foldchange <1 compared to all other samples. An injector malfunction is therefore suspected. Besides these samples, D3396, B9324 and D0167 were additionally removed as outliers as they showed a light hemolytic trend and had above-average intra-group variation (see size of boxes and Whiskers in A).

Text T-2. Power calculation after outlier removal.

Due to the high number of outliers, G*Power 3.1 (Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). *Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. Behavior Research Methods, 41, 1149-1160.*) was used to estimate statistical power with the residual sample sizes (ESI⁻: n = 7 per group after outlier removal). For an exemplary true positive finding showing a minimum foldchange of 0.6 or 1.4 and intra-group CV of 20 % (as obtained from the data), a statistical power of 0.93 can be generated at a significance level of $\alpha = 0.05$ for a two-tailed Student's *t*-test. Regarding the limited availability of mouse plasma samples and the observed effect sizes in true positive findings, statistical power was deemed acceptable for hypothesis generation in untargeted LC-MS lipidomics.

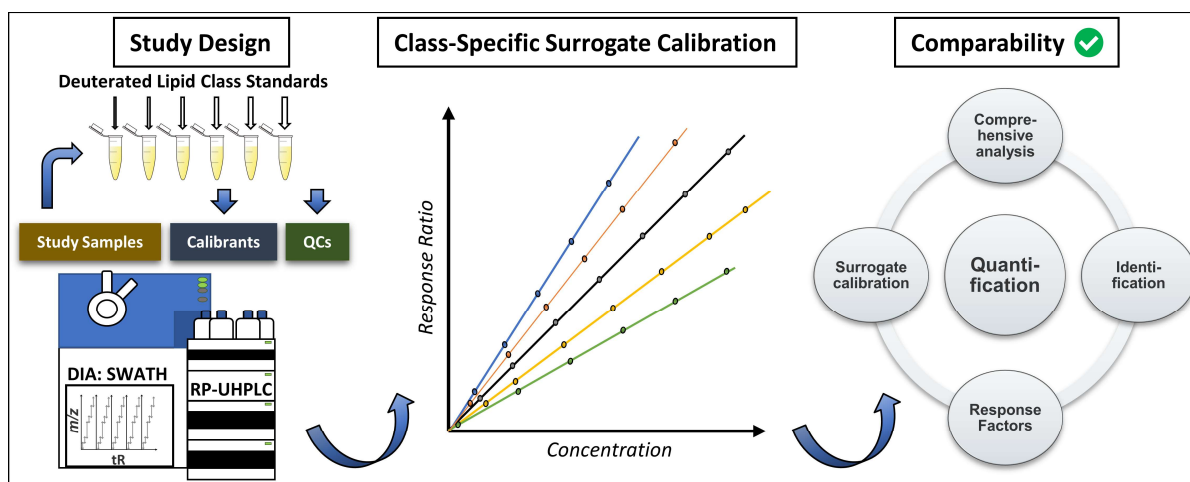
5.3. Comprehensive Lipidomics of Mouse Plasma using Class-Specific Surrogate Calibrants and SWATH Acquisition for Untargeted Quantification.

Bernhard Drotleff[†], Julia Illison[§], Jörg Schlotterbeck[†], Robert Lukowski[§],
Michael Lämmerhofer[†]

[†]Institute of Pharmaceutical Sciences, Pharmaceutical (Bio-)Analysis, University of
Tübingen, Tübingen, Germany

[§]Institute of Pharmaceutical Sciences, Pharmacology, Toxicology and Clinical Pharmacy,
University of Tübingen, Tübingen Germany

Submitted to *Analytica Chimica Acta*, Status: Resubmitted after Revision



Graphical Abstract

5.3.1. Abstract

With continuous advances in mass spectrometry and its ever-increasing applications in numerous scientific fields, also lipidomics has gained rising attention in the recent years. Several strategies for lipidomic profiling are developed, with targeted analysis of selected lipid species, typically utilized for lipid quantification via low-resolution triple quadrupoles, and untargeted analysis by high-resolution instruments, focusing on hypothesis generation for prognostic, diagnostic and/or disease-relevant biomarker discovery. The latter methodologies generally yield relative data with limited inter-assay comparability. In this work we aimed to combine untargeted analysis and absolute quantification to enhance data quality and to obtain independent results for optimum comparability to previous studies or database entries. For the lipidomic analysis of mouse plasma, RP-UHPLC hyphenated to a high-resolution quadrupole TOF mass spectrometer in comprehensive data-independent SWATH acquisition mode was employed. This way, quantifiable data on the MS and the MS/MS level were recorded, which increases assay specificity and quantitative performance. Due to the lack of an appropriate blank matrix for untargeted lipidomics, we herein established a sophisticated strategy for lipid class-specific calibration with stable isotope labeled standards (surrogate calibrants). Acceptable values for accuracy and precision well below $\pm 15\%$ bias were reached for the majority of surrogate calibrants. However, to achieve sufficient accuracy for target lipids, response factors to corresponding surrogate calibrants are required. An approach to estimate response factors via a standard reference material (NIST SRM 1950) was therefore conducted. Furthermore, a useful workflow for post-acquisition re-calibration, involving response factor determination and iteratively built libraries, is suggested. In comparison to single-point calibration, the presented surrogate calibrant method was shown to yield results with improved accuracy that are largely in accordance with standard addition. Quantitative results of real samples were then compared to two previously published dietary mouse plasma studies that provided absolute lipid levels.

5.3.2. Introduction

Due to emerging insights in biological pathways of endogenous lipids and their importance in disease progression,¹ lipidomic profiling and related analytical methods have evolved to become a key field in analytical chemistry. With the steady development of sophisticated mass spectrometric methods, the number of published articles covering lipidomics has been continuously increasing in the recent years.² Many advances in lipidomic biomarker discovery are being made,³⁻⁵ but also methodological progress regarding study design,⁶ databases,^{7, 8} software applications and data processing⁹ is rapidly deployed. Studies aiming for new results in hypothesis generation of potential biomarkers typically utilize high-resolution mass spectrometry for untargeted data acquisition and focus on the observation of relative fold-changes of detected compounds between distinct groups of interest and control groups. These metrics of relative quantification, however, limit inter-study, inter-batch or even inter-sequence comparability and do not allow the comparison of found biomarker levels to reference levels that are covered in databases, such as the Human Metabolome Database.¹⁰ The ultimate approach to overcome these limitations is absolute quantification of compounds of interest. Although accurate determination of absolute levels for hundreds or thousands of features in complex matrices is difficult to accomplish, approaches towards quantitative, untargeted lipidomics must be pursued. In this context, various difficulties concerning calibration, normalization via internal standards (ISs), control of matrix effects and requirements for validation arise.

Most approaches towards quantification in untargeted lipidomics employed shotgun analysis methods in combination with high-resolution instruments.¹¹⁻¹⁵ Its main advantage is the simultaneous ionization of lipids with added ISs for optimum normalization, yet, the enhanced concurring ionization processes can lead to ion suppression, which results in decreased sensitivity and impeded detection of low abundant lipids. Moreover, direct infusion adds an increased risk for compromised assay specificity by in-source fragmentation and higher probability for signal-interferences from isomers and isobars.¹⁶ Other promising results have already been achieved by using SFC-MS¹⁷ or HILIC-MS.¹⁸ Here, lipids are separated according to lipid class polarities and class-specific ISs are co-eluted. Efforts to utilize lipid species separation via RP-LC-MS for quantitative purposes are mainly limited to a reduced number of target compounds¹⁹⁻²² or require global lipid labeling.¹⁵ Furthermore, the majority of these assays is conducted with low resolution triple quadrupole instruments.^{23, 24} Due to the study design of many untargeted methods, which often implement single-point calibration with class-specific ISs,^{25, 26} results are typically semi-quantitative²⁷ as absolute quantification in accordance to quantification guidelines (e.g. FDA guidelines for bioanalytical method validation²⁸) requires multi-level calibration for each analyte. On the other hand, following the

existing guidelines is not always possible for the untargeted analysis of endogenous compounds, and no comparable instructions exist for this purpose. Alternative approaches that comply as far as achievable, thus need to be developed and investigated.

In this work, an approach to achieve absolute quantification of selected lipid classes in an untargeted lipidomic RP-LC-MS assay is presented. By using stable isotope labeled lipids of various classes for matrix-matched surrogate calibration, class-specific quantification of compounds of interest can be executed retrospectively post-analysis. Via SWATH acquisition (sequential window acquisition of all theoretical fragment ion mass spectra),²⁹ comprehensive lipid analysis could be achieved, which enables quantification on TOF-MS or SWATH-MS/MS level (using precursor or product ions). SWATH is a data-independent acquisition technique that offers many advantages like full coverage of MS/MS fragments for enhanced selectivity and higher identification rates than data-dependent acquisition. In addition, it provides high sensitivity that is comparable to that of multiple reaction monitoring (MRM) with triple quadrupole instruments.

Though, an additional determination and application of response factors between target analytes and surrogate calibrants as well as knowledge about linear ranges is required to achieve true absolute quantification.¹⁸ The performance of surrogate calibration was compared to single-point (i.e. 1-point) calibration and standard addition. Furthermore, method-specific response factors for various lipid species were calculated using consensus values obtained from certified reference material for human plasma (NIST SRM 1950).

5.3.3. Experimental Section

Materials

Acetonitrile (MeCN, Ultra LC-MS grade), methanol (MeOH, Ultra LC-MS grade), 2-propanol (IPA, Ultra LC-MS grade) and formic acid (98 %, w/v, ACS grade) were supplied by Carl Roth (Karlsruhe, Germany). Ammonium formate was purchased from Sigma–Aldrich (Saint Louis, MO, USA). SPLASH LipidoMIX (Lipidomix), 14:0-14:0 phosphatidylcholine (PC), 16:0-16:0 PC, 18:0-18:0 PC, 17:1 lysophosphatidylcholine (LPC) and 20:0 LPC were purchased from Avanti Polar Lipids (Alabaster, AL, USA). Arachidonic-acid(d8), α -linolenic-acid(d14) and linoleic-acid(d4) were acquired from Cayman Chemical (Ann Arbor, MI, USA). Oleoylethanolamide was acquired from abcr GmbH (Karlsruhe, Germany). Type I purity water was obtained from a Purelab Ultra purification system (ELGA LabWater, Celle, Germany). Standard reference material (SRM) of human plasma (SRM 1950) for response factor evaluation of lipids was acquired from the National Institute of Standards and Technology (NIST, Gaithersburg, MD, USA). Plasma samples of mice were collected during a previous

study³⁰ with permission of the local authorities and conducted in accordance with the German legislation on the protection of animals.

Sample preparation

Blood was collected by cardiac puncture from experimental mice after 18 weeks of dietary feeding under deep terminal anaesthesia induced by xylazine (10 µg/g body weight (BW)), ketamine (80 µg/g BW) and a 2 - 4 % isoflurane in oxygen inhalation. Upon disposal of the 27-gauge needle, blood was transferred from the syringe into 0.5 mL EDTA coated tubes that were chilled on ice and gently mixed. Samples were centrifuged for 15 min at 3,000 rpm at 4 °C and stored as aliquots of 50 µL plasma per individual at -80 °C upon further processing. After thawing on ice, aliquots of 25 µL were used for sample preparation, respectively, whereas residual volumes were pooled for the preparation of calibration and QC samples (see section below). IPA-based protein precipitation for untargeted lipid extraction^{31, 32} was obtained by addition of 55 µL IPA and 20 µL MeOH to the individual plasma aliquots. This ratio was chosen to achieve uniform solvent composition in all samples (see preparation of calibration and QC samples). After precipitation and subsequent vortexing, the samples were centrifuged for 10 min at 15,000 × g and 4 °C with a 5415R microcentrifuge (Eppendorf, Hamburg, Germany). The supernatant was transferred to a 250 µL conical glass insert in a 1.5 mL glass vial, which was immediately sealed with a crimp cap and stored at 4 °C in the autosampler for the time of analysis. Samples were analyzed as soon as possible after preparation and the analytical sequence was started within 2 h after the final centrifugation.

In order to perform 1-point calibration, Lipidomix and labeled fatty acids (arachidonic-acid(d8), α-linolenic-acid(d14), linoleic-acid(d4)) were spiked into the methanolic portion of the precipitation solvent. Addition of 1-point calibrants also enabled the application of various IS-based normalization techniques for untargeted data processing. Final concentrations of spiked standards in study samples are listed in Table S-1.

Preparation of calibration and QC samples - Quantitative study design

In accordance to regular targeted, quantitative assays, calibration and QC samples were prepared to assess linear dynamic range, precision and accuracy. For matrix-matched calibration, labeled lipid standards (Lipidomix and fatty acids) were spiked into a mouse plasma pool (prepared from aliquots of all study samples) in differing concentrations to serve as class-specific surrogate calibrants.³³⁻³⁵

The Lipidomix contains quantitative amounts of deuterated lipids to relatively reflect the ratios in human plasma. The following lipids are covered: 18:1(d7) cholesteryl ester (CE), 15:0-18:1(d7)diacylglycerol (DAG), 18:1(d7) lysophosphatidylcholine (LPC), 18:1(d7) lysophosphatidylethanolamine (LPE), 18:1(d7) monoacylglycerol (MAG), 15:0-18:1(d7)

phosphatidylcholine (PC), 15:0-18:1(d7) phosphatidylethanolamine (PE), 15:0-18:1(d7)phosphatidylglycerol (PG), 15:0-18:1(d7) phosphatidylinositol (PI), 15:0-18:1(d7) phosphatidylserine (PS), d18:1-18:1(d9) sphingomyelin (SM), 15:0-18:1(d7)-15:0 triacylglycerol (TAG) as well as cholesterol(d7) (see Table S-1). In addition, also 15:0-18:1(d7) phosphatidic acid is present in the Lipidomix, but was not further considered for this study due to its poor peak shape.

As the concentrations of the individual lipids in the Lipidomix could not be altered and since detection sensitivity for each lipid class representative is distinct, regular 6-point-calibration (minimally required by international guidelines) would not have been suitable for universal coverage of the linear ranges of each surrogate calibrant. Therefore, 11 calibration samples (plus an additional true matrix blank, i.e. unspiked mouse plasma pool) were prepared by serial dilution of the spiked methanolic portion of the extraction solvent (see sample preparation). The content of MeOH in the precipitation mix (IPA:MeOH, 2.75:1, v/v) was selected, as the Lipidomix is provided in a methanolic solution by the manufacturer. Consequently, sample preparation was adjusted to the composition of the highest calibration sample (calibration 11: 25 μ L plasma, 55 μ L IPA, 20 μ L Lipidomix and SIL fatty acids in MeOH). To evaluate and control for intra-sequence (i.e. intra-assay) precision and accuracy, 5 quantitative QCs (QC_{quant}), which were spiked to yield concentrations at 1 %, 5 %, 25 %, 50 % and 80 % of the highest calibration sample, were established. For internal standardization, 80 ng mL⁻¹ of 17:1 LPC were spiked to each sample before preparation. Surrogate calibrant concentrations in calibration and QC samples are listed in Table S-2 and Table S-3.

The sequence of analysis was designed to cover calibration and QC_{quant} samples at the beginning, middle and end of the batch. In-between, study samples with embedded system QCs (QC_{syst} ; after each block of 5 samples) were incorporated in a randomized manner. QC_{syst} samples were independently prepared but had identical surrogate calibrant concentrations to QC_{quant} level 3. This way, besides being used for monitoring of instrument stability and normalization (e.g. LOWESS³⁶), QC_{syst} samples were also used to control for stability of quantitative performance throughout the sequence. The principal scheme for the analytical batch can be seen in Figure 1 (for more details see Table S-4).



Figure 1. Measurement scheme of the analytical sequence. QC_{syst} samples were repeatedly analyzed after each 5 real sample measurements.

LC-method

Instrumental analysis was performed based on the method of Tsugawa et. al.⁹ Chromatography was carried out on a 1290 Infinity UHPLC system (Agilent Technologies, Waldbronn, Germany) via an Acquity UPLC CSH C18 column (100 mm × 2.1 mm, 1.7 μm, 130 Å) with a VanGuard Acquity UPLC CSH C18 pre-column (5 mm × 2.1 mm, 1.7 μm, 130 Å) (Waters, Milford, MA, USA). Mobile phase A consisted of 60:40 MeCN:H₂O (v/v) with 0.1 % formic acid (v/v) and 10 mM ammonium formate. Mobile phase B consisted of 90:9:1 IPA:MeCN:H₂O (v/v/v) with 0.1 % formic acid (v/v) and 10 mM ammonium formate. The gradient (0.0 min, 15 % B; 2.0 min, 30 % B; 2.5 min, 48 % B; 11.00 min, 82 % B; 11.50 min, 99 % B; 12.00 min, 99 % B; 12.10 min, 15 % B, 15.00 min, 15 % B) was operated at a flowrate of 0.6 mL min⁻¹ and a constant oven temperature of 65 °C. Injection volume of a connected PAL HTC-xt autosampler (CTC Analytics, Zwingen, Switzerland) was set to 3 μL in positive and 5 μL in negative ionization mode (to increase feature detection due to generally lower sensitivity in negative mode).

MS-method

The chromatographic system was hyphenated to a TripleTOF 5600+ mass spectrometer and operated with the ESI-probe of a DuoSpray source (Sciex, Framingham, MA, USA). Ion source parameters were as follows: curtain gas (N₂) 35 psi; nebulizer gas (N₂) 60 psi; heater gas (N₂) 60 psi, ion source voltage floating +5,500 V (positive mode) and -4,500 V (negative mode), declustering potential: ±80 V, source temperature 350 °C. For fragmentation in untargeted screening, collision energy was set to 35 V with a spread of ±15 V for each SWATH-MS/MS experiment, respectively. An accumulation time of 200 ms was assigned to the TOF-MS experiment for precursor detection in the mass range of *m/z* 50 – 1,250. Every ionization mode-specific method covered 25 SWATH-MS/MS experiments with a respective accumulation time of 20 ms. Resolving power on TOF-MS level was over 30,000 (FWHM @ *m/z* 829.5393) and over 15,000 (FWHM @ *m/z* 397.2122) on SWATH-MS/MS level in high sensitivity mode. Total cycle time summed up to 750 ms, which yielded a minimum of 10 points per peak for an average peak width at base of 8 s. Ionization mode-dependent selection of SWATH window widths was done using swathTUNER.³⁷ The initial input data was generated from a QC_{sys}t sample that was analyzed with a preliminary TOF-MS method, which was operated in an information-dependent acquisition (IDA) mode. SWATH window settings are listed in Table S-5. Samples were first analyzed in positive and subsequently in negative mode. Mass calibration was achieved via infusion of sodium acetate (0.1 mg mL⁻¹ in MeCN:H₂O, 1:1, v/v) every 10th injection. The analytical system was controlled by Analyst 1.7 TF software (Sciex).

Validation

In general, it was pursued to follow existing guidelines for bioanalytical method validation as far as possible. Therefore, selectivity of surrogate calibrants was assessed via the analysis of 6 individual blank mouse plasma samples in positive and negative mode. In contrast to the deuterated precursors or fragment moieties of surrogate calibrants, endogenous target lipids are not always interference-free and do not exclusively yield selective products (e.g. due to overlapping fragmentation in the case of closely eluting precursors that are carrying identical fatty acid moieties). Therefore, special care has to be taken and selectivity must be verified prior to surrogate quantification, i.e. by searching for signs of interference like peak shoulders etc. In the case of insufficient assay specificity, an alternative mode (positive / negative, TOF-MS / SWATH-MS/MS) for quantification has to be chosen, if applicable. When an interference is suspected, the sample can be re-run with an elongated gradient and it can be investigated if the peak is split up into two or more peaks.

Furthermore, ≥ 6 -point calibration and multi-level QCs were established for each surrogate calibrant to evaluate the linear range, precision and accuracy. The acceptance criteria for the inclusion of surrogate calibration samples were adopted from the FDA guidelines, i.e. non-zero calibrants were ± 15 % of the nominal concentrations (except for LLOQ where ± 20 % were accepted) and ≥ 75 % of the included surrogate calibrants met the criteria. For the linear ranges, high similarities between surrogate calibrants and corresponding target analytes must be assumed for correct quantification. If this assumption is in doubt, its validity can be verified post-acquisition via standard addition of the target analyte of interest.

Matrix effects were elaborated by continuous post-column infusion⁴⁴ of surrogate calibrants into the regular analytical flow of blank mouse plasma samples (i.e. devoid of deuterated surrogate calibrants) via a T-piece. This way, matrix effects could be monitored and estimated across the whole retention time interval of the class, instead of for just one peak via post-extraction spiking experiments.⁴³ In contrast to post-column infusion, these experiments would have also required an exceeding volume of the mouse plasma samples. In this approach, potential matrix compounds, that are causing an increase (ion enhancement) or decrease (ion suppression) in the extracted ion chromatograms (EICs) of the surrogate calibrants, are exhibited. To verify the absence of matrix effects, surrogate calibrant EICs should show a constant signal during relevant retention time (t_R) intervals of corresponding lipid classes. Here, it is assumed that lipid species are exposed similar behavior in terms of matrix effects within a lipid class and that deviations in chain length and saturation lead to identical results. This assumption was also investigated by post-column infusion (see Figure S-6).

Data processing

The LC-MS setup enabled the usage of a well-established workflow for lipidomic analysis³⁸ with rapid data processing via MS-DIAL⁹ (version 3.20), which is covering peak finding, alignment, deconvolution, identification (score-based on t_R and MS/MS similarity to the LipidBlast library⁸), and normalization of SWATH data. After conversion of the recorded raw data (.wiff extension) into Analysis Base Files (.abf extension) via the ABF converter (Reifycs, Tokyo, Japan), MS-DIAL projects were created separately for each ionization mode. Processing parameters were adjusted to the following settings: peak finding between 0.3 – 13 min; precursor m/z range from 50 – 1,250; TOF-MS tolerance: m/z 0.01; SWATH-MS/MS tolerance: m/z 0.025; smoothing level: 2; minimum number of points per peak: 5; minimum peak height: 500 cps; t_R tolerance for LipidBlast⁸ based identification: 1.0 min; identification score cut-off: 80 %. Peak alignment was based on the 3rd QC_{sys}t sample with a t_R tolerance of 0.1 min, TOF-MS tolerance of m/z 0.02 and a detection frequency of at least 70 % in one group. Blank subtraction was exerted for signals that had a foldchange <5 in the average samples compared to the average blank signals. The final alignment files covered the following feature counts: positive mode, 2083 features after blank subtraction including 529 identified lipids; negative mode, 1103 features after blank subtraction including 179 identified lipids. For quantitative data processing, PeakView 2.2 (Sciex) and MultiQuant 3.0 (Sciex) were utilized. Here, peak areas were extracted with a \pm 10 mDa mass window in the associated mass spectrometric experiment. Further settings were automated integration by a MQIII algorithm, Gaussian smoothing (width: 2 data points), noise percentage of 90 %, baseline subtraction window of 0.1 min and a peak splitting factor of 2. Moreover, Excel 2019 (Microsoft, Redmond, WA, USA), SPSS Statistics 23 (IBM, Armonk, NY, USA) and Origin 2019 (OriginLab, Northampton, MA, USA) were used for additional data evaluation.

Study samples

Plasma samples were derived from control mice and mice lacking *BK* in various adipocyte populations (both on a C57Bl/6N strain background). Adipocyte-specific controls (genotype: *adiponectin-CreERT2*^{tg/+}; *BK*^{+/*L2*} (CTR group)) and pre-mutant *BK* animals (genotype: *adiponectin-CreERT2*^{tg/+}; *BK*^{L1/*L2*} (KO group)) were generated as previously described.^{30, 39, 40} Dietary feeding protocols were performed with *adipoqBK-CTR* and *adipoqBK-KO* mice that either received a high-fat-diet (HFD) or a control diet (CD) for 18 weeks.³⁹ To avoid sex-dependent effects only male mice were designated to the dietary feeding at an age of 10 weeks. Body weight gain, fat masses and non-fat components of the body, food intake, body core temperature and numerous other parameters of the CD- and HFD-exposed *adipoqBK-CTR* and *adipoqBK-KO* mice were reported by Illison et al.³⁰

5.3.4. Results and Discussion

Method characteristics

The employed RPLC-MS/MS assay is a lipid-species separation method. It allows separation of potential isotopic interferences ($M+2$ isotopologues of lipids with 1 additional double bond⁴¹), of many isomeric lipid species and leads to reasonable spread of the lipids over the chromatogram as to minimize matrix effects. Data-independent acquisition with SWATH was utilized for data generation. It results in comprehensive MS and MS/MS data over the entire chromatogram and across all samples with the benefit that EICs for quantitative analysis can be retrieved post-acquisition from MS and MS/MS data, whichever is more selective or more sensitive. Other selective ion traces can then serve for intra-assay cross validation of assay specificity. Lipidomix standards added to the samples before preparation are usually used as ISs for single point calibration. This standard mix has the advantage that these lipids elute in the middle of the lipid species distribution so that all lipids quantified with this single lipid class specific standard elute relatively close-by. However, herein we test a complementary calibration approach using the Lipidomix calibrant series for matrix-matched surrogate calibration and compare it with single point lipid class specific calibration in terms of assay accuracy.

Selectivity and IS selection

After analysis of 6 individual blank mouse plasma samples, no interfering peaks were detected for surrogate calibrant mass traces in relevant t_R intervals (see Table S-1 and Table S-6).

For optimum internal standardization, addition of a complementary set of labeled lipids (with a mass shift of ≥ 3 Da to surrogate calibrants and unlabeled analytes) would have been advantageous. However, commercial availability of such standards is limited or not given. As an alternative, odd-chain lipid species, which were shown to be of explicitly lower content in human plasma than even-chain lipids,⁴² can be suitable as ISs as long as endogenous concentrations are below detectable levels in study samples. Therefore, the measured blank plasma samples were screened for potentially suitable odd-chain lipid ISs by checking endogenous background signals for mass traces of odd-chain lipid standards. The odd-chain lipid species that were detected in these samples are shown in Table S-11. Eventually, only one odd-chain lipid species that was available at short notice in our lab, 17:1 LPC, was found acceptable and could subsequently be used as a single IS (IS_{quant}) for surrogate calibration.

Linear ranges and intra-assay precision and accuracy

Following the scheme in Figure 1, intra-sequence (i.e. intra-assay) precision and accuracy were determined for independent QC_{quant} ($n = 3$ per level) and QC_{sys} ($n = 11$) samples. Results

were obtained in positive and negative mode for the most sensitive and interference-free adduct of each surrogate calibrant (see Table S-1). A chromatogram of all surrogate calibrants is presented in Figure 2. Due to the comprehensive nature of SWATH acquisition, fully quantifiable EICs on MS/MS level could also be obtained for the evaluation of precision and accuracy, presuming selective fragmentation in associated SWATH windows is present.

Linear ranges, coefficients of determination of calibration functions (R^2) and estimated LODs for the mode of favourable performance are shown in Figure 3 or are presented in more detail for all modes in Table S-8. Results for precision and accuracy are listed in Table 1. They indicate that thresholds of $\pm 15\%$ bias, which are generally accepted for targeted assays, can be reached for the majority of surrogate calibrants. Overall, good precision, with most values being below 15% for the favourable modes of the respective lipids, is achieved. Accuracy is in accepted ranges as well and only a limited number exceeds $\pm 15\%$ bias for certain QC levels in the best-performing modes. For poor-performing calibrants that did not pass the acceptance criteria according to the FDA guidelines, the reason could be the lack of suitable ISs, as one early eluting IS might not be able to sufficiently reflect analytical behavior of all lipid classes throughout the run. With an upgraded IS design, if suitable standards are made available, further improved results can be expected for future studies. Due to the comprehensive design of the MS method, general settings of declustering potential and collision energy were chosen. Enhanced and refined results for lipids or lipid classes of special interest can therefore certainly be reached when working with optimized MS and MS/MS parameters.

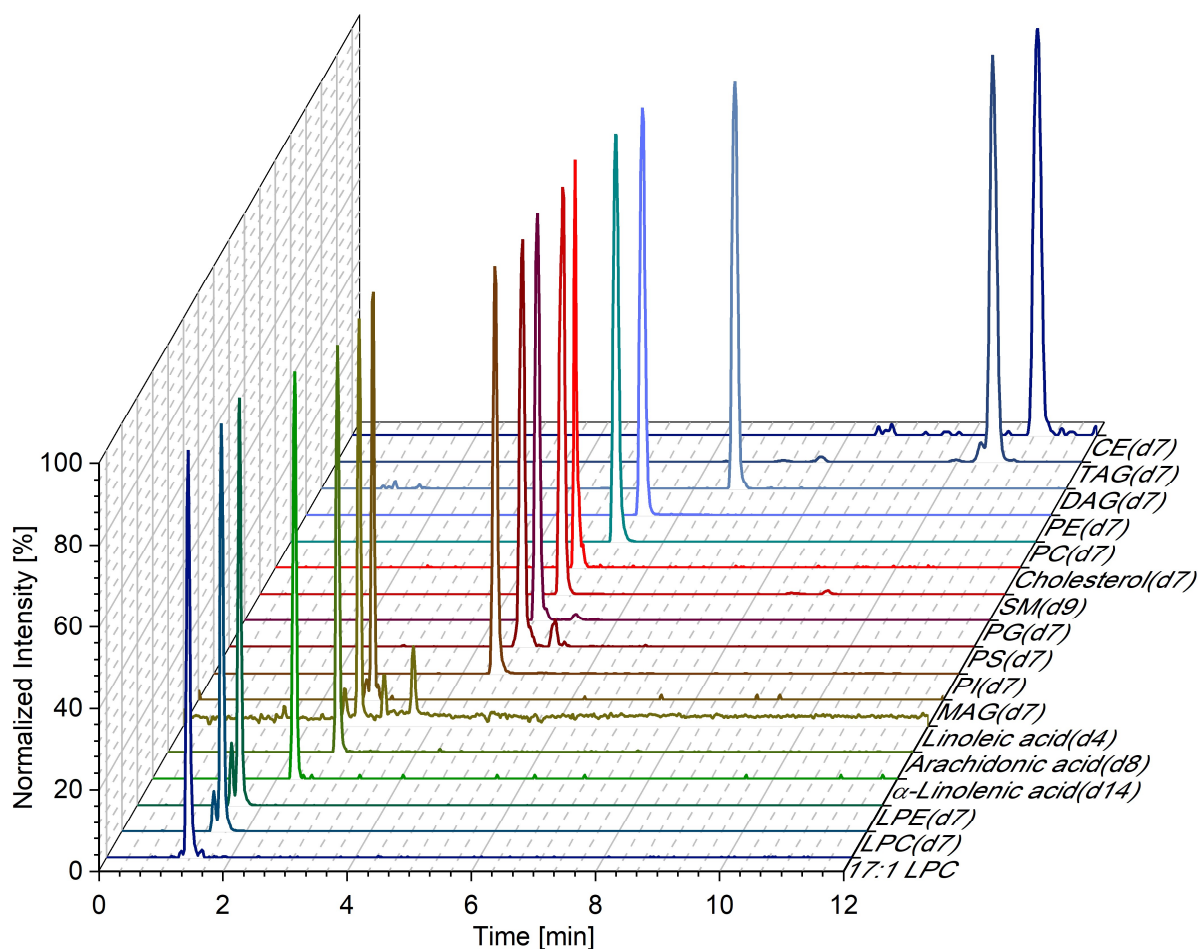


Figure 2. Chromatogram of surrogate calibrants. Results from calibration 11 sample in mode of favourable performance (see Table 1 or the marked mode of favourable performance in Figure 3.) is shown. The specific m/z values of the presented EICs are presented in Table S-6. In addition, the positive TOF-MS signal of 17:1 LPC (IS_{quant}) is drawn.

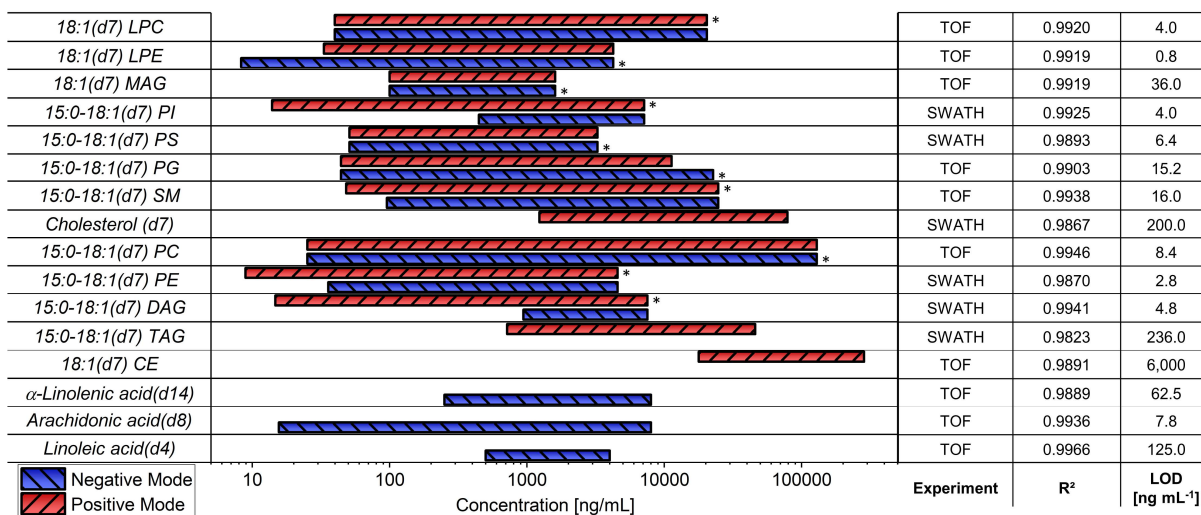


Figure 3. Linear ranges of surrogate calibrants. Polarity that showed favourable performance, concerning linear range, precision and accuracy, is marked with a *. The MS level of best performance, as well R² values derived from calibration results are listed in the table. For more detailed results of all MS modes see Table S-8. LOD results are estimated via results for signal-to-noise ratio of ≈ 3 .

Interestingly, matrix effects again appear to be reduced on SWATH-MS/MS level (Figure S-3) and can only be significantly detected for PC(d7) (Figure S-3G). Regarding the principles of the ionization process, this finding seems paradox since similar extents of matrix effects to TOF-MS experiments are expected. A contributing factor for this finding could be the gated ion transmission control (ITC), which is usually applied in TOF-MS to avoid detector saturation in case of high ion load.⁴⁵ Depending on total ion current and individual peak intensities, ion transmission is modulated cycle-by-cycle and a correction factor is applied to the output signal to account for ITC fluctuations. Whereas ITC is steadily regulated in TOF-MS, ion transmission is permanently set to 100 % for SWATH-MS/MS experiments due to the significantly lower ion load after precursor isolation and the greatly reduced risk for detector saturation. Total ion chromatograms (TICs) and ITC progression from the same experiments for positive and negative mode TOF-MS are shown in Figure S-4. Plots for comparison of TOF-MS and SWATH-MS/MS matrix effects for exemplary surrogate calibrants are displayed in Figure S-5. Ultimately, special care has to be taken for lipid classes affected by matrix effects, in particular when surrogate calibrant and target analyte do not undergo identical degrees of matrix effects. For critical lipid classes, the number of surrogate calibrants and ISs should be maximized. Concerning the already discussed lack of suitable lipid standards, another approach could be the determination of response factors in representative matrix like a QC sample. However, enhanced uncertainty of results is given if high inter-sample variability of matrix effects is observed, since response factors are dependent on the underlying matrix and may change from lot-to-lot.

Table 1. Results for precision and accuracy of surrogate calibrants in mouse plasma (pooled QC).^a

Surrogate Calibrant	Experiment	QC _{quant 1}		QC _{quant 2}		QC _{quant 3}		QC _{quant 4}		QC _{quant 5}	
		Prec. [%]	Acc. [%]	Prec. [%]	Acc. [%]	Prec. [%]	Acc. [%]	Prec. [%]	Acc. [%]	Prec. [%]	Acc. [%]
LPC(d7)	*TOF ⁺	8.7	88.7	2.7	88.5	5.5	93.8	0.6	113.8	1.8	97.5
	SWATH ⁺	30.9	81.9	6.7	90.7	7.4	92.4	1.7	116.2	2.2	101.7
	TOF ⁻	11.1	93.0	4.5	88.8	2.4	92.6	5.4	116.8	7.7	108.6
LPE(d7)	SWATH ⁻	2.6	92.8	9.7	91.0	6.5	86.5	9.1	102.2	0.8	85.2
	TOF ⁺	22.6	92.2	12.9	87.2	6.3	93.8	3.7	113.3	3.5	92.0
	SWATH ⁺	2.1	113.0	9.8	98.3	7.6	96.5	2.6	114.5	4.9	99.3
MAG(d7)	*TOF ⁻	11.9	100.1	1.8	87.0	6.2	93.4	0.8	114.0	0.4	99.1
	SWATH ⁻	-	-	-	-	1.5	86.8	16.3	83.9	8.6	86.4
	TOF ⁺	-	-	-	-	-	-	-	-	-	-
PI(d7)	SWATH ⁺	-	-	-	-	17.1	95.6	13.1	113.4	5.2	109.1
	*TOF ⁻	-	-	-	-	16.8	90.5	6.9	110.8	6.3	92.0
	SWATH ⁻	-	-	-	-	-	-	-	-	-	-
PS(d7)	TOF ⁺	-	-	7.2	89.4	7.4	90.1	4.8	114.4	1.0	98.2
	*SWATH ⁺	18.0	88.0	2.5	85.8	13.6	93.2	1.8	115.2	5.6	101.5
	TOF ⁻	-	-	-	-	-	-	-	-	-	-
PG(d7)	SWATH ⁻	-	-	-	-	9.8	82.3	7.0	108.6	4.8	97.6
	TOF ⁺	-	-	-	-	7.4	97.8	1.0	115.0	2.9	103.7
	SWATH ⁺	-	-	17.3	73.5	14.6	80.9	13.2	96.6	7.2	81.4
SM(d9)	TOF ⁻	10.6	108.4	7.7	86.6	3.7	87.6	4.5	113.8	1.0	98.4
	*SWATH ⁻	8.0	105.2	12.9	90.1	10.3	91.4	8.9	99.0	4.4	90.3
	TOF ⁺	-	-	-	-	-	-	6.0	108.6	4.8	79.0
Cholesterol(d7)	SWATH ⁺	9.3	99.9	4.6	93.7	8.4	101.8	5.1	97.2	-	-
	*TOF ⁻	7.0	85.9	5.9	81.0	8.7	98.0	1.2	114.3	4.9	101.9
	SWATH ⁻	-	-	20.9	89.4	12.4	92.7	10.5	119.4	14.6	101.5
Cholesterol(d7)	TOF ⁺	4.9	87.2	3.3	75.1	11.0	88.3	2.7	106.9	3.8	87.8
	SWATH ⁺	4.7	101.7	8.2	95.0	7.7	90.6	3.9	108.9	2.7	92.4
	TOF ⁻	23.7	71.7	10.8	75.0	8.6	88.9	6.6	121.6	4.5	105.2
Cholesterol(d7)	SWATH ⁻	4.2	96.1	20.6	103.4	7.0	98.7	6.8	128.7	10.8	107.2
	TOF ⁺	-	-	-	-	-	-	-	-	-	-
	*SWATH ⁺	-	-	20.1	84.1	10.6	96.8	7.3	104.0	1.4	103.0
Cholesterol(d7)	TOF ⁻	-	-	-	-	-	-	-	-	-	-
	SWATH ⁻	-	-	-	-	-	-	-	-	-	-

Surrogate Calibrant	Experiment	QC _{quant 1}		QC _{quant 2}		QC _{quant 3}		QC _{quant 4}		QC _{quant 5}	
		Prec. [%]	Acc. [%]	Prec. [%]	Acc. [%]	Prec. [%]	Acc. [%]	Prec. [%]	Acc. [%]	Prec. [%]	Acc. [%]
PC(d7)	TOF ⁺	3.5	84.8	2.3	73.9	7.7	84.0	3.4	120.7	2.7	94.8
	SWATH ⁺	7.9	94.4	4.0	79.7	11.4	92.1	4.4	110.1	3.2	103.1
	^a TOF ⁻	6.6	84.4	3.0	75.0	5.9	87.7	6.1	123.9	9.4	100.6
	SWATH ⁻	15.4	70.7	22.7	85.3	9.0	81.1	10.7	100.8	-	-
PE(d7)	TOF ⁺	22.0	88.1	4.9	79.7	4.7	89.6	0.9	115.6	1.7	98.1
	^a SWATH ⁺	0.7	100.6	1.7	89.0	9.6	96.5	6.8	114.4	8.8	95.4
	TOF ⁻	12.6	94.7	3.3	75.7	4.3	89.2	4.8	117.2	2.2	101.2
	SWATH ⁻	-	-	50.0	83.0	6.6	83.6	8.0	115.0	5.9	99.1
DAG(d7)	TOF ⁺	7.3	111.6	2.7	74.4	7.4	85.4	4.5	119.4	3.8	103.3
	^a SWATH ⁺	11.8	84.8	10.4	73.8	10.6	87.1	7.3	126.7	6.0	112.2
	TOF ⁻	-	-	-	-	22.8	74.0	21.3	70.8	24.1	95.8
	SWATH ⁻	-	-	-	-	-	-	-	-	-	-
TAG(d7)	TOF ⁺	-	-	-	-	-	-	-	-	-	-
	^a SWATH ⁺	-	-	15.6	84.5	7.8	86.7	1.8	122.0	9.9	93.6
	TOF ⁻	-	-	-	-	-	-	-	-	-	-
	SWATH ⁻	-	-	-	-	-	-	-	-	-	-
CE(d7)	^a TOF ⁺	-	-	-	-	25.8	114.1	17.9	150.8	35.7	110.2
	SWATH ⁺	-	-	-	-	-	-	-	-	-	-
	TOF ⁻	-	-	-	-	-	-	-	-	-	-
	SWATH ⁻	-	-	-	-	-	-	-	-	-	-
Arachidonic acid(d8)	TOF ⁺	-	-	-	-	-	-	-	-	-	-
	SWATH ⁺	-	-	-	-	-	-	-	-	-	-
	^a TOF ⁻	3.3	58.6	3.8	56.7	5.0	60.9	6.7	132.6	4.8	112.8
	SWATH ⁻	-	-	28.9	102.6	17.9	57.4	13.4	89.8	6.9	93.3
α -Linolenic acid(d14)	TOF ⁺	-	-	-	-	-	-	-	-	-	-
	SWATH ⁺	-	-	-	-	-	-	-	-	-	-
	^a TOF ⁻	-	-	-	-	9.4	57.6	4.8	122.9	3.2	102.5
	SWATH ⁻	-	-	-	-	-	-	-	-	-	-
Linoleic acid(d4)	TOF ⁺	-	-	-	-	-	-	-	-	-	-
	SWATH ⁺	-	-	-	-	-	-	-	-	-	-
	^a TOF ⁻	-	-	-	-	6.0	54.1	1.5	102.4	2.4	85.9
	SWATH ⁻	-	-	-	-	-	-	-	-	-	-

^aResults for QC_{quant} samples were obtained from 3 technical replicates except for QC_{quant 3}. Here results were combined from 3 technical replicates of a QC_{quant 3} sample and 11 technical replicates of a QC_{sys} sample. Concentrations for calibrants and QCs are listed in Table S-2 and Table S-3. The mode that shows favourable performance in terms of linear range, precision and accuracy is marked with *.

Response factors of target analytes to surrogate calibrants

Most studies dealing with class-specific quantification of lipids have been shotgun approaches,¹¹⁻¹³ which require de-isotoping and high resolution instruments with resolving power >100,000 to rule out signal interferences of co-eluting compounds. In addition, t_R as orthogonal information for compound identification is not available and partial misannotation due to in-source fragmentation (e.g. LPC can decompose to FA or LPE) has been shown.^{46, 47} For shotgun assays, instrument responses of lipid species within polar lipid classes were reported to be mostly identical due to the main effect of the polar head group on the ionization efficiency.¹⁹ For the other lipid classes, responses between lipid species are dependent on chain length and the degree of saturation.¹² Other factors affecting detector response were total lipid concentration, instrument settings and solvent composition. Furthermore, studies utilizing surrogate calibrants found differing ionization efficiencies between stable isotope labeled compounds and unlabeled target analytes.^{33, 34} Accordingly, the determination of response factors between surrogate calibrants and corresponding target analytes is essential to ensure accurate quantification. To take matrix effects (and recovery) into account, response factors are ideally determined in representative sample matrix and after the analytes have undergone the identical sample preparation. Owing to limited sample volume, evaluation of response factors in neat solution may be an acceptable compromise in regard to general global profiling methods aiming primarily at hypothesis generation.

For proof-of-principle, several lipid standards (14:0-14:0 PC, 16: 16:0 PC, 18:0-18:0 PC, 20:0 LPC) were acquired. Standards, surrogate calibrants and 17:1 LPC (IS_{quant}) were spiked into MeOH, followed by serial dilution (with IS_{quant} spiked MeOH) to 5 different concentration levels. These samples were analyzed with the identical method that was used for previous study measurements. Response factors were calculated via the slopes ($slope_{lipid\ standard} / slope_{surrogate\ calibrant}$) and results are shown in Table 2. For shotgun approaches, linear relationships between acyl chain length and response factors could be observed for PCs.¹² Here, neither using mass concentrations nor molarities, an apparent relationship between chain length and response factor could be constructed. This implies that for gradient elution in RP chromatography, which results in differing retention times for lipid species, estimation of response factors is aggravated due to the changing solvent composition. A simple extrapolation factor that is depending on structural characteristics can therefore not be readily determined. The same applies to SWATH-MS/MS data, in which the additional factor of fragmentation efficiency is further complicating the matter.

Assessment of response factors for numerous lipid species is labor intensive and associated to high costs for standard compounds. Yet, method-specific response factor libraries could be iteratively established, as the stability of response factors has been previously demonstrated for other mass spectrometric instruments.¹⁸ A productive workflow for future studies could be

to acquire standards for compounds of interest after final data processing and to subsequently determine their response factors to their corresponding surrogate calibrants. This way, significantly regulated lipids and/or potential biomarkers can be quantified post-acquisition for universal comparability. This workflow (i.e. post-acquisition re-calibration) can be of special interest when sample volume of individual samples is limited, and standard addition is hence not applicable.

Table 2. Response factors of lipid standards to surrogate calibrants in neat solution.

		14:0-14:0 PC	16:0-16:0 PC	18:0-18:0 PC	20:0 LPC
TOF-MS ⁺	mass	1.013	1.022	0.665	0.966
	molarity	0.912	0.996	0.698	1.008
SWATH-MS/MS ⁺	mass	2.383	0.990	0.724	0.849
	molarity	2.145	0.965	0.760	0.886
TOF-MS ⁻	mass	1.127	1.093	0.612	1.058
	molarity	1.015	1.065	0.642	1.104
SWATH-MS/MS ⁻	mass	1.934	1.789	0.848	0.922
	molarity	1.741	1.744	0.890	0.962
t_R [min]		4.32	5.61	7.02	2.88

Response factors were calculated on the basis of mass concentration and molarity. Results are based on peak area as this parameter was also used for quantification. However, also peak height did not reveal any linear relationship between carbon chain length of lipid species and response factors. 17:1 LPC was used as IS for all compounds

Comparison of quantification between standard addition, surrogate calibration and 1-point calibration

In general, standard addition has several drawbacks that limit its routine implementation into untargeted studies: (i) it is not suitable if the concentration in the sample is close to the upper limit of the linear range, since additional spiking will result in a nonlinear increase of the signal; (ii) it requires a significant amount of additional laboratory work and analysis time for each sample; (iii) it is not applicable if the sample volume is limited; (iv) the samples are exposed to ageing until analytes of interest are evaluated and standard addition can be prepared. However, since it is accepted as a valid approach for quantification, it was used as a reference method for the comparison of surrogate calibration and 1-point calibration.

Given the linear calibration functions of the respective surrogate calibrant (see Table S-8) and the response factor of the target lipid species (Table 2), absolute quantification via surrogate calibration can be executed. For cross-validation purposes, a 5-level standard addition of 14:0-14:0 PC, 16:0-16:0 PC, 18:0-18:0 PC and 20:0 LPC into pooled QC samples was conducted (see Table S-7). The results for standard addition were considered as the most accurate, thus,

they were compared to results obtained from surrogate calibration (established via post-acquisition re-calibration) and 1-point calibration (Table S-1 and Table S-3) of the QC pool. Precision for standard addition was calculated via the error of the y-intercept. Accuracy was determined via the unspiked sample, after adjusting for the endogenous lipid concentration calculated by extrapolation of the standard addition curve. For surrogate calibration and 1-point calibration, precision was determined via the 14 replicates of QC_{sys} + QC_{quant} 3 samples. Response factors are generally mandatory for 1-point calibration, too. Accordingly they were also applied to this method. A comparison of the three quantification methods is given in Table 3.

Except for 14:0-14:0 PC, acceptable agreement between standard addition and surrogate calibration was obtained in positive TOF-MS mode by this post-acquisition re-calibration. Relatively high deviations in 14:0-14:0 PC quantification could be due to the neglected matrix effects (Figure 4H and Figure S-2I) (note that the response factor was determined in MeOH), or the relatively low concentration close to the lower end of the linear range of the surrogate calibrant 15:0-18:1(d7) PC (LLOQ: 25.11 ng mL⁻¹). Excellent results were achieved for 20:0 LPC throughout all modes for surrogate calibration. This finding could further imply the need for improved internal standardization as 20:0 LPC was quantified with a class-specific IS (IS_{quant} 17:1 LPC). In most cases, 1-point calibration led to overestimation of target lipid concentrations and to systematically higher concentration values than for surrogate calibration. Major drawbacks, which are resulting in increased uncertainty for 1-point calibration, are: (i) the defective calibration function that is automatically forced through the origin; (ii) the inability to apply weighted regression; (ii) the requirement for the calibrant concentration to be in the linear range, which requires preliminary experiments. In contrast to the partially unacceptable results obtained from 1-point calibration (despite considering response factors), surrogate calibration involving post-acquisition re-calibration yielded absolute concentration values comparable to standard addition results. By assessment of matrix-matched response factors and proper internal standardization, performance of surrogate calibration is likely to be further improved.

Table 3. Results of standard addition, surrogate calibration and 1-point calibration of mouse plasma samples (pooled QC).

Method	14:0-14:0 PC			16:0-16:0 PC			18:0-18:0 PC			20:0 LPC		
	Conc. [ng mL ⁻¹]	Acc. [%]	Prec. [%]	Conc. [ng mL ⁻¹]	Acc. [%]	Prec. [%]	Conc. [ng mL ⁻¹]	Acc. [%]	Prec. [%]	Conc. [ng mL ⁻¹]	Acc. [%]	Prec. [%]
TOF-MS - positive mode												
Standard addition	87.6	102.9	2.6	13,288	102.3	3.0	789.1	99.0	3.5	1,298.4	103.5	0.8
Surrogate calibration	125.2	142.9	6.0	14,701	110.6	8.0	813.6	103.1	6.7	1,468.9	113.1	4.8
1-point calibration	148.7	169.7	9.9	17,534	132.0	2.2	973.1	123.3	8.5	1,571.1	121.0	7.7
SWATH-MS/MS - positive mode												
Standard addition	-	-	-	15,910	90.9	3.0	-	-	-	1,543.3	98.0	2.1
Surrogate calibration	-	-	-	12,689	79.8	6.1	-	-	-	1,720.1	111.5	9.9
1-point calibration	-	-	-	13,644	85.8	12.9	-	-	-	1,864.2	120.8	12.4
TOF-MS - negative mode												
Standard addition	96.8	100.3	5.2	14,420	91.1	0.4	759.9	102.0	5.4	1,404.9	100.5	0.9
Surrogate calibration	140.5	145.1	9.3	12,398	86.0	5.8	595.5	78.4	2.3	1,417.6	100.9	3.5
1-point calibration	161.6	166.9	11.8	14,013	97.2	9.9	624.4	82.2	7.7	1,511.3	107.6	4.1
SWATH-MS/MS - negative mode												
Standard addition	-	-	-	-	-	-	677.6	98.1	3.2	1,324.9	95.2	2.6
Surrogate calibration	-	-	-	-	-	-	717.1	105.8	11.2	1,472.4	111.1	6.4
1-point calibration	-	-	-	-	-	-	1,180.5	174.2	13.5	1,680.0	126.8	8.7

Cells marked with * indicate a peak interference in the respective mode, leading to inaccurate integration and quantification.

Estimation of (matrix-matched) response factors using human plasma NIST SRM 1950

A brief description of the NIST SRM 1950 standard reference material and further information about associated consensus values can be found in the Supplementary Material, Text S1. For the anticipated response factor evaluation, surrogate calibrants and 17:1 LPC (IS_{quant}) were spiked to NIST SRM 1950 plasma. The samples were subsequently prepared with the identical protocol that was also used for the experimental mouse plasma samples. The supernatant was then diluted with IS_{quant}-spiked precipitation solvent to yield a 5-point dilution series. In order to avoid bias by underestimation of matrix effects, only moderate dilutions of the supernatant were prepared (1:1.33, 1:2, 1:4, 1:8; v:v). Equivalent to the above described response factor evaluation procedure for standard reference analytes (i.e. post-acquisition re-calibration approach), slopes of SRM 1950 lipids were divided by the slope of the corresponding surrogate calibrant. Only lipids with a COD <30 % (see Text S-1), that could be detected free of interference, were considered for response factor evaluation. As lipid concentrations were reported for total sum compositions (e.g. PC 36:0), it had to be assumed that different isomer lipid species (e.g. PC 18:0-18:0 and PC 16:0-20:0) contribute equally to the resulting response factor. If individual species should be quantified, a specific MS/MS signal has to be used for calibration, response factor determination and re-calibration. Results are listed in Table S-9 and the distribution of response factors in positive and negative mode are visualized in Figure S-7.

To achieve reliable results, peak integration has to be consistent throughout the respective lipid class. As some lipid signals are close to detector saturation and others are close to LOQ, further care has to be taken to verify working in correct linear ranges. Although high similarities in the matrix compositions of mouse and human plasma can be expected, potential response factor inaccuracies should be kept in mind due to the usage of plasma from distinct species. One major drawback of this approach is the partially high uncertainty and variance of inter-laboratory results, and the observed discrepancies to values determined by the LIPID MAPS consortium (Table S-9). Nevertheless, reasonable results, especially for LPCs, for which response factors are closely distributed around 1, could be obtained. As LPCs are one of the most polar lipid classes, instrument response is less affected by differences in chain length or saturation. Also matrix effects were shown to be of minor extent or at least uniform throughout the elution interval for LPCs (see Figure 4 and Figure S-2). If only a weak influence of the mobile phase composition is assumed, a class-wide, matrix-matched response factor that closely resembles the response factor determined in neat solution (Table 2) can be expected.

Inter-study comparison.

The comparison of lipidomic profiles of rodents that received a high fat diet versus rodents that received a standard (control) diet has been described in the literature. For two dietary mouse studies,^{48, 49} absolute concentration values of various lipid species were reported. These results were compared to the quantitative results in this study, which were obtained via surrogate calibration. Due to differences like genetic background, age, diet composition, duration of feeding and other factors, which can highly influence the plasma lipidome (see Table S-10), unrestricted inter-study comparability is not given and perfect similarity of lipid levels cannot be expected. Nevertheless, quantitative results are suitable to investigate the impact of these differences on lipid plasma levels and to identify unaffected lipid species. Comprehensive results for inter-study comparison are listed in Table S-10.

In conclusion, levels of 11 lipid species showed less than 30 % deviation when compared to results of CD-fed mice from Eisinger et al.⁴⁹ (for HFD-fed mice 5 lipids showed less than 30 % deviation). Compared to results from Barber et al.,⁴⁸ 4 lipids in CD-fed mice and only 3 lipids in HFD-fed mice had lower than 30 % deviation. Furthermore, when comparing the two previously published studies, also only 5 lipids in CD-fed mice and 6 lipids in HFD-fed mice showed a plasma level deviation below 30 %. Overall, the majority of lipids were elevated, compared to the previously published studies (72.2 % of all lipids for Barber et al.⁴⁸, 83.8% of all lipids for Eisinger et al.⁴⁹), which might be related to the time point of sampling or the duration of the feeding (Table S-10). It should also be noted that 32.4 % of the obtained results from surrogate calibration exceeded the linear range and are likely to be overestimated. However, the study design enables the choice of a less sensitive polarity (e.g. ESI⁻ instead of ESI⁺) or MS/MS level quantification with a less sensitive fragment ion, which can yield an enhanced linear range for improved quantification.

5.3.5. Conclusions

With the presented work, an alternative approach towards class-specific quantification in an untargeted RP-LC-MS lipidomic assay is suggested. Surrogate calibrant methods have been shown highly suitable for quantification of endogenous analytes, especially when no true blank matrix is available. This principle has been transferred to untargeted high-resolution MS in combination with comprehensive, quantifiable SWATH-acquisition, which offers potential improvements regarding selectivity when compared to TOF. Moreover, SWATH is able to generate enhanced sensitivity, as MS/MS generally yields beneficial signal-to-noise ratios. Nevertheless, SWATH did not automatically provide the best results for all lipid classes and several surrogate calibrants showed superior performance on the TOF-level. This might be due to the lack of specific and sensitive fragments, yet, MS parameters like collision energy

can be individually tailored for each SWATH experiment to optimize the sensitivity. After all, it was demonstrated that sufficient precision and accuracy can be obtained with SWATH (and TOF) and that previously determined response factors are needed for accurate quantification. For the majority of surrogate calibrants, sufficient results for precision and accuracy, complying with proposed thresholds for targeted assays, were obtained. The presented approach was shown to yield improved results compared to 1-point calibration and was in acceptable agreement with standard addition for tested lipids. Moreover, issues concerning response factors, which were shown to be essential for accurate quantification, were addressed by post-acquisition re-calibration via analysis of authentic standards (in methanolic solution) and NIST SRM 1950 (matrix-matched but species mismatched). Productive workflows, which comprise response factor determination for analytes of interest post-analysis, or method-specific response factor libraries emphasize the potential application of this approach.

Ultimately, the target goal to obtain quantitative results that enhance inter-study, inter-batch or database comparability was demonstrated. Yet, future challenges remain, primarily with the persistent lack of suitable ISs. An alternative approach to account for this issue could be the use of low abundant odd-chain lipid species for surrogate calibration in combination with stable isotope labeled lipids as ISs for interference-free normalization. Regarding the potential advantages, further studies addressing absolute quantification are anticipated as they can aid to maximize the extent and quality of the information that can be extracted from untargeted lipidomic assays.

Acknowledgements

We acknowledge the financial support by the “Struktur- und Innovationsfonds Baden-Württemberg (SI-BW)”, the German Science Funds (DFG no. INST 37/821-1 FUGG) and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project number 374031971 – TRR 240.

Table 3. Results of standard addition, surrogate calibration and 1-point calibration of mouse plasma samples (pooled QC).

Method	14:0-14:0 PC			16:0-16:0 PC			18:0-18:0 PC			20:0 LPC		
	Conc. [ng mL ⁻¹]	Acc. [%]	Prec. [%]	Conc. [ng mL ⁻¹]	Acc. [%]	Prec. [%]	Conc. [ng mL ⁻¹]	Acc. [%]	Prec. [%]	Conc. [ng mL ⁻¹]	Acc. [%]	Prec. [%]
TOF-MS - positive mode												
Standard addition	87.6	102.9	2.6	13,288	102.3	3.0	789.1	99.0	3.5	1,298.4	103.5	0.8
Surrogate calibration	125.2	142.9	6.0	14,701	110.6	8.0	813.6	103.1	6.7	1,468.9	113.1	4.8
1-point calibration	148.7	169.7	9.9	17,534	132.0	2.2	973.1	123.3	8.5	1,571.1	121.0	7.7
SWATH-MS/MS - positive mode												
Standard addition	-	-	-	15,910	90.9	3.0	-	-	-	1,543.3	98.0	2.1
Surrogate calibration	-	-	-	12,689	79.8	6.1	-	-	-	1,720.1	111.5	9.9
1-point calibration	-	-	-	13,644	85.8	12.9	-	-	-	1,864.2	120.8	12.4
TOF-MS - negative mode												
Standard addition	96.8	100.3	5.2	14,420	91.1	0.4	759.9	102.0	5.4	1,404.9	100.5	0.9
Surrogate calibration	140.5	145.1	9.3	12,398	86.0	5.8	595.5	78.4	2.3	1,417.6	100.9	3.5
1-point calibration	161.6	166.9	11.8	14,013	97.2	9.9	624.4	82.2	7.7	1,511.3	107.6	4.1
SWATH-MS/MS - negative mode												
Standard addition	-	-	-	-	-	-	677.6	98.1	3.2	1,324.9	95.2	2.6
Surrogate calibration	-	-	-	-	-	-	717.1	105.8	11.2	1,472.4	111.1	6.4
1-point calibration	-	-	-	-	-	-	1,180.5	174.2	13.5	1,680.0	126.8	8.7

Cells marked with * indicate a peak interference in the respective mode, leading to inaccurate integration and quantification.

5.3.6. References

- [1] D.J. Stephenson, L.A. Hoeflerlin, C.E. Chalfant, *Lipidomics in translational research and the clinical significance of lipid-based biomarkers*, *Transl. Res.*, 189 (2017) 13-29.
- [2] K. Yang, X. Han, *Lipidomics: Techniques, Applications, and Outcomes Related to Biomedical Sciences*, *Trends. Biochem. Sci.*, 41 (2016) 954-969.
- [3] F.R. Yan, H. Zhao, Y.M. Zeng, *Lipidomics: a promising cancer biomarker*, *Clin. Transl. Med.*, 7 (2018).
- [4] Y.Y. Zhao, X.L. Cheng, R.C. Lin, F. Wei, *Lipidomics applications for disease biomarker discovery in mammal models*, *Biomark. Med.*, 9 (2015) 153-168.
- [5] Y.Y. Zhao, X.L. Cheng, R.C. Lin, *Lipidomics Applications for Discovering Biomarkers of Diseases in Clinical Chemistry*, *Int. Rev. Cel. Mol. Bio.*, 313 (2014) 1-26.
- [6] T. Cajka, O. Fiehn, *Toward Merging Untargeted and Targeted Methods in Mass Spectrometry-Based Metabolomics and Lipidomics*, *Anal. Chem.*, 88 (2016) 524-545.
- [7] M. Sud, E. Fahy, D. Cotter, A. Brown, E.A. Dennis, C.K. Glass, A.H. Merrill, R.C. Murphy, C.R.H. Raetz, D.W. Russell, S. Subramaniam, *LMSD: LIPID MAPS structure database*, *Nucleic Acids Res.*, 35 (2007) D527-D532.
- [8] T. Kind, K.H. Liu, D.Y. Lee, B. DeFelice, J.K. Meissen, O. Fiehn, *LipidBlast in silico tandem mass spectrometry database for lipid identification*, *Nat. Methods.*, 10 (2013) 755-758.
- [9] H. Tsugawa, T. Cajka, T. Kind, Y. Ma, B. Higgins, K. Ikeda, M. Kanazawa, J. VanderGheynst, O. Fiehn, M. Arita, *MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis*, *Nat. Methods.*, 12 (2015) 523-526.
- [10] D.S. Wishart, Y.D. Feunang, A. Marcu, A.C. Guo, K. Liang, R. Vazquez-Fresno, T. Sajed, D. Johnson, C.R. Li, N. Karu, Z. Sayeeda, E. Lo, N. Assempour, M. Berjanskii, S. Singhal, D. Arndt, Y.J. Liang, H. Badran, J. Grant, A. Serra-Cayuela, Y.F. Liu, R. Mandal, V. Neveu, A. Pon, C. Knox, M. Wilson, C. Manach, A. Scalbert, *HMDB 4.0: the human metabolome database for 2018*, *Nucleic Acids Res.*, 46 (2018) D608-D617.
- [11] M.A. Surma, R. Herzog, A. Vasilj, C. Klose, N. Christinat, D. Morin-Rivron, K. Simons, M. Masoodi, J.L. Sampaio, *An automated shotgun lipidomics platform for high throughput, comprehensive, and quantitative analysis of blood plasma intact lipids*, *Eur. J. Lipid. Sci. Tech.*, 117 (2015) 1540-1549.
- [12] M. Koivusalo, P. Haimi, L. Heikinheimo, R. Kostianen, P. Somerharju, *Quantitative determination of phospholipid compositions by ESI-MS: effects of acyl chain length, unsaturation, and lipid concentration on instrument response*, *J. Lipid. Res.*, 42 (2001) 663-672.
- [13] M. Wang, C.Y. Wang, X.L. Han, *Selection of Internal Standards for Accurate Quantification of Complex Lipid Species in Biological Extracts by Electrospray Ionization Mass Spectrometry-What, How and Why?*, *Mass Spectrom. Rev.*, 36 (2017) 693-714.
- [14] B. Burla, M. Arita, M. Arita, A.K. Bendt, A. Cazenave-Gassiot, E.A. Dennis, K. Ekroos, X. Han, K. Ikeda, G. Liebisch, M.K. Lin, T.P. Loh, P.J. Meikle, M. Oresic, O. Quehenberger, A. Shevchenko, F. Torta, M.J.O. Wakelam, C.E. Wheelock, M.R. Wenk, *MS-based lipidomics of human blood plasma: a community-initiated position paper to develop accepted guidelines*, *J. Lipid. Res.*, 59 (2018) 2001-2017.
- [15] E. Rampler, C. Coman, G. Hermann, A. Sickmann, R. Ahrends, G. Koellensperger, *LILY-lipidome isotope labeling of yeast: in vivo synthesis of C-13 labeled reference lipids for quantification by mass spectrometry*, *Analyst*, 142 (2017) 1891-1899.
- [16] T. Cajka, O. Fiehn, *Comprehensive analysis of lipids in biological systems by liquid chromatography-mass spectrometry*, *Trac-Trend. Anal. Chem.*, 61 (2014) 192-206.
- [17] M. Lisa, M. Holcapek, *High-Throughput and Comprehensive Lipidomic Analysis Using Ultrahigh-Performance Supercritical Fluid Chromatography-Mass Spectrometry*, *Anal. Chem.*, 87 (2015) 7187-7195.
- [18] E. Cifkova, M. Holcapek, M. Lisa, M. Ovcacikova, A. Lycka, F. Lynen, P. Sandra, *Nontargeted quantitation of lipid classes using hydrophilic interaction liquid chromatography-electrospray ionization mass spectrometry with single internal standard and response factor approach*, *Anal. Chem.*, 84 (2012) 10064-10070.
- [19] M. Wang, C. Wang, X. Han, *Selection of internal standards for accurate quantification of complex lipid species in biological extracts by electrospray ionization mass spectrometry-What, how and why?*, *Mass Spectrom. Rev.*, 36 (2017) 693-714.

- [20] K. Contrepolis, S. Mahmoudi, B.K. Ubhi, K. Papsdorf, D. Hornburg, A. Brunet, M. Snyder, Cross-Platform Comparison of Untargeted and Targeted Lipidomics Approaches on Aging Mouse Plasma, *Sci Rep-Uk*, 8 (2018).
- [21] E.D. Tague, B.M. Woodall, J.R. Harp, A.T. Farmer, E.M. Fozo, S.R. Campagna, Expanding lipidomics coverage: effective ultra performance liquid chromatography-high resolution mass spectrometer methods for detection and quantitation of cardiolipin, phosphatidylglycerol, and lysyl-phosphatidylglycerol, *Metabolomics*, 15 (2019) 53.
- [22] C. Hinz, S. Liggi, G. Mocciaro, S.M. Jung, I. Induruwa, M.C.D.A. Pereira, C.E. Bryant, S.W. Meckelmann, V.B. O'Donnell, R.W. Farndale, J.C. Fjeldsted, J.L. Griffin, A comprehensive UHPLC ion mobility QTOF method for profiling and quantification of eicosanoids, other oxylipins and fatty acids, *Anal. Chem.*, (2019).
- [23] J.A. Bowden, A. Heckert, C.Z. Ulmer, C.M. Jones, J.P. Koelmel, L. Abdullah, L. Ahonen, Y. Alnouti, A.M. Armando, J.M. Asara, T. Bamba, J.R. Barr, J. Bergquist, C.H. Borchers, J. Brandsma, S.B. Breitkopf, T. Cajka, A. Cazenave-Gassiot, A. Checa, M.A. Cinel, R.A. Colas, S. Cremers, E.A. Dennis, J.E. Evans, A. Fauland, O. Fiehn, M.S. Gardner, T.J. Garrett, K.H. Gotlinger, J. Han, Y. Huang, A.H. Neo, T. Hyotylainen, Y. Izumi, H. Jiang, H. Jiang, J. Jiang, M. Kachman, R. Kiyonami, K. Klavins, C. Klose, H.C. Kofeler, J. Kolmert, T. Koal, G. Koster, Z. Kuklennyik, I.J. Kurland, M. Leadley, K. Lin, K.R. Maddipati, D. McDougall, P.J. Meikle, N.A. Mellett, C. Monnin, M.A. Moseley, R. Nandakumar, M. Oresic, R. Patterson, D. Peake, J.S. Pierce, M. Post, A.D. Postle, R. Pugh, Y. Qiu, O. Quehenberger, P. Ramrup, J. Rees, B. Rembiesa, D. Reynaud, M.R. Roth, S. Sales, K. Schuhmann, M.L. Schwartzman, C.N. Serhan, A. Shevchenko, S.E. Somerville, L. St John-Williams, M.A. Surma, H. Takeda, R. Thakare, J.W. Thompson, F. Torta, A. Triebel, M. Trotsmuller, S.J.K. Ubhayasekera, D. Vuckovic, J.M. Weir, R. Welti, M.R. Wenk, C.E. Wheelock, L. Yao, M. Yuan, X.H. Zhao, S. Zhou, Harmonizing lipidomics: NIST interlaboratory comparison exercise for lipidomics using SRM 1950-Metabolites in Frozen Human Plasma, *J. Lipid. Res.*, 58 (2017) 2275-2288.
- [24] O. Quehenberger, A.M. Armando, A.H. Brown, S.B. Milne, D.S. Myers, A.H. Merrill, S. Bandyopadhyay, K.N. Jones, S. Kelly, R.L. Shaner, C.M. Sullards, E. Wang, R.C. Murphy, R.M. Barkley, T.J. Leiker, C.R. Raetz, Z. Guan, G.M. Laird, D.A. Six, D.W. Russell, J.G. McDonald, S. Subramaniam, E. Fahy, E.A. Dennis, Lipidomics reveals a remarkable diversity of lipids in human plasma, *J. Lipid. Res.*, 51 (2010) 3299-3305.
- [25] T. Cajka, J.T. Smilowitz, O. Fiehn, Validating Quantitative Untargeted Lipidomics Across Nine Liquid Chromatography-High-Resolution Mass Spectrometry Platforms, *Anal. Chem.*, 89 (2017) 12360-12368.
- [26] A. Triebel, M. Trotsmuller, J. Hartler, T. Stojakovic, H.C. Kofeler, Lipidomics by ultrahigh performance liquid chromatography-high resolution mass spectrometry and its application to complex biological samples, *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.*, 1053 (2017) 72-80.
- [27] T. Hyotylainen, L. Ahonen, P. Poho, M. Oresic, Lipidomics in biomedical research-practical considerations, *Bba-Mol. Cell. Biol. L.*, 1862 (2017) 800-803.
- [28] *Bioanalytical Method Validation - Guidance for Industry*, in: U.F.a.D. Administration (Ed.), 2018.
- [29] L.C. Gillet, P. Navarro, S. Tate, H. Rost, N. Selevsek, L. Reiter, R. Bonner, R. Aebersold, Targeted Data Extraction of the MS/MS Spectra Generated by Data-independent Acquisition: A New Concept for Consistent and Accurate Proteome Analysis, *Mol. Cell. Proteomics*, 11 (2012) O111.016717.
- [30] J. Illison, L. Tian, H. McClafferty, M. Werno, L.H. Chamberlain, V. Leiss, A. Sassmann, S. Offermanns, P. Ruth, M.J. Shipston, R. Lukowski, Obesogenic and Diabetogenic Effects of High-Calorie Nutrition Require Adipocyte BK Channels, *Diabetes*, 65 (2016) 3621-3635.
- [31] M.H. Sarafian, M. Gaudin, M.R. Lewis, F.P. Martin, E. Holmes, J.K. Nicholson, M.E. Dumas, Objective Set of Criteria for Optimization of Sample Preparation Procedures for Ultra-High Throughput Untargeted Blood Plasma Lipid Profiling by Ultra Performance Liquid Chromatography-Mass Spectrometry, *Anal. Chem.*, 86 (2014) 5766-5774.
- [32] C. Calderón, C. Sanwald, J. Schlotterbeck, B. Drotleff, M. Lämmerhofer, Comparison of simple monophasic versus classical biphasic extraction protocols for comprehensive UHPLC-MS/MS lipidomic analysis of HeLa cells, *Anal. Chim. Acta*, 1048 (2018) 66-74.
- [33] B. Drotleff, M. Hallschmid, M. Lämmerhofer, Quantification of steroid hormones in plasma using a surrogate calibrant approach and UHPLC-ESI-QTOF-MS/MS with SWATH-acquisition combined with untargeted profiling, *Analytica Chimica Acta*, 1022 (2018) 70-80.
- [34] B.R. Jones, G.A. Schultz, J.A. Eckstein, B.L. Ackermann, Surrogate matrix and surrogate analyte approaches for definitive quantitation of endogenous biomolecules, *Bioanalysis*, 4 (2012) 2343-2356.

- [35] W.L. Li, L.H. Cohen, Quantitation of endogenous analytes in biofluid without a true blank matrix, *Anal. Chem.*, 75 (2003) 5854-5859.
- [36] W.S. Cleveland, Lowess - a Program for Smoothing Scatterplots by Robust Locally Weighted Regression, *Am. Stat.*, 35 (1981) 54-54.
- [37] Y. Zhang, A. Bilbao, T. Bruderer, J. Luban, C. Strambio-De-Castilla, F. Lisacek, G. Hopfgartner, E. Varesio, The Use of Variable Q1 Isolation Windows Improves Selectivity in LC-SWATH-MS Acquisition, *J. Proteome Res.*, 14 (2015) 4359-4371.
- [38] J. Schlotterbeck, M. Chatterjee, M. Gawaz, M. Lammerhofer, Comprehensive MS/MS profiling by UHPLC-ESI-QTOF-MS/MS using SWATH data-independent acquisition for the study of platelet lipidomes in coronary artery disease, *Anal. Chim. Acta*, 1046 (2019) 1-15.
- [39] A. Sassmann, S. Offermanns, N. Wettschureck, Tamoxifen-inducible Cre-mediated recombination in adipocytes, *Genesis*, 48 (2010) 618-625.
- [40] M. Sausbier, C. Arntz, I. Bucurenciu, H. Zhao, X.B. Zhou, U. Sausbier, S. Feil, S. Kamm, K. Essin, C.A. Sailer, U. Abdullah, P. Krippeit-Drews, R. Feil, F. Hofmann, H.G. Knaus, C. Kenyon, M.J. Shipston, J.F. Storm, W. Neuhuber, M. Korth, R. Schubert, M. Gollasch, P. Ruth, Elevated blood pressure linked to primary hyperaldosteronism and impaired vasodilation in BK channel-deficient mice, *Circulation*, 112 (2005) 60-68.
- [41] M. Wang, Y.Y. Huang, X.L. Han, Accurate mass searching of individual lipid species candidates from high-resolution mass spectra for shotgun lipidomics, *Rapid Commun. Mass. Sp.*, 28 (2014) 2201-2210.
- [42] B. Jenkins, J.A. West, A. Koulman, A Review of Odd-Chain Fatty Acid Metabolism and the Role of Pentadecanoic Acid (C15:0) and Heptadecanoic Acid (C17:0) in Health and Disease, *Molecules*, 20 (2015) 2425-2444.
- [43] B.K. Matuszewski, M.L. Constanzer, C.M. Chavez-Eng, Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on HPLC-MS/MS, *Anal. Chem.*, 75 (2003) 3019-3030.
- [44] P.J. Taylor, Matrix effects: the Achilles heel of quantitative high-performance liquid chromatography-electrospray-tandem mass spectrometry, *Clin. Biochem.*, 38 (2005) 328-334.
- [45] G.L. Andrews, B.L. Simons, J.B. Young, A.M. Hawkridge, D.C. Muddiman, Performance characteristics of a new hybrid quadrupole time-of-flight tandem mass spectrometer (TripleTOF 5600), *Anal. Chem.*, 83 (2011) 5442-5446.
- [46] R.M. Gathungu, P. Larrea, M.J. Sniatynski, V.R. Marur, J.A. Bowden, J.P. Koelmel, P. Starke-Reed, V.S. Hubbard, B.S. Kristal, Optimization of Electrospray Ionization Source Parameters for Lipidomics To Reduce Misannotation of In-Source Fragments as Precursor Ions, *Anal. Chem.*, 90 (2018) 13523-13532.
- [47] J.P. Koelmel, C.Z. Ulmer, C.M. Jones, R.A. Yost, J.A. Bowden, Common cases of improper lipid annotation using high resolution tandem mass spectrometry data and corresponding limitations in biological interpretation (vol 1862, pg 766, 2017), *Bba-Mol. Cell. Biol. L.*, 1862 (2017) 1024-1024.
- [48] M.N. Barber, S. Risis, C. Yang, P.J. Meikle, M. Staples, M.A. Febbraio, C.R. Bruce, Plasma lysophosphatidylcholine levels are reduced in obesity and type 2 diabetes, *Plos One*, 7 (2012) e41456.
- [49] K. Eisinger, G. Liebisch, G. Schmitz, C. Aslanidis, S. Krautbauer, C. Buechler, Lipidomic analysis of serum from high fat diet induced obese mice, *Int. J. Mol. Sci.*, 15 (2014) 2991-3002.

5.3.7. Supplementary Material

Table S-1. Concentration of single-point calibrants in study samples.^a

Compound name	Conc. [ng mL ⁻¹]	t _R [min]	Adduct type	
			pos	neg
17:1 LPC [*]	80	1.31 ± 0.00	[M+H] ⁺	[M+FA-H] ⁻
18:1(d7) LPC [§]	1,360	1.61 ± 0.01	[M+H] ⁺	[M+FA-H] ⁻
18:1(d7) LPE [§]	283	1.65 ± 0.00	[M+H] ⁺	[M-H] ⁻
18:1(d7) MAG [§]	107	2.81 ± 0.01	-	[M+FA-H] ⁻
15:0-18:1(d7) PI [§]	476	4.53 ± 0.00	[M+NH ₄] ⁺	[M-H] ⁻
15:0-18:1(d7) PS [§]	218	4.69 ± 0.03	[M+H] ⁺	[M-H] ⁻
15:0-18:1(d7) PG [§]	1,507	4.78 ± 0.04	[M+H] ⁺	[M-H] ⁻
d18:1-18:1(d9) SM [§]	1,648	4.87 ± 0.00	[M+H] ⁺	[M+FA-H] ⁻
Cholesterol(d7) [§]	5,248	4.83 ± 0.02	[M-H ₂ O+H] ⁺	-
15:0-18:1(d7) PC [§]	8,571	5.24 ± 0.00	[M+H] ⁺	[M+FA-H] ⁻
15:0-18:1(d7) PE [§]	304	5.42 ± 0.00	[M+H] ⁺	[M-H] ⁻
15:0-18:1(d7) DAG [§]	501	6.67 ± 0.00	[M+NH ₄] ⁺	[M+FA-H] ⁻
15:0-18:1(d7)-15:0 TAG [§]	3,056	10.59 ± 0.00	[M+Na] ⁺	-
18:1(d7) CE [§]	18,992	11.04 ± 0.05	[M+NH ₄] ⁺	-
Arachidonic acid(d8)	533	2.73 ± 0.01	-	[M-H] ⁻
α-Linolenic(d14)	533	2.28 ± 0.01	-	[M-H] ⁻
Linoleic acid(d4)	533	2.84 ± 0.01	-	[M-H] ⁻

^aExtracted masses of precursors and fragments are listed in Table S-6. ^{*}17:1 LPC served as IS for IS_{quant} and was spiked to all samples. [§]Lipid species that are covered in the Lipidomix.

Table S-2. Concentration of surrogate calibrants in calibration samples.^b

Surrogate calibrant	Cal. 1	Cal. 2	Cal. 3	Cal. 4	Cal. 5	Cal. 6	Cal. 7	Cal. 8	Cal. 9	Cal. 10	Cal. 11
Dilution factor to Cal. 11 ^a	5120	512	256	128	64	32	16	8	4	2	-
18:1(d7) LPC	3.98	39.84	79.69	159.38	318.75	637.5	1275	2550	5100	10200	20400
18:1(d7) LPE	0.83	8.28	16.56	33.13	66.25	132.5	265	530	1060	2120	4240
18:1(d7) MAG	0.31	3.13	6.25	12.50	25.00	50.0	100	200	400	800	1600
15:0-18:1(d7) PI	1.39	13.93	27.87	55.73	111.46	222.9	446	892	1783	3567	7134
15:0-18:1(d7) PS	0.64	6.38	12.75	25.51	51.01	102.0	204	408	816	1632	3265
15:0-18:1(d7) PG	4.42	44.16	88.32	176.64	353.28	706.6	1413	2826	5652	11305	22610
d18:1-18:1(d9) SM	4.83	48.28	96.56	193.13	386.25	772.5	1545	3090	6180	12360	24720
Cholesterol(d7)	15.38	153.75	307.50	615.00	1230.00	2460.0	4920	9840	19680	39360	78720
15:0-18:1(d7) PC	25.11	251.09	502.19	1004.38	2008.75	4017.5	8035	16070	32140	64280	128560
15:0-18:1(d7) PE	0.89	8.91	17.81	35.63	71.25	142.5	285	570	1140	2280	4560
15:0-18:1(d7) DAG	1.47	14.69	29.38	58.75	117.50	235.0	470	940	1880	3760	7520
15:0-18:1(d7)-15:0 TAG	8.95	89.53	179.06	358.13	716.25	1432.5	2865	5730	11460	22920	45840
18:1(d7) CE	55.64	556.41	1112.81	2225.63	4451.25	8902.5	17805	35610	71220	142440	284880
Arachidonic acid(d8)	1.56	15.63	31.25	62.50	125.00	250.0	500	1000	2000	4000	8000
α -Linolenic acid(d14)	1.56	15.63	31.25	62.50	125.00	250.0	500	1000	2000	4000	8000
Linoleic acid(d4)	1.56	15.63	31.25	62.50	125.00	250.0	500	1000	2000	4000	8000

^bConcentrations are given in ng mL⁻¹. ^aSamples were prepared by serial dilution of the spiked methanolic portion of the precipitation mix. Dilution factors are given compared to the highest calibrant.

Table S-3. Concentration of calibrants (surrogate or 1-point calibrants) in QC samples.^c

Compound name	QC_{quant 1}	QC_{quant 2}	QC_{quant 3} / QC_{syst}	QC_{quant 4}	QC_{quant 5}
18:1(d7) LPC	204.0	1020	5100	10200	16320
18:1(d7) LPE	42.4	212	1060	2120	3392
18:1(d7) MAG	16.0	80	400	800	1280
15:0-18:1(d7) PI	71.3	356.7	1783.4	3566.8	5706.9
15:0-18:1(d7) PS	32.6	163.2	816.2	1632.4	2611.9
15:0-18:1(d7) PG	226.1	1130.5	5652.4	11304.9	18087.8
d18:1-18:1(d9) SM	247.2	1236	6180	12360	19776
Cholesterol(d7)	787.2	3936	19680	39360	62976
15:0-18:1(d7) PC	1285.6	6428	32140	64280	102848
15:0-18:1(d7) PE	45.6	228	1140	2280	3648
15:0-18:1(d7) DAG	75.2	376	1880	3760	6016
15:0-18:1(d7)-15:0 TAG	458.4	2292	11460	22920	36672
18:1(d7) CE	2848.8	14244	71220	142440	227904
Arachidonic acid(d8)	80.0	400	2000	4000	6400
α -Linolenic acid(d14)	80.0	400	2000	4000	6400
Linoleic acid(d4)	80.0	400	2000	4000	6400

^cConcentrations are given in ng mL⁻¹.

Table S-4. Detailed sequence of analysis.^d

Sample Name	Class	Analytical order	Sample Name	Class	Analytical order	Sample Name	Class	Analytical order
Matrix Blank	-	0	QC _{syst}	QC _{syst}	35	D3396	KO-HFD	70
Calibration 0	Cal	1	D3417	CTR-HFD	36	QC _{syst}	QC _{syst}	71
Calibration 1	Cal	2	D0185	CTR-CD	37	D3394	KO-HFD	72
Calibration 2	Cal	3	D0167	CTR-HFD	38	D3380	CTR-CD	73
Calibration 3	Cal	4	D3385	KO-CD	39	D3305	CTR-HFD	74
Calibration 4	Cal	5	D3435	KO-CD	40	D3397	CTR-HFD	75
Calibration 5	Cal	6	QC _{syst}	QC _{syst}	41	D3370	CTR-HFD	76
Calibration 6	Cal	7	B9278	CTR-HFD	42	QC _{syst}	QC _{syst}	77
Calibration 7	Cal	8	D3353	KO-HFD	43	D0139	CTR-CD	78
Calibration 8	Cal	9	B9258	KO-HFD	44	D0168	KO-CD	79
Calibration 9	Cal	10	D0181	KO-HFD	45	D0176	CTR-HFD	80
Calibration 10	Cal	11	D3371	KO-HFD	46	B9345	CTR-CD	81
Calibration 11	Cal	12	QC _{syst}	QC _{syst}	47	B9324	CTR-HFD	82
QC _{quant} 1	QC _{quant}	13	Matrix Blank	-	48	QC _{syst}	QC _{syst}	83
QC _{quant} 2	QC _{quant}	14	Calibration 0	Cal	49	D0142	KO-CD	84
QC _{quant} 3	QC _{quant}	15	Calibration 1	Cal	50	QC _{syst}	QC _{syst}	85
QC _{quant} 4	QC _{quant}	16	Calibration 2	Cal	51	QC _{syst}	QC _{syst}	86
QC _{quant} 5	QC _{quant}	17	Calibration 3	Cal	52	QC _{syst}	QC _{syst}	87
D0163	CTR-CD	18	Calibration 4	Cal	53	Matrix Blank	-	88
D0138	CTR-CD	19	Calibration 5	Cal	54	Calibration 0	Cal	89
D3432	KO-CD	20	Calibration 6	Cal	55	Calibration 1	Cal	90
D0143	CTR-CD	21	Calibration 7	Cal	56	Calibration 2	Cal	91
D3381	CTR-CD	22	Calibration 8	Cal	57	Calibration 3	Cal	92
QC _{syst}	QC _{syst}	23	Calibration 9	Cal	58	Calibration 4	Cal	93
D3377	KO-CD	24	Calibration 10	Cal	59	Calibration 5	Cal	94
B9256	KO-HFD	25	Calibration 11	Cal	60	Calibration 6	Cal	95
D0137	CTR-CD	26	QC _{quant} 1	QC _{quant}	61	Calibration 7	Cal	96
B9305	CTR-HFD	27	QC _{quant} 2	QC _{quant}	62	Calibration 8	Cal	97
B9343	KO-CD	28	QC _{quant} 3	QC _{quant}	63	Calibration 9	Cal	98
QC _{syst}	QC _{syst}	29	QC _{quant} 4	QC _{quant}	64	Calibration 10	Cal	99
B9257	KO-HFD	30	QC _{quant} 5	QC _{quant}	65	Calibration 11	Cal	100
D0191	KO-CD	31	D0186	KO-CD	66	QC _{quant} 1	QC _{quant}	101
D0183	KO-HFD	32	D3379	CTR-CD	67	QC _{quant} 2	QC _{quant}	102
D0182	KO-HFD	33	D3357	CTR-HFD	68	QC _{quant} 3	QC _{quant}	103
D3392	KO-HFD	34	D3383	CTR-CD	69	QC _{quant} 4	QC _{quant}	104
						QC _{quant} 5	QC _{quant}	105

^dSample D3371 was removed from further data processing as a classification error was suspected. Labeled lipid concentrations of QC_{syst} samples were identical to QC_{quant} 3 samples but independently prepared.

Table S-5. MS and MS/MS experiment settings in positive and negative ionization mode.^e

Experiment	Positive mode		Negative mode	
	<i>m/z</i> start	<i>m/z</i> end	<i>m/z</i> start	<i>m/z</i> end
TOF-MS	50	1,250	50	1,250
SWATH-MS/MS 1	50	163.6	50	119.4
SWATH-MS/MS 2	162.6	214.6	118.4	135.4
SWATH-MS/MS 3	213.6	256.8	134.4	160.4
SWATH-MS/MS 4	255.8	310.3	159.4	175.5
SWATH-MS/MS 5	309.3	360.3	174.5	194.6
SWATH-MS/MS 6	359.3	427.4	193.6	214.6
SWATH-MS/MS 7	426.4	497.9	213.6	243.4
SWATH-MS/MS 8	496.9	541.9	242.4	283.7
SWATH-MS/MS 9	540.9	572.9	282.7	301.7
SWATH-MS/MS 10	571.9	622.5	300.7	344.6
SWATH-MS/MS 11	621.5	666.1	343.6	377.8
SWATH-MS/MS 12	665.1	701.1	376.8	395.6
SWATH-MS/MS 13	700.1	731.1	394.6	437.4
SWATH-MS/MS 14	730.1	755.1	436.4	481.8
SWATH-MS/MS 15	754.1	771.2	480.8	519.9
SWATH-MS/MS 16	770.2	786.2	518.9	579.4
SWATH-MS/MS 17	785.2	800.9	578.4	636.8
SWATH-MS/MS 18	799.9	814.1	635.8	689.2
SWATH-MS/MS 19	813.1	828.6	688.2	730
SWATH-MS/MS 20	827.6	841.2	729	775
SWATH-MS/MS 21	840.2	864.1	774	815
SWATH-MS/MS 22	863.1	892.8	814	853.1
SWATH-MS/MS 23	891.8	927.3	852.1	887
SWATH-MS/MS 24	926.3	1,044.2	886	953.2
SWATH-MS/MS 25	1,043.2	1,250	952.2	1,250

^eSWATH window settings were determined using IDA data of a QC_{sys}t sample. After peak-finding on TOF-MS level via PeakView, data was transferred to swathTUNER. SWATH design was based on optimized precursor density.

Table S-6. Sum formulas and extracted masses for surrogate calibrants and ISs.^f

Compound name	Sum Formula	Positive mode		Negative mode	
		TOF extraction mass	SWATH extraction mass	TOF extraction mass	SWATH extraction mass
17:1 LPC	C ₂₅ H ₅₀ NO ₇ P	508.3398	508.3398	552.3307	492.3096
18:1(d7) LPC	C ₂₆ H ₄₅ ² H ₇ NO ₇ P	529.3994	346.3333	573.3903	288.2925
18:1(d7) LPE	C ₂₃ H ₃₉ ² H ₇ NO ₇ P	487.3524	346.3333	485.3379	288.2925
18:1(d7) MAG	C ₂₁ H ₃₃ ² H ₇ O ₄	-	272.2965	408.3348	-
15:0-18:1(d7) PI	C ₄₂ H ₇₂ ² H ₇ O ₁₃ P	847.6036	570.5473	828.5625	828.5625
15:0-18:1(d7) PS	C ₃₉ H ₆₇ ² H ₇ NO ₁₀ P	755.5563	570.5473	753.5417	288.2925
15:0-18:1(d7) PG	C ₃₉ H ₆₈ ² H ₇ O ₁₀ P	742.5610	570.5473	740.5465	288.2925
d18:1-18:1(d9) SM	C ₄₁ H ₇₂ ² H ₉ N ₂ O ₆ P	738.6470	720.6364	782.6379	722.6368
Cholesterol(d7)	C ₂₇ H ₃₉ ² H ₇ O	376.3955	376.3955	-	-
15:0-18:1(d7) PC	C ₄₁ H ₇₃ ² H ₇ NO ₈ P	753.6134	570.5473	797.6043	288.2925
15:0-18:1(d7) PE	C ₃₈ H ₆₇ ² H ₇ NO ₈ P	711.5664	570.5473	709.5519	288.2925
15:0-18:1(d7) DAG	C ₃₆ H ₆₁ ² H ₇ O ₅	605.5844	346.3333	632.5488	-
15:0-18:1(d7)-15:0 TAG	C ₅₁ H ₈₉ ² H ₇ O ₆	834.7539	834.7539	-	-
18:1(d7) CE	C ₄₅ H ₇₁ ² H ₇ O ₂	675.6780	-	-	-
Arachidonic acid(d8)	C ₂₀ H ₂₄ ² H ₈ O ₂	-	-	311.2832	267.2933
α-Linolenic(d14)	C ₁₈ H ₁₆ ² H ₁₄ O ₂	-	-	291.3052	-
Linoleic acid(d4)	C ₁₈ H ₂₈ ² H ₄ O ₂	-	-	283.2581	-

^fIn some cases SWATH and TOF masses are identical due to insufficient fragmentation or improved data quality for precursors in SWATH-MS/MS (see 17:1 LPC, Cholesterol(d7), PI(d7)).

Table S-7. Spiked lipid concentrations for standard addition in mouse plasma (pooled QC).^g

Sample	Spiked concentration [ng mL ⁻¹]			
	14:0-14:0 PC	16:0-16:0 PC	18:0-18:0 PC	20:0 LPC
Std.add. 0	0	0	0	0
Std.add. 1	150	15,000	800	1,500
Std.add. 2	300	30,000	1,600	3,000
Std.add. 3	450	45,000	2,400	4,500
Std.add. 4	600	60,000	3,200	6,000

^gMouse plasma was spiked and subsequently processed according to the described sample preparation protocol.

Table S-8. Calibration results for surrogate calibrants in mouse plasma (pooled QC).^h

Surrogate Calibrant	Experiment	Linear range [ng mL ⁻¹] (LLOQ – ULOQ)	Weighting	R ²	Slope	Intercept	IS _{quant}
LPC(d7)	*TOF ⁺	39.84 – 20,400	1/x ²	0.9920	5.91E-04	-1.75E-04	√
	SWATH ⁺	39.84 – 20,400	1/x ²	0.9951	9.19E-05	-4.17E-04	√
	TOF ⁻	39.84 – 20,400	1/x ²	0.9908	7.85E-04	4.94E-04	√
LPE(d7)	SWATH ⁻	159.38 – 10,200	1/x	0.9938	6.72E-04	1.74E-02	√
	TOF ⁺	33.13 – 4,240	1/x	0.9974	1.25E-04	-6.71E-04	√
	SWATH ⁺	33.13 – 4,240	1/x ²	0.9908	3.73E-04	2.70E-03	√
MAG(d7)	*TOF ⁻	8.28 – 4,240	1/x ²	0.9919	1.12E-03	-2.06E-03	√
	SWATH ⁻	8.28 – 1,060	1/x	0.9975	2.01E-03	9.72E-03	√
	TOF ⁺	-	-	-	-	-	-
PI(d7)	SWATH ⁺	100 – 1,600	1/x	0.9801	5.06E-05	-1.10E-03	√
	*TOF ⁻	100 – 1,600	1/x	0.9919	3.55E-05	-1.37E-03	√
	SWATH ⁻	-	-	-	-	-	-
PS(d7)	TOF ⁺	111.46 – 7,134	1/x	0.9918	3.19E-05	1.27E-04	√
	*SWATH ⁺	13.93 – 7,134	1/x ²	0.9925	1.27E-03	4.50E-03	√
	TOF ⁻	-	-	-	-	-	-
PG(d7)	SWATH ⁻	445.85 – 7,134	1/x	0.9832	7.46E-04	1.43E-01	√
	TOF ⁺	204.06 – 3,265	1/x	0.9959	8.97E-05	2.37E-03	√
	SWATH ⁺	51.01 – 3,265	1/x ²	0.9856	8.44E-05	2.34E-05	√
SM(d9)	TOF ⁻	51.01 – 3,265	1/x ²	0.9912	4.75E-04	-7.47E-03	√
	*SWATH ⁻	51.01 – 3,265	1/x ²	0.9893	1.99E-03	-1.04E-02	√
	TOF ⁺	5652.44 – 22,610	none	0.9549	4.31E-06	1.57E-02	√
Cholesterol(d7)	SWATH ⁺	44.16 – 11,305	1/x ²	0.9857	1.60E-03	1.66E-02	√
	*TOF ⁻	44.16 – 22,610	1/x ²	0.9903	4.60E-04	6.04E-03	√
	SWATH ⁻	353.28 – 22,610	1/x	0.9930	2.81E-04	-1.52E-02	√
Cholesterol(d7)	*TOF ⁺	48.28 – 24,720	1/x ²	0.9938	2.17E-04	-6.21E-04	√
	SWATH ⁺	193.13 – 24,720	1/x ²	0.9890	2.81E-05	1.15E-03	√
	TOF ⁻	96.56 – 24,720	1/x ²	0.9871	1.04E-04	3.23E-03	√
Cholesterol(d7)	SWATH ⁻	95.56 – 24,720	1/x ²	0.9892	2.55E-04	2.08E-02	√
	TOF ⁺	-	-	-	-	-	-
	*SWATH ⁺	1,230 – 78,720	1/x ²	0.9867	3.87E-06	2.21E-04	√
Cholesterol(d7)	TOF ⁻	-	-	-	-	-	-
	SWATH ⁻	-	-	-	-	-	-

Surrogate Calibrant	Experiment	Linear range [ng mL ⁻¹] (LLOQ – ULOQ)	Weighting	R ²	Slope	Intercept	IS _{quant}
PC(d7)	TOF ⁺	25.11 – 128,560	1/x ²	0.9947	4.75E-04	-5.87E-04	✓
	SWATH ⁺	502.2 – 128,560	1/x	0.9988	3.86E-05	-1.94E-03	✓
	TOF ⁻	25.11 – 128,560	1/x ²	0.9946	2.79E-04	1.35E-03	✓
	SWATH ⁻	502.2 – 64,280	1/x ²	0.9897	1.60E-04	4.22E-02	✓
PE(d7)	TOF ⁺	8.91 – 4,560	1/x ²	0.9933	1.48E-04	6.36E-04	✓
	SWATH ⁺	8.91 – 4,560	1/x ²	0.9870	2.35E-03	6.61E-03	✓
	TOF ⁻	35.63 – 4,560	1/x ²	0.9891	4.49E-04	-3.06E-03	✓
	SWATH ⁻	71.25 – 4,560	1/x ²	0.9901	6.74E-04	9.16E-03	✓
DAG(d7)	TOF ⁺	58.75 – 7,520	1/x ²	0.9981	9.38E-05	3.20E-03	✓
	SWATH ⁺	14.69 – 7,520	1/x	0.9941	4.19E-04	-1.22E-03	✓
	TOF ⁻	940 – 7,520	1/x ²	0.9770	7.16E-06	2.02E-03	✓
	SWATH ⁻						
TAG(d7)	TOF ⁺						
	SWATH ⁺	716.25 – 45,840	1/x ²	0.9823	1.15E-04	-1.24E-02	✓
	TOF ⁻						
	SWATH ⁻						
CE(d7)	TOF ⁺	17,805 – 284,880	none	0.9891	7.49E-02	2.22E+02	X
	SWATH ⁺						
	TOF ⁻						
	SWATH ⁻						
Arachidonic acid(d8)	TOF ⁺						
	SWATH ⁺						
	TOF ⁻	15.63 – 8,000	1/x ²	0.9936	7.40E-04	-3.22E-03	✓
	SWATH ⁻	250 – 4,000	1/x ²	0.9841	5.91E-05	-5.10E-03	✓
α-Linolenic acid(d14)	TOF ⁺						
	SWATH ⁺						
	TOF ⁻	250 – 8,000	1/x ²	0.9889	7.96E-05	-3.27E-03	✓
	SWATH ⁻						
Linoleic acid(d4)	TOF ⁺						
	SWATH ⁺						
	TOF ⁻	500 – 4,000	1/x ²	0.9966	1.29E-04	-8.92E-03	✓
	SWATH ⁻						

^hIS_{quant} column indicates if 17:1 LPC was used as IS for calibration. The mode that shows favourable performance in terms of linear range, precision and accuracy is marked with *.

Table S-9. Response factor determination using NIST SRM 1950 plasma.¹

Analyte	Sum formula	t_r [min]	Consensus location (MEDM) [ng mL ⁻¹]	Standard uncertainty (u) [ng mL ⁻¹]	COD [%]	Δ to LIPID MAPS consortium [%]	TOF-MS response factor (mass dependent)	
							pos	neg
CE 16:0	C ₄₃ H ₇₆ O ₂	11.11	131,263	36,254	28	8	1.70*	-
CE 16:1	C ₄₃ H ₇₄ O ₂	10.52	62,305	16,822	27	-8	14.91*	-
CE 17:1	C ₄₄ H ₇₆ O ₂	10.82	5,224	637	13	-73	15.48*	-
CE 18:0	C ₄₅ H ₈₀ O ₂	11.43	9,797	2417	25	-74	5.64*	-
CE 18:1	C ₄₅ H ₇₈ O ₂	11.09	292,995	71,621	25	-15	10.11*	-
CE 18:2	C ₄₅ H ₇₆ O ₂	10.62	1,103,442	279,106	26	-9	16.64*	-
CE 18:3	C ₄₅ H ₇₄ O ₂	10.2	54,354	15,530	28	-43	47.13*	-
CE 20:4	C ₄₇ H ₇₆ O ₂	10.36	235,587	39,040	17	46	57.06*	-
CE 20:5	C ₄₇ H ₇₄ O ₂	9.92	25,501	5,771	23	-	58.73*	-
CE 22:6	C ₄₉ H ₇₆ O ₂	10.11	25,794	6,623	26	16	51.61*	-
Cholesterol	C ₂₇ H ₄₆ O	4.89	297,723	42,532	14	-6	1.05	-
FA 17:0	C ₁₇ H ₃₄ O ₂	3.53	460	62	14	-	-	0.86
FA 17:1	C ₁₇ H ₃₂ O ₂	3.06	231	40	18	-	-	1.04
FA 18:3	C ₁₈ H ₃₀ O ₂	2.42	807	173	21	152	-	1.21
FA 20:1	C ₂₀ H ₃₈ O ₂	3.89	497	53	9	-	-	0.68
FA 20:5	C ₂₀ H ₃₆ O ₂	2.31	127	17	13	-3	-	1.53
FA 22:6	C ₂₂ H ₃₂ O ₂	2.7	493	56	11	55	-	1.84
LPC 14:0	C ₂₂ H ₄₆ NO ₇ P	0.97	468	94	19	-	1.04	1.19
LPC 15:0	C ₂₃ H ₄₈ NO ₇ P	1.21	250	53	22	-	1.10	1.35
LPC 16:0	C ₂₄ H ₅₀ NO ₇ P	1.53	36,181	5,452	15	145	1.12	1.21
LPC 16:1	C ₂₄ H ₄₈ NO ₇ P	1.09	1,185	173	15	-37	1.17	1.34
LPC 17:0	C ₂₅ H ₅₂ NO ₇ P	1.93	714	122	18	-	0.98	1.10
LPC 18:0	C ₂₆ H ₅₄ NO ₇ P	2.39	14,139	1,728	12	15	1.04	1.14
LPC 18:1	C ₂₆ H ₅₂ NO ₇ P	1.7	9,390	1,200	13	24	1.04	1.18
LPC 18:2	C ₂₆ H ₅₀ NO ₇ P	1.24	11,432	1,507	13	31	-	1.28
LPC 18:3	C ₂₆ H ₄₈ NO ₇ P	0.97	228	67	30	-	1.02	1.09
LPC 20:1	C ₂₈ H ₅₆ NO ₇ P	2.55	104	13	12	-	1.04	0.91
LPC 20:2	C ₂₈ H ₅₄ NO ₇ P	1.9	126	24	19	-	1.01	1.02
LPC 20:3	C ₂₈ H ₅₂ NO ₇ P	1.46	982	142	15	-46	1.03	1.08
LPC 20:4	C ₂₈ H ₅₀ NO ₇ P	1.21	3,262	326	10	5	1.00	0.96

LPC 20:5	C ₂₈ H ₄₈ NO ₇ P	0.93	179	50	28	-	0.97	0.97
LPC 22:6	C ₃₀ H ₅₀ NO ₇ P	1.15	437	79	18	-51	1.09	1.06
LPC O-16:0	C ₂₄ H ₅₂ NO ₆ P	1.84	265	77	29	11	1.14	1.27
LPC O-20:0	C ₂₈ H ₆₀ NO ₆ P	3.35	13	3	22	-	1.76	-
LPC P-16:0	C ₂₄ H ₅₀ NO ₆ P	1.78	221	62	27	-74	1.32	1.36
LPE 16:0	C ₂₁ H ₄₄ NO ₇ P	1.57	413	122	29	-74	0.94	1.18
LPE 18:2	C ₂₃ H ₄₄ NO ₇ P	1.27	907	267	30	-76	1.05	1.34
PC 30:0	C ₃₈ H ₇₆ NO ₈ P	4.9	1,130	226	20	-	1.30	0.95
PC 32:0	C ₄₀ H ₈₀ NO ₈ P	5.61	5,285	734	14	-37	1.27	1.53
PC 32:1	C ₄₀ H ₇₈ NO ₈ P	4.99	9,516	1,391	15	-55	1.08	0.96
PC 34:0	C ₄₂ H ₈₄ NO ₈ P	6.35	1,600	282	18	-73	1.01	1.76
PC 34:1	C ₄₂ H ₈₂ NO ₈ P	5.68	91,209	15,962	17	33	1.33	1.14
PC 36:1	C ₄₄ H ₈₆ NO ₈ P	6.42	20,491	3,625	17	-74	1.09	1.40
PC 36:2	C ₄₄ H ₈₄ NO ₈ P	5.85	110,056	19,653	17	-44	1.37	1.08
PC 36:3	C ₄₄ H ₈₂ NO ₈ P	5.31	78,410	10,977	14	-39	1.12	0.96
PC 36:4	C ₄₄ H ₈₀ NO ₈ P	5.02	117,312	21,898	19	-14	0.96	0.91
PC 36:5	C ₄₄ H ₇₈ NO ₈ P	4.59	8,581	1,404	17	-15	0.89	1.49
PC 38:2	C ₄₆ H ₈₈ NO ₈ P	6.52	1,873	163	9	-94	1.09	1.82
PC 38:3	C ₄₆ H ₈₆ NO ₈ P	6.05	21,116	4,223	20	-	0.99	1.27
PC 38:4	C ₄₆ H ₈₄ NO ₈ P	5.74	68,051	11,342	17	-67	0.89	0.58
PC 38:5	C ₄₆ H ₈₂ NO ₈ P	5.08	33,941	6,384	19	-51	-	0.29
PC 38:6	C ₄₆ H ₈₀ NO ₈ P	4.85	33,050	3,547	11	-36	1.17	0.33
PC 40:4	C ₄₈ H ₈₈ NO ₈ P	6.26	2,431	310	13	-92	0.96	1.37
PC 40:6	C ₄₈ H ₈₄ NO ₈ P	5.54	11,678	2,169	19	-82	0.89	0.89
PC 40:7	C ₄₈ H ₈₂ NO ₈ P	4.92	2,912	632	21	-85	1.01	0.52
PC 40:8	C ₄₈ H ₈₀ NO ₈ P	4.43	606	166	28	-97	0.83	1.25
PE 34:1	C ₃₉ H ₇₆ NO ₈ P	5.87	862	122	14	-79	0.99	0.45
PE 34:2	C ₃₉ H ₇₄ NO ₈ P	5.31	1,575	186	12	-82	0.94	0.70
PE 36:1	C ₄₁ H ₈₀ NO ₈ P	6.07	970	194	20	-85	0.90	0.68
PE 36:2	C ₄₁ H ₇₈ NO ₈ P	5.81	4,985	588	12	-77	0.75	0.67
PE 36:3	C ₄₁ H ₇₆ NO ₈ P	5.43	1,781	282	16	-85	1.05	0.72
PE 36:4	C ₄₁ H ₇₄ NO ₈ P	5.2	2,294	289	13	-86	0.84	0.48
PE 38:3	C ₄₃ H ₈₀ NO ₈ P	5.92	732	154	21	-86	0.66	0.60
PE 38:4	C ₄₃ H ₇₈ NO ₈ P	5.43	6,221	922	15	-83	-	0.51
PE 38:5	C ₄₃ H ₇₆ NO ₈ P	5.3	2,068	360	17	-87	1.07	-

PE 38:6	C43H74NO8P	5	2,445	451	19	-84	0.73	0.24
PE 40:6	C45H78NO8P	5.72	1,426	285	20	-82	0.83	0.27
PI 32:1	C41H77O13P	4.3	453	89	19	-25	1.29	1.06
PI 34:1	C43H81O13P	4.93	2,009	352	17	37	2.21	-
PI 34:2	C43H79O13P	4.44	2,338	317	14	4	1.22	0.93
PI 36:1	C45H85O13P	5.61	1,817	510	28	20	1.23	0.69
PI 36:2	C45H83O13P	5.09	6,646	803	12	-	1.22	0.31
PI 36:3	C45H81O13P	4.56	1,894	250	14	84	0.95	0.81
PI 36:4	C45H79O13P	4.36	2,577	412	16	135	0.86	0.82
PI 38:3	C47H85O13P	5.26	3,023	480	16	84	1.03	-
PI 38:4	C47H83O13P	4.99	16,855	1,952	11	74	1.59	0.30
PI 38:5	C47H81O13P	4.48	2,213	389	18	95	0.90	0.75
PI 38:6	C47H79O13P	4.21	283	27	10	-35	0.91	0.84
PI 40:4	C49H87O13P	5.48	275	38	14	-33	1.21	-
PI 40:6	C49H83O13P	4.84	765	146	19	56	1.40	0.21
SM d31:1	C36H73N2O6P	3.98	126	32	25	-	0.88	1.20
SM d32:1	C37H75N2O6P	4.23	5,670	945	17	-8	0.90	2.62
SM d33:1	C38H77N2O6P	4.51	3,238	441	14	-25	0.86	2.54
SM d34:0	C39H81N2O6P	5.16	4,089	917	22	-66	0.77	1.00
SM d34:1	C39H79N2O6P	4.79	70,303	10,545	15	26	1.02	1.01
SM d34:2	C39H77N2O6P	4.25	11,216	1,542	14	2136	0.93	2.46
SM d35:1	C40H81N2O6P	5.22	1,793	416	23	-34	0.94	1.22
SM d36:0	C41H85N2O6P	5.80	1,466	359	24	-51	0.69	0.85
SM d36:1	C41H83N2O6P	5.49	14,622	2,705	18	25	1.02	1.55
SM d36:2	C41H81N2O6P	4.97	6,999	1,094	16	-11	1.10	0.94
SM d37:1	C42H85N2O6P	5.92	745	171	23	12	1.45	2.56
SM d38:1	C43H87N2O6P	6.49	8,350	2,353	27	14	1.01	1.95
SM d38:2	C43H85N2O6P	5.61	3,937	984	25	-21	0.90	0.98
SM d39:1	C44H89N2O6P	6.65	2,783	773	29	-17	0.92	2.17
SM d40:1	C45H91N2O6P	7.01	15,744	4,015	25	33	1.40	1.90
SM d40:2	C45H89N2O6P	6.37	9,422	2,198	24	-24	1.44	2.38
SM d41:1	C46H93N2O6P	7.06	6,169	1,683	27	9	1.22	2.28
SM d41:2	C46H91N2O6P	6.76	4,635	1,119	24	-1	1.13	2.28
SM d42:1	C47H95N2O6P	7.68	16,305	4,402	28	58	0.78	1.28
SM d42:2	C47H93N2O6P	6.96	35,782	8,945	25	33	1.05	1.35

SM d42:3	C ₄₇ H ₉₁ N ₂ O ₆ P	6.40	13,791	3,813	27	-18	1.26	1.78
SM d43:2	C ₄₈ H ₉₅ N ₂ O ₆ P	7.19	827	240	29	-40	1.30	-
TAG 48:0	C ₅₁ H ₉₈ O ₆	10.97	3,633	969	26	-	0.63	-
TAG 48:1	C ₅₁ H ₉₆ O ₆	10.54	10,469	2,577	24	-51	0.95	-
TAG 48:2	C ₅₁ H ₉₄ O ₆	10.13	12,853	2,249	18	-22	1.72	-
TAG 48:4	C ₅₁ H ₉₀ O ₆	9.27	1,039	184	18	-	2.65	-
TAG 49:1	C ₅₂ H ₉₈ O ₆	10.75	1,639	344	21	-	2.59	-
TAG 50:0	C ₅₃ H ₁₀₂ O ₆	11.3	3,174	693	22	-67	0.37	-
TAG 50:1	C ₅₃ H ₁₀₀ O ₆	10.97	31,668	8,334	26	-40	1.06	-
TAG 50:2	C ₅₃ H ₉₈ O ₆	10.59	39,073	9,976	26	-42	0.94	-
TAG 50:3	C ₅₃ H ₉₆ O ₆	10.17	19,074	5,474	29	-60	2.45	-
TAG 51:1	C ₅₄ H ₁₀₂ O ₆	11.14	1,525	407	27	-	1.44	-
TAG 51:2	C ₅₄ H ₁₀₀ O ₆	10.79	4,058	930	22	-	2.00	-
TAG 52:1	C ₅₅ H ₁₀₄ O ₆	11.31	12,060	2,498	20	-52	1.44	-
TAG 52:3	C ₅₅ H ₁₀₀ O ₆	10.65	85,738	24,864	28	-52	1.28	-
TAG 53:2	C ₅₆ H ₁₀₄ O ₆	11.17	1,660	358	21	-	2.37	-
TAG 53:3	C ₅₆ H ₁₀₂ O ₆	10.81	3,224	959	29	-	1.40	-
TAG 54:7	C ₅₇ H ₉₆ O ₆	9.53	4,913	1,316	26	-	1.02	-
TAG 56:3	C ₅₉ H ₁₀₈ O ₆	11.31	1,279	128	10	-	1.36	-
TAG 58:9	C ₆₁ H ₁₀₀ O ₆	9.95	1,115	251	22	-	2.28	-

ⁱResponse factors are calculated based on consensus values from at least 5 laboratories for lipids with a COD <30 %. Response factors for FAs were determined based on arachidonic acid(d8) as corresponding surrogate calibrant. *Response factors for CEs are (partially) extremely elevated and assume an erroneous determination, most probably due to not working in the correct linear range (CE levels in NIST SRM 1950 exceeded the linear range of the surrogate calibrant). Other factors could be stability issues in the plasma or species-dependent in-source fragmentation. In addition, also the deviation of results obtained from previous analyses by the LIPID MAPS consortium²² to consensus values is listed. SWATH-MS/MS response factors were not evaluated but can be potentially extracted for lipids of special interest and when interferences are observed in TOF-MS. Lipid nomenclature is based on the system proposed by LIPID MAPS.

Table S-10. Inter-study comparison of lipid concentrations.¹

	Results				Results of Barber et al. [44]				Results of Eisinger et al. [45]			
	male C57Bl/6N (<i>adiponectin-CreERT2^{9/+}, BK^{-f/L2} / adiponectin-CreERT2^{9/+}, BK^{-f/L2}</i>)				male C57Bl/6J				male C57Bl/6			
	plasma (EDTA tubes) RP-LC ESI-MS(/MS) (quadrupole time-of-flight, QTOF 5600+, Sciex)				plasma (EDTA tubes) RP-LC ESI+-MS/MS (triple quadrupole, QTRAP 4000, Sciex)				serum Direct flow injection ESI+-MS/MS (triple quadrupole, Quattro Ultima, Micromass)			
Age	10 weeks				8 weeks				14 weeks			
Feeding duration	18 weeks				12 weeks				14 weeks			
Sampling Quantification												
Analyte	CD		HFD		CD		HFD		CD		HFD	
	Mean [ng mL ⁻¹]	CV [%]	Mean [ng mL ⁻¹]	CV [%]	Mean [ng mL ⁻¹]	CV [%]	Mean [ng mL ⁻¹]	CV [%]	Mean [ng mL ⁻¹]	CV [%]	Mean [ng mL ⁻¹]	CV [%]
CE 16:0	1.1E+04*	47.3	1.8E+04*	30.6	-	-	-	-	3.4E+04	11.5	3.5E+04	19.4
CE 16:1	6.3E+05*	36.5	2.1E+05*	34.3	-	-	-	-	1.1E+05	10.9	8.3E+04	24.1
CE 18:1	4.8E+05*	35.4	2.5E+05*	31.6	-	-	-	-	8.0E+04	12.5	1.1E+05	22.7
CE 18:2	4.2E+06*	40.5	7.5E+06*	33.3	-	-	-	-	4.5E+05	5.8	4.6E+05	23.9
CE 18:3	5.0E+05*	52.0	5.2E+05*	40.4	-	-	-	-	4.5E+04	6.7	3.2E+04	24.1
CE 20:4	1.9E+07*	47.4	2.8E+07*	33.9	-	-	-	-	1.3E+06	8.5	2.1E+06	26.7
CE 20:5	8.9E+05*	51.7	1.6E+06*	53.8	-	-	-	-	1.6E+04	13.1	1.9E+04	23.7
CE 22:6	4.9E+06*	61.2	8.0E+06*	40.0	-	-	-	-	1.8E+05	7.8	2.3E+05	30.0
Chol-esterol	379,004	27.0	439,293	29.8	-	-	-	-	117,620	10.1	155,396	22.9
LPC 14:0	890	11.7	1,038	20.1	200	6.0	133	7.5	-	-	-	-
LPC 15:0	592	12.2	981	13.4	214	4.2	140	4.3	231	10.4	265	18.1
LPC 16:0	102,602	12.6	141,677	12.4	17,731	3.5	15,242	2.8	34,615	5.9	29,748	21.7
LPC 16:1	14,858	12.0	4,462	13.4	2,293	9.3	1,046	3.0	4,620	7.2	2,720	17.1
LPC 18:0	45,817	19.5	111,144	11.3	9,166	2.2	9,787	2.0	14,014	9.7	20,607	20.7
LPC 18:1	99,131	12.8	78,874	9.0	13,292	5.2	10,161	2.6	43,158	6.8	43,919	19.3
LPC 18:3	1,145	16.9	1,504	14.5	-	-	-	-	932	6.1	466	17.8
LPC 20:0	1,066	14.9	2,119	20.0	416	11.3	310	9.0	1,087	11.9	642	21.0
LPC 20:1	3,036	14.1	1,130	24.1	722	4.4	384	10.4	-	-	-	-
LPC 20:2	2,155	22.3	1,742	18.5	347	11.8	344	13.1	-	-	-	-
LPC 20:3	8,326	18.3	10,535	27.2	3,263	11.7	2,748	14.0	1,206	19.9	2,374	15.9
LPC 20:4	14,458	24.7	28,185	19.2	10,865	3.2	12,334	2.1	11,499	12.3	14,135	10.9

LPC 20:5	662	11.3	1,361	21.1	806	18.2	358	10.3	260	18.8	309	13.9
LPC 22:6	7,268	15.8	19,624	18.2	7,582	5.2	7,820	4.8	3,775	8.7	3,934	17.6
PC 30:0	755	26.6	2,572	33.4	165	9.7	195	6.7	-	-	-	-
PC 32:0	12,184	28.6	33,037	22.2	1,461	6.2	1,584	3.8	9,535	10.9	9,601	20.5
PC 32:1	18,529	20.2	9,349	22.0	1,708	2.5	1,679	2.1	7,928	9.4	6,669	25.8
PC 34:0	5,040	34.5	23,601	22.8	1,597	0.6	1,588	0.9	1,524	55.0	2,492	48.9
PC 34:1	315,816	9.7	324,981	10.1	1,776	1.4	1,704	4.7	138,600	4.2	185,694	30.1
PC 36:1	154,676	17.3	224,512	17.1	2,031	2.1	2,099	6.7	16,157	4.6	29,247	17.3
PC 36:2	212,668	16.5	492,805	12.7	1,681	4.2	1,549	3.2	59,886	9.7	60,625	12.3
PC 36:3	65,030	20.8	172,291	26.0	1,684	4.6	2,166	5.0	31,560	5.6	28,502	16.2
PC 36:4	64,824	17.4	146,885	13.8	2,771	3.8	3,147	4.0	65,632	8.5	81,352	16.0
PC 36:5	11,531	17.5	32,129	21.8	2,302	1.4	2,377	7.7	3,151	21.5	2,980	14.4
PC 38:2	28,776	25.7	33,310	17.9	805	45.0	658	9.9	4,274	11.2	4,934	34.8
PC 38:3	40,159	20.3	84,969	25.5	1,714	7.9	1,686	11.9	4,962	18.5	11,281	19.4
PC 38:4	15,994	24.4	68,497	16.8	1,372	15.8	1,295	6.6	28,484	13.1	58,435	11.2
PC 38:5	6,338	22.7	5,579	19.1	3,607	3.0	1,634	25.2	16,065	10.5	22,142	17.6
PC 38:6	16,745	12.4	41,020	16.6	1,209	2.2	2,246	18.6	34,332	7.0	36,242	22.7
PC 40:4	2,457	21.9	2,780	20.0	-	-	-	-	3,017	10.0	3,956	34.5
PC 40:6	11,998	13.3	59,842	18.2	2,518	4.0	2,849	2.5	11,561	7.4	16,208	17.1
PC 40:7	5,351	18.6	4,649	23.9	1,833	6.0	2,003	2.8	8,047	13.0	7,972	17.4
PE 34:1	529	27.8	989	25.0	160	13.1	119	10.1	-	-	-	-
PE 34:2	1,829	28.2	3,553	26.9	380	16.3	282	10.6	1,224	16.3	537	22.7
PE 36:1	527	43.8	1,148	36.7	334	14.4	492	12.6	425	35.1	574	23.3
PE 36:2	2,835	40.2	9,170	33.9	1,051	16.7	1,558	14.1	1,510	13.3	1,317	20.9
PE 36:3	2,494	26.4	4,839	31.1	379	12.1	342	13.5	1,113	23.4	497	17.9
PE 36:4	3,678	21.2	3,840	24.2	560	19.5	554	6.3	1,835	19.0	1,791	12.4
PE 38:3	332	62.7	1,297	45.4	322	12.7	447	5.1	-	-	-	-
PE 38:5	5,847	18.3	5,810	18.4	736	9.9	736	3.8	6,488	27.4	7,147	17.7
PE 38:6	2,798	19.3	4,892	22.5	1,167	12.9	1,442	17.9	10,307	28.3	9,367	27.7
PE 40:6	534	21.3	2,260	20.2	431	9.7	679	19.6	2,772	31.4	2,812	22.5
PI 34:2	3,145	23.7	7,860	15.9	-	-	-	-	568	8.8	301	16.6
PI 36:1	840	23.2	3,408	33.0	-	-	-	-	216	4.2	303	8.6
PI 36:2	1,106	34.5	3,133	27.7	-	-	-	-	1,088	11.1	665	9.0
PI 36:4	7,042	15.3	12,250	9.9	-	-	-	-	2,045	8.4	1,546	19.5
PI 38:4	7,698	10.8	13,328	16.0	-	-	-	-	15,480	5.8	22,320	24.8

PI 38:5	7,956	19.0	7,771	19.2	-	-	-	-	-	1,797	12.3	1,806	16.2
PI 38:6	601	3.5	2,067	9.8	-	-	-	-	-	185	9.7	159	16.4
PI 40:6	212	10.8	602	20.3	-	-	-	-	-	246	11.0	219	12.3
SM 32:1	1,489	11.7	4,762	24.0	4.3E+04	7.9	1.1E+05	7.1	-	310	54.5	459	20.5
SM 33:1	2,235	6.9	6,123	22.4	1.8E+05	8.3	2.6E+05	6.9	-	441	21.8	537	30.7
SM 34:1	31,478	11.7	67,364	14.1	1.2E+06	4.9	1.2E+06	3.2	-	5,554	8.5	8,296	18.9
SM 34:2	6,826	19.5	11,799	18.1	5.1E+05	4.7	5.5E+05	8.2	-	876	8.0	1,037	12.8
SM 36:1	2,368	21.4	6,542	19.2	2.3E+05	11.7	4.6E+05	11.5	-	738	23.7	1,382	23.3
SM 36:2	1,633	22.8	3,920	24.9	1.2E+05	10.8	2.3E+05	14.3	-	313	53.7	634	28.7
SM 38:1	1,567	22.8	8,847	25.2	1.4E+06	5.0	1.4E+06	7.1	-	-	-	-	-
SM 38:2	736	22.8	1,746	28.1	8.1E+05	13.6	6.8E+05	6.6	-	-	-	-	-
SM 40:1	13,939	27.7	77,085	25.5	2.2E+06	5.5	2.0E+06	6.0	-	1,976	40.6	2,889	26.4
SM 40:2	18,863	20.8	27,314	31.4	1.9E+06	11.1	1.8E+06	6.7	-	1,044	58.6	1,782	22.4
SM 41:1	3,119	22.6	7,561	22.9	-	-	-	-	-	873	25.7	969	34.8
SM 41:2	3,757	14.7	5,650	29.9	-	-	-	-	-	615	31.2	679	16.5
SM 42:1	3,529	24.7	6,984	24.3	5.1E+05	8.4	4.2E+05	11.4	-	1,304	10.0	1,353	33.1
SM 42:2	26,952	20.4	31,432	29.0	1.9E+06	5.3	1.3E+06	3.9	-	5,676	5.7	6,595	20.5
SM 42:3	9,894	15.2	9,983	27.6	1.9E+06	4.5	2.0E+06	3.5	-	1,744	21.8	1,330	31.7
SM 43:2	319	31.3	404	35.6	3.4E+04	6.2	3.3E+04	12.4	-	-	-	-	-
TAG 48:0	899	39.0	1,103	49.9	317	31.9	1,108	23.6	-	-	-	-	-
TAG 48:1	5,620	38.7	9,726	29.0	10,631	54.2	13,322	24.7	-	-	-	-	-
TAG 48:2	89,292	34.9	101,268	31.9	1,679	42.0	3,686	33.8	-	-	-	-	-
TAG 50:0	687	40.2	880	51.8	1,336	22.8	3,819	18.3	-	-	-	-	-
TAG 50:1	39,955	32.5	59,978	20.5	11,673	31.6	34,722	26.9	-	-	-	-	-
TAG 50:2	107,165	26.6	102,243	19.3	18,664	26.9	29,900	20.9	-	-	-	-	-
TAG 50:3	201,489	30.4	193,182	22.5	5,581	28.2	9,413	28.4	-	-	-	-	-
TAG 52:1	17,677	49.6	47,656	33.7	11,753	21.3	52,535	23.5	-	-	-	-	-
TAG 52:3	161,955	26.4	199,885	18.5	34,203	16.1	75,350	16.4	-	-	-	-	-

Cholesterol is determined as free cholesterol. SD: standard deviation. Following modes were used for quantification: positive mode TOF-MS; CEs, LPCs, PEs, SMs; positive mode SWATH-MS/MS; cholesterol, TAGs; negative mode TOF-MS; PCs, negative mode SWATH-MS/MS; Pls. For values marked in red, the concentration exceeded the linear range of the surrogate calibrant (prior to response factor adjustment) and values are prone to overestimation. Although linear range can be possibly higher for some lipid classes, validation of higher linear ranges is aggravated as the spiking volume of the Lipidmix is limited. If the linear range is exceeded for target analytes, alternative modes (different polarity or SWATH-MS/MS) can be selected for quantification. *Due to erroneous determination of response factors (see Table S-9), concentrations of CEs are given without response factor adjustment (response factor = 1).

Table S-11. Odd chain lipid species detected and identified in mouse plasma.^k

FA 17:1	PC 17:2-17:2	SM d15:0/18:2	TAG 15:1-16:0-18:2
PC 13:0-13:0	PC 17:2-17:2	SM d15:0/28:2	TAG 15:1-17:1-19:1
PC 14:0e/3:0	PC 17:2-20:4	SM d15:2/20:2	TAG 16:0-17:0-18:1
PC 14:1e/3:0	PC 18:3e/15:0	SM d15:2/26:2	TAG 16:0-17:1-18:1
PC 14:1e/5:0	PC 18:5e/17:0	SM d15:3/28:2	TAG 16:0-17:2-20:0
PC 15:0-15:0	PC 19:0-18:1	SM d17:0/18:2	TAG 16:1-16:1-19:1
PC 15:0-16:0	PC 19:0-18:2	SM d17:0/22:2	TAG 16:1-17:0-18:2
PC 15:0-18:1	PC 19:0-18:2	SM d18:1/23:0	TAG 16:1-18:1-21:1
PC 15:0-18:2	PC 19:0-20:3	SM d19:0/22:2	TAG 17:0-17:1-19:1
PC 15:0-18:2	PC 19:0-20:4	SM d19:2/23:0	TAG 17:0-17:1-19:2
PC 15:0-20:4	PC 19:0-20:4	TAG 12:0-14:0-17:0	TAG 17:0-17:1-19:3
PC 15:0-22:6	PC 20:0-19:2	TAG 12:0-15:0-18:1	TAG 17:0-18:1-18:2
PC 15:1-15:1	PC 21:1-21:1	TAG 12:0-16:1-17:1	TAG 17:0-18:2-18:2
PC 16:0-17:1	PC 21:2-20:4	TAG 13:0-13:0-14:1	TAG 17:1-18:0-18:1
PC 16:2e/3:0	PE 22:4e/18:1	TAG 13:0-13:0-18:3	TAG 18:1-18:1-19:1
PC 17:0-18:0	SM d14:0/19:1	TAG 14:0-15:0-16:0	LPC 15:0
PC 17:0-18:1	SM d14:0/19:1	TAG 14:0-15:0-18:1	LPC 15:0e
PC 17:0-18:1	SM d14:0/21:1	TAG 14:0-16:0-17:0	LPC 17:0
PC 17:0-18:2	SM d14:0/22:1	TAG 14:1-20:1-21:1	LPC 17:2
PC 17:0-20:3	SM d14:0/22:2	TAG 15:0-15:0-21:4	LPC 19:0
PC 17:0-20:4	SM d14:1/25:0	TAG 15:0-15:0-21:4	LPC 19:1
PC 17:0-20:4	SM d14:1/25:0	TAG 15:0-15:1-21:1	LPC 21:0
PC 17:0-22:6	SM d14:1/27:0	TAG 15:0-16:0-18:1	LPC 23:0
PC 17:1-18:1	SM d14:1/29:0	TAG 15:0-16:0-20:2	LPC 25:0
PC 17:1-18:2	SM d14:2/22:1	TAG 15:0-16:1-18:1	LPE 17:0
PC 17:1-20:4	SM d15:0/16:1	TAG 15:1-16:0-16:0	

^kAligned results for the 6 “blank” plasma samples, which were analysed to validate selectivity, are shown (positive and negative mode are merged).

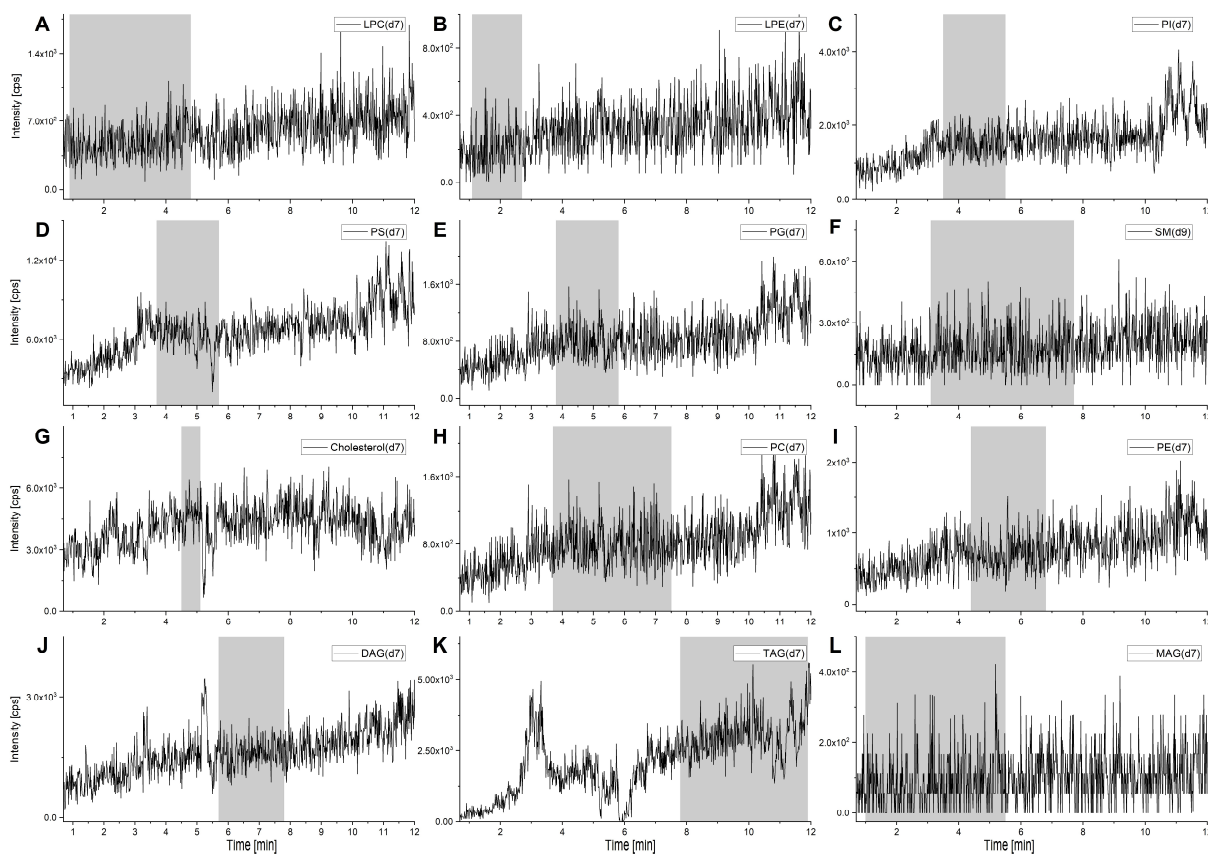


Figure S-1. Matrix effect evaluation by post-column infusion – positive mode SWATH-MS. SWATH-MS/MS mass traces of surrogate calibrants in positive mode during analysis of blank plasma samples and simultaneous post-column infusion of surrogate calibrants via a T-piece. Shaded areas indicate t_R intervals in which species of corresponding lipid classes were detected and identified. Ion suppression can be observed for PS(d7) and TAG(d7). As TAGs are one of the most abundant lipids in plasma, this effect can be explained by co-eluting TAGs that concur for ionization. A: LPC(d7), B: LPE(d7), C: PI(d7), D: PS(d7), E: PG(d7), F: SM(d9), G: Cholesterol(d7), H: PC(d7), I: PE(d7), J: DAG(d7), K: TAG(d7), L: MAG(d7).

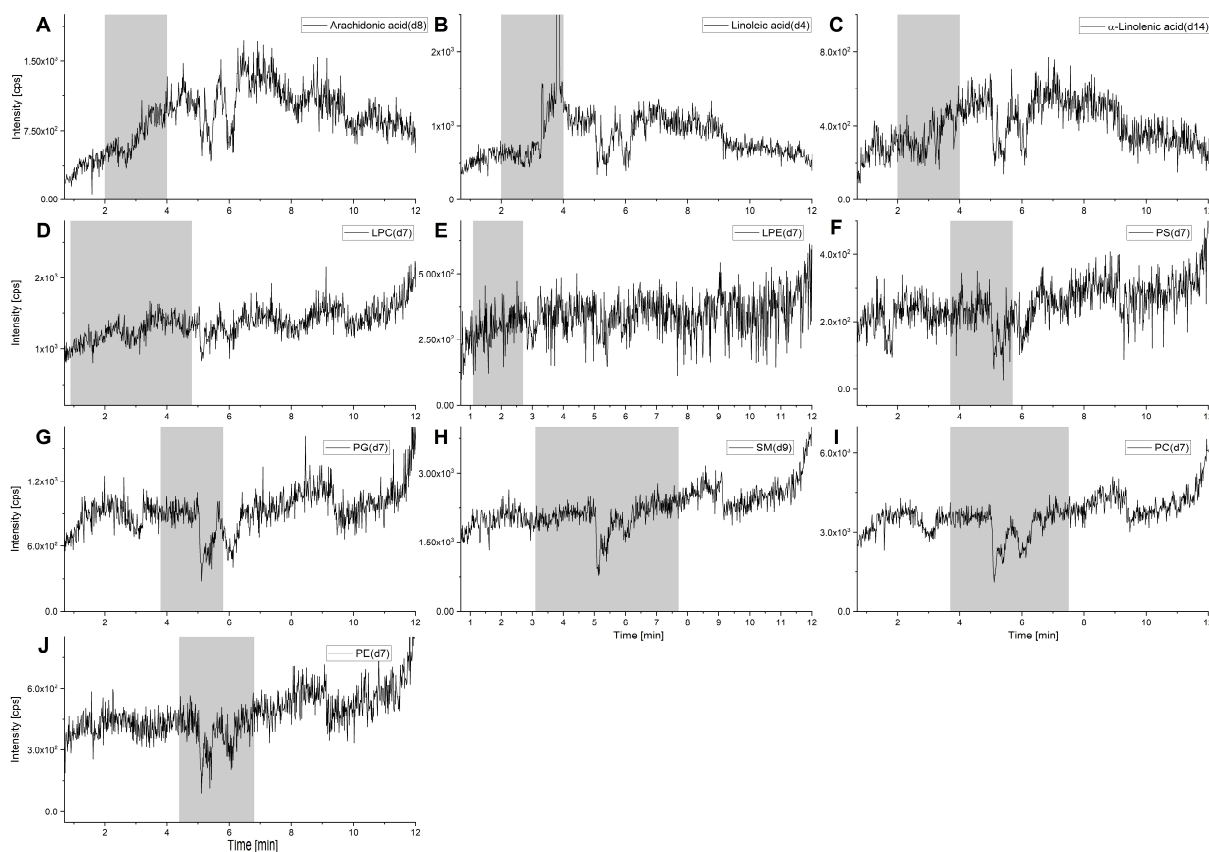


Figure S-2. Matrix effect evaluation by post-column infusion – negative mode TOF-MS. TOF-MS mass traces of surrogate calibrants in negative mode during analysis of blank plasma samples and simultaneous post-column infusion of surrogate calibrants via a T-piece. Shaded areas indicate t_R intervals in which species of corresponding lipid classes were detected and identified. Signals for MAG(d7) and DAG(d7) were too low for assessment of matrix effects. Significant ion suppression can be spotted at the t_R interval from 5 – 6.5 minutes. This effect is most probably derived from a concurring ionization effect of eluting PCs, which are highly abundant in plasma. A: Arachidonic acid(d8), B: Linoleic acid(d4), C: α -Linolenic acid(d14), D: LPC(d7), E: LPE(d7), F: PS(d7), G: PG(d7), H: SM(d9), I: PC(d7), J: PE(d7).

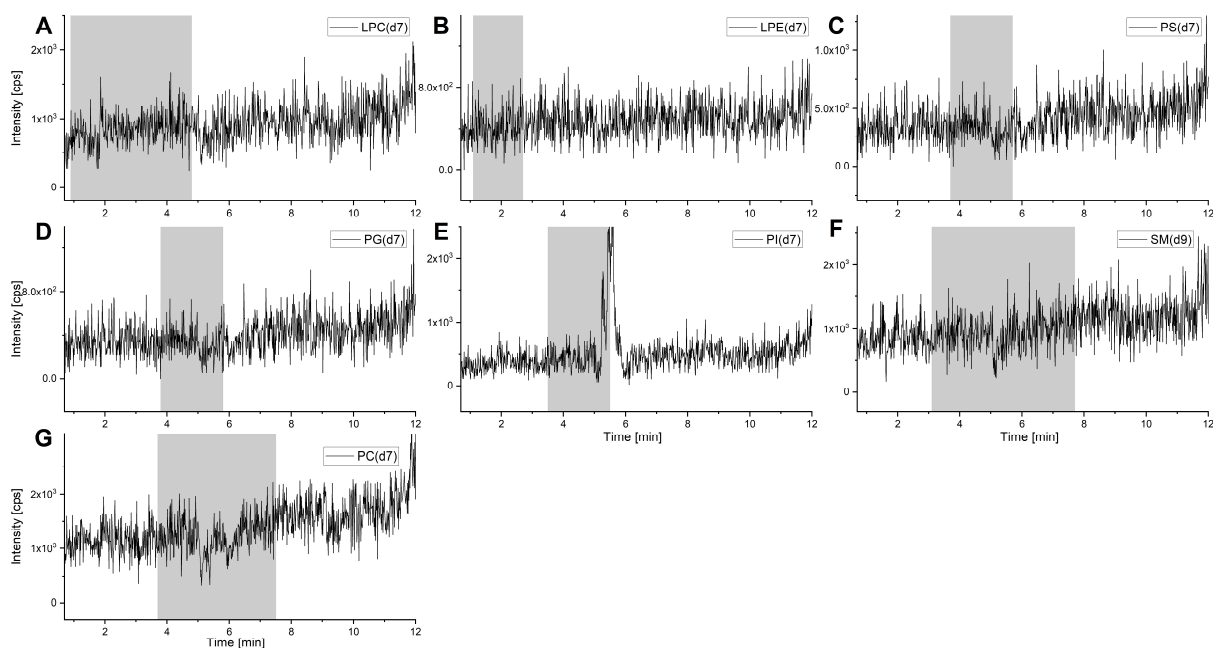


Figure S-3. Matrix effect evaluation by post-column infusion – negative mode SWATH-MS. SWATH-MS/MS mass traces of surrogate calibrants in negative mode during analysis of blank plasma samples and simultaneous post-column infusion of surrogate calibrants via a T-piece. Shaded areas indicate t_R intervals in which species of corresponding lipid classes were detected and identified. Signals for PE(d7) and arachidonic acid(d8) were too low for assessment of matrix effects. For PI(d7) (panel E), the increase of the XIC trace is not derived from an ion enhancement effect but rather from a signal-interference of a closely eluting compound. A: LPC(d7), B: LPE(d7), C: PS(d7), D: PG(d7), E: PI(d7), F: SM(d9), G: PC(d7).

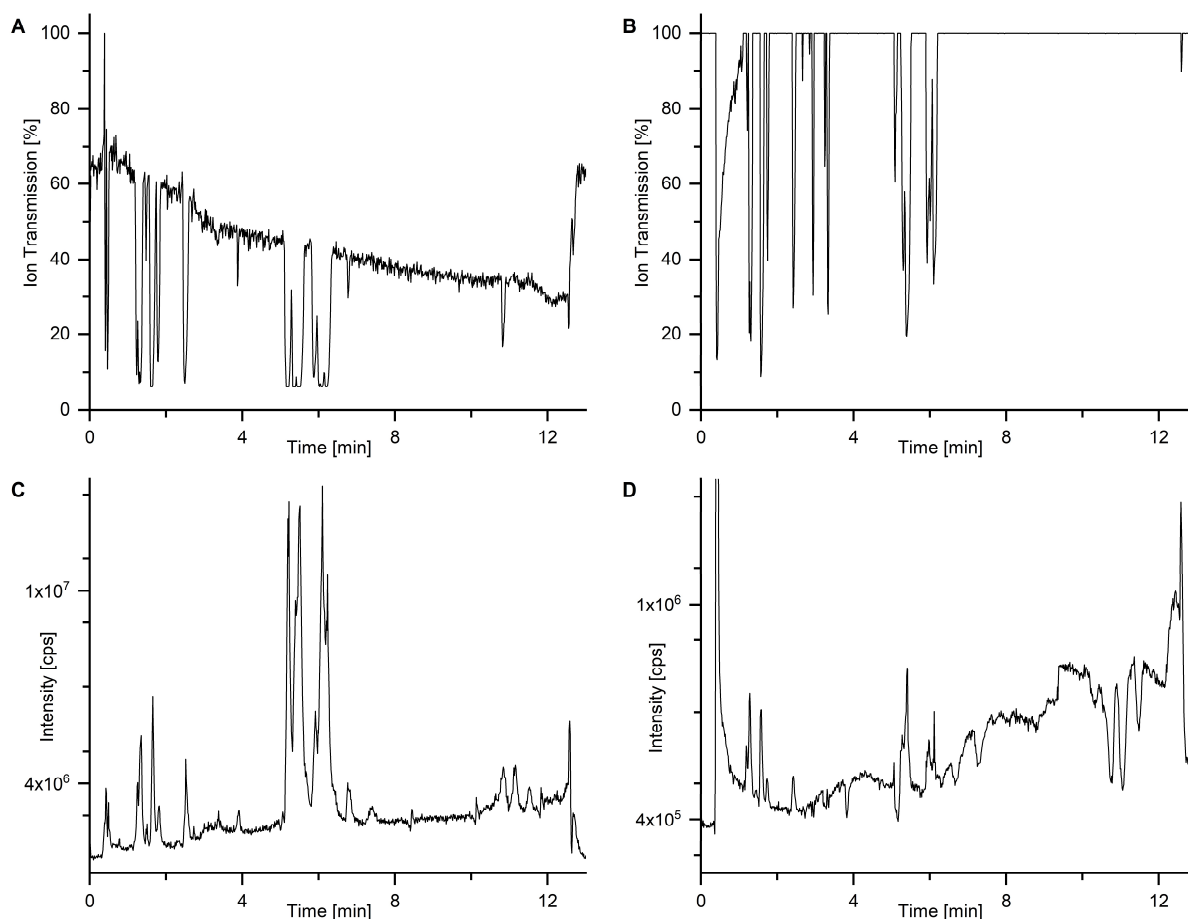


Figure S-4. TOF-MS ion transmission and TICs of a QC_{sys} sample. A: Ion transmission control in positive mode TOF-MS, B: Ion transmission control in negative mode TOF-MS, C: positive mode TOF-MS TIC, D: negative mode TOF-MS TIC. TIC intensity is about 10x higher in positive mode TOF-MS (panel C) than in negative mode TOF-MS (panel D). To avoid detector saturation, ion transmission never reaches 100 % for positive mode TOF-MS and most severe short-term ion transmission regulation is about 40 – 50 % (panel A). In negative mode TOF-MS, ion transmission control is only applied during elution intervals of highly abundant lipids (panel B). Due to the overall 10x lower ion load, ion transmission is often at 100 %, but regulation with rapid changes of 80 – 90 % in ion transmission occur during elution. For SWATH-MS/MS experiments ITC is permanently set to 100 %.

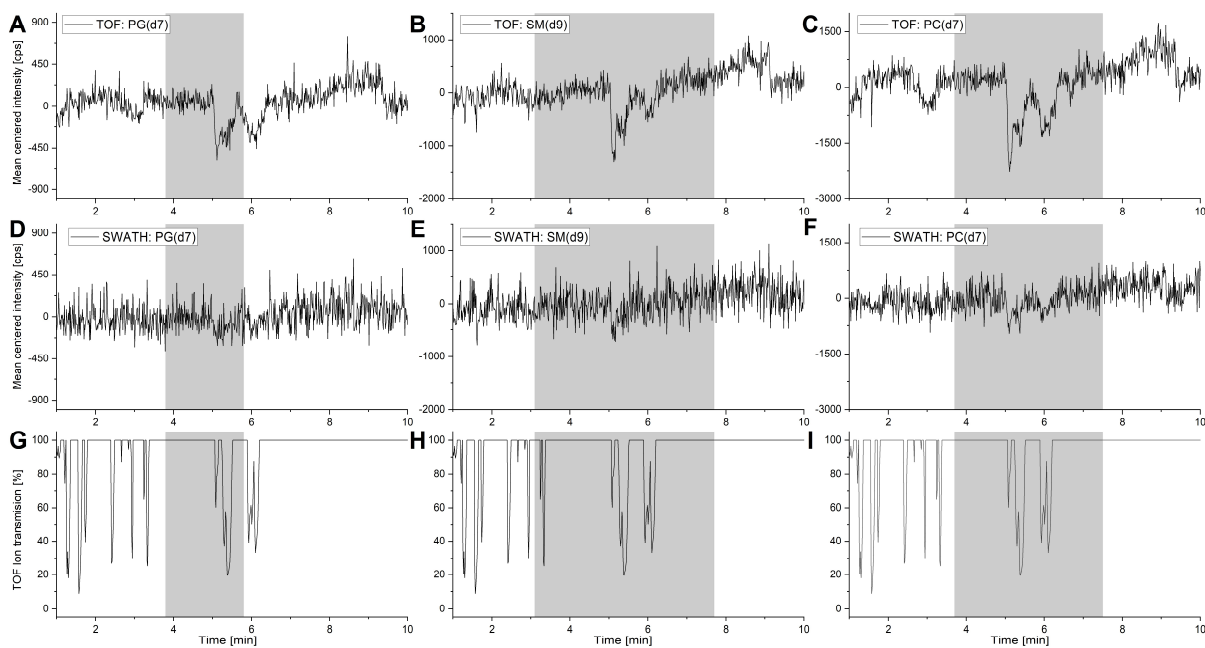


Figure S-5. Comparison of matrix effects for negative TOF-MS and SWATH-MS/MS (post-column infusion). For better comparability of matrix effects on differing MS-levels, exemplary negative mode data of PG(d7), SM(d9) and PC(d7) is shown. Due to differences in absolute signal intensities between TOF-MS and SWATH-MS/MS EICs data were mean centered. A: TOF EIC of PG(d7), B: TOF EIC of SM(d9), C: TOF EIC of PC(d7), D: SWATH EIC of PG(d7), E: SWATH EIC of SM(d9), F: SWATH EIC of PC(d7), G-I: TOF ITC progression (negative mode).

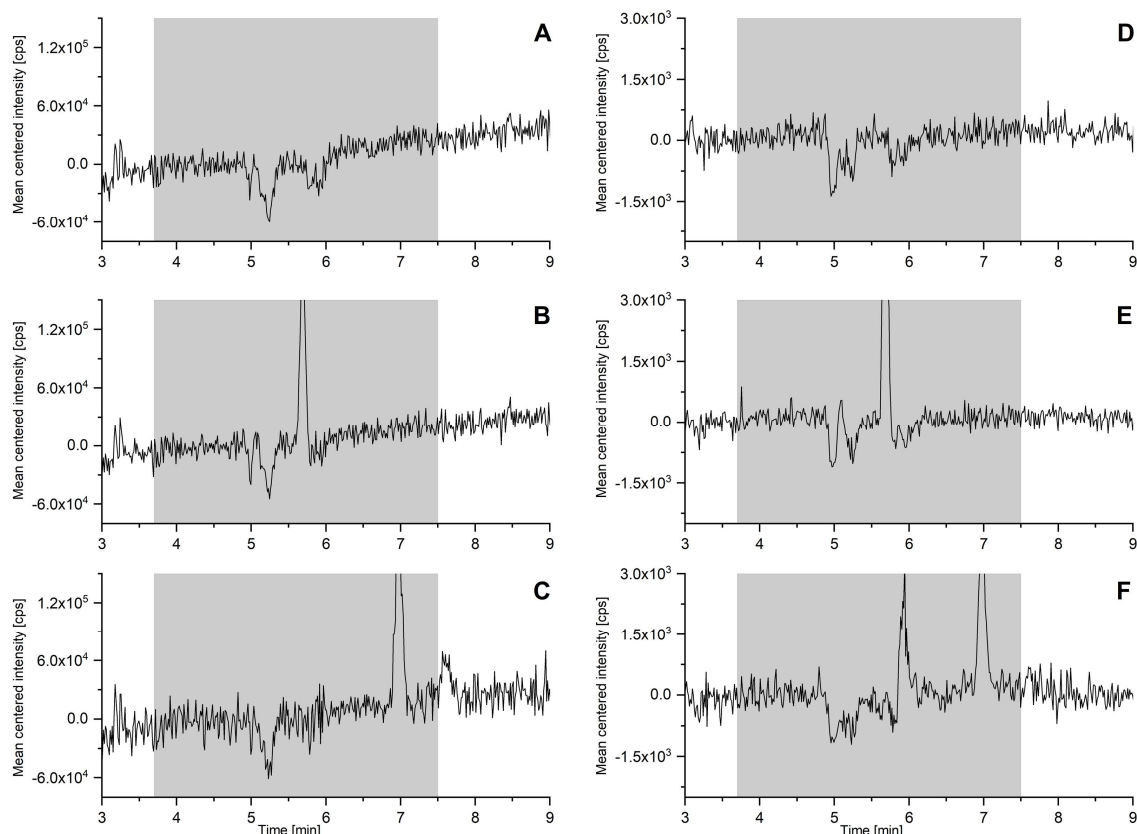


Figure S-6. Comparison of matrix effects for different PC standards via post-column infusion.^f ^fEICs of mass traces for A, 14:0-14:0 PC - positive mode; B, 16:0-16:0 PC - positive mode; C, 18:0-18:0 PC - positive mode; D, 14:0-14:0 PC - negative mode; E, 16:0-16:0 PC - negative mode; F, 18:0-18:0 PC - negative mode; are shown. Due to the relatively high abundance of 16:0-16:0 PC and 18:0-18:0 PC in mouse plasma (see Table 3), peaks are visible in the respective mass traces during their corresponding retention time. For 14:0-14:0 PC the plasma levels are too low and no peak is spotted in post-column equilibration. All traces were mean-centered for best comparability. The respective matrix effect profiles are highly similar for each PC species.

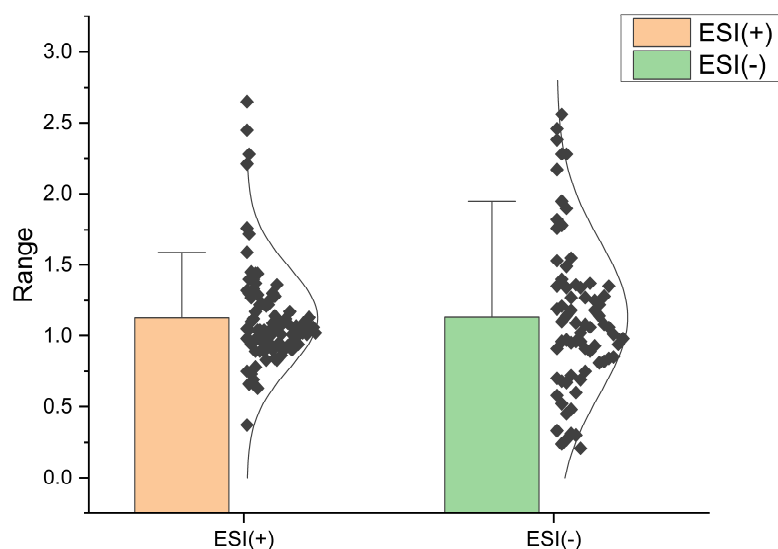


Figure S-7. Response factor distribution for the evaluation via NIST SRM 1950.

Supplementary Text S-1.

In 2006, NIST developed a human plasma reference material (SRM 1950) derived from 100 individuals to represent the average US population (for further information see <https://srm1950.nist.gov>). Lipidomic analysis of this material, including quantification of absolute levels for several lipid species, was executed by a consortium organized by LIPID MAPS.[see reference 24 in the main manuscript] Later, a harmonizing initiative by Bowden et al.[see reference 23 in the main manuscript] established consensus values for SRM 1950, by summarizing lipid concentrations quantified by a minimum of 5 independent laboratories, with most of them utilizing targeted (triple-quadrupole) mass-spectrometric methods. Consensus values were reported as medians of means (MEDM) from inter-laboratory results, together with a standard uncertainty (u), calculated via median absolute deviations (for details see Bowden et al.[see reference 2 in the main manuscript]). Analogous to CVs, these two metrics were used for calculation of coefficients of dispersion (COD) to express inter-laboratory variation of results. NIST also lists certified lipid concentrations in the certificate of analysis, which, however, could not be used since they represent values for total cholesterol and total fatty acids.

6. Appendix

6.1. Supplementary Accepted Manuscripts

6.1.1. Insulin and Estrogen Independently and Differentially Reduce Macronutrient Intake in Healthy Men

*Rosemarie Krug[§], Linda Mohwinkel[‡], Bernhard Drotleff[¶], Jan Born^{§,‡,‡},
Manfred Hallschmid^{§,‡,‡}*

[§]Department of Medical Psychology and Behavioral Neurobiology, University of Tübingen,
Tübingen, Germany

[‡]Department of Neuroendocrinology, University of Lübeck, Lübeck, Germany

[¶]Institute of Pharmaceutical Sciences, Pharmaceutical (Bio-)Analysis, University of Tübingen,
Tübingen, Germany

[‡]German Center for Diabetes Research (DZD), Munich-Neuherberg, Germany

[#]Institute for Diabetes Research and Metabolic Diseases of the Helmholtz Center Munich at
the University of Tübingen (IDM), Tübingen, Germany

**Reprinted with permission from The Journal of Clinical Endocrinology & Metabolism,
Volume 103, Pages 1393-1401,
DOI: 10.1210/jc.2017-01835**

Copyright (2018) Endocrine Society

Insulin and Estrogen Independently and Differentially Reduce Macronutrient Intake in Healthy Men

Rosemarie Krug,¹ Linda Mohwinkel,² Bernhard Drotleff,³ Jan Born,^{1,4,5} and Manfred Hallschmid^{1,4,5}

¹Department of Medical Psychology and Behavioral Neurobiology, University of Tübingen, 72076 Tübingen, Germany; ²Department of Neuroendocrinology, University of Lübeck, 23538 Lübeck, Germany; ³Institute of Pharmaceutical Sciences, University of Tübingen, 72076 Tübingen, Germany; ⁴German Center for Diabetes Research (DZD), 85764 München-Neuherberg, Germany; and ⁵Institute for Diabetes Research and Metabolic Diseases of the Helmholtz Center Munich at the University of Tübingen (IDM), 72076 Tübingen, Germany

Context: Insulin administration to the central nervous system inhibits food intake, but this effect has been found to be less pronounced in female compared with male organisms. This sex-specific pattern has been suggested to arise from a modulating influence of estrogen signaling on the insulin effect.

Objective: We assessed in healthy young men whether pretreatment with transdermal estradiol interacts with the hypophagic effect of central nervous insulin administration via the intranasal pathway.

Design, Setting, Participants, and Intervention: According to a 2×2 design, two groups of men (n = 16 in each group) received a 3-day transdermal estradiol (100 µg/24 h) or placebo pretreatment and on two separate mornings were intranasally administered 160 IU regular human insulin or placebo.

Main Outcome Measures: We assessed free-choice *ad libitum* calorie intake from a rich breakfast buffet and relevant blood parameters in samples collected before and after breakfast.

Results: Estrogen treatment induced a 3.5-fold increase in serum estradiol concentrations and suppressed serum testosterone concentrations by 70%. Independent of estradiol administration, intranasal insulin reduced the intake of carbohydrates during breakfast, attenuating in particular the consumption of sweet, palatable foods. Estradiol treatment *per se* decreased protein consumption. We did not find indicators of eating-related interactions between both hormones.

Conclusions: Results indicate that, in an acute setting, estrogen does not interact with central nervous insulin signaling in the control of eating behavior in healthy men. Insulin and estradiol rather exert independent inhibiting effects on macronutrient intake. (*J Clin Endocrinol Metab* 103: 1393–1401, 2018)

The pancreatic hormone insulin, in addition to its peripheral effects, modulates central nervous functions, including the control of energy metabolism (1, 2). The direct application of insulin to the brain via intracerebroventricular infusion in animals and via intranasal administration in humans (3) has been shown to decrease food intake and body weight in mice (4), rats (5), baboons (6), and men (7, 8). The hypophagic effect of central

nervous insulin appears to display a preponderance in male compared with female organisms (7–9). In humans, intranasal insulin administration acutely curbs food intake in men but not in women (7) and during long-term treatment reduces body fat content in male but not age-matched female subjects (8). Similarly, in contrast to male rats, intact female rats do not reduce their food intake upon intracerebroventricular insulin treatment (9).

ISSN Print 0021-972X ISSN Online 1945-7197

Printed in USA

Copyright © 2018 Endocrine Society

Received 17 August 2017. Accepted 9 January 2018.

First Published Online 12 January 2018

Abbreviations: ANOVA, analysis of variance; FSH, follicle-stimulating hormone; LH, luteinizing hormone; SEM, standard error of the mean.

Sensitivity to the anorexigenic effect of intracerebroventricular insulin, however, can be induced in female animals by ovariectomy associated with a reduction in plasma estradiol concentrations; *vice versa*, estrogen-treated male rats are no longer susceptible to the hypophagic effect of the hormone (10), suggesting that sex-related differences in estrogen signaling modulate the impact of central nervous insulin on eating behavior. Nevertheless, although in women intranasal insulin does not affect food intake in the fasted state, it inhibits snack intake when administered postprandially (11), indicating that the peptide can decrease calorie consumption also in female subjects.

Against this background and considering that estrogen has been reported to attenuate energy intake and body weight in animals (12), we investigated whether estrogen and insulin acutely interact in the regulation of eating behavior in humans. We assessed the effect of intranasal insulin on food intake in healthy young men who were pretreated for 3 days with transdermal estradiol or placebo, hypothesizing that increasing the circulating concentrations of estrogen decreases the susceptibility of men to the anorexigenic effect of insulin.

Subjects and Methods

Subjects, design, and procedure

Thirty-two healthy men aged between 18 and 31 years (mean age, 23.94 ± 0.52 years; mean body mass index, 22.80 ± 0.36 kg/m²) participated in the experiment. Current illness, vegetarianism, and habitual dietary idiosyncrasies (*e.g.*, because of allergy) were excluded by clinical examination. All subjects were free of medication and were nonsmokers. They gave written informed consent to the study, which conformed to the Declaration of Helsinki as revised in 2008 and was approved by the local Ethics Committee on Research Involving Humans.

Study design and experimental procedures are summarized in Fig. 1. According to a 2×2 design, subjects were randomly

assigned to two groups of participants ($n = 16$ in each group) who were treated with either estradiol (“estrogen patch” group; 24.38 ± 0.93 years, 22.62 ± 0.50 kg/m²) or placebo (“placebo patch” group; 23.50 ± 0.49 years, $P > 0.41$; 22.98 ± 0.54 kg/m², $P > 0.61$) each time before participating in two individual experimental sessions where they received intranasal insulin or placebo. Three days before each test session, subjects attended our laboratory at 5:00 PM. In the participants of the estrogen patch group, two transdermal estradiol patches (Estradot 50; Novartis Pharma, Nuremberg, Germany) were applied to the abdomen, delivering a total dose of 100 µg estradiol per 24 hours according to the manufacturer. Participants of the placebo patch group received two patches that looked identical to the estradiol patches but did not contain the hormone. The patches were renewed by the experimenters after 24 and 48 hours (*i.e.*, the third pair of patches was attached on the day before and removed directly after the experiment proper). Subjects and experimenters were blinded to the patches and the intranasal treatment. Experimental sessions were separated by at least 3 weeks, and the order of conditions was balanced across subjects.

The experimental procedure on each test day was similar to our previous experiments on the acute effects of intranasal insulin on food intake (7, 13). All subjects remained fasted and abstained from drinking caloric beverages after 10:00 PM on the evening before testing. After arrival at the laboratory at around 8:00 AM, a venous cannula was inserted into each subject’s nondominant arm for the collection of venous blood and the determination of blood glucose (HemoCue B-Glucose-Analyzer; HemoCue AB, Angelholm, Sweden). Sessions started with a 60-minute baseline period, which included blood sampling at 8:15, 8:30, and 8:45 AM and ratings of mood and hunger. At 9:00 AM, subjects were intranasally administered 16 puffs (0.1 mL, eight per nostril) of insulin or placebo at 30-second intervals, amounting to a total dose of 1.6 mL insulin (160 IU) (Insulin Actrapid; Novo Nordisk, Mainz, Germany), or vehicle. At 10:25 AM, after postadministration blood sampling at 10- to 20-minute intervals and further assessments of mood and hunger, a standardized free-choice breakfast buffet was offered comprising a variety of food choices (Table 1) from which subjects ate *ad libitum* during the subsequent 30 minutes. Subjects were not aware that their food intake was measured by weighing buffet components before and after breakfast. This

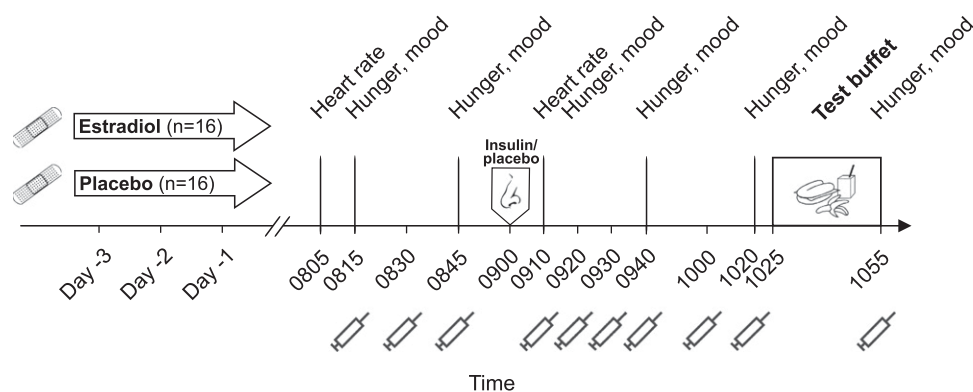


Figure 1. Experimental procedure. Two groups of 16 healthy men who had been pretreated with transdermal estradiol (100 µg/24 h for 3 days) or placebo participated in two experimental sessions. After a baseline period of around 60 minutes, subjects were intranasally administered 160 IU insulin or, in the other condition, placebo at 9:00 AM before a free-choice test breakfast buffet was offered around 85 minutes later. Self-rated hunger, thirst, tiredness, and mood were repeatedly assessed, and blood samples for the determination of glucose and hormone concentrations were obtained (syringe symbols). Heart rate and blood pressure were assessed twice.

Table 1. Composition of the Breakfast Test Buffet

Food	Weight (g)	Energy (kcal)	Carbohydrate (g)	Fat (g)	Protein (g)
Neutral					
Whole wheat bread	165	329	63.9	2	12.1
Wheat rolls	300	857	167.6	5.4	30
White bread	30	73	14.7	0.4	2.5
Butter	75	580	0.5	62.4	0.5
Whole milk	750	495	35.3	26.8	25.4
Condensed milk	40	54	3.9	3	2.6
Sweet					
Jam	50	140	34.2	0	0.1
Hazelnut spread	40	218	22.7	12.4	2.7
Honey	40	123	30	0	0.2
Sugar	24	98	24	0	0
Fruit curd	150	148	24.8	1.2	8.8
Banana	190	167	38.1	0.3	2.2
Apple	120	71	17	0.1	0.4
Pear	190	105	23.5	0.6	0.9
Orange juice	400	173	36	1	4
Savory					
Poultry sausage	40	74	0.1	4.3	8.3
Cervelat sausage	34	138	0.1	11.8	6.9
Sliced cheese	100	198	0	23	20.8
Cream cheese (natural)	33	87	1.1	8.2	1.8
Cream cheese (herbs)	40	84	1.2	7.2	3.2
Total	2811	4312	538	170	133

Breakfast was served with coffee or tea as requested by the participant.

procedure has been shown to enable the precise assessment of food intake in the fasted state (7, 13, 14). Throughout the experiments, subjects repeatedly underwent a battery of cognitive tests unrelated to the topic of the current study (data not shown). After final blood sampling and another assessment of mood and hunger, subjects were asked in a short interview which patch (estrogen/placebo) and which spray (insulin/placebo) they thought they had received. In these interviews, none of the subjects reported adverse side effects.

Hormonal and psychometric assessments

Blood samples were centrifuged, and plasma and serum were stored at -80°C . Serum concentrations of insulin, C-peptide, and cortisol (all sampling time points) and of luteinizing hormone (LH) and follicle-stimulating hormone (FSH) (first and third baseline time points) were determined by Immulite (DPC, Los Angeles, CA). Plasma concentrations of estradiol (*i.e.*, 17β -estradiol) and testosterone were determined for every other sampling time point (Fig. 2A and 2B) by ultra-high-performance liquid chromatography and subsequent mass spectrometry validated according to FDA guidelines. Samples were precipitated with 5% H_3PO_4 , purified, and concentrated in methanol before chromatographic separation was performed on an ultra-high-performance liquid chromatography instrument (1290 UHPL; Agilent Technologies, Waldbronn, Germany). Analyte detection was carried out on a hyphenated TripleTOF 5600+ mass spectrometer (Sciex, Concord, Ontario, Canada) in positive ionization mode. For quantification, a surrogate calibrant method using $^{13}\text{C}_3$ -estradiol and $^{13}\text{C}_3$ -testosterone in true plasma matrix was established. Internal standardization was obtained by spiking of d_5 -estradiol and d_5 -testosterone. Quantifiable ranges for estradiol and testosterone were 10 to 1000 pg/mL and 20 to 15,000 pg/mL, respectively.

In a simultaneous mass spectrometric survey scan, we recorded precursor ion data for an untargeted profiling of plasma samples that yielded the relative concentrations of additional steroids of interest [*i.e.*, epitestosterone (17α -testosterone), dihydrotestosterone (androstanolone), androstenedione, dehydroepiandrosterone (androstenolone), progesterone, and hydroxyprogesterone].

Hunger, thirst, and tiredness were rated on nine-point scales twice during baseline, at 20- to 30-minute intervals after spray administration, and after the test breakfast. In parallel, mood was assessed with five-point scales covering the categories good/bad mood, alertness/sleepiness, and calmness/agitation [Mehrdimensionaler Befindlichkeitsfragebogen (MDBF); ref. 15]. Blood pressure and heart rate were measured before and ~ 10 minutes after spray administration.

Statistical analyses

Analyses were performed with SPSS[®] Statistics Version 21 (IBM, Armonk, NY) and based on repeated-measures analyses of variance (ANOVA) with the between-subjects factor "Group" (estrogen patch vs placebo patch) and the within-subject factors "Treatment" (insulin vs placebo), "Time," "Macronutrient," and "Taste" (*i.e.*, neutral/sweet/savory) as appropriate. Significant ANOVA interactions were specified by Student's *t* tests. All data are presented as means \pm standard error of the mean (SEM). A *P* value < 0.05 was considered significant.

Results

Hormonal parameters

Transdermal estrogen in comparison with placebo treatment induced a 3.5-fold increase in baseline plasma

estradiol concentrations [F(1,30) = 75.38, $P < 0.0001$] and a 70% decrease in testosterone [F(1,30) = 88.19, $P < 0.0001$ for Group] (Fig. 2A and 2B). Both estradiol and testosterone displayed a postprandial drop after breakfast intake ($P < 0.0001$ for Time) that was more pronounced in the groups pretreated with estrogen (both $P < 0.002$ for Group \times Time). Intranasal insulin did not display a

modulatory influence on these parameters (all $P > 0.23$). Concentrations of LH and FSH measured during baseline were strongly suppressed after estrogen treatment (both $P < 0.01$ for Group) (Fig. 2C). Supplemental analyses of steroid hormones (Table 2) indicated that transdermal estrogen administration roughly halved plasma concentrations of epitestosterone and induced 27% and

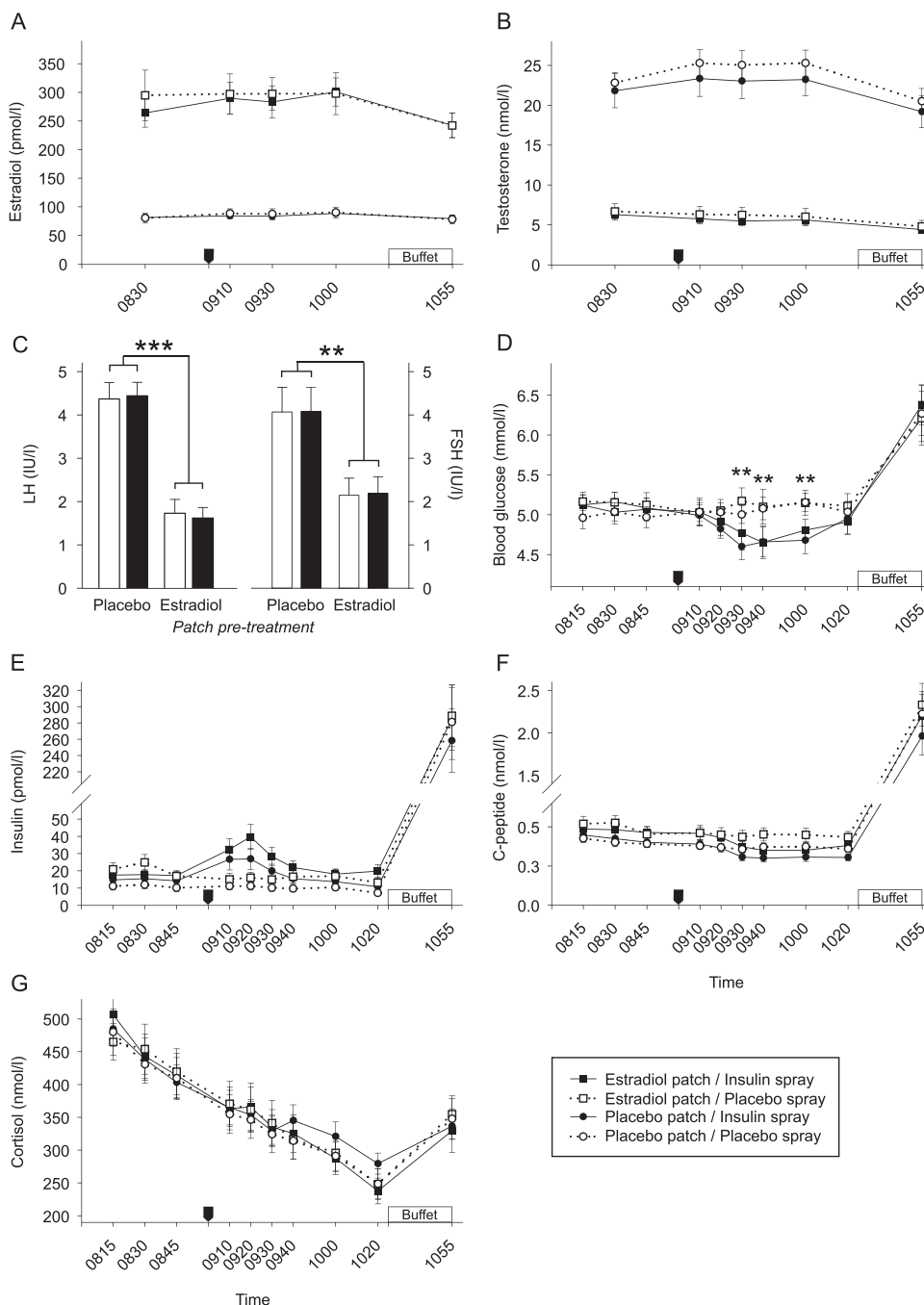


Figure 2. Endocrine parameters. Plasma concentrations of (A) 17β -estradiol and (B) testosterone; (C) serum concentrations of LH and FSH; (D) blood glucose concentrations; and serum concentrations of (E) insulin, (F), C-peptide, and (G) cortisol. Experiments were performed in two groups of 16 men each who had received 3 days of transdermal estradiol (100 μ g/24 h; squares) or placebo pretreatment (circles) before participating in experimental sessions starting with baseline measurements followed by the intranasal spray administration of 160 IU insulin (filled symbols, solid lines) or placebo (empty symbols, dashed lines), respectively, at 9:00 AM (arrow mark). LH and FSH represent the average of the 8:15 and 8:45 AM baseline measurements. Values are means \pm SEM. $**P < 0.01$ for the ANOVA factor Treatment (D); $**P < 0.01$ and $***P < 0.001$ for the ANOVA factor Group (C).

Table 2. Supplemental Steroid Measurements

	Placebo Patch		Estradiol Patch		ANOVA Result ^a	
	Placebo Spray	Insulin Spray	Placebo Spray	Insulin Spray	Group	Treatment
Epitestosterone (17 α -testosterone)	41.42 \pm 2.62	40.81 \pm 3.95	18.97 \pm 2.23	18.84 \pm 2.50	F(1,29) = 43.91, <i>P</i> < 0.001	F(1,29) = 0.90, <i>P</i> = 0.35
Dihydrotestosterone (Androstanolone)	133.77 \pm 9.62	124.55 \pm 9.30	100.30 \pm 7.71	88.61 \pm 6.00	F(1,30) = 15.33, <i>P</i> < 0.001	F(1,30) = 1.87, <i>P</i> = 0.18
Androstenedione	1216.63 \pm 108.87	1113.46 \pm 95.55	1033.52 \pm 87.05	979.30 \pm 97.84	F(1,29) = 1.10, <i>P</i> = 0.30	F(1,29) = 6.73, <i>P</i> < 0.02
Dehydroepiandrosterone (Androstenolone)	741.41 \pm 63.87	694.04 \pm 52.62	817.51 \pm 81.19	832.87 \pm 78.65	F(1,29) = 1.44, <i>P</i> = 0.24	F(1,29) = 0.18, <i>P</i> = 0.67
Progesterone	619.90 \pm 88.90	500.20 \pm 94.13	538.26 \pm 92.01	432.08 \pm 42.90	F(1,29) = 0.43, <i>P</i> = 0.52	F(1,29) = 3.02, <i>P</i> = 0.09
Hydroxyprogesterone	754.86 \pm 69.49	641.86 \pm 53.90	242.43 \pm 34.47	238.78 \pm 34.61	F(1,30) = 53.27, <i>P</i> < 0.001	F(1,30) = 3.02, <i>P</i> = 0.09

Relative plasma levels as derived from response ratios (peak area of the analyte/peak area of the internal standard) obtained in a mass spectrometric survey scan (relative quantification) and expressed as areas under the curve of the main experimental period (8:30–10:55 AM). Experiments were performed in two groups of men who received 3 days of transdermal estradiol (100 μ g/24 h) or placebo before participating in experimental sessions including intranasal treatment with 160 IU insulin or placebo.

^aResults for the ANOVA factors Group and Treatment; respective interactions were not significant (*n* = 31 or 32).

66% reductions in, respectively, dihydrotestosterone and hydroxyprogesterone, all independent of insulin or placebo administration. Intranasal insulin compared with placebo induced a mild, estradiol patch-independent decrease in androstenedione.

Parameters of glucose metabolism and cortisol concentrations did not differ between conditions during baseline (all *P* > 0.09) and were generally not affected by estrogen treatment (all *P* > 0.10 for respective interactions). Intranasal insulin administration induced a slight decrease in blood glucose concentrations [F(3,98) = 3.81, *P* < 0.02 for Treatment \times Time) that remained within the euglycemic range (Fig. 2D). Corresponding changes in serum insulin and C-peptide after intranasal insulin in comparison with placebo administration failed to reach statistical significance in ANOVA [F(2,37) = 2.42,

P < 0.12, and F(1,33) = 3.01, *P* < 0.09 for Treatment \times Time) (Fig. 2E and 2F). However, supplemental area-under-the-curve analyses covering the time period between the final baseline (8:45 AM) and the final pre-breakfast sample (10:22 AM) indicated significant respective increases in serum insulin [F(1,30) = 15.15, *P* < 0.01] and decreases in serum C-peptide [F(1,30) = 4.91, *P* < 0.04]. Cortisol concentrations showed the expected circadian decline but were not affected by any of the hormonal interventions (*P* > 0.41) (Fig. 2G).

Food intake

Across groups, insulin in comparison with placebo specifically reduced the intake of carbohydrates from the test buffet [F(1,30) = 5.60, *P* < 0.03; F(2,48) = 4.39, *P* < 0.03 for Treatment \times Macronutrient] (Fig. 3A) against the

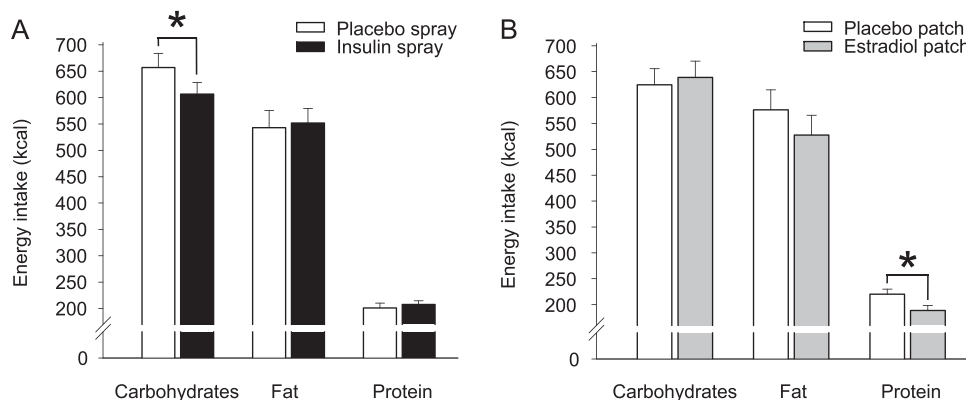


Figure 3. Food intake from the test buffet. Intake of macronutrients (kcal) from a standardized free-choice breakfast buffet presented 85 minutes after intranasal spray administration of 160 IU insulin or placebo in healthy men who had received 3 days of transdermal estradiol (100 μ g/24 h) or placebo pretreatment (*n* = 16 in each group) before the experimental day. (A) Macronutrient intake in the insulin (black bars) and the placebo spray conditions (white bars) collapsed across the estrogen and placebo patch groups. (B) Macronutrient intake in the estradiol (gray bars) and the placebo patch (white bars) groups collapsed across the insulin and placebo spray conditions. Values are means \pm SEM. **P* < 0.05 for the ANOVA factors Treatment (A) and Group (B).

background of comparable total food intake in both conditions [$F(1,30) = 0.58, P > 0.45$] (Table 3), whereas the intake of fat and protein was not affected by insulin (all $P > 0.48$) (Fig. 3A). The suppressive effect of intranasal insulin on carbohydrate intake was confirmed in covariance analyses correcting for the difference between conditions in pre-breakfast concentrations of blood glucose [$F(1,29) = 5.38, P < 0.03$; $F(2,47) = 3.91, P < 0.04$ for Treatment \times Macronutrient] and serum insulin [$F(1,29) = 6.87, P < 0.02$ and $F(2,48) = 7.08, P < 0.01$; values expressed as areas under the curve as defined previously]. The reduction in the consumption of carbohydrates was also reflected by slight decreases and increases in the consumption of food items with sweet and savory taste, respectively [$F(2,54) = 3.88, P < 0.04$ for Treatment \times Taste] (Table 3). Estrogen administration in the 3 days preceding the experiment did not modulate the effect of intranasal insulin on food intake regarding carbohydrate consumption [$F(1,30) = 0.03, P > 0.87$ for Treatment \times Group] or total intake and consumption of fat, protein, and sweet vs savory foods (all $P > 0.52$).

Independent of the intranasal treatment estrogen *per se* attenuated the intake of protein from the test buffet [$F(1,30) = 5.12, P = 0.03$ for Group] (Fig. 3B). Total intake and the intake of fat and carbohydrate remained unaffected by estrogen (all $P > 0.47$). Although protein specificity of estrogen's anorexigenic effect was not statistically confirmed [$F(2,48) = 0.72, P > 0.46$ for Group \times Macronutrient], estrogen treatment in particular reduced the intake of savory food items [estrogen vs placebo patch, 297.45 ± 22.28 vs 381.95 ± 22.28 kcal; $F(1,30) = 7.19, P < 0.02$], such as sliced and natural cream cheese [138.72 ± 16.11 vs 184.93 ± 16.11 kcal, $F(1,30) = 4.11, P < 0.06$; 22.87 ± 7.97 vs 44.91 ± 7.97 kcal, $F(1,30) = 3.83, P < 0.06$, respectively] in favor of an increase in the intake of items like hazelnut spread [100.11 ± 17.26 vs

50.31 ± 17.26 kcal; $F(1,30) = 4.18, P = 0.05$]. Hunger and thirst ratings were not affected by estrogen treatment or insulin administration (all $P > 0.13$).

Control parameters

Self-rated mood and alertness according to the MDBF adjective scale generally improved during the experiment (both $P < 0.002$ for Time) but were not affected by estrogen treatment or insulin administration (all $P > 0.11$). Accordingly, tiredness rated on nine-point scales decreased between morning and noon ($P < 0.001$). Neither tiredness nor calmness/agitation ratings showed differences between conditions or groups (all $P > 0.10$). Cardiovascular parameters were not affected by estrogen or insulin administration (all $P > 0.21$). In the estrogen patch group, systolic/diastolic blood pressure measured 10 minutes after intranasal insulin in comparison with placebo administration was $118.33 \pm 3.45/72.40 \pm 2.36$ vs $123.88 \pm 2.71/72.44 \pm 2.59$ mm Hg and $124.56 \pm 2.83/71.06 \pm 2.94$ vs $120.13 \pm 3.69/72.25 \pm 2.32$ mm Hg in the placebo patch group. Heart rate in the estrogen patch group was (insulin vs placebo) 59.67 ± 1.80 vs 59.06 ± 2.14 beats per minute and 58.63 ± 2.24 vs 59.88 ± 2.71 beats per minute in the placebo patch group. In the post-experimental interviews, subjects were not able to correctly indicate whether they had received estrogen or placebo patches ($P > 0.53$) and insulin or placebo sprays ($P > 0.71$; χ^2 tests).

Discussion

Central nervous insulin administration exerts stronger acute (7) and long-term (8) catabolic effects in male subjects than in female subjects; here we investigated whether estrogen signaling contributes to this sex-specific pattern. We found that strongly increasing circulating

Table 3. Food Intake From the Test Buffet

Food Intake (kcal)	Placebo	Insulin	P Value ^a
Total	1401.29 \pm 55.84	1365.82 \pm 45.61	0.45
Neutral food	743.34 \pm 40.59	696.78 \pm 38.33	0.18
Wheat rolls	361.24 \pm 27.81	301.79 \pm 29.24	0.06
Sweet food	343.50 \pm 25.58	304.08 \pm 23.66	0.09
Jam	37.38 \pm 6.82	25.76 \pm 7.44	0.03
Hazelnut spread	89.20 \pm 14.01	61.23 \pm 13.01	0.03
Honey	17.32 \pm 5.17	8.85 \pm 3.61	0.09
Sugar	10.69 \pm 2.46	6.74 \pm 2.38	0.07
Savory food	314.45 \pm 21.02	364.95 \pm 18.44	0.04
Cervelat sausage	67.95 \pm 8.37	83.03 \pm 8.17	0.10
Sliced cheese	147.46 \pm 13.94	176.19 \pm 11.96	0.03

Total food intake, food intake according to taste, and consumption of specific food items (all in kcal). All neutral, savory, and sweet foods contained in the test buffet are listed in Table 1. Values are means \pm SEM calculated across experimental groups (placebo and estrogen patch) for the experimental conditions (placebo and insulin spray).

^aP values for the ANOVA factor Treatment (n = 32).

estrogen concentrations in healthy young men by means of transdermal estradiol patches does not alter the suppressive effect on carbohydrate intake of intranasal insulin. This outcome stands in some contrast to findings in animals indicating that estrogen action interferes with the anorexigenic effect of brain insulin (9, 10). Estrogen administration *per se* was revealed to induce a small but discernible reduction in protein consumption, indicating that both insulin and estrogen, but in an independent fashion, induce restraining effects on the intake of macronutrients in men.

The pretreatment of our subjects with estradiol patches worn for 3 consecutive days proved to be highly effective, as evidenced by the 3.5-fold increase in circulating estrogen, whereas the concentrations of LH and FSH were roughly halved and serum testosterone dropped by about 70%. Estradiol administration markedly reduced plasma levels of epitestosterone, dihydrotestosterone, and hydroxyprogesterone, further indicating a pronounced impact of our intervention on steroid signaling. The estrogen-induced reduction in protein intake from the test breakfast buffet fits with animal experiments, indicating that centrally administered estrogen, similar to the adiposity signal leptin (16), inhibits food intake (17, 18). This effect is likely mediated via estrogen receptors expressed in the hypothalamic arcuate and ventromedial nuclei and the nucleus of the solitary tract in the hindbrain (19) but also in the reward-processing ventral tegmental area (20). Daily food intake in naturally cycling women reaches its nadir during the peri-ovulatory phase when estradiol concentrations are maximal (21, 22). Protein consumption has been found to be less pronounced during this phase as compared with the mid-luteal phase (23), although findings on cycle-dependent fluctuations in the intake of specific macronutrients are not unanimous (24, 25). Estrogen-induced reductions in the concentrations of testosterone and dihydrotestosterone might have contributed to decreased calorie consumption (26–28). Most recently, 17 α -estradiol, an enantiomer of 17 β -estradiol, has been suggested to induce centrally mediated catabolic effects (29). Although in our study the estrogen effect was evident for protein rather than fat or carbohydrate intake and appeared to focus on savory foods, it underlines the potential of estrogen delivery to restrain food intake (30).

Irrespective of estrogen pretreatment, intranasal insulin reduced carbohydrate consumption. This finding supports previous observations that intranasal insulin acutely decreases free-choice breakfast intake in healthy young men (7), although overall calorie intake was not reduced in the present experiments. *Post hoc* analyses of intense eaters displaying total calorie intake above the median of the respective placebo spray conditions

revealed an insulin-driven, estrogen-independent decrease also in overall food intake ($n = 8$ per estrogen and placebo patch group, respectively; data not shown). These results support the notion that insulin transported to the central nervous system acts as a negative feedback signal in the control of eating behavior (2, 6). It appears unlikely that the moderate insulin-induced reduction in androstenedione levels was involved in this effect (31). Intranasal insulin also induced a slight increase in serum insulin and a euglycemic decrease in blood glucose concentrations that presumably stemmed from a small ratio of exogenous insulin entering the circulation via the nasal mucosa (7, 13, 32). Insulin's effect on food intake was confirmed in analyses corrected for these subtle changes in glucoregulation, so that a peripheral mediation of the decrease in carbohydrate consumption may be excluded.

Our observation that insulin restrains carbohydrate intake and the ingestion of sweet food items ties in with previous findings of reduced calorie and, in particular, carbohydrate intake after pre-sleep intranasal insulin administration (33) and of an insulin-induced decrease in the intake of chocolate cookies in nonfasted women (11). Studies in rats, in contrast, have indicated that central insulin administration predominantly reduces fat intake (34) but have also found respective reductions in sucrose self-administration (35). In humans, intranasal insulin acutely reduces the responsiveness to food stimuli of the ventral tegmentum and nucleus accumbens of the brain reward circuit as well as rated food palatability in men and women (36). Although animal experiments in general confirm that insulin inhibits the reward-related consumption of palatable foods (37, 38), conflicting data exist on the effect of insulin on dopaminergic signaling (39, 40).

Contrary to our expectation, estrogen pretreatment did not modulate the reduction in carbohydrate intake induced by intranasal insulin. In male rats, the peripheral administration of estradiol at a dosage of 2 μg administered every fourth day for 1 month completely blunted the reduction in 24-hour food intake and body weight observed in control animals after insulin injection into the third cerebral ventricle (10). In these animals, estradiol treatment increased plasma estradiol concentrations by 60%, averaging peri-ovulatory peak concentrations of female animals (41), whereas increasing serum estradiol concentrations by a factor of 3.5 in our male subjects yielded concentrations of around 300 pmol/L; women typically achieve ovulatory concentrations of 100 to 600 pmol/L. In principle, extended administration periods and higher dosages of estradiol might modulate the hypophagic effect of insulin, but estrogen's impact on eating behavior in the present paradigm is clearly indicated

by its suppressive effect on protein intake. Thus, it seems safe to conclude that inducing peripheral estrogen concentrations in healthy young men that approximate the situation regularly found in women and that induce strong reductions in the concentration of testosterone and related steroid hormones does not alter the brain's sensitivity to the anorexigenic impact of insulin. Different expression patterns of estrogen receptors in the male and female brain (42, 43) or basic genetic sex differences might contribute to altered estrogen-insulin interactions in men compared with women. However, in previous experiments restricted to women (13), the difference in estrogen levels between young women receiving ethinyl estradiol-dominant contraceptives and postmenopausal women, in accordance with the present data, was not associated with differences in the response to intranasal insulin.

Conclusion

In healthy men, central nervous insulin delivery via the intranasal route decreases carbohydrate consumption regardless of whether concurrent circulating estrogen concentrations are normal or elevated. This pattern indicates that estrogen, which displays moderate suppressive effects on protein intake, and insulin do not acutely interact in the regulation of eating behavior in humans. Further investigations, which should also cover longer time scales, are needed to gain insight into neurobiological mechanisms underlying the stronger anorexigenic effect of central insulin in male than female subjects reported in animals (9) and humans (7, 8). Such studies might contribute to the development of sex-specific, individually tailored approaches in the treatment of eating disorders.

Acknowledgments

We thank Kirstin Nordhausen (Department of Internal Medicine I, University of Lübeck, Germany) and Heidi Ruf and Martina Grohs (Department of Neuroendocrinology, University of Lübeck, Lübeck, Germany) for invaluable laboratory work. Aero Pump (Hochheim, Germany) provided the precision nasal air pumps.

Financial Support: This work was supported by grants from Deutsche Forschungsgemeinschaft (SFB 654), the German Federal Ministry of Education and Research (BMBF) to the German Center for Diabetes Research (DZD e.V.; 01GI0925), and the Helmholtz Alliance ICAMED—Imaging and Curing Environmental Metabolic Diseases, through the Initiative and Network Fund of the Helmholtz Association.

Clinical Trial Information: Deutsches Register Klinischer Studien (German Clinical Trials Register) DRKS00007175 (7 November 2014).

Author Contributions: R.K., J.B., and M.H. designed research; R.K., L.M., B.D., and M.H. conducted research and analyzed data; R.K. and M.H. wrote the paper; all authors read and approved the final manuscript. M.H. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Correspondence and Reprint Requests: Manfred Hallschmid, PhD, Department of Medical Psychology and Behavioral Neurobiology, University of Tübingen, Otfried-Müller-Str. 25, 72076 Tübingen, Germany. E-mail: manfred.hallschmid@uni-tuebingen.de.

Disclosure Summary: The authors have nothing to disclose.

References

- Vogt MC, Brüning JC. CNS insulin signaling in the control of energy homeostasis and glucose metabolism - from embryo to old age. *Trends Endocrinol Metab*. 2013;24(2):76–84.
- Clemmensen C, Müller TD, Woods SC, Berthoud HR, Seeley RJ, Tschöp MH. Gut-brain cross-talk in metabolic control. *Cell*. 2017;168(5):758–774.
- Born J, Lange T, Kern W, McGregor GP, Bickel U, Fehm HL. Sniffing neuropeptides: a transnasal approach to the human brain. *Nat Neurosci*. 2002;5(6):514–516.
- Brown LM, Clegg DJ, Benoit SC, Woods SC. Intraventricular insulin and leptin reduce food intake and body weight in C57BL/6J mice. *Physiol Behav*. 2006;89(5):687–691.
- Chavez M, Kaiyala K, Madden LJ, Schwartz MW, Woods SC. Intraventricular insulin and the level of maintained body weight in rats. *Behav Neurosci*. 1995;109(3):528–531.
- Woods SC, Lotter EC, McKay LD, Porte D, Jr. Chronic intracerebroventricular infusion of insulin reduces food intake and body weight of baboons. *Nature*. 1979;282(5738):503–505.
- Benedict C, Kern W, Schultes B, Born J, Hallschmid M. Differential sensitivity of men and women to anorexigenic and memory-improving effects of intranasal insulin. *J Clin Endocrinol Metab*. 2008;93(4):1339–1344.
- Hallschmid M, Benedict C, Schultes B, Fehm HL, Born J, Kern W. Intranasal insulin reduces body fat in men but not in women. *Diabetes*. 2004;53(11):3024–3029.
- Clegg DJ, Riedy CA, Smith KA, Benoit SC, Woods SC. Differential sensitivity to central leptin and insulin in male and female rats. *Diabetes*. 2003;52(3):682–687.
- Clegg DJ, Brown LM, Woods SC, Benoit SC. Gonadal hormones determine sensitivity to central leptin and insulin [published correction appears in *Diabetes* 2007;56(10):2649]. *Diabetes*. 2006;55(4):978–987.
- Hallschmid M, Higgs S, Thienel M, Ott V, Lehnert H. Postprandial administration of intranasal insulin intensifies satiety and reduces intake of palatable snacks in women. *Diabetes*. 2012;61(4):782–789.
- Mauvais-Jarvis F, Clegg DJ, Hevener AL. The role of estrogens in control of energy balance and glucose homeostasis. *Endocr Rev*. 2013;34(3):309–338.
- Krug R, Benedict C, Born J, Hallschmid M. Comparable sensitivity of postmenopausal and young women to the effects of intranasal insulin on food intake and working memory. *J Clin Endocrinol Metab*. 2010;95(12):E468–E472.
- Ott V, Finlayson G, Lehnert H, Heitmann B, Heinrichs M, Born J, Hallschmid M. Oxytocin reduces reward-driven food intake in humans. *Diabetes*. 2013;62(10):3418–3425.
- Steyer R, Schwenkmezger P, Notz P, Eid M. *Der mehrdimensionale Befindlichkeitsfragebogen (MDBF)*. Handanweisung. Göttingen, Germany: Hogrefe; 1997.

16. Gao Q, Horvath TL. Cross-talk between estrogen and leptin signaling in the hypothalamus. *Am J Physiol Endocrinol Metab.* 2008; **294**(5):E817–E826.
17. Brown LM, Gent L, Davis K, Clegg DJ. Metabolic impact of sex hormones on obesity. *Brain Res.* 2010; **1350**:77–85.
18. Xu Y, Nedungadi TP, Zhu L, Sobhani N, Irani BG, Davis KE, Zhang X, Zou F, Gent LM, Hahner LD, Khan SA, Elias CF, Elmquist JK, Clegg DJ. Distinct hypothalamic neurons mediate estrogenic effects on energy homeostasis and reproduction. *Cell Metab.* 2011; **14**(4):453–465.
19. Thammacharoen S, Lutz TA, Geary N, Asarian L. Hindbrain administration of estradiol inhibits feeding and activates estrogen receptor- α -expressing cells in the nucleus tractus solitarius of ovariectomized rats. *Endocrinology.* 2008; **149**(4):1609–1617.
20. Richard JE, López-Ferreras L, Anderberg RH, Olandersson K, Skibicka KP. Estradiol is a critical regulator of food-reward behavior. *Psychoneuroendocrinology.* 2017; **78**:193–202.
21. Roney JR, Simmons ZL. Ovarian hormone fluctuations predict within-cycle shifts in women's food intake. *Horm Behav.* 2017; **90**:8–14.
22. Lyons PM, Truswell AS, Mira M, Vizzard J, Abraham SF. Reduction of food intake in the ovulatory phase of the menstrual cycle. *Am J Clin Nutr.* 1989; **49**(6):1164–1168.
23. Gorczyca AM, Sjaarda LA, Mitchell EM, Perkins NJ, Schliep KC, Wactawski-Wende J, Mumford SL. Changes in macronutrient, micronutrient, and food group intakes throughout the menstrual cycle in healthy, premenopausal women. *Eur J Nutr.* 2016; **55**(3):1181–1188.
24. Dye L, Blundell JE. Menstrual cycle and appetite control: implications for weight regulation. *Hum Reprod.* 1997; **12**(6):1142–1151.
25. Bowen DJ, Grunberg NE. Variations in food preference and consumption across the menstrual cycle. *Physiol Behav.* 1990; **47**(2):287–291.
26. Asarian L, Geary N. Modulation of appetite by gonadal steroid hormones. *Philos Trans R Soc Lond B Biol Sci.* 2006; **361**(1471):1251–1263.
27. Kanaya N, Vonderfecht S, Chen S. Androgen (dihydrotestosterone)-mediated regulation of food intake and obesity in female mice. *J Steroid Biochem Mol Biol.* 2013; **138**:100–106.
28. Nunez AA, Siegel LI, Wade GN. Central effects of testosterone on food intake in male rats. *Physiol Behav.* 1980; **24**(3):469–471.
29. Stout MB, Steyn FJ, Jurczak MJ, Camporez JG, Zhu Y, Hawse JR, Jurk D, Palmer AK, Xu M, Pirtskhalava T, Evans GL, de Souza Santos R, Frank AP, White TA, Monroe DG, Singh RJ, Casaclang-Verzosa G, Miller JD, Clegg DJ, LeBrasseur NK, von Zglinicki T, Shulman GI, Tchkonian T, Kirkland JL. 17α -estradiol alleviates age-related metabolic and inflammatory dysfunction in male mice without inducing feminization. *J Gerontol A Biol Sci Med Sci.* 2017; **72**(1):3–15.
30. Finan B, Yang B, Ottaway N, Stemmer K, Müller TD, Yi CX, Habegger K, Schriever SC, García-Cáceres C, Kabra DG, Hembree J, Holland J, Raver C, Seeley RJ, Hans W, Irmler M, Beckers J, de Angelis MH, Tiano JP, Mauvais-Jarvis F, Perez-Tilve D, Pfluger P, Zhang L, Gelfanov V, DiMarchi RD, Tschöp MH. Targeted estrogen delivery reverses the metabolic syndrome. *Nat Med.* 2012; **18**(12):1847–1856.
31. Hargrave KR, Wright BE, Svec F, Porter JR. Dietary and hypothalamic changes in delta 4-androstenedione-treated Zucker rats. *Physiol Behav.* 1997; **61**(4):619–626.
32. Benedict C, Brede S, Schiöth HB, Lehnert H, Schultes B, Born J, Hallschmid M. Intranasal insulin enhances postprandial thermogenesis and lowers postprandial serum insulin levels in healthy men. *Diabetes.* 2011; **60**(1):114–118.
33. Santiago JC, Hallschmid M. Central nervous insulin administration before nocturnal sleep decreases breakfast intake in healthy young and elderly subjects. *Front Neurosci.* 2017; **11**:54.
34. Chavez M, Riedy CA, Van Dijk G, Woods SC. Central insulin and macronutrient intake in the rat. *Am J Physiol.* 1996; **271**(3 Pt 2):R727–R731.
35. Figlewicz DP, Bennett JL, Aliakbari S, Zavosh A, Sipols AJ. Insulin acts at different CNS sites to decrease acute sucrose intake and sucrose self-administration in rats. *Am J Physiol Regul Integr Comp Physiol.* 2008; **295**(2):R388–R394.
36. Tiedemann LJ, Schmid SM, Hettel J, Giesen K, Francke P, Büchel C, Brassen S. Central insulin modulates food valuation via mesolimbic pathways. *Nat Commun.* 2017; **8**:16052.
37. Mebel DM, Wong JC, Dong YJ, Borgland SL. Insulin in the ventral tegmental area reduces hedonic feeding and suppresses dopamine concentration via increased reuptake. *Eur J Neurosci.* 2012; **36**(3):2336–2346.
38. Könner AC, Klöckener T, Brüning JC. Control of energy homeostasis by insulin and leptin: targeting the arcuate nucleus and beyond. *Physiol Behav.* 2009; **97**(5):632–638.
39. Labouèbe G, Liu S, Dias C, Zou H, Wong JC, Karunakaran S, Clee SM, Phillips AG, Boutrel B, Borgland SL. Insulin induces long-term depression of ventral tegmental area dopamine neurons via endocannabinoids. *Nat Neurosci.* 2013; **16**(3):300–308.
40. Stouffer MA, Woods CA, Patel JC, Lee CR, Witkovsky P, Bao L, Machold RP, Jones KT, de Vaca SC, Reith MEA, Carr KD, Rice ME. Insulin enhances striatal dopamine release by activating cholinergic interneurons and thereby signals reward. *Nat Commun.* 2015; **6**:8543.
41. Asarian L, Geary N. Cyclic estradiol treatment normalizes body weight and restores physiological patterns of spontaneous feeding and sexual receptivity in ovariectomized rats. *Horm Behav.* 2002; **42**(4):461–471.
42. Kruijver FP, Balesar R, Espila AM, Unmehopa UA, Swaab DF. Estrogen receptor- α distribution in the human hypothalamus in relation to sex and endocrine status. *J Comp Neurol.* 2002; **454**(2):115–139.
43. Zuloaga DG, Zuloaga KL, Hinds LR, Carbone DL, Handa RJ. Estrogen receptor β expression in the mouse forebrain: age and sex differences. *J Comp Neurol.* 2014; **522**(2):358–371.

6.1.2. Comparison of Simple Monophasic versus Classical Biphase Extraction Protocols for Comprehensive UHPLC-MS/MS Lipidomic Analysis of HeLa Cells

*Carlos Calderón, Corinna Sanwald, Jörg Schlotterbeck, Bernhard Drotleff,
Michael Lämmerhofer*

Institute of Pharmaceutical Sciences, Pharmaceutical (Bio-)Analysis, University of Tübingen,
Tübingen, Germany

**Reprinted with permission from *Analytica Chimica Acta*, Volume 1048, Pages 66-74,
DOI: 10.1016/j.aca.2018.10.035**

Copyright (2019) Elsevier B.V.



Contents lists available at ScienceDirect

Analytica Chimica Acta

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

Comparison of simple monophasic versus classical biphasic extraction protocols for comprehensive UHPLC-MS/MS lipidomic analysis of Hela cells

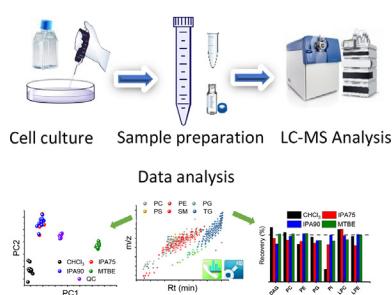
Carlos Calderón, Corinna Sanwald, Jörg Schlotterbeck, Bernhard Drotleff, Michael Lämmerhofer*

Institute of Pharmaceutical Sciences, Pharmaceutical (Bio-)Analysis, University of Tübingen, Auf der Morgenstelle 8, 72076, Tübingen, Germany

HIGHLIGHTS

- Systematic comparison of four extraction protocols for lipidomic analysis.
- 292 identified lipids in ESI (+) and 206 identified lipids in ESI (-) analyzed
- Monophasic mixture IPA:H₂O 90:10 v/v shows robust performance for lipid analysis.
- Analysis with MS-Dial allow identification of ~15% of detected features.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 27 July 2018

Received in revised form

12 October 2018

Accepted 16 October 2018

Available online xxx

Keywords:

Extraction

Lipidomics

Mass spectrometry

Data-independent acquisition

Hela cells

Lipid profiling

ABSTRACT

In this study, two monophasic isopropanol-water mixtures (IPA:H₂O 75:25 v/v and IPA:H₂O 90:10 v/v) were compared with traditionally employed biphasic methods of Bligh & Dyer and Matyash et al. as extraction systems for lipidomics analysis in Hela cells. Samples were analyzed by UHPLC-ESI-QTOF-MS/MS in positive and negative mode using sequential window acquisition of all theoretical fragment ion spectra (SWATH) and a relatively new software (MS-DIAL) was employed for the processing of the data which includes detection of peaks, MS/MS spectra deconvolution, identification of detected lipids and alignment of peaks through the analyzed samples.

The studied performance parameters such as precision, recoveries of isotopically labeled internal standards and endogenous lipids, number of extracted lipids, and complexity of employed procedure showed that extraction with IPA:H₂O 90:10 v/v performs similar to the Matyash protocol and better than Bligh & Dyer as well as IPA:H₂O 75:25 v/v. However, less complex monophasic protocol which is simpler to implement and can be executed in plastic rather than glass, make the monophasic IPA:H₂O 90:10 v/v protocol an excellent alternative to the classical biphasic protocols for reversed phase LC-MS lipidomics studies.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

Liquid chromatography-mass spectrometry (LC-MS) has become the most widely used analytical tool for lipidomics analysis

* Corresponding author. Pharmaceutical (Bio-)Analysis, Institute of Pharmaceutical Sciences, University of Tübingen, Auf der Morgenstelle, 8, 72076, Tübingen, Germany.

E-mail address: michael.laemmerhofer@uni-tuebingen.de (M. Lämmerhofer).

Abbreviations

BMP	Bismonoacylglycerophosphate	(L)PC	(Lyso)phosphatidylcholine
CE	Cholesteryl ester	(L)PE	(Lyso)phosphatidylethanolamine
Cer-NDS	Ceramide non-hydroxyfatty acid-dihydrosphingosine	(L)PG	(Lyso)phosphatidylglycerol
Cer-NS	Ceramide non-hydroxyfatty acid-sphingosine	(L)PI	(Lyso)phosphatidylinositol
DG	Diacylglycerol	(L)PS	(Lyso)phosphatidylserine
EP	Extraction protocol	MG	Monoacylglycerol
EtherPC	Ether-linked phosphatidylcholine	MTBE	Methyl <i>tert</i> -butyl ether
EtherPE	Ether-linked phosphatidylethanolamine	OxFA	Oxidized fatty acid
FA	Fatty acid	OxPC	Oxidized phosphatidylcholine
HexCer-NDS	Hexosylceramide non-hydroxyfatty acid-dihydrosphingosine	OxPE	Oxidized phosphatidylethanolamine
HexCer-NS	Hexosylceramide non-hydroxyfatty acid-sphingosine	OxPG	Oxidized phosphatidylglycerol
IPA	isopropanol	OxPI	Oxidized phosphatidylinositol
(L)PA	(Lyso)phosphatidic acid	OxPS	Oxidized phosphatidylserine
		SM	Sphingomyelin
		SWATH	Sequential window acquisition of all theoretical fragment ion spectra
		TG	Triacylglycerol
		TIC	Total ion current

in the last few years [1–4]. Despite many reported publications on this topic, there is no consensus about the most adequate protocol to follow. One of the most critical points is the sample preparation and therefore the choice of a solvent or solvent mixture for the lipid extraction, since this will determine which lipid classes are mostly recovered from the sample [5]. Extractions with the system chloroform-methanol-water (CHCl_3 -MeOH- H_2O), at specific ratios, were introduced more than six decades ago, with the pioneering works of Folch [6] and Bligh & Dyer [7], and have frequently been described as “gold standards” [1,8–13]. Hexane-isopropanol-water (Hexane-IPA- H_2O) was proposed in 1978 by Hara et al. as a less toxic option [14]. However, its performance in terms of wide coverage of distinct lipid classes is modest [15–17]. Recently, two methods using methyl *tert*-butyl ether-methanol-water (MTBE-MeOH- H_2O) [10] and butanol-MeOH- H_2O [18] were introduced and they have become popular because of their similar or even better performance, less tedious procedure for preparing the samples and less toxicity with respect to the chloroform mixtures.

Solvent mixtures have been suggested for lipid extraction after considering that most of the biological samples are composed of some amount of water and that the lipids are mainly soluble in organic solvents. Thus, the strategy for extractions of lipids has traditionally been based on two steps, a first one where a miscible solvent mixture (considering the water present in the sample) is added to the sample, which allows good interaction between solvent and sample matrix, and a second one where more aqueous or organic solvent or both are added to the original mixture in order to create a biphasic system which separates the extracted lipids, in the organic layer, from the rest of the sample, in the aqueous layer. However, considering that nowadays many workflows for lipidomics analysis include a reversed phase liquid chromatographic separation and polar interferences elute in the first minutes of the separation, the use of a biphasic system is not strictly required [19]. In 2017 Jurowski et al. [19] reviewed the use of some monophasic mixtures for lipidomics studies. Among the cited examples are CHCl_3 :MeOH (2:1 v/v) [20,21], 1-butanol/methanol (1:1 v/v) [22] for plasma samples, CHCl_3 :MeOH (1:2 v/v) for sphingolipids analysis in mammalian cells [23,24] and aqueous isopropanol for extraction of lipids on microalgae [25]. In 2014, Sarafian et al. [26] published a study showing a monophasic mixture IPA- H_2O as a good choice for lipid extraction of plasma samples with good recoveries for most of the lipid classes.

In this study, we evaluated the performance of monophasic

isopropanol extraction in comparison to biphasic extraction protocols in detail for lipid extraction from HeLa Cells. No information about the suitability of this monophasic extraction protocol for mammalian cells and how it compares to classical extraction protocols was available. Thus, two IPA- H_2O mixtures (75:25 v/v and 90:10 v/v) were compared with the biphasic extraction systems: CHCl_3 -MeOH- H_2O (2:2:1.8 v/v/v, Bligh & Dyer) [7] and MTBE:-MeOH: H_2O (10:3:2.5, v/v/v, Matyash) [10], which currently are two of the most widely employed protocols for lipid analysis [8,11,16,17,27–30].

2. Materials and methods

2.1. Materials

Mobile phases were prepared with solvents of LC-MS grade. Methanol (MeOH), acetonitrile (ACN) and isopropanol (IPA) were supplied by Roth (Karlsruhe, Germany). As additive, formic acid (FA, 98%) was obtained by Carl Roth (Karlsruhe, Germany) and ammonium formate was purchased from Sigma-Aldrich (Steinheim, Germany). Water was purified by a water filtration system from Elga (High Wycombe, United Kingdom).

Solvents for extraction were of HPLC grade: chloroform (CHCl_3 , $\geq 99.8\%$) and *tert*-butyl methyl ether (MTBE, anhydrous, 99.8%) from Sigma-Aldrich.

SPLASH™ Lipidomix® solution containing the following isotopically labeled internal standards (ILIS): 15:0–18:1(d7) PC, 15:0–18:1(d7) PE, 15:0–18:1(d7) PS, 15:0–18:1(d7) PG, 15:0–18:1(d7) PI, 15:0–18:1(d7) PA, 18:1(d7) LPC, 18:1(d7) LPE, 18:1(d7) Chol Ester, 18:1(d7) MG, 15:0–18:1(d7) DG, 15:0–18:1(d7)-15:0 TG, 18:1(d9) SM, Cholesterol (d7) was obtained from Avanti Polar Lipids (Alabama, USA) (See Table A1 for more information about internal standards).

2.2. Cell culture

The human cervical cancer HeLa cells adapted to serum free conditions (AC free, ECACC 08011102) were grown in T75-flask and EX-CELL HeLa serum free media (Sigma Aldrich) supplemented with 2 mM L-glutamine (Sigma Aldrich), until they reached a density of around $2\text{--}3 \times 10^6$ cells/mL. Afterwards they were transferred to 50 mL falcon tubes and centrifuged for 5 min at 700 rcf, before the supernatant was discarded. The cell pellet was resuspended in

15 mL ice-cold Dulbecco's Phosphate Buffered Saline (PBS, Sigma Aldrich) for washing and centrifuged again for 5 min at 700 rcf. The washing was repeated twice and after the last resuspension cells were counted in triplicate with a hemocytometer. According to the mean of the counted cells, aliquots containing approximately 8×10^5 cells were transferred randomly to 15 mL falcon tubes (for extraction with isopropanol mixtures) and 15 mL glass tubes (for extractions with CHCl_3 and MTBE). Samples were centrifuged at 700 rcf for 5 min and the pellets were stored at -80°C until extraction.

2.3. Extraction protocols

Extraction solvents were kept on ice before their addition to the samples. For each extraction protocol, 10 samples (pellets containing 8×10^5 cells) were extracted. In order to estimate the recovery of some lipid classes, 5 of these samples were spiked before the extraction with 100 μL of 5% v/v SPLASH Lipidomix solution in methanol and were resuspended with 100 μL MeOH before the LC-MS measurement (pre-extraction spiking). For the other 5 samples 100 μL of MeOH were added before the extraction and they were resuspended with 100 μL of 5% v/v SPLASH Lipidomix solution before the LC-MS measurement (post-extraction spiking). Recoveries were calculated as ratio of average intensities for internal standards in the pre-extraction and post extraction spiked samples.

2.3.1. Extraction with MTBE:MeOH:H₂O (10:3:2.5, v/v/v, "MTBE")

This EP was based on Matyash et al. [10]. Either methanol or solution of 5% SPLASH Lipidomix in MeOH (100 μL) was added to the pellet. Then 1.4 mL of methanol and 5 mL of MTBE were added. Vortexing (30 s), ultrasonication (2 min) and vortexing (30 s) cycle was applied to disrupt the pellet. Samples were incubated on ice while shaking (500 rpm, 60 min). After the extraction, 1.25 mL of H₂O was added and samples were incubated on ice for another 10 min. Afterwards, centrifugation at 3500 rcf for 10 min was applied and the upper layer was transferred to a glass tube. Samples were dried in an evaporator (Genevac EZ-2; Warminster, Pennsylvania, USA) for 10 h under nitrogen protection. The lipid extract was resuspended in 100 μL of either methanol or solution of 5% SPLASH Lipidomix while sonication (2 min) and vortexing (30 s) were applied. Centrifugation at 3500 rcf for 10 min was applied and the supernatant was transferred to vials for MS-measurements. During the last step, 10 μL aliquot of each sample were transferred to a separate vial to prepare a pooled QC sample.

2.3.2. Extraction with IPA:H₂O (75:25 v/v, "IPA75")

Either methanol or solution of 5% SPLASH Lipidomix in MeOH (100 μL) was added to the pellet. Then 5.0 mL of IPA:H₂O (75:25 v/v) were added. Vortexing (30 s), ultrasonication (2 min) and vortexing (30 s) cycle was applied to disrupt the pellet. Samples were incubated on ice while shaking (500 rpm, 60 min). After the extraction, centrifugation at 3500 rcf for 10 min was applied and supernatant was transferred to a 15 mL falcon tube. Samples were dried in an evaporator for 10 h under nitrogen protection. The lipid extract was resuspended in 100 μL of either methanol or solution of 5% SPLASH Lipidomix while sonication (2 min) and vortexing (30 s) were applied. Centrifugation at 3500 rcf for 10 min was applied and the supernatant was transferred to vials for MS-measurements. During the last step, 10 μL aliquot of each sample were transferred to a separate vial to prepare a pooled QC sample.

2.3.3. Extraction with IPA:H₂O (90:10 v/v, "IPA90")

As described before for extraction with IPA75, only that IPA:H₂O (90:10 v/v) was used instead of IPA:H₂O (75:25 v/v).

2.3.4. Extraction with CHCl₃-MeOH-H₂O (2:2:1.8 v/v/v, "CHCl₃")

This EP was based on Bligh & Dyer [7]. Either methanol or solution of 5% SPLASH Lipidomix in MeOH (100 μL) was added to the pellet. Then, 0.8 mL of H₂O, 1.90 mL of MeOH and 1.0 mL of CHCl_3 were added. Vortexing (30 s), ultrasonication (2 min) and vortexing (30 s) cycle was applied to disrupt the pellet. Samples were incubated on ice while shaking (500 rpm, 60 min). After the extraction 1.0 mL CHCl_3 and 1.0 mL of H₂O were added and samples were incubated on ice for another 10 min. Afterwards, centrifugation at 3500 rcf for 10 min was applied and the upper layer was transferred to a glass tube. Samples were dried in an evaporator for 10 h under nitrogen protection. The lipid extract was resuspended in 100 μL of either methanol or solution of 5% SPLASH Lipidomix while sonication (2 min) and vortexing (30 s) were applied. Centrifugation at 3500 rcf for 10 min was applied and the supernatant was transferred to vials for MS-measurements. During the last step, 10 μL aliquot of each sample were transferred to a separate vial to prepare a pooled QC sample.

2.3.5. Blank extractions for all extraction protocols

Seven blank extraction replicates for each extraction protocol were performed following the same steps indicated above, only that extraction solvents were added to empty falcon or glass tubes. For further analysis, an additional IPA90 blank extraction protocol (also 7 replicates) was performed replacing plasticware with glassware. The results from these blank samples were used to correct the result from the cell extractions.

2.3.6. Extractions without internal standards

One cell extract for each EP was prepared following the same steps described before, but no ILIS were added. These samples were used to validate assay specificity for internal standards i.e. to check that no significant signals are present at *m/z* and retention times corresponding to the ILIS. For all EPs no signals i.e. no interferences were found.

2.4. LC-MS measurement

All analyses were performed on an Agilent 1290 Series UHPLC instrument (Agilent, Waldbronn, Germany) coupled to a Sciex TripleTOF 5600 + mass spectrometer (Sciex, Concord, Ontario; Canada). Duospray ion source for ESI in positive and negative ion mode was used. Sample injections were done in randomized order with a Pal HTC-XS autosampler from CTC (Zwingen, Switzerland). QC samples were run at the beginning of the sequence, during the sequence (every five samples) and at the end of the sequence.

Chromatographic separation was performed according to conditions published by Tsugawa et al. [31] in order to enable retention time scoring for peak identification with MS-Dial software. Briefly an Acquity UPLC CSH C18 Column, 130 Å, 1.7 μm , 2.1 mm \times 100 mm with an Acquity UPLC CSH C18 VanGuard Pre-column, 130 Å, 1.7 μm , 2.1 mm \times 5 mm (Waters, Eschborn, Germany) was used. The mobile phase was composed of 10 mM ammonium formate and 0.1% formic acid in A) 60:40 ACN:H₂O (v/v) and B) 90:10 (v/v) IPA:ACN. The following gradient profile was used: 0.00 min, 15% B; 2.00 min, 30% B; 2.50 min, 48% B; 11.00 min, 82% B; 11.50 min, 99% B; 12.00 min, 99% B; 12.10 min, 15% B, 15.00 min, 15% B. Flow rate was 600 $\mu\text{L}/\text{min}$ and column temperature was 65°C . The injection volumes were 3 μL and 5 μL for positive and negative mode, respectively.

The following MS-settings of the mass spectrometer were used: Curtain gas (CUR) 35 psi, nebulizer gas (GS1) 60 psi, drying gas (GS2) 60 psi, ion-spray voltage floating (ISVF) +5500 V in positive and -4500 V in negative mode, source temperature (T) 350°C , collision energy 45 V, collision energy spread 15 V, declustering

potential (DP) 80 V, mass range m/z 50–1250 in ESI (+) and 50–1050 in ESI (–), and RF Transmission (RF) m/z 40: 50% and m/z 120: 50%. An external mass calibration was performed every five samples (see Table A2).

MS data was obtained by using sequential window acquisition of all theoretical fragment ion spectra (SWATH) [32,33], after optimizing Q1 windows size (See Table A3) with Swath Tuner software [34].

2.5. MS data processing

MS-Dial software (RIKEN, version 3.06) [31] was used to process the MS data (see parameters in Table A4). This included detection of peaks, MS² data deconvolution, compound identification and alignment of peaks through all the samples. For identification a cut off value of 80% was selected: This value is based on 6 different similarity scores: 1 for retention time, 1 for m/z , 1 for isotopic pattern and 3 for MS/MS (dot product, dot product reversed and presence). An important condition established in MS-Dial was that a peak was selected for alignment only when it was present in at least 51% of the samples of one extraction protocol. Features which were relatively close (m/z difference less than 0.03 Da and retention time difference less than 0.1 min) in the alignment file of MS-Dial were visually inspected in order to determine if they are effectively corresponding to more than one feature, otherwise the repeated feature was removed. List of aligned peaks from MS-Dial were further evaluated with Multiquant 3.0 (Sciex). Intensities were processed for principal component analysis (PCA) with MarkerView (Sciex) and exported to Excel for statistical analysis. A feature was considered for comparison between the different extraction protocols when it was present in at least 90% of the samples of one extraction protocol having a CV less than 30% for the 10 extraction replicates of that protocol. Furthermore blank extraction samples were used to exclude features that have m/z difference less than 0.01 Da, retention time difference less than 0.5 min and fold change less than 5 between the cell extraction and

blank extraction replicates (List of detected features in extraction blanks are in Appendix D).

Peaks corresponding to internal standards were removed from MS-Dial detected features and were analyzed directly with Multiquant to evaluate the recoveries.

3. Results and discussion

In order to compare the different EPs some modifications were introduced to the originally published protocols. Thus, extraction volumes were modified to be as similar as possible for all EPs while keeping the solvent ratios for each EP as they were published. Variables like temperature, vortexing time, vortexing intensity, centrifugation time and centrifugation speed were kept the same for all extraction protocols. For the same reason no re-extractions were done for any of the extraction protocols. The solvent evaporation step was performed overnight to save time and it was kept at 10 h even when the time required for the evaporation of each solvent ranges from less than 3 h, in the case of CHCl₃ and MTBE, to approximately 6 h, in the case of IPA75. Thus, all the extracts were kept under nitrogen atmosphere until resuspension to minimize possible oxidation of lipids.

3.1. Chromatograms and principal component analysis (PCA)

Total ion current (TIC) chromatograms in positive and negative mode for extracted samples with different EPs (Fig. 1) show a very similar profile with only some slight differences, especially during the first minutes. After processing the data with MS-Dial and Multiquant, peak intensities for detected features were analyzed by PCA (Fig. 2). Score plots with the first two principal components, in both ESI (+) and ESI (–) mode (Fig. 2A and C), show a clear separation between the samples extracted with each protocol, except for extractions with isopropanol 75% and isopropanol 90%, which are overlapped. This result indicates that the detected features and their corresponding intensities show

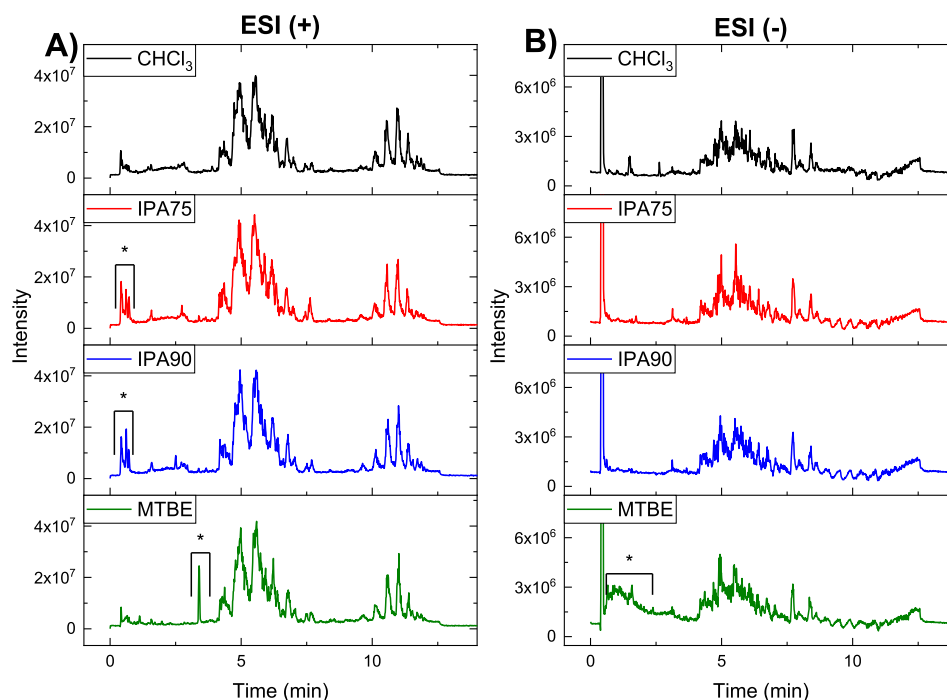


Fig. 1. Representative TICs for samples extracted with four different EPs using A) ESI (+) and B) ESI (–) mode. Asterisks indicate major differences between the chromatograms.

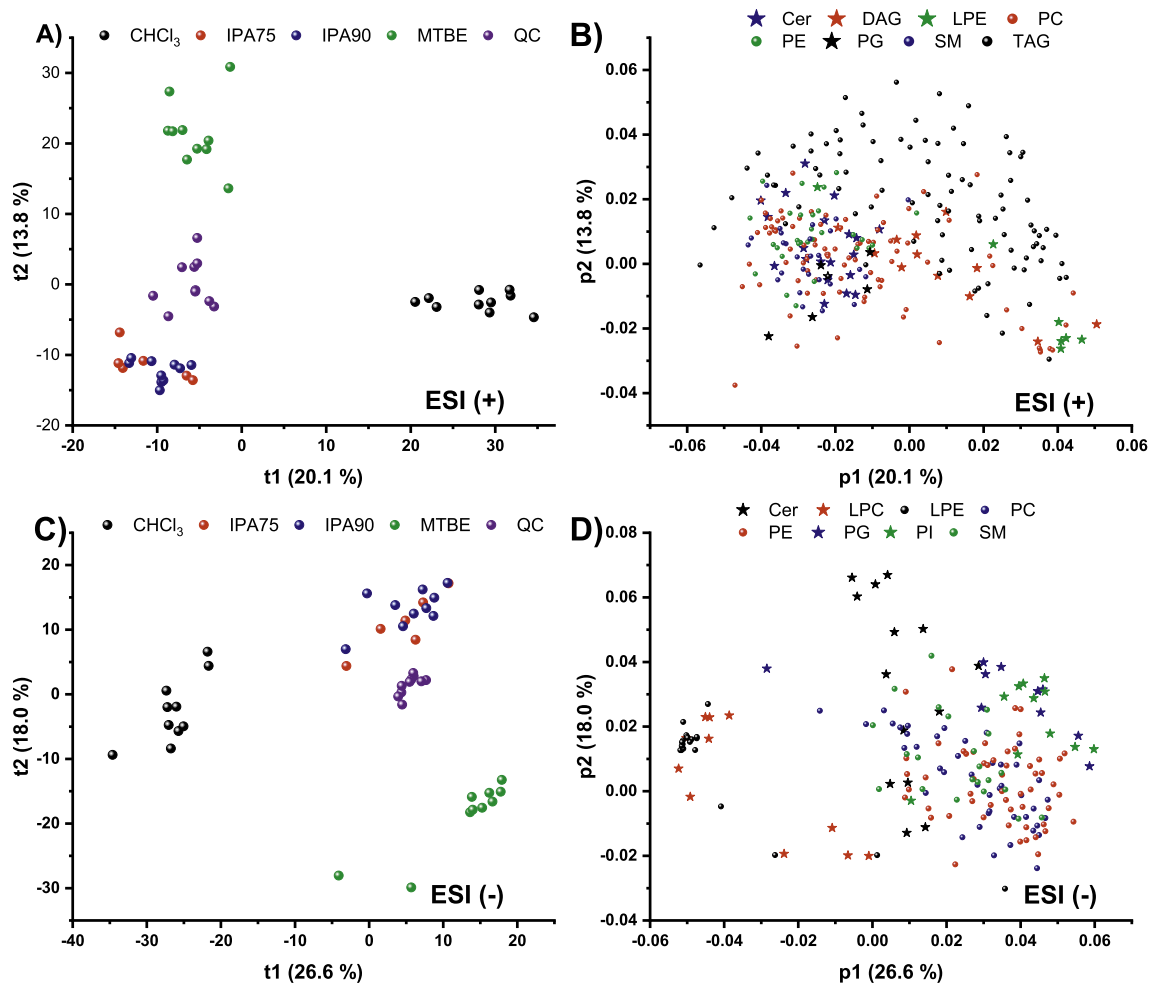


Fig. 2. PCA plots for intensities of detected features. A) Scores plot in ESI (+), B) Loadings plot including only lipid classes with more than 5 identified features ESI (+) C) Scores plot in ESI (-), D) Loadings plot including only lipid classes with more than 5 identified features ESI (-).

significant alterations between the different extraction protocols. Loadings plots, with the first two principal components, are shown in Fig. 2B and D, for positive and negative mode respectively (for a better visualization, unknown features and lipid classes with less than five identified features were excluded from each loadings plot). In Fig. 2B it is possible to note influences of some lipid classes to the shown differentiation of EPs in Fig. 2A, for example LPE are oriented in the direction of CHCl_3 and PGs are oriented in the direction of the EPs IPA75 and IPA90, which means that those lipid classes are better extracted with the mentioned EPs. In the same manner, in Fig. 2D, LPCs and LPEs are oriented in the same direction of the EP CHCl_3 and Cers are oriented in the direction of the EPs IPA75 and IPA90.

3.2. Number of detected features

Table 1 and Fig. 3 show the numbers and distribution of features detected in ESI (+) and ESI (-) for the EPs (Specific data about detected features can be observed in Appendix B). Features list obtained after processing with MS-Dial was reprocessed with Multiquant. In this manner we are combining the capabilities of MS-Dial for recognizing features and the identification of approximately 15% of them with the capabilities of Multiquant for a more controlled integration, making it easier to determine whether a feature is present or not in a group of samples. After the processing with Multiquant and the feature filtration (only those features

which are present in at least 90% of the samples of one EP and have a CV less than 30% for the replicates of that EP were selected) the number of features was reduced from 1872 to 1382 in ESI (+) and from 1541 to 1024 in ESI (-) but the number of features which were found with each EP was higher (See Table 1). Similar procedures with MS-Dial and Multiquant were applied to blank extraction replicates and peak exclusion lists were used to filter only the features coming from the cell pellets (1167 features in ESI (+) and 842 features in ESI (-)). In terms of the total number of detected features, all extractions protocols show similar performance in ESI (+) (maximum difference is 4% between IPA90 and IPA75) and a slight greater amount is obtained with CHCl_3 in ESI (-) (difference of 9% respect to the MTBE protocol). Venn diagrams in Fig. 3 show the distribution of detected features in ESI (+) and ESI (-) modes. It is possible to see that 72% of the features in ESI (+) and 77% in negative mode were detected with all EPs. In ESI (+) the amount of features that can be exclusively extracted with each extraction protocol is very similar (6.7% with CHCl_3 , 5.7% in common with IPA75 and IPA90 and 5.1% with MTBE).

In ESI (-), CHCl_3 protocol extracts exclusively 9.7% of total detected features in comparison to 3.4% of MTBE and 2.6% of IPA75 and IPA90 (in common). These features detected with only one of the specific EP (or two of them in the case of IPA75 and IPA90) correspond to not identified features which are spread through the whole studied range of m/z and retention time and for this reason no specific lipid class can be assigned to them.

Table 1
Description of processed features with MS-Dial and Multiquant softwares.

Description	ESI (+)					ESI (-)				
	Total	CHCl ₃	IPA75	IPA90	MTBE	Total	CHCl ₃	IPA75	IPA90	MTBE
Detected features in MS-Dial ^a	1872	904	893	928	1038	1541	708	674	718	1015
Detected features after processing with Multiquant ^b	1382	1110	1054	1118	1183	1074	824	759	771	906
Detected features after correction with blanks	1167	991 (85%)	955 (82%)	1003 (86%)	985 (84%)	842	777 (92%)	713 (85%)	725 (86%)	701 (83%)
Identified features	292	289 (99%)	285 (97.6%)	290 (99.3%)	291 (99.7%)	206	205 (99.5%)	201 (97.6%)	205 (99.5%)	200 (97.1%)
• SM	29	29	28	29	29	22	22	22	22	22
• PG	6	6	5	6	6	10	10	10	10	10
• PE	29	29	29	29	29	55	55	54	55	55
• PC	88	88	88	87	87	43	43	43	43	43
• LPE	7	7	7	7	7	22	22	20	22	20
• LPC	4	4	4	4	4	11	11	11	11	10
• Cer-NS	12	12	10	12	12	6	6	6	6	6
• Cer-NDS	5	5	5	5	5	2	2	2	2	2
• TG	92	89	89	91	92					
• HexCer-NS	3	3	3	3	3	1	1	1	1	1
• DG	12	12	12	12	12					
• CE	2	2	2	2	2					
• BMP	3	3	3	3	3					
• PS	-	-	-	-	-	1	1	1	1	1
• PI	-	-	-	-	-	11	11	10	11	11
• OxPS	-	-	-	-	-	1	1	1	1	1
• OxPG	-	-	-	-	-	1	1	1	1	1
• OxPE	-	-	-	-	-	6	6	5	5	5
• OxPC	-	-	-	-	-	3	3	3	3	3
• LPI	-	-	-	-	-	2	2	2	2	2
• LPG	-	-	-	-	-	2	2	2	2	2
• HexCer-NDS	-	-	-	-	-	5	5	5	5	5
• FA	-	-	-	-	-	2	1	2	2	0

^a Present in at least 51% of samples of one group.^b Present in at least 90% of samples of one EP having a CV < 30%.

3.3. Identified lipids and relative recovery of endogenous lipids

Table 1 also shows the distribution of lipids identified with each EP (Specific data about identified features can be observed in Appendix C). Only IPA75 and MTBE extraction allowed the identification of a few less features (2% less in each mode for IPA75 and 2% less in negative mode for MTBE), it means that in terms of the number of identified features the four studied EPs have similar performance. However, Fig. 2B and D, already showed that even when similar number of lipids were detected with each EP, their intensities were significantly different for some lipid classes.

In order to describe the relative ability of each EP to extract a particular lipid class, a relative recovery of endogenous extracted

lipids was calculated using IPA90 as a reference. Thus, the average intensity of each identified lipid extracted with CHCl₃, IPA75 and MTBE was normalized with respect to the average intensity of the same lipid after extraction with IPA90. Afterwards, the normalized values were averaged for the lipids belonging to the same lipid class (Fig. 4, Table A5). The results show similar intensities for some lipid classes independently of which protocol was employed. However, in some other cases significant differences are noted. In ESI (-) mode, for example, significantly lower intensities were obtained for polar lipids LPG, LPI, PG, PI, PS and FA when CHCl₃ and MTBE extractions were employed. Also, higher intensities are achieved for LPE and LPC with CHCl₃ protocol while lower ones are obtained with IPA75 and MTBE. This higher recovery for LPC and LPE with

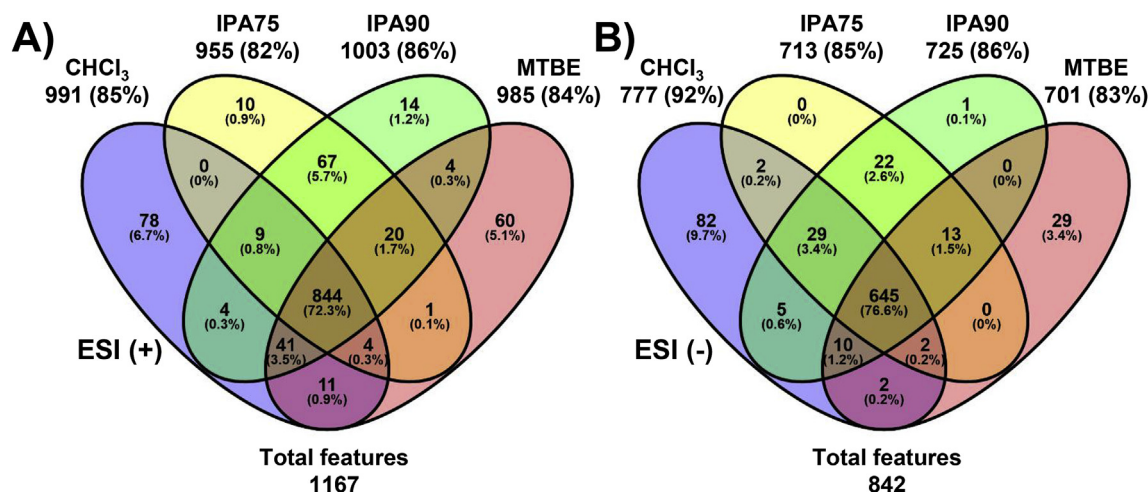


Fig. 3. Venn diagrams showing the distribution of detected features with each EP and their overlapping selectivities. A) Results from ESI (+) and B) ESI (-) modes.

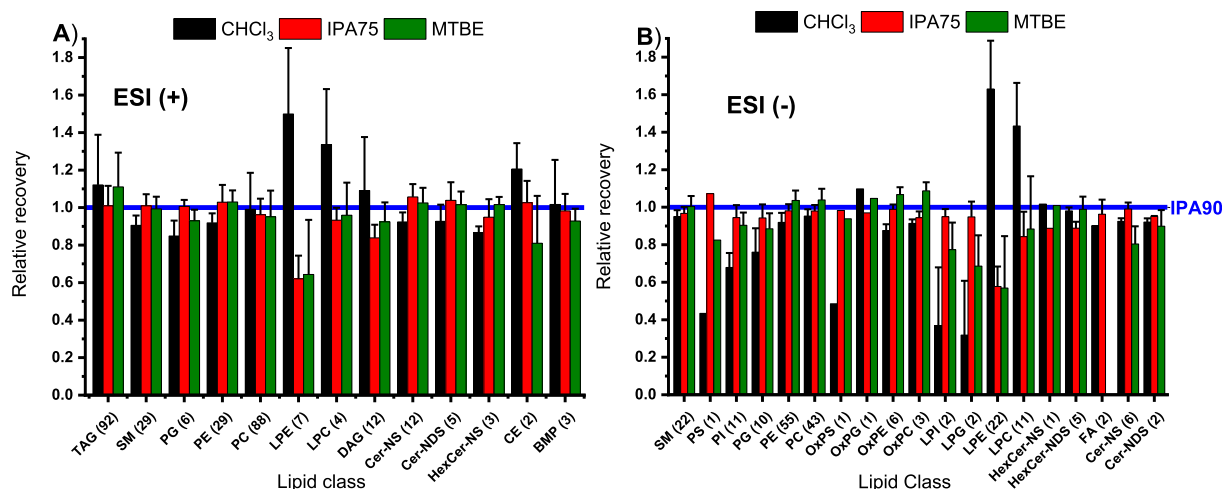


Fig. 4. Relative lipid class recoveries of endogenous lipids obtained with different EPs and using IPA90 as a reference. A) Results from ESI (+) and B) ESI (-) modes. Numbers in parenthesis indicate the number of identified lipids from each lipid class that were normalized and averaged.

CHCl_3 has to be further investigated considering that these are two of the most polar lipid classes and it is expected that polar mixtures IPA:H₂O can extract better these substances, as it has been indicated previously [26].

3.4. Lipid recovery of internal standards

To determine the absolute recovery of lipid classes with each EP, pre-extraction and post-extraction spiking of the samples with a mixture of ILIS was performed. Fig. 5 and Table A6 show the % recovery for each ILIS. As an average MTBE and IPA90 have the best performance. In the case of CHCl_3 , its performance is significantly higher than other protocols for the recovery of TG. IPA75 shows recoveries in most of the cases lower than the ones that can be reached with IPA90. CHCl_3 protocols shows significantly lower recoveries for polar lipid classes PA, PI and PS.

Here, it is important to highlight that beside the fact that ILIS were added to the pellet just before the addition of extraction solvent, which means they have less interaction with cellular matrix and more direct contact with extraction solvent than endogenous lipids, a good correlation was observed for the recoveries that were calculated with the ILIS in comparison with the relative

recoveries of endogenous lipids described in section 3.3. Only exception for this finding was the anomalous mentioned case of LPEs and LPCs.

Also, good correlation was observed between the obtained recoveries and reported results by Sarafian et al. [26] in plasma samples after extraction with isopropanol and relative comparison with protocols based on MTBE (Matyash) and CHCl_3 (Bligh and Dyer, Folch). However, only recovery of odd-chained internal standards was reported at that moment and no comparison with endogenous lipids was done.

Higher recoveries obtained with IPA90 respect to IPA75, especially for the most abundant lipid classes are also in good concordance with comparison done by Yao et al. [25] in microalgae and soybeans, where aqueous isopropanol mixtures with 88% and 95% of isopropanol yielded higher oil extraction efficiency than mixtures with 50% and 70% isopropanol.

3.5. Precision

For a comparison of precisions obtained with the EPs, CV % for each detected feature was calculated. A profile of the CVs for the features that were found with each EP (1167 features in ESI (+) and

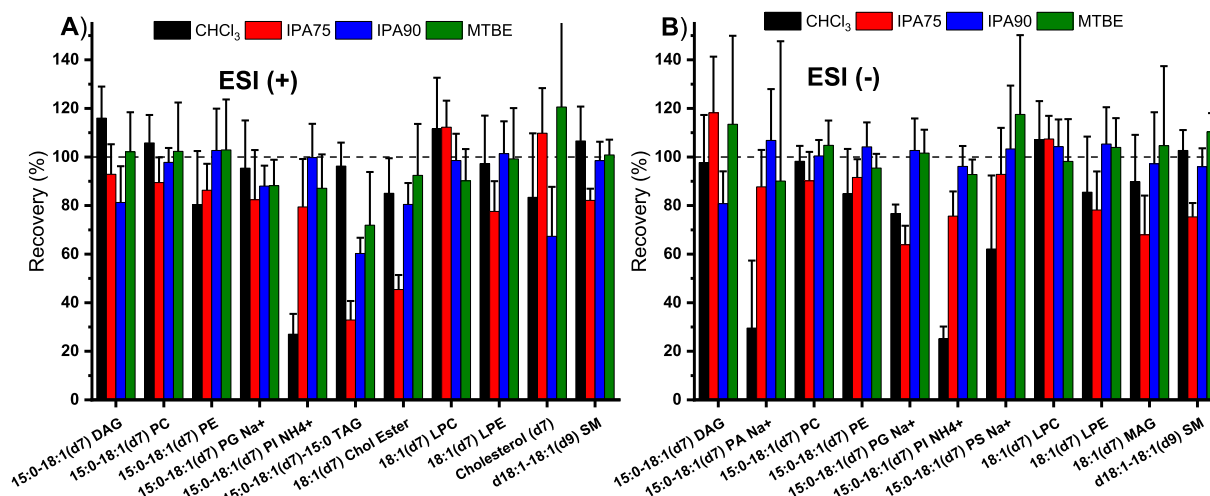


Fig. 5. Recoveries of isotopically labeled internal standards after pre and post-extraction spiking. A) Results from ESI (+) and B) ESI (-) modes.

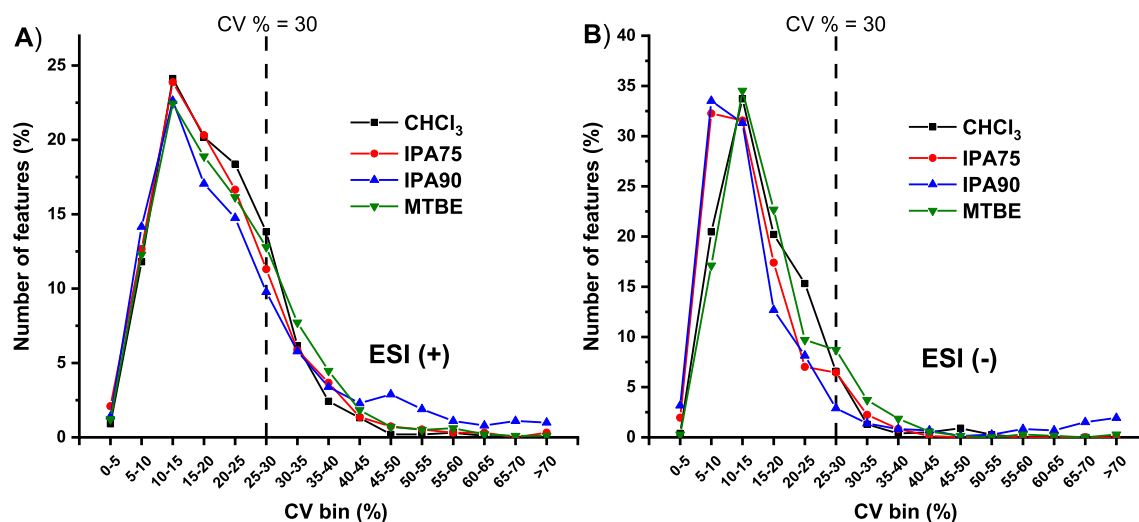


Fig. 6. Distribution of CVs (%) obtained for precision evaluation of peak intensities of features detected in all four EPs. A) Results from ESI (+) (1167 analyzed features) and B) ESI (-) modes (842 analyzed features).

842 features in ESI (-), see Fig. 3) is shown in Fig. 6 and Table A7. For features detected in ESI (+), CV profiles show to be very similar having a maximum CV % bin from 10 to 15% and around 80% of the features with CV less than 30% independently of the employed EP. For features detected with ESI (-), IPA75 and IPA90 have a maximum CV bin from 5 to 10%, while CHCl_3 and MTBE have a maximum CV bin from 10 to 15%. All EPs have more than 90% of detected features with a CV below 30%.

3.6. Protocol complexity

Many publications have already made emphasis about the advantage of using extraction protocols where the organic layer is the upper phase (as in MTBE protocol) and not the lower one (as in CHCl_3 protocol) of a two-phase partitioning system. The reason for this is the more tedious removal of the lipid containing organic phase when this one is in the bottom, especially because a layer of protein is located between the organic and aqueous layer. In the case of tested EPs IPA75 and IPA90, the fact of having a monophasic mixture makes this process even easier because only a separation of the supernatant from the solid residue is required. A possible disadvantage of employing monophasic mixtures for extraction relies on the presence of salts as part of the extract. However, in reversed phase LC-MS this is not necessarily a problem because these salts elute during the first minutes of chromatographic run. As part of this research, QC samples run at the beginning, in between and at the end of measured sequences in ESI (+) and ESI (-), did not show significant differences in terms of intensity (see Fig. A1).

Another advantage for IPA75 and IPA90 extractions, is that they are compatible with plasticware, which does not happen with MTBE and CHCl_3 . In our study, extraction blanks were performed to remove those features which are not coming from cell pellets (lists of features detected in extractions blanks are in Appendix D). Additionally, an extra extraction blank for IPA90 using glassware, yielded a higher amount of detected features than the corresponding one using plasticware (See Fig A2 and A3). This higher amount of detected features could be related with the cleaning process employed for the glass tubes. Therefore, in the case that glassware is employed for these extractions, either new glassware has to be utilized for each new extraction, which is very expensive or a very strict cleaning protocol has to be employed, which will

demand extra time and make the process more tedious. Consequently, it can be said that the extraction protocols IPA75 and IPA90, employing plasticware, are less time consuming, cheaper and easy to automatize.

4. Concluding remarks

After comparison of the performance for extraction of lipids, it is possible to conclude that there was no significant difference with the number of identified lipids with each EP. In terms of recoveries for different lipid classes, extraction with IPA90 showed similar results as MTBE for most of the lipid classes, better results than CHCl_3 for the more polar lipid classes and better results than IPA75 for the less polar lipid classes. Precision with IPA mixtures showed to be slightly better in ESI (-) and similar in ESI (+) than the precision obtained with MTBE and CHCl_3 . In terms of complexity, monophasic extractions with IPA offered a simpler, less time consuming and cheaper protocol. Also, MS signal intensities did not show any decrease after samples extracted with IPA:H₂O mixtures were measured, which could be corroborated with the reproducibility of measured intensities for QC samples through the whole study. Considering all these aspects, extraction with IPA90 represents an excellent alternative as a solvent for developing reversed-phase LC-MS lipidomics studies.

Conflicts of interest

None.

Acknowledgements

C.C. is grateful to the DAAD for a doctoral fellowship (DAAD no. 57129429). We are grateful to the "Struktur-und Innovationsfonds für die Forschung (SI-BW)" by the regional government of Baden-Württemberg (Ministry of Science, Research and Arts) and the German Science Foundation (DFG no. INST 37/821-1 FUGG) for financial support of infrastructure.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.aca.2018.10.035>.

References

- [1] T. Cajka, O. Fiehn, Comprehensive analysis of lipids in biological systems by liquid chromatography-mass spectrometry, *TrAC Trends Anal. Chem.* (Reference Ed.) 61 (2014) 192–206, <https://doi.org/10.1016/j.trac.2014.04.017>.
- [2] X. Han, Lipidomics for studying metabolism, *Nat. Rev. Endocrinol.* 12 (2016) 668–679, <https://doi.org/10.1038/nrendo.2016.98>.
- [3] M. Wang, C. Wang, R.H. Han, X. Han, Novel advances in shotgun lipidomics for biology and medicine, *Prog. Lipid Res.* 61 (2016) 83–108, <https://doi.org/10.1016/j.plipres.2015.12.002>.
- [4] M. Holcapek, Lipidomics, *Anal. Bioanal. Chem.* 407 (2015) 4971–4972, <https://doi.org/10.1007/s00216-015-8740-0>.
- [5] S. Tumanov, J.J. Kamphorst, Recent advances in expanding the coverage of the lipidome, *Curr. Opin. Biotechnol.* 43 (2017) 127–133, <https://doi.org/10.1016/j.copbio.2016.11.008>.
- [6] J. Folch, M. Lees, G.H. Sloane Stanley, A simple method for the isolation and purification of total lipides from animal tissues, *J. Biol. Chem.* 226 (1957) 497–509, <http://www.ncbi.nlm.nih.gov/pubmed/13671378>.
- [7] E.G. Bligh, W.J. Dyer, A rapid method of total lipid extraction and purification, *Can. J. Biochem. Physiol.* 37 (1959) 911–917, <https://doi.org/10.1139/o59-099>.
- [8] C. Breil, M. Abert Vian, T. Zemb, W. Kunz, F. Chemat, “Bligh and Dyer” and Folch methods for solid–liquid–liquid extraction of lipids from microorganisms. Comprehension of solvation mechanisms and towards substitution with alternative solvents, *Int. J. Mol. Sci.* 18 (2017) 708, <https://doi.org/10.3390/ijms18040708>.
- [9] T.A. Lydic, J. V Busik, G.E. Reid, A monophasic extraction strategy for the simultaneous lipidome analysis of polar and nonpolar retina lipids, *J. Lipid Res.* 55 (2014) 1797–1809, <https://doi.org/10.1194/jlr.D050302>.
- [10] V. Matyash, G. Liebisch, T.V. Kurzchalia, A. Shevchenko, D. Schwudke, Lipid extraction by methyl-tert-butyl ether for high-throughput lipidomics, *J. Lipid Res.* 49 (2008) 1137–1146, <https://doi.org/10.1194/jlr.D700041-JLR200>.
- [11] R.E. Patterson, A.J. Ducrocq, D.J. McDougall, T.J. Garrett, R.A. Yost, Comparison of blood plasma sample preparation methods for combined LC–MS lipidomics and metabolomics, *J. Chromatogr. B* 1002 (2015) 260–266, <https://doi.org/10.1016/j.jchromb.2015.08.018>.
- [12] C.Z. Ulmer, R.A. Yost, T.J. Garrett, Global UHPLC/HRMS Lipidomics Workflow for the Analysis of Lymphocyte Suspension Cultures, 2017, pp. 175–185, https://doi.org/10.1007/978-1-4939-6946-3_13.
- [13] H. Zhang, Y. Gao, J. Sun, S. Fan, X. Yao, X. Ran, C. Zheng, M. Huang, H. Bi, Optimization of lipid extraction and analytical protocols for UHPLC-ESI-HRMS-based lipidomic analysis of adherent mammalian cancer cells, *Anal. Bioanal. Chem.* 409 (2017) 5349–5358, <https://doi.org/10.1007/s00216-017-0483-7>.
- [14] A. Hara, N.S. Radin, Lipid extraction of tissues with a low-toxicity solvent, *Anal. Biochem.* 90 (1978) 420–426, [https://doi.org/10.1016/0003-2697\(78\)90046-5](https://doi.org/10.1016/0003-2697(78)90046-5).
- [15] T.W.T. Rupasinghe, Lipidomics: Extraction Protocols for Biological Matrices, 2013, pp. 71–80, https://doi.org/10.1007/978-1-62703-577-4_6.
- [16] A. Reis, A. Rudnitskaya, G.J. Blackburn, N.M. Fauzi, A.R. Pitt, C.M. Spickett, A comparison of five lipid extraction solvent systems for lipidomic studies of human LDL, *J. Lipid Res.* 54 (2013) 1812–1824, <https://doi.org/10.1194/jlr.M034330>.
- [17] J. Sheng, R. Vannela, B.E. Rittmann, Evaluation of methods to extract and quantify lipids from *Synechocystis* PCC 6803, *Bioresour. Technol.* 102 (2011) 1697–1703, <https://doi.org/10.1016/j.biortech.2010.08.007>.
- [18] L. Löfgren, M. Ståhlman, G.-B. Forsberg, S. Saarinen, R. Nilsson, G.I. Hansson, The BUMÉ method: a novel automated chloroform-free 96-well total lipid extraction method for blood plasma, *J. Lipid Res.* 53 (2012) 1690–1700, <https://doi.org/10.1194/jlr.D023036>.
- [19] K. Jurowski, K. Kochan, J. Walczak, M. Barańska, W. Piekoszewski, B. Buszewski, Comprehensive review of trends and analytical strategies applied for biological samples preparation and storage in modern medical lipidomics: state of the art, *TrAC Trends Anal. Chem.* (Reference Ed.) 86 (2017) 276–289, <https://doi.org/10.1016/j.trac.2016.10.014>.
- [20] J.M. Weir, G. Wong, C.K. Barlow, M.A. Greeve, A. Kowalczyk, L. Almasy, A.G. Comuzzie, M.C. Mahaney, J.B.M. Jowett, J. Shaw, J.E. Curran, J. Blangero, P.J. Meikle, Plasma lipid profiling in a large population-based cohort, *J. Lipid Res.* 54 (2013) 2898–2908, <https://doi.org/10.1194/jlr.P035808>.
- [21] P.J. Meikle, G. Wong, D. Tsorotes, C.K. Barlow, J.M. Weir, M.J. Christopher, G.L. MacIntosh, B. Goudey, L. Stern, A. Kowalczyk, I. Haviv, A.J. White, A.M. Dart, S.J. Duffy, G.L. Jennings, B.A. Kingwell, Plasma lipidomic analysis of stable and unstable coronary artery disease, *Arterioscler. Thromb. Vasc. Biol.* 31 (2011) 2723–2732, <https://doi.org/10.1161/ATVBAHA.111.234096>.
- [22] Z.H. Alshehry, C.K. Barlow, J.M. Weir, Y. Zhou, M.J. McConville, P.J. Meikle, An efficient single phase method for the extraction of plasma lipids, *Metabolites* 5 (2015) 389–403, <https://doi.org/10.3390/metabo5020389>.
- [23] R.L. Shaner, J.C. Allegood, H. Park, E. Wang, S. Kelly, C.A. Haynes, M.C. Sullards, A.H. Merrill, Quantitative analysis of sphingolipids for lipidomics using triple quadrupole and quadrupole linear ion trap mass spectrometers, *J. Lipid Res.* 50 (2009) 1692–1707, <https://doi.org/10.1194/jlr.D800051-JLR200>.
- [24] M.C. Sullards, Y. Liu, Y. Chen, A.H. Merrill, Analysis of mammalian sphingolipids by liquid chromatography tandem mass spectrometry (LC-MS/MS) and tissue imaging mass spectrometry (TIMS), *Biochim. Biophys. Acta - Mol. Cell Biol. Lipids.* 1811 (2011) 838–853, <https://doi.org/10.1016/j.bbalip.2011.06.027>.
- [25] L. Yao, S.L. Lee, T. Wang, J.A. Gerde, Comparison of lipid extraction from microalgae and soybeans with aqueous isopropanol, *J. Am. Oil Chem. Soc.* 90 (2013) 571–578, <https://doi.org/10.1007/s11746-012-2197-5>.
- [26] M.H. Sarafian, M. Gaudin, M.R. Lewis, F.-P. Martin, E. Holmes, J.K. Nicholson, M.-E. Dumas, Objective set of criteria for optimization of sample preparation procedures for ultra-high throughput untargeted blood plasma lipid profiling by ultra performance liquid chromatography–mass spectrometry, *Anal. Chem.* 86 (2014) 5766–5774, <https://doi.org/10.1021/ac500317c>.
- [27] J. Sostare, R. Di Guida, J. Kirwan, K. Chalal, E. Palmer, W.B. Dunn, M.R. Viant, Comparison of modified Matyash method to conventional solvent systems for polar metabolite and lipid extractions, *Anal. Chim. Acta.* 1037 (2018) 301–315, <https://doi.org/10.1016/j.aca.2018.03.019>.
- [28] C. Pizarro, I. Arenzana-Rámila, N. Pérez-del-Notario, P. Pérez-Matute, J.M. González-Sáiz, Plasma lipidomic profiling method based on ultrasound extraction and liquid chromatography mass spectrometry, *Anal. Chem.* 85 (2013) 12085–12092, <https://doi.org/10.1021/ac403181c>.
- [29] S.K. Jensen, Improved Bligh and dyer extraction procedure, *Lipid Technol.* 20 (2008) 280–281, <https://doi.org/10.1002/lite.200800074>.
- [30] M.L. Dória, Z. Cotrim, B. Macedo, C. Simões, P. Domingues, L. Helguero, M.R. Domingues, Lipidomic approach to identify patterns in phospholipid profiles and define class differences in mammary epithelial and breast cancer cells, *Breast Canc. Res. Treat.* 133 (2012) 635–648, <https://doi.org/10.1007/s10549-011-1823-5>.
- [31] H. Tsugawa, T. Cajka, T. Kind, Y. Ma, B. Higgins, K. Ikeda, M. Kanazawa, J. VanderGheynst, O. Fiehn, M. Arita, MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis, *Nat. Methods* 12 (2015) 523–526, <https://doi.org/10.1038/nmeth.3393>.
- [32] G. Hopfgartner, D. Tonoli, E. Varesio, High-resolution mass spectrometry for integrated qualitative and quantitative analysis of pharmaceuticals in biological matrices, *Anal. Bioanal. Chem.* 402 (2012) 2587–2596, <https://doi.org/10.1007/s00216-011-5641-8>.
- [33] L.C. Gillet, P. Navarro, S. Tate, H. Röst, N. Selevsek, L. Reiter, R. Bonner, R. Aebersold, Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis, *Mol. Cell. Proteomics* 11 (2012), <https://doi.org/10.1074/mcp.O111.016717>.
- [34] Y. Zhang, A. Bilbao, T. Bruderer, J. Luban, C. Strambio-De-Castilla, F. Lisacek, G. Hopfgartner, E. Varesio, The use of variable Q1 isolation windows improves selectivity in LC-SWATH-MS acquisition, *J. Proteome Res.* 14 (2015) 4359–4371, <https://doi.org/10.1021/acs.jproteome.5b00543>.

6.2. Data Processing Script for R

6.2.1. Selection of Internal Standard-based Normalization for MS-Data

The following script, including an exemplary dataset, was published on GitHub (<https://github.com/LaemmerhoferLab/Selection-of-IS-Normalization>). It yields a summary table and a graph to compare intra-group metrics of variation between IS-based normalization methods.

6.2.1.1. Readme.md

This code performs various internal standard-based normalization methods (B-MIS, CCMN, NOMIS, RUVrandom) on MS-datasets and extracts statistical parameters and metrics of intra-group variance for comparison. Ultimately, the aim is to provide a valid and simplified tool for selection of the optimum normalization strategy. This script was created using RStudio (Version 1.1.383)

Title: "Selection of Internal Standard based Normalization" Author: "B. Drotleff et al."

Date: "May 22th, 2019"

Following R packages are required to run the script: *MetNorm*, *NormalizeMets*, *metabolomics*, *qvalue*, *ggplot2*, *sgof*, *caTools*, *ropls*, *tibble*.

To ensure correct operation, the input data file must fulfill the following requirements:

- The R-project file must be in the same folder as the input data file
- The input data file must be named "dataset.csv" with the file type .csv
- Columns must have headers --> column 1: "Samples", column 2: "Class"
 - Rows in column 1 must contain individual sample names
- Rows in column 2 must contain numeric numbers that represent the corresponding class affiliation
- The first two classes should be the samples of interest that are compared against each other (e.g. class 1 = treated samples, class 2 = control samples)
 - Quality control samples (QCs) must be included as class 3
 - Starting from column 3, sample-wise response values for internal standards (ISs) must be listed

- Subsequent columns must contain sample-wise response values for aligned features
- Empty cells should be avoided (execute missing value imputation prior to this script)
- Dataset-specific input parameters that have to be adjusted by the user are:
 - Vector that describes columns that contain IS response values (including 0 <- non-normalization for B-MIS) e.g.: 0 stands for non-normalization, values 1-5 indicate 5 IS in the first 5 rows `ISvect<-as.vector(c(0,1,2,3,4,5))`
 - Factor k for RUVrandom normalization e.g.: 3 factors of unwanted variation `k=3`

An exemplary dataset (`dataset.csv`) is provided. This dataset was derived from open source data in the `NormalizeMets` R package. For further description of this dataset see: De Livera, Alysha M, M. Aho-Sysi, Laurent Jacob, J. Gagnon-Bartch, Sandra Castillo, J.A. Simpson, and Terence P. Speed. 2015. Statistical methods for handling unwanted variation in metabolomics data. *Analytical Chemistry* 87 (7). American Chemical Society: 3606-15. (DOI: 10.1021/ac502439y)

6.2.1.2. Script Code

```
#Script was created using RStudio (Version 1.1.383)

#Title: "Selection of Internal Standard based Normalization"
#Author: "B. Drotleff et al."
#Date: "May 22th, 2019"

#load required packages
suppressMessages(library(MetNorm))
suppressMessages(library(NormalizeMets))
suppressMessages(library(metabolomics))
suppressMessages(library(qvalue))
suppressMessages(library(ggplot2))
suppressMessages(library(sgof))
suppressMessages(library(caTools))
suppressMessages(library(ropls))
suppressMessages(library(tibble))

#R-project file must be in the same folder as the dataset csv file. Check by typing
getwd()

#To ensure correct operation, the input data file must fulfill the following requirements:
#The R-project file must be in the same folder as the input data file
#Columns must have headers --> column 1: "Samples", column 2: "Class"
#Rows in column 1 must contain individual sample names
#Rows in column 2 must contain numeric numbers that represent the corresponding class affiliation
#The first two classes should be the samples of interest that are compared against each other (e.g. class 1 = treated
samples, class 2 = control samples)
#Quality control samples (QCs) must be included as class 3
#Starting from column 3, sample-wise response values for internal standards (ISs) must be listed
#Subsequent columns must contain sample-wise response values for aligned features
#Empty cells should be avoided (execute missing value imputation prior to this script)
#Dataset-specific input parameters that have to be adjusted by the user are:
#Vector that describes columns that contain IS response values (including 0 <- non-normalization for B-MIS)
#(see below)
#Factor k for RUVrandom normalization (see below)
#The file must be named "dataset.csv" and be of the file type .csv
data.raw <- read.table("dataset.csv", sep = ",", header = TRUE, row.names = 1, check.names = FALSE)

#set columns that contain IS information. IS columns must be the first response value containing columns in the dataset. The
value "0" has to be kept in the vector/dataframe as it represents non-normalization (use of raw data) for B-MIS.
#here, e.g.: 0 stands for non-normalization, values 1-5 indicate 5 IS in the first 5 rows.
ISvect<-as.vector(c(0,1,2,3,4,5))
IS<-data.frame(t(ISvect))

#set factors of unwanted variation (k) for RUVrandom normalization (here k=3)
k=3

#Create new folder for statistical analysis of raw data
dir.create("raw data")
setwd("raw data")
#log-transform the dataset matrix
data.raw.log<-LogTransform(data.raw, base=exp(1))$output

#create across-group RLA plot
tiff("RLAPlot_ag_raw.tiff", width=8, height=8, units="in", res=300)
RlaPlots(data.raw.log, "ag", outline=FALSE, ylim=c(-1,1))
dev.off()

#create within-group RLA plot
tiff("RLAPlot_wg_raw.tiff", width=8, height=8, units="in", res=300)
RlaPlots(data.raw.log, "wg", outline=FALSE, ylim=c(-1,1))
dev.off()

#compute p-values (Student's t-test, if Welch's t-test is anticipated use "var.equal = FALSE")
p.val <- matrix(as.character(combn(3,2)), ncol = 3)
p.values.raw <- data.frame(row.names=paste(p.val[1,],"vs",p.val[2,]))
for(i in 2:ncol(data.raw.log))
  {stat.test <- function(ttest) t.test(data.raw.log[,i] ~ data.raw.log[,1], data = data.raw.log, subset =
  data.raw.log[,1] %in% ttest, paired = FALSE, exact = TRUE, var.equal=TRUE)
  p.values.raw[,i-1] <- as.numeric(sapply(apply(p.val,2,stat.test),"[,3]))}
colnames(p.values.raw)<-colnames(data.raw.log[-1])
```

```

#plot p-value histogram
tiff("p-value_histogram_raw.tiff", width=8, height=8, units="in", res=300)
  hist(matrix(as.numeric(p.values.raw[1,])), main=row.names(p.values.raw[1,]), breaks = 20, xlab = "p-value",
  las = 1)
dev.off()

#compute q-values for p-value adjustment (usually only suitable for large datasets)
plotfile <- "q-values_raw.tiff"
  q.values.raw <- qvalue(matrix(as.numeric(p.values.raw[1,])), lambda=seq(0,0.95,0.001), fdr.level = 0.05, p.fdr
  = TRUE)
tiff("q-values_raw.tiff", width=8, height=8, units="in", res=300)
plot(q.values.raw)
dev.off()

  #create dataframe with p- and q-values and sort according to ascending p-values for subsequent p-value
  adjustment methods
  p.val.stat.raw<-data.frame(cbind(ID
  c(1:length(p.values.raw[1,])),round(matrix(as.numeric(p.values.raw[1,])),digits=4),
  round(q.values.raw$qvalues,digits=4)))
rownames(p.val.stat.raw)<-rownames(t(p.values.raw))
colnames(p.val.stat.raw)<-cbind("ID","p-value","q-value")
p.val.stat.raw <- p.val.stat.raw[order(p.val.stat.raw[,2]),]

#compute Benjamini-Hochberg p-value adjustment
p.val.stat.raw$BH<-p.adjust(p.val.stat.raw[,2], method="BH")

#compute SGoF p-value adjustment
SGoF.raw<-SGoF(p.val.stat.raw[,2], alpha=0.05, gamma=0.05)
p.val.stat.raw$SGoF<-SGoF.raw$Adjusted.pvalues

#compute ROC statistics including AUC values
  #get IDs of 25 most significant samples (or other suitable number depending on the dataset)
  ID.pval.raw<-p.val.stat.raw[1:25,1]
data.raw.log.samples<-subset(data.raw.log, data.raw.log[,1]==1 | data.raw.log[,1]==2)
  #calculate mean AUC for 25 most significant samples
  meanAUC.raw<-rowMeans(colAUC(data.raw.log.samples[,ID.pval.raw+1], data.raw.log.samples[,1]))
  #calculate number of features with an AUC >0.8
  AUC.raw<-colAUC(data.raw.log.samples[,2:length(data.raw.log.samples)], data.raw.log.samples[,1])
  nAUC.raw<-length(AUC.raw[(AUC.raw[1,]>0.8)])

#compute PCA and PLS-DA
ropls.class<-as.factor(data.raw.log.samples[,1])
ropls.data.raw.log<-data.raw.log.samples[, 2:dim(data.raw.log.samples)[2]]
  #PCA
  data.multiv.stat.raw.PCA <- oppls(roppls.data.raw.log)
  print(plot(data.multiv.stat.raw.PCA,parAsColFcvn=roppls.class))
  savePlot("PCA_raw.wmf")
  #PLS-DA
  #if the first predictive component is already not significant you will get an error message and PLS-
  DA will not be computed
  data.multiv.stat.raw.PLSDA <- oppls(roppls.data.raw.log, roppls.class, pred1 = NA)
  print(plot(data.multiv.stat.raw.PLSDA))
  savePlot("PLS-DA_raw.wmf")
  #Extract R2Y and Q2Y of PLS-DA (only possible if PLS-DA was computed)
  try(R2YQ2Y.raw<-getSummaryDF(data.multiv.stat.raw.PLSDA),silent=TRUE)
  R2YQ2Y.raw<-c(R2YQ2Y.raw[,2],R2YQ2Y.raw[,3],R2YQ2Y.raw[,5])

#compute CVs (geometric), median absolute deviation (MAD) and variance (Var) for experimental classes
  #CV(geometric, calculation via standard deviation for log-transformed data)
  CV.raw<-data.frame(matrix(ncol=3, nrow=ncol(data.raw.log)-1))
  colnames(CV.raw)<-cbind("CV.raw QCs","CV.raw Class1","CV.raw Class2")
  Subset<-subset(data.raw.log, data.raw.log[,1]==1)
  CV.raw[,2]<-apply(Subset[,2:ncol(Subset)], 2, FUN=sd)
  Subset<-subset(data.raw.log, data.raw.log[,1]==2)
  CV.raw[,3]<-apply(Subset[,2:ncol(Subset)], 2, FUN=sd)
  Subset<-subset(data.raw.log, data.raw.log[,1]==3)
  CV.raw[,1]<-apply(Subset[,2:ncol(Subset)], 2, FUN=sd)
  #formula to translate standard deviation to geometric CV (only valid for log-transformed data)
  CV.raw<-sqrt(exp(CV.raw^2)-1)*100

  #MAD
  MAD.raw<-data.frame(matrix(ncol=3, nrow=ncol(data.raw.log)-1))
  colnames(MAD.raw)<-cbind("MAD.raw QCs","MAD.raw Class1","MAD.raw Class2")
  Subset<-subset(data.raw.log, data.raw.log[,1]==1)
  MAD.raw[,2]<-apply(Subset[,2:ncol(Subset)], 2, FUN=mad)
  Subset<-subset(data.raw.log, data.raw.log[,1]==2)

```



```

MAD.raw[,3]<-apply(Subset[,2:ncol(Subset)], 2, FUN=mad)
Subset<-subset(data.raw.log, data.raw.log[,1]==3)
MAD.raw[,1]<-apply(Subset[,2:ncol(Subset)], 2, FUN=mad)

#Var
Var.raw<-data.frame(matrix(ncol=3, nrow=ncol(data.raw.log)-1))
colnames(Var.raw)<-cbind("Var.raw QCs", "Var.raw Class1", "Var.raw Class2")
Subset<-subset(data.raw.log, data.raw.log[,1]==1)
Var.raw[,2]<-apply(Subset[,2:ncol(Subset)], 2, FUN=var)
Subset<-subset(data.raw.log, data.raw.log[,1]==2)
Var.raw[,3]<-apply(Subset[,2:ncol(Subset)], 2, FUN=var)
Subset<-subset(data.raw.log, data.raw.log[,1]==3)
Var.raw[,1]<-apply(Subset[,2:ncol(Subset)], 2, FUN=var)

VarStat.raw<-cbind(CV.raw,MAD.raw,Var.raw)

setwd(".")
#create summary of statistical parameters for comparison of normalization
options(digits=4)
Summary<-data.frame(matrix(ncol=18,nrow=1))
Summary<-
t(c(sum(p.val.stat.raw[,2]<0.05),sum(p.val.stat.raw[,3]<0.05),sum(p.val.stat.raw[,4]<0.05),sum(p.val.stat.raw
[,5]<0.05),meanAUC.raw,nAUC.raw,median(CV.raw[,1]),median(CV.raw[,2]),median(CV.raw[,3]),median(M
AD.raw[,1]),median(MAD.raw[,2]),median(MAD.raw[,3]),median(Var.raw[,1]),median(Var.raw[,2]),median(Va
r.raw[,3]),
ifelse(exists("R2YQ2Y.raw"),R2YQ2Y.raw[1,NA],ifelse(exists("R2YQ2Y.raw"),R2YQ2Y.raw[2,NA],ifelse(ex
ists("R2YQ2Y.raw"),R2YQ2Y.raw[3,NA])))

colnames(Summary)<-cbind("p-value <0.05", "q-value <0.05", "Benjamini-Hochberg adj. p-value", "SGOF p-
value <0.05", "mean AUC (#25)", "# AUC >0.8", "median CV(geom) in QCs", "median CV(geom) Class1",
"median CV(geom) Class2", "median MAD QCs", "median MAD Class1", "median MAD Class2", "median Var
QCs", "median Var Class1", "median Var Class2", "PLS-DA R2Y", "PLS-DA Q2Y", "#Optimum Components")

#Create new folder for statistical analysis of B-MIS normalized data. Here, as stated by Drotleff et al., B-MIS was executed without
consideration of cut-offs for minimum CV in raw data or CV improvement compared to raw data.
dir.create("B-MIS")
setwd("B-MIS")

#compute normalization with all IS
#add column with value 1 for all samples for non-normalization
data.raw.BMIS<-add_column(data.raw,"noIS"=1:length(data.raw$Class),.after="Class")
data.raw.BMIS$noIS<-1

#calculate CVs for all potential ISs (via mean and standard deviation)
#class-specific mean after normalizatin with all IS
mean.BMIS<-data.frame(matrix(nrow=length(IS), ncol=ncol(data.raw.BMIS)-
1))
for (i in 1:length(IS))
{mean.BMIS.norm <- paste("IS", i, sep = "")
assign(mean.BMIS.norm,
data.frame(c(data.raw.BMIS[1],data.raw.BMIS[2:ncol(data.raw.BMIS)]/data.r
aw.BMIS[,IS[,i]+2]), check.names = FALSE))
#change value for Class here if other class than QC (Class=3) should be
processed
Subset<-subset(get(mean.BMIS.norm), Class==3)
mean.BMIS[i,]<-(apply(Subset[,2:ncol(Subset)], 2, FUN=mean))}
#class-specific standard deviation after normalizatin with all IS
sd.BMIS<-data.frame(matrix(nrow=length(IS), ncol=ncol(data.raw.BMIS)-1))
for (i in 1:length(IS))
{sd.BMIS.norm <- paste("IS", i, sep = "")
assign(sd.BMIS.norm,
data.frame(c(data.raw.BMIS[1],data.raw.BMIS[2:ncol(data.raw.BMIS)]/data.r
aw.BMIS[,IS[,i]+2]), check.names = FALSE))
#change value for Class here if other class than QC (Class=3) should be
processed
Subset<-subset(get(sd.BMIS.norm), Class==3)
sd.BMIS[i,]<-(apply(Subset[,2:ncol(Subset)], 2, FUN=sd))}
CV.norm.BMIS<-sd.BMIS/mean.BMIS*100
colnames(CV.norm.BMIS)<-colnames(data.raw.BMIS[-1])

#Find IS that produces minimum class-specific CV
BMIS.assignment<-data.frame(matrix(nrow=2, ncol=ncol(data.raw.BMIS)-1))
BMIS.assignment[1,]<-apply(CV.norm.BMIS[,1:ncol(CV.norm.BMIS)], 2, FUN=min)
for (i in 1:ncol(BMIS.assignment))
{BMIS.assignment[2,i]<-
rownames(CV.norm.BMIS)[which.min(apply(CV.norm.BMIS[i,],MARGIN=1,min))]}

```

```

rownames(BMIS.assignment) <- c("minimum CV in %", "IS ID")
colnames(BMIS.assignment) <- colnames(data.raw.BMIS[-1])
write.table(t(BMIS.assignment), "B-MIS_assignment.csv", sep = ",", quote = FALSE,
append = FALSE, row.names = TRUE, col.names = NA)

#Normalize each feature with assigned B-MIS
BMIS.assignment<-cbind(1,BMIS.assignment)
colnames(BMIS.assignment)<-colnames(data.raw.BMIS)
data.BMIS<-data.matrix(rbind(BMIS.assignment[2,],data.raw.BMIS))

for (i in length(IS):(length(data.BMIS[1,])-2))
{for (z in 1:length(IS))
{
if(data.BMIS[1,i+2]==z)
{data.BMIS[i+2]<-data.BMIS[i+2]/data.BMIS[z+1]}}
data.BMIS<-data.frame(data.BMIS[-1,-2])
colnames(data.BMIS)<-colnames(data.raw)
write.table(data.BMIS, "B-MIS.csv", sep = ",", quote = FALSE, append = FALSE, row.names =
TRUE, col.names = NA)

#statistical analysis of B-MIS normalized dataset
#log-transform the dataset matrix
data.BMIS.log<-LogTransform(data.BMIS, base=exp(1))$output

#create across-group RLA plot
tiff("RLAPlot_ag_BMIS.tiff", width=8, height=8, units="in", res=300)
RlaPlots(data.BMIS.log, "ag", outline=FALSE, ylim=c(-1,1))
dev.off()

#create within-group RLA plot
tiff("RLAPlot_wg_BMIS.tiff", width=8, height=8, units="in", res=300)
RlaPlots(data.BMIS.log, "wg", outline=FALSE, ylim=c(-1,1))
dev.off()

#compute p-values (Student's t-test, if Welch's t-test is anticipated use "var.equal =
FALSE")
p.val <- matrix(as.character(combn(3,2)),ncol = 3)
p.values.BMIS <- data.frame(row.names=paste(p.val[1,],"vs",p.val[2,]))
for(i in 2:ncol(data.BMIS.log))
{stat.test <- function(ttest) t.test(data.BMIS.log[,i] ~ data.BMIS.log[,1], data =
data.BMIS.log, subset = data.BMIS.log[,1] %in% ttest, paired = FALSE, exact = TRUE,
var.equal=TRUE)
p.values.BMIS[,i-1] <- as.numeric(sapply(apply(p.val,2,stat.test),"",3))}
colnames(p.values.BMIS)<-colnames(data.BMIS.log[-1])

#plot p-value histogram
tiff("p-value_histogram_BMIS.tiff", width=8, height=8, units="in", res=300)
hist(matrix(as.numeric(p.values.BMIS[1,])), main=row.names(p.values.BMIS[1,]),breaks
= 20, xlab = "p-value", las = 1)
dev.off()

#compute q-values for p-value adjustment
plotfile <- "q-values_BMIS.tiff"
q.values.BMIS <- data.frame(row.names=paste(p.val[1,],"vs",p.val[2,]))
qvalue(matrix(as.numeric(p.values.BMIS[1,])),lambda=seq(0,0.95,0.001), fdr.level =
0.05,pfdr = TRUE)
tiff("q-values_BMIS.tiff", width=8, height=8, units="in", res=300)
plot(q.values.BMIS)
dev.off()

#create dataframe with p- and q-values and sort according to ascending p-values for
subsequent p-value adjustment methods
p.val.stat.BMIS<-data.frame(cbind(ID =
c(1:length(p.values.BMIS[1,])),round(matrix(as.numeric(p.values.BMIS[1,])),digits=4),
round(q.values.BMIS$qvalues,digits=4)))
rownames(p.val.stat.BMIS)<-rownames(t(p.values.BMIS))
colnames(p.val.stat.BMIS)<-cbind("ID","p-value","q-value")
p.val.stat.BMIS <- p.val.stat.BMIS[order(p.val.stat.BMIS[,2]),]

#compute Benjamini-Hochberg p-value adjustment
p.val.stat.BMIS$BH<-p.adjust(p.val.stat.BMIS[,2], method="BH")

#compute SGoF p-value adjustment
SGoF.BMIS<-SGoF(p.val.stat.BMIS[,2], alpha=0.05, gamma=0.05)
p.val.stat.BMIS$SGoF<-SGoF.BMIS$Adjusted.pvalues

```

```

#compute ROC statistics including AUC values
  #get IDs of 25 most significant samples (or other suitable number depending
  on the dataset)
  ID.pval.BMIS<-p.val.stat.BMIS[1:25,1]
  data.BMIS.log.samples<-subset(data.BMIS.log, data.BMIS.log[,1]==1 |
  data.BMIS.log[,1]==2)
#calculate mean AUC for 25 most significant samples
  meanAUC.BMIS<-
  rowMeans(colAUC(data.BMIS.log.samples[ID.pval.BMIS+1],
  data.BMIS.log.samples[,1]))
#calculate number of features with an AUC >0.8
  AUC.BMIS<-
  colAUC(data.BMIS.log.samples[,2:length(data.BMIS.log.samples[1,])],
  data.BMIS.log.samples[,1])
  nAUC.BMIS<-length(AUC.BMIS[(AUC.BMIS[1,]>0.8)])

#compute PCA and PLS-DA
ropls.class<-as.factor(data.BMIS.log.samples[,1])
ropls.data.BMIS.log<-data.BMIS.log.samples[,2:dim(data.BMIS.log.samples)[2]]
#PCA
  data.multiv.stat.BMIS.PCA <- opl(ropls.data.BMIS.log)
  print(plot(data.multiv.stat.BMIS.PCA,parAsColFcVn=ropls.class))
  savePlot("PCA_BMIS.wmf")
#PLS-DA
  #if the first predictive component is already not significant you will get an error
  message and PLS-DA will not be computed
  data.multiv.stat.BMIS.PLSDA <- opl(ropls.data.BMIS.log, roples.class, pred1 =
  NA)
  print(plot(data.multiv.stat.BMIS.PLSDA))
  savePlot("PLS-DA_BMIS.wmf")
#Extract R2Y and Q2Y of PLS-DA (only possible if PLS-DA was computed)
  try(R2YQ2Y.BMIS<-getSummaryDF(data.multiv.stat.BMIS.PLSDA),silent=TRUE)
  R2YQ2Y.BMIS<-c(R2YQ2Y.BMIS[,2],R2YQ2Y.BMIS[,3],R2YQ2Y.BMIS[,5])

#compute CVs (geometric), median absolute deviation (MAD) and variance (Var) for experimental
classes
#CV(geometric, calculation via standard deviation for log-transformed data)
  CV.BMIS<-data.frame(matrix(ncol=3, nrow=ncol(data.BMIS.log)-1))
  colnames(CV.BMIS)<-cbind("CV.BMIS QCs","CV.BMIS Class1","CV.BMIS Class2")
  Subset<-subset(data.BMIS.log, data.BMIS.log[,1]==1)
  CV.BMIS[,2]<-apply(Subset[,2:ncol(Subset)], 2, FUN=sd)
  Subset<-subset(data.BMIS.log, data.BMIS.log[,1]==2)
  CV.BMIS[,3]<-apply(Subset[,2:ncol(Subset)], 2, FUN=sd)
  Subset<-subset(data.BMIS.log, data.BMIS.log[,1]==3)
  CV.BMIS[,1]<-apply(Subset[,2:ncol(Subset)], 2, FUN=sd)
  #formula to translate standard deviation to geometric CV (only valid for log-transformed
  data)
  CV.BMIS<-sqrt(exp(CV.BMIS^2)-1)*100

#MAD
  MAD.BMIS<-data.frame(matrix(ncol=3, nrow=ncol(data.BMIS.log)-1))
  colnames(MAD.BMIS)<-cbind("MAD.BMIS QCs","MAD.BMIS Class1","MAD.BMIS Class2")
  Subset<-subset(data.BMIS.log, data.BMIS.log[,1]==1)
  MAD.BMIS[,2]<-apply(Subset[,2:ncol(Subset)], 2, FUN=mad)
  Subset<-subset(data.BMIS.log, data.BMIS.log[,1]==2)
  MAD.BMIS[,3]<-apply(Subset[,2:ncol(Subset)], 2, FUN=mad)
  Subset<-subset(data.BMIS.log, data.BMIS.log[,1]==3)
  MAD.BMIS[,1]<-apply(Subset[,2:ncol(Subset)], 2, FUN=mad)

#Var
  Var.BMIS<-data.frame(matrix(ncol=3, nrow=ncol(data.BMIS.log)-1))
  colnames(Var.BMIS)<-cbind("Var.BMIS QCs","Var.BMIS Class1","Var.BMIS Class2")
  Subset<-subset(data.BMIS.log, data.BMIS.log[,1]==1)
  Var.BMIS[,2]<-apply(Subset[,2:ncol(Subset)], 2, FUN=var)
  Subset<-subset(data.BMIS.log, data.BMIS.log[,1]==2)
  Var.BMIS[,3]<-apply(Subset[,2:ncol(Subset)], 2, FUN=var)
  Subset<-subset(data.BMIS.log, data.BMIS.log[,1]==3)
  Var.BMIS[,1]<-apply(Subset[,2:ncol(Subset)], 2, FUN=var)

  VarStat.BMIS<-cbind(CV.BMIS,MAD.BMIS,Var.BMIS)

setwd("..")
#create summary of statistical parameters for comparison of normalization
options(digits=4)

```

```

Summary<-
rbind(Summary,t(c(sum(p.val.stat.BMIS[,2]<0.05),sum(p.val.stat.BMIS[,3]<0.05),sum(p.
val.stat.BMIS[,4]<0.05),sum(p.val.stat.BMIS[,5]<0.05),meanAUC.BMIS,nAUC.BMIS,me
dian(CV.BMIS[,1]),median(CV.BMIS[,2]),median(CV.BMIS[,3]),median(MAD.BMIS[,1]),
median(MAD.BMIS[,2]),median(MAD.BMIS[,3]),median(Var.BMIS[,1]),median(Var.BMI
S[,2]),median(Var.BMIS[,3]),ifelse(exists("R2YQ2Y.BMIS"),R2YQ2Y.BMIS[1,NA],ifelse
(exists("R2YQ2Y.BMIS"),R2YQ2Y.BMIS[2,NA],ifelse(exists("R2YQ2Y.BMIS"),R2YQ2
Y.BMIS[3,NA])))

#Create new folder for statistical analysis of CCMN normalized data.
dir.create("CCMN")
setwd("CCMN")

#normalize the dataset with CCMN
data.raw.NormMets<-data.raw.log[-1]
data.raw.CCMN <- NormQcmets(data.raw.NormMets, data.raw.log[1], method = "ccmn", qcmets = ISvect[-
1],check.names=FALSE)
data.CCMN<-data.frame(matrix(nrow=nrow(data.raw.CCMN$featuredata),
ncol=ncol(data.raw.CCMN$featuredata))
data.CCMN<-data.raw.CCMN$featuredata
colnames(data.CCMN)<-colnames(data.raw[(length(IS)+1):length(data.raw)])
data.CCMN<-add_column(data.CCMN,"Class"=1:length(data.raw$Class),.before=1)
data.CCMN$Class<-data.raw$Class
write.table(data.CCMN, "CCMN.csv", sep = ",", quote = FALSE, append = FALSE, row.names = TRUE,
col.names = NA)

#statistical analysis of CCMN normalized dataset

#create across-group RLA plot
tiff("RLAPlot_ag_CCMN.tiff", width=8, height=8, units="in", res=300)
RlaPlots(data.CCMN, "ag", outline=FALSE, ylim=c(-1,1))
dev.off()

#create within-group RLA plot
tiff("RLAPlot_wg_CCMN.tiff", width=8, height=8, units="in", res=300)
RlaPlots(data.CCMN, "wg", outline=FALSE, ylim=c(-1,1))
dev.off()

#compute p-values (Student's t-test, if Welch's t-test is anticipated use "var.equal =
FALSE")
p.val <- matrix(as.character(combn(3,2)),ncol = 3)
p.values.CCMN <- data.frame(row.names=paste(p.val[1,],"vs",p.val[2,]))
for(i in 2:ncol(data.CCMN))
{stat.test <- function(ttest) t.test(data.CCMN[,i] ~ data.CCMN[,1], data = data.CCMN, subset =
data.CCMN[,1] %in% ttest, paired = FALSE, exact = TRUE, var.equal=TRUE)
p.values.CCMN[,i-1] <- as.numeric(sapply(apply(p.val,2,stat.test),"",3))}
colnames(p.values.CCMN)<-colnames(data.CCMN[-1])

#plot p-value histogram
tiff("p-value_histogram_CCMN.tiff", width=8, height=8, units="in", res=300)
hist(matrix(as.numeric(p.values.CCMN[1,])), main=row.names(p.values.CCMN[1,]),breaks = 20,
xlab = "p-value", las = 1)
dev.off()

#compute q-values for p-value adjustment
plotfile <- "q-values_CCMN.tiff"
q.values.CCMN <- qvalue(matrix(as.numeric(p.values.CCMN[1,])),lambda=seq(0,0.95,0.001), fdr.level =
0.05,pfdr = TRUE)
tiff("q-values_CCMN.tiff", width=8, height=8, units="in", res=300)
plot(q.values.CCMN)
dev.off()

#create dataframe with p- and q-values and sort according to ascending p-values for subsequent p-value
adjustment methods
p.val.stat.CCMN<-data.frame(cbind(ID
c(1:length(p.values.CCMN[1,])),round(matrix(as.numeric(p.values.CCMN[1,]),digits=4),
round(q.values.CCMN$qvalues,digits=4)))
rownames(p.val.stat.CCMN)<-rownames(t(p.values.CCMN))
colnames(p.val.stat.CCMN)<-cbind("ID","p-value","q-value")
p.val.stat.CCMN <- p.val.stat.CCMN[order(p.val.stat.CCMN[,2]),]

#compute Benjamini-Hochberg p-value adjustment
p.val.stat.CCMN$BH<-p.adjust(p.val.stat.CCMN[,2], method="BH")

```

```

#compute SGoF p-value adjustment
SGoF.CCMN<-SGoF(p.val.stat.CCMN[,2], alpha=0.05, gamma=0.05)
p.val.stat.CCMN$SGoF<-SGoF.CCMN$Adjusted.pvalues

#compute ROC statistics including AUC values
#get IDs of 25 most significant samples (or other suitable number depending on the
dataset)
ID.pval.CCMN<-p.val.stat.CCMN[1:25,1]
data.CCMN.samples<-subset(data.CCMN, data.CCMN[,1]==1 | data.CCMN[,1]==2)
#calculate mean AUC for 25 most significant samples
meanAUC.CCMN<-rowMeans(colAUC(data.CCMN.samples[,ID.pval.CCMN+1],
data.CCMN.samples[,1]))
#calculate number of features with an AUC >0.8
AUC.CCMN<-colAUC(data.CCMN.samples[,2:length(data.CCMN.samples[,1])],
data.CCMN.samples[,1])
nAUC.CCMN<-length(AUC.CCMN[(AUC.CCMN[,1]>0.8)])

#compute PCA and PLS-DA
ropls.class<-as.factor(data.CCMN.samples[,1])
ropls.data.CCMN<-data.CCMN.samples[, 2:dim(data.CCMN.samples)[2]]
#PCA
data.multiv.stat.CCMN.PCA <- oplS(ropls.data.CCMN)
print(plot(data.multiv.stat.CCMN.PCA,parAsColFcVn=ropls.class))
savePlot("PCA_CCMN.wmf")
#PLS-DA
#if the first predictive component is already not significant you will get an error message and PLS-DA will
not be computed
data.multiv.stat.CCMN.PLSDA <- oplS(ropls.data.CCMN, ropls.class, pred1 = NA)
print(plot(data.multiv.stat.CCMN.PLSDA))
savePlot("PLS-DA_CCMN.wmf")
#Extract R2Y and Q2Y of PLS-DA (only possible if PLS-DA was computed)
try(R2YQ2Y.CCMN<-getSummaryDF(data.multiv.stat.CCMN.PLSDA),silent=TRUE)
R2YQ2Y.CCMN<-c(R2YQ2Y.CCMN[,2],R2YQ2Y.CCMN[,3],R2YQ2Y.CCMN[,5])

#compute CVs (geometric), median absolute deviation (MAD) and variance (Var) for experimental
classes
#CV(geometric, calculation via standard deviation for log-transformed data)
CV.CCMN<-data.frame(matrix(ncol=3, nrow=ncol(data.CCMN)-1))
colnames(CV.CCMN)<-cbind("CV.CCMN QCs","CV.CCMN Class1","CV.CCMN
Class2")
Subset<-subset(data.CCMN, data.CCMN[,1]==1)
CV.CCMN[,2]<-apply(Subset[,2:ncol(Subset)], 2, FUN=sd)
Subset<-subset(data.CCMN, data.CCMN[,1]==2)
CV.CCMN[,3]<-apply(Subset[,2:ncol(Subset)], 2, FUN=sd)
Subset<-subset(data.CCMN, data.CCMN[,1]==3)
CV.CCMN[,1]<-apply(Subset[,2:ncol(Subset)], 2, FUN=sd)
#formula to translate standard deviation to geometric CV (only valid for log-transformed
data)
CV.CCMN<-sqrt(exp(CV.CCMN^2)-1)*100

#MAD
MAD.CCMN<-data.frame(matrix(ncol=3, nrow=ncol(data.CCMN)-1))
colnames(MAD.CCMN)<-cbind("MAD.CCMN QCs","MAD.CCMN Class1","MAD.CCMN
Class2")
Subset<-subset(data.CCMN, data.CCMN[,1]==1)
MAD.CCMN[,2]<-apply(Subset[,2:ncol(Subset)], 2, FUN=mad)
Subset<-subset(data.CCMN, data.CCMN[,1]==2)
MAD.CCMN[,3]<-apply(Subset[,2:ncol(Subset)], 2, FUN=mad)
Subset<-subset(data.CCMN, data.CCMN[,1]==3)
MAD.CCMN[,1]<-apply(Subset[,2:ncol(Subset)], 2, FUN=mad)

#Var
Var.CCMN<-data.frame(matrix(ncol=3, nrow=ncol(data.CCMN)-1))
colnames(Var.CCMN)<-cbind("Var.CCMN QCs","Var.CCMN Class1","Var.CCMN
Class2")
Subset<-subset(data.CCMN, data.CCMN[,1]==1)
Var.CCMN[,2]<-apply(Subset[,2:ncol(Subset)], 2, FUN=var)
Subset<-subset(data.CCMN, data.CCMN[,1]==2)
Var.CCMN[,3]<-apply(Subset[,2:ncol(Subset)], 2, FUN=var)
Subset<-subset(data.CCMN, data.CCMN[,1]==3)
Var.CCMN[,1]<-apply(Subset[,2:ncol(Subset)], 2, FUN=var)

```

```

VarStat.CCMN<-cbind(CV.CCMN,MAD.CCMN,Var.CCMN)

setwd("../")
#create summary of statistical parameters for comparison of normalization
options(digits=4)
Summary<-
rbind(Summary,t(c(sum(p.val.stat.CCMN[,2]<0.05),sum(p.val.stat.CCMN[,3]<0.05),sum(p.val.stat.CCMN[,4]<0.05),sum(p.val.stat.CCMN[,5]<0.05),meanAUC.CCMN,nAUC.CCMN,median(CV.CCMN[,1]),median(CV.CCMN[,2]),median(CV.CCMN[,3]),median(MAD.CCMN[,1]),median(MAD.CCMN[,2]),median(MAD.CCMN[,3]),median(Var.CCMN[,1]),median(Var.CCMN[,2]),median(Var.CCMN[,3]),ifelse(exists("R2YQ2Y.CCMN"),R2YQ2Y.CCMN[1,NA]),ifelse(exists("R2YQ2Y.CCMN"),R2YQ2Y.CCMN[2,NA]),ifelse(exists("R2YQ2Y.CCMN"),R2YQ2Y.CCMN[3,NA])))

#Create new folder for statistical analysis of NOMIS normalized data.
dir.create("NOMIS")
setwd("NOMIS")

#normalize the dataset with NOMIS
data.raw.NOMIS <- NormQcmets(data.raw.NormMets, method = "nomis", qcmets = ISvect[1],check.names=FALSE)
data.NOMIS<-data.frame(matrix(nrow=nrow(data.raw.NOMIS$featuredata),ncol=ncol(data.raw.NOMIS$featuredata))
data.NOMIS<-data.raw.NOMIS$featuredata
colnames(data.NOMIS)<-colnames(data.raw[(length(IS)+1):length(data.raw)])
data.NOMIS<-add_column(data.NOMIS,"Class"=1:length(data.raw$Class),.before=1)
data.NOMIS$Class<-data.raw$Class
write.table(data.NOMIS, "NOMIS.csv", sep = ",", quote = FALSE, append = FALSE, row.names = TRUE, col.names = NA)

#statistical analysis of NOMIS normalized dataset

#create across-group RLA plot
tiff("RLAPlot_ag_NOMIS.tiff", width=8, height=8, units="in", res=300)
RlaPlots(data.NOMIS, "ag", outline=FALSE, ylim=c(-1,1))
dev.off()

#create within-group RLA plot
tiff("RLAPlot_wg_NOMIS.tiff", width=8, height=8, units="in", res=300)
RlaPlots(data.NOMIS, "wg", outline=FALSE, ylim=c(-1,1))
dev.off()

#compute p-values (Student's t-test, if Welch's t-test is anticipated use "var.equal = FALSE")
p.val <- matrix(as.character(combn(3,2)),ncol = 3)
p.values.NOMIS <- data.frame(row.names=paste(p.val[1,],"vs",p.val[2,])
for(i in 2:ncol(data.NOMIS))
{stat.test <- function(ttest) t.test(data.NOMIS[,i] ~ data.NOMIS[,1], data = data.NOMIS, subset = data.NOMIS[,1] %in% ttest, paired = FALSE, exact = TRUE, var.equal=TRUE)
p.values.NOMIS[,i-1] <- as.numeric(sapply(apply(p.val,2,stat.test),"[",3))}
colnames(p.values.NOMIS)<-colnames(data.NOMIS[-1])

#plot p-value histogram
tiff("p-value_histogram_NOMIS.tiff", width=8, height=8, units="in", res=300)
hist(matrix(as.numeric(p.values.NOMIS[1,])), main=row.names(p.values.NOMIS[1,]),breaks = 20, xlab = "p-value", las = 1)
dev.off()

#compute q-values for p-value adjustment
plotfile <- "q-values_NOMIS.tiff"
q.values.NOMIS <- qvalue(matrix(as.numeric(p.values.NOMIS[1,])),lambda=seq(0,0.95,0.001), fdr.level = 0.05,pfdr = TRUE)
tiff("q-values_NOMIS.tiff", width=8, height=8, units="in", res=300)
plot(q.values.NOMIS)
dev.off()

#create dataframe with p- and q-values and sort according to ascending p-values for subsequent p-value adjustment methods
p.val.stat.NOMIS<-data.frame(cbind(ID =
c(1:length(p.values.NOMIS[1,])),round(matrix(as.numeric(p.values.NOMIS[1,])),digits=4),
round(q.values.NOMIS$qvalues,digits=4)))
rownames(p.val.stat.NOMIS)<-rownames(t(p.values.NOMIS))
colnames(p.val.stat.NOMIS)<-cbind("ID","p-value","q-value")
p.val.stat.NOMIS <- p.val.stat.NOMIS[order(p.val.stat.NOMIS[,2]),]

#compute Benjamini-Hochberg p-value adjustment

```

```

p.val.stat.NOMIS$BH<-p.adjust(p.val.stat.NOMIS[,2], method="BH")

#compute SGoF p-value adjustment
SGoF.NOMIS<-SGoF(p.val.stat.NOMIS[,2], alpha=0.05, gamma=0.05)
p.val.stat.NOMIS$SGoF<-SGoF.NOMIS$Adjusted.pvalues

#compute ROC statistics including AUC values
#get IDs of 25 most significant samples (or other suitable number depending on the
dataset)
ID.pval.NOMIS<-p.val.stat.NOMIS[1:25,1]
data.NOMIS.samples<-subset(data.NOMIS, data.NOMIS[,1]==1 | data.NOMIS[,1]==2)
#calculate mean AUC for 25 most significant samples
meanAUC.NOMIS<-rowMeans(colAUC(data.NOMIS.samples[,ID.pval.NOMIS+1],
data.NOMIS.samples[,1]))
#calculate number of features with an AUC >0.8
AUC.NOMIS<-colAUC(data.NOMIS.samples[,2:length(data.NOMIS.samples[,1,]),
data.NOMIS.samples[,1])
nAUC.NOMIS<-length(AUC.NOMIS[(AUC.NOMIS[1,]>0.8)])

#compute PCA and PLS-DA
ropls.class<-as.factor(data.NOMIS.samples[,1])
ropls.data.NOMIS<-data.NOMIS.samples[, 2:dim(data.NOMIS.samples)[2]]
#PCA
data.multiv.stat.NOMIS.PCA <- opls(ropls.data.NOMIS)
print(plot(data.multiv.stat.NOMIS.PCA,parAsColFcvn=ropls.class))
savePlot("PCA_NOMIS.wmf")
#PLS-DA
#if the first predictive component is already not significant you will get an error message and PLS-DA will
not be computed
data.multiv.stat.NOMIS.PLSDA <- opls(ropls.data.NOMIS, ropls.class, pred1 = NA)
print(plot(data.multiv.stat.NOMIS.PLSDA))
savePlot("PLS-DA_NOMIS.wmf")
#Extract R2Y and Q2Y of PLS-DA (only possible if PLS-DA was computed)
try(R2YQ2Y.NOMIS<-getSummaryDF(data.multiv.stat.NOMIS.PLSDA),silent=TRUE)
R2YQ2Y.NOMIS<-c(R2YQ2Y.NOMIS[,2],R2YQ2Y.NOMIS[,3],R2YQ2Y.NOMIS[,5])

#compute CVs (geometric), median absolute deviation (MAD) and variance (Var) for experimental
classes
#CV(geometric, calculation via standard deviation for log-transformed data)
CV.NOMIS<-data.frame(matrix(ncol=3, nrow=ncol(data.NOMIS)-1))
colnames(CV.NOMIS)<-cbind("CV.NOMIS_QCs","CV.NOMIS Class1","CV.NOMIS
Class2")
Subset<-subset(data.NOMIS, data.NOMIS[,1]==1)
CV.NOMIS[,2]<-apply(Subset[,2:ncol(Subset)], 2, FUN=sd)
Subset<-subset(data.NOMIS, data.NOMIS[,1]==2)
CV.NOMIS[,3]<-apply(Subset[,2:ncol(Subset)], 2, FUN=sd)
Subset<-subset(data.NOMIS, data.NOMIS[,1]==3)
CV.NOMIS[,1]<-apply(Subset[,2:ncol(Subset)], 2, FUN=sd)
#formula to translate standard deviation to geometric CV (only valid for log-transformed
data)
CV.NOMIS<-sqrt(exp(CV.NOMIS^2)-1)*100

#MAD
MAD.NOMIS<-data.frame(matrix(ncol=3, nrow=ncol(data.NOMIS)-1))
colnames(MAD.NOMIS)<-cbind("MAD.NOMIS_QCs","MAD.NOMIS
Class1","MAD.NOMIS Class2")
Subset<-subset(data.NOMIS, data.NOMIS[,1]==1)
MAD.NOMIS[,2]<-apply(Subset[,2:ncol(Subset)], 2, FUN=mad)
Subset<-subset(data.NOMIS, data.NOMIS[,1]==2)
MAD.NOMIS[,3]<-apply(Subset[,2:ncol(Subset)], 2, FUN=mad)
Subset<-subset(data.NOMIS, data.NOMIS[,1]==3)
MAD.NOMIS[,1]<-apply(Subset[,2:ncol(Subset)], 2, FUN=mad)

#Var
Var.NOMIS<-data.frame(matrix(ncol=3, nrow=ncol(data.NOMIS)-1))
colnames(Var.NOMIS)<-cbind("Var.NOMIS_QCs","Var.NOMIS Class1","Var.NOMIS
Class2")
Subset<-subset(data.NOMIS, data.NOMIS[,1]==1)
Var.NOMIS[,2]<-apply(Subset[,2:ncol(Subset)], 2, FUN=var)
Subset<-subset(data.NOMIS, data.NOMIS[,1]==2)
Var.NOMIS[,3]<-apply(Subset[,2:ncol(Subset)], 2, FUN=var)
Subset<-subset(data.NOMIS, data.NOMIS[,1]==3)
Var.NOMIS[,1]<-apply(Subset[,2:ncol(Subset)], 2, FUN=var)

VarStat.NOMIS<-cbind(CV.NOMIS,MAD.NOMIS,Var.NOMIS)

```

```

setwd("../")
#create summary of statistical parameters for comparison of normalization
options(digits=4)
Summary<-
rbind(Summary,t(c(sum(p.val.stat.NOMIS[,2]<0.05),sum(p.val.stat.NOMIS[,3]<0.05),sum(p.val.stat.NOMIS[,4]<0.05),sum(p.val.stat.NOMIS[,5]<0.05),meanAUC.NOMIS,nAUC.NOMIS,median(CV.NOMIS[,1]),median(CV.NOMIS[,2]),median(CV.NOMIS[,3]),median(MAD.NOMIS[,1]),median(MAD.NOMIS[,2]),median(MAD.NOMIS[,3]),median(Var.NOMIS[,1]),median(Var.NOMIS[,2]),median(Var.NOMIS[,3]),ifelse(exists("R2YQ2Y.NOMIS"),R2YQ2Y.NOMIS[1],NA),ifelse(exists("R2YQ2Y.NOMIS"),R2YQ2Y.NOMIS[2],NA),ifelse(exists("R2YQ2Y.NOMIS"),R2YQ2Y.NOMIS[3],NA))))

#Create new folder for statistical analysis of RUVrand normalized data.
dir.create("RUVrand")
setwd("RUVrand")

#normalize the dataset with RUVrand
data.raw.RUVrand <- NormQcmets(data.raw.NormMets, method = "ruvrand", qcmets = ISvect[-1], k=k)
data.RUVrand<-as.data.frame(data.raw.RUVrand$featuredata)
data.RUVrand<-add_column(data.RUVrand,"Class"=1:length(data.raw$Class),.before=1)
data.RUVrand$Class<-data.raw$Class
write.table(data.RUVrand, "RUVrand.csv", sep = ",", quote = FALSE, append = FALSE, row.names = TRUE, col.names = NA)

#statistical analysis of RUVrand normalized dataset

#create across-group RLA plot
tiff("RLAPlot_ag_RUVrand.tiff", width=8, height=8, units="in", res=300)
RlaPlots(data.RUVrand, "ag", outline=FALSE, ylim=c(-1,1))
dev.off()

#create within-group RLA plot
tiff("RLAPlot_wg_RUVrand.tiff", width=8, height=8, units="in", res=300)
RlaPlots(data.RUVrand, "wg", outline=FALSE, ylim=c(-1,1))
dev.off()

#compute p-values (Student's t-test, if Welch's t-test is anticipated use "var.equal = FALSE")
p.val <- matrix(as.character(combn(3,2)),ncol = 3)
p.values.RUVrand <- data.frame(row.names=paste(p.val[1,],"vs",p.val[2,]))
for(i in 2:ncol(data.RUVrand))
{stat.test <- function(ttest) t.test(data.RUVrand[,i] ~ data.RUVrand[,1], data = data.RUVrand, subset = data.RUVrand[,1] %in% ttest, paired = FALSE, exact = TRUE, var.equal=TRUE)
p.values.RUVrand[,i-1] <- as.numeric(sapply(apply(p.val,2,stat.test),"[",3))}
colnames(p.values.RUVrand)<-colnames(data.RUVrand[-1])

#plot p-value histogram
tiff("p-value_histogram_RUVrand.tiff", width=8, height=8, units="in", res=300)
hist(matrix(as.numeric(p.values.RUVrand[1,])), main=row.names(p.values.RUVrand[1,]),breaks = 20, xlab = "p-value", las = 1)
dev.off()

#compute q-values for p-value adjustment
plotfile <- "q-values_RUVrand.tiff"
q.values.RUVrand <- qvalue(matrix(as.numeric(p.values.RUVrand[1,])),lambda=seq(0,0.95,0.001), fdr.level = 0.05,pfdr = TRUE)
tiff("q-values_RUVrand.tiff", width=8, height=8, units="in", res=300)
plot(q.values.RUVrand)
dev.off()

#create dataframe with p- and q-values and sort according to ascending p-values for subsequent p-value adjustment methods
p.val.stat.RUVrand<-data.frame(cbind(ID
=
c(1:length(p.values.RUVrand[1,]),round(matrix(as.numeric(p.values.RUVrand[1,]),digits=4),
round(q.values.RUVrand$qvalues,digits=4)))
rownames(p.val.stat.RUVrand)<-rownames(t(p.values.RUVrand))
colnames(p.val.stat.RUVrand)<-cbind("ID","p-value","q-value")
p.val.stat.RUVrand <- p.val.stat.RUVrand[order(p.val.stat.RUVrand[,2]),]

#compute Benjamini-Hochberg p-value adjustment
p.val.stat.RUVrand$BH<-p.adjust(p.val.stat.RUVrand[,2], method="BH")

#compute SGoF p-value adjustment
SGoF.RUVrand<-SGoF(p.val.stat.RUVrand[,2], alpha=0.05, gamma=0.05)
p.val.stat.RUVrand$SGoF<-SGoF.RUVrand$Adjusted.pvalues

```



```

#compute ROC statistics including AUC values
  #get IDs of 25 most significant samples (or other suitable number depending on the
  dataset)
  ID.pval.RUVrand<-p.val.stat.RUVrand[1:25,1]
data.RUVrand.samples<-subset(data.RUVrand, data.RUVrand[,1]==1 | data.RUVrand[,1]==2)
  #calculate mean AUC for 25 most significant samples
  meanAUC.RUVrand<-rowMeans(colAUC(data.RUVrand.samples[,ID.pval.RUVrand+1],
  data.RUVrand.samples[,1]))
  #calculate number of features with an AUC >0.8
  AUC.RUVrand<-colAUC(data.RUVrand.samples[,2:length(data.RUVrand.samples[,1,]),
  data.RUVrand.samples[,1])
  nAUC.RUVrand<-length(AUC.RUVrand[(AUC.RUVrand[1,]>0.8)])

#compute PCA and PLS-DA
ropls.class<-as.factor(data.RUVrand.samples[,1])
ropls.data.RUVrand<-data.RUVrand.samples[, 2:dim(data.RUVrand.samples)[2]]
  #PCA
  data.multiv.stat.RUVrand.PCA <- opl(ropls.data.RUVrand)
  print(plot(data.multiv.stat.RUVrand.PCA,parAsColFvN=ropls.class))
  savePlot("PCA_RUVrand.wmf")
  #PLS-DA
#if the first predictive component is already not significant you will get an error message and PLS-DA will
not be computed
  data.multiv.stat.RUVrand.PLSDA <- opl(ropls.data.RUVrand, roppls.class, pred1 = NA)
  print(plot(data.multiv.stat.RUVrand.PLSDA))
  savePlot("PLS-DA_RUVrand.wmf")
  #Extract R2Y and Q2Y of PLS-DA (only possible if PLS-DA was computed)
  try(R2YQ2Y.RUVrand<-
  getSummaryDF(data.multiv.stat.RUVrand.PLSDA),silent=TRUE)
  R2YQ2Y.RUVrand<-
  c(R2YQ2Y.RUVrand[,2],R2YQ2Y.RUVrand[,3],R2YQ2Y.RUVrand[,5])

#compute CVs (geometric), median absolute deviation (MAD) and variance (Var) for experimental
classes
  #CV(geometric, calculation via standard deviation for log-transformed data)
  CV.RUVrand<-data.frame(matrix(ncol=3, nrow=ncol(data.RUVrand)-1))
  colnames(CV.RUVrand)<-cbind("CV.RUVrand          QCs", "CV.RUVrand
  Class1", "CV.RUVrand Class2")
  Subset<-subset(data.RUVrand, data.RUVrand[,1]==1)
  CV.RUVrand[,2]<-apply(Subset[,2:ncol(Subset)], 2, FUN=sd)
  Subset<-subset(data.RUVrand, data.RUVrand[,1]==2)
  CV.RUVrand[,3]<-apply(Subset[,2:ncol(Subset)], 2, FUN=sd)
  Subset<-subset(data.RUVrand, data.RUVrand[,1]==3)
  CV.RUVrand[,1]<-apply(Subset[,2:ncol(Subset)], 2, FUN=sd)
  #formula to translate standard deviation to geometric CV (only valid for log-transformed
  data)
  CV.RUVrand<-sqrt(exp(CV.RUVrand^2)-1)*100

  #MAD
  MAD.RUVrand<-data.frame(matrix(ncol=3, nrow=ncol(data.RUVrand)-1))
  colnames(MAD.RUVrand)<-cbind("MAD.RUVrand          QCs", "MAD.RUVrand
  Class1", "MAD.RUVrand Class2")
  Subset<-subset(data.RUVrand, data.RUVrand[,1]==1)
  MAD.RUVrand[,2]<-apply(Subset[,2:ncol(Subset)], 2, FUN=mad)
  Subset<-subset(data.RUVrand, data.RUVrand[,1]==2)
  MAD.RUVrand[,3]<-apply(Subset[,2:ncol(Subset)], 2, FUN=mad)
  Subset<-subset(data.RUVrand, data.RUVrand[,1]==3)
  MAD.RUVrand[,1]<-apply(Subset[,2:ncol(Subset)], 2, FUN=mad)

  #Var
  Var.RUVrand<-data.frame(matrix(ncol=3, nrow=ncol(data.RUVrand)-1))
  colnames(Var.RUVrand)<-cbind("Var.RUVrand          QCs", "Var.RUVrand
  Class1", "Var.RUVrand Class2")
  Subset<-subset(data.RUVrand, data.RUVrand[,1]==1)
  Var.RUVrand[,2]<-apply(Subset[,2:ncol(Subset)], 2, FUN=var)
  Subset<-subset(data.RUVrand, data.RUVrand[,1]==2)
  Var.RUVrand[,3]<-apply(Subset[,2:ncol(Subset)], 2, FUN=var)
  Subset<-subset(data.RUVrand, data.RUVrand[,1]==3)
  Var.RUVrand[,1]<-apply(Subset[,2:ncol(Subset)], 2, FUN=var)

  VarStat.RUVrand<-cbind(CV.RUVrand,MAD.RUVrand,Var.RUVrand)

setwd("../")
#create summary of statistical parameters for comparison of normalization

```

```

options(digits=4)
Summary<-
rbind(Summary,t(c(sum(p.val.stat.RUVrand[,2]<0.05),sum(p.val.stat.RUVrand[,3]<0.05),sum(p.val.stat.RUVr
and[,4]<0.05),sum(p.val.stat.RUVrand[,5]<0.05),meanAUC.RUVrand,nAUC.RUVrand,median(CV.RUVrand[,
1]),median(CV.RUVrand[,2]),median(CV.RUVrand[,3]),median(MAD.RUVrand[,1]),median(MAD.RUVrand[,2]
),median(MAD.RUVrand[,3]),median(Var.RUVrand[,1]),median(Var.RUVrand[,2]),median(Var.RUVrand[,3]),if
else(exists("R2YQ2Y.RUVrand"),R2YQ2Y.RUVrand[1,NA),ifelse(exists("R2YQ2Y.RUVrand"),R2YQ2Y.RUV
rand[2,NA),ifelse(exists("R2YQ2Y.RUVrand"),R2YQ2Y.RUVrand[3,NA))))
rownames(Summary)<-rbind("raw data","B-MIS","CCMN","NOMIS","RUVrandom")

graphics.off()
write.table(Summary, "Summary.csv", sep = ",", quote = FALSE, append = FALSE, row.names = TRUE, col.names = NA)

#create plot to compare intra-group metrics of variation between normalization methods
#for comparison all datasets must have the same length ->remove ISs from raw data, BMIS and RUVrandom)
#compute lists with intra-group metrics of variation for each normalization
VarStat.raw<-VarStat.raw[-(1:length(IS)-1),]
VarStat.BMIS<-VarStat.BMIS[-(1:length(IS)-1),]
VarStat.RUVrand<-VarStat.RUVrand[-(1:length(IS)-1),]
VarStat.List<-cbind(VarStat.RUVrand,VarStat.NOMIS,VarStat.CCMN,VarStat.BMIS,VarStat.raw)

VarData<-list()
for (i in 1:(length(VarStat.List)-36))
{VarData[[i]]<-cbind(VarStat.List[i],VarStat.List[i+9],VarStat.List[i+18],VarStat.List[i+27],VarStat.List[i+36])}

#compute plot
plotnames<-c("CV (QC)","CV (Class1)","CV (Class2)","MAD (QC)","MAD (Class1)","MAD (Class2)","Var
(QC)","Var (Class1)","Var (Class2)")
normnames<-c("RUVrand","NOMIS","CCMN", "B-MIS","raw data")

tiff("Comparison_metrics_of_variation.tiff", width=17, height=8.5, units="cm", res=600)
par(mfrow=c(3,3),cex.lab=2, oma = c(0,0.2,0.1,0), mar = c(1.5,3.7,1.4,0.4),mgp=c(3,0.2,0),cex=0.4)
for (i in 1:length(plotnames))
{boxplot(VarData[[i]],main=plotnames[i],notch=TRUE,outline=FALSE, horizontal=TRUE,xaxt="n",yaxt="n")
axis(side=1,tck=-0.01,las=1)
axis(side=2,tck=0,las=1,labels=normnames, at=(1:length(normnames)))
abline(v=median(VarData[[i]][,5]), col="darkred", lwd=1.5)
grid(5, ny = FALSE, col = "black", lty = 3, lwd = 0.5)}
dev.off()

```

7. Acknowledgements

I would like to thank all the people who contributed to this thesis in many different ways throughout the last years.

First of all, I deeply thank my supervisor Prof. Dr. Michael Lämmerhofer for giving me the opportunity to work in his group and for his endless support throughout my PhD thesis. He managed to provide a sophisticated working environment and many highly interesting research projects. Even more outstanding was his ambitious mentoring including the countless hours of discussing, motivating, revising and advising at all times. I am very grateful for his professional guidance and in particular also for his trust in my work.

I would also like to thank Prof. Dr. Stefan Laufer for the second supervision and the evaluation of my thesis.

Many thanks go to Dr. Jeannie Horak, Dr. Stefan Neubauer and Dr. Jörg Schlotterbeck, for sharing their vast knowledge about analytical chemistry with me over the past years. They were always open for various scientific debates and never hesitated to offer their help when it was needed.

Especially Jörg is highly acknowledged, as he spent hundreds of hours with me in the mass spec laboratory, where we were trying to decipher the functions of the instrument. I often consulted him for his smart scientific advice and he was a great partner to share moments of success but also frustration when things did not work out.

I also want to thank Carlos Calderón for the many fruitful discussions about mass spectrometric workflows, preferably in the late evenings in the office. Carlos was always open for new ideas and a great critic that helped me to rethink many aspects of my work.

I am very grateful for my former colleagues Corinna Sanwald, Markus Höldrich, Adrian Sievers-Engler, Dr. Alexandra Zimmermann, Dr. Stefan Polnick, Mike Kaupert, Siyao Liu, Ulrich Woiwode, and Tomáš Pluháček as well as my current colleagues Stefanie Bäurer, Ryan Karongo, Kristina Schmitt, Ece Aydin, Malgorzata Cebo, Christian Geibel, Feiyang Li, Peng Li and Marc Wolter. It has been great to cross paths and some good times were shared in the lab and at the many conferences that we were able to attend. Furthermore, I would like to appreciate the work of our secretaries, Eveline Wachendorfer and Ingrid Straub, who take care of many duties to keep our workgroup running.

Many thanks also to Prof. Dr. Robert Lukowski and Prof. Dr. Manfred Hallschmid for the successful collaboration and the provision of the interesting real samples.

Special thanks also to all my close friends who took my mind off from time to time and always had an open ear.

I would like to express my gratitude to my partner and friend Steffi. During all the years I could always rely on her support and her love. I must also apologize to her for the many lonely nights she had to spend while I was preparing this thesis.

Finally, I wish to thank my parents, Johann and Johanna, for their unbelievable support in so many ways since the beginning of my studies. Many thanks also to my sister Sabine. They all were always there for me and made me feel very privileged.

“The imagination of nature is far, far greater than the imagination of man”

Richard Feynman

Curriculum Vitae

Personal Information

Name: Bernhard Drotleff
Date of birth: 19.10.1987
Place of birth: Mediasch / Siebenbürgen (Rumänien)

Dissertation

04/2015 – present: **PhD student at Eberhard Karls Universität Tübingen**
Institute of Pharmaceutical Sciences
Pharmaceutical (Bio-) Analysis

Thesis title: Development and Optimization of Targeted/Untargeted Lipidomic Screening Methodologies in Pharmaceutical and Clinical Bioanalysis by Liquid Chromatography-High Resolution-Mass Spectrometry

Supervision: Prof. Dr. Michael Lämmerhofer

Education

10/2011 – 04/2014 **Master of Science in Toxicology (M.Sc.)**
Charité – Universitätsmedizin Berlin, Germany

Thesis title: Development of a liquid chromatography-tandem mass spectrometry based method for therapeutic drug monitoring of antiretroviral substances in human plasma

Supervision: Dr. Torsten Binscheck-Domaß (Labor Berlin – Charité Vivantes GmbH)

04/2008 – 10/2011 **Bachelor of Science in Pharmaceutical Sciences (B.Sc.)**
Ludwig-Maximilians-Universität München, Germany

Thesis title: Cloning of AAV-vectors for the viral gene therapy of retinitis pigmentosa

Supervision: Prof. Dr. Martin Biel