DISSERTATION

Submitted to the
Faculty of Biology and Chemistry
of the Justus Liebig University Giessen
for the degree of
Doctor rerum naturalium (Dr. rer. nat.)

presented by

staatl. geprüfte Dipl. LMChem Sabrina Liesenfeld

2021

Novel non-targeted LC-HRMS approaches in veterinary drug residue analysis and environmental analysis

Dean:
Prof. Dr. Thomas Wilke

Referees:

1. Prof. Dr. Gerd Hamscher
   Institute of Food Chemistry and Food Biotechnology,
   Justus Liebig University Giessen

2. Prof. Dr. Mirko Bunzel
   Department of Food Chemistry and Phytochemistry,
   Karlsruhe Institute of Technology

Selbstständigkeitserklärung

Ich erkläre: Ich habe die vorgelegte Dissertation selbstständig und ohne unerlaubte fremde Hilfe und nur mit den Hilfen angefertigt, die ich in der Dissertation angegeben habe. Alle Textstellen, die wörtlich oder sinngemäß aus veröffentlichten Schriften entnommen sind, und alle Angaben, die auf mündlichen Auskünften beruhen, sind als solche kenntlich gemacht. Ich stimme einer evtl. Überprüfung meiner Dissertation durch eine Antiplagiat-Software zu. Bei den von mir durchgeführten und in der Dissertation erwähnten Untersuchungen habe ich die Grundsätze guter wissenschaftlicher Praxis, wie sie in der „Satzung der Justus-Liebig-Universität Gießen zur Sicherung guter wissenschaftlicher Praxis" niedergelegt sind, eingehalten.

_____ _____

Datum Unterschrift

This dissertation contains text sections, tables and figures, which have previously been published and are reproduced with permission from:

Liesenfeld, S.; Steliopoulos, P.; Hamscher, G., Comprehensive metabolomics analysis of non-targeted LC-HRMS data provides valuable insights regarding the origin of veterinary drug residues. *Journal of Agricultural and Food Chemistry*, **2020**, 68 (44), 12493-12502. Copyright 2020 American Chemical Society

Liesenfeld, S.; Steliopoulos, P.; Wenig, S.; Gottstein, V.; Hamscher, G., Comprehensive LC-HRMS metabolomics analyses for the estimation of environmental inputs of altrenogest in pig breeding. *Chemosphere*, **2022**, 287, in-press (https://doi.org/10.1016/j.chemosphere.2021.132353).

Furthermore, some text sections concerning NMR analysis were provided by colleagues. These text sections were clearly marked as such.

To my family

Table of contents

Abstract

Liquid chromatography coupled with high resolution mass spectrometry (LC-HRMS) enables the implementation of non-targeted screening analysis and complements standard triple−quadrupole mass spectrometry in veterinary drug residue control as well as in environmental analysis. Metabolomics workflows offer the opportunity to investigate and identify metabolites and biomarkers for metabolical change in various biological samples. The challenge in evaluating HRMS data is working with the huge amount of data which is acquired in a non-targeted HRMS setting, rendering several thousand mass spectrometric features per sample.

In this thesis, a workflow for processing and evaluation of non-targeted LC-HRMS data was established. The versatile use of non-targeted metabolomics (metabolic fingerprinting) was demonstrated by applying this approach in two different analytical fields, i.e., official veterinary drug residue control and environmental analysis.

The established metabolomics workflow consists of (1) a data processing procedure which renders an aligned peak list of all analytes in all samples, (2) chemometric analyses in order to distinguish different test groups from each other, (3) biomarker candidate selection by applying univariate and bivariate statistical methods and (4) pathway mapping as a complement to manual biomarker identification. A robust data processing procedure was developed and optimized using a statistical design of experiments (DoE) with a fractional factorial design. Minimum time span, minimum height and the amplitude of noise chosen for chromatogram deconvolution were identified as factors of significance for reliable peak detection, integration and alignment. The results of the DoE were further used to optimize the final parameter settings. This experimental set-up highlights the benefits of DoE, since a manageable number of experiments rendered valuable information for workflow development and optimization.

The aim of the first practical application of this new procedure was to gain further information about the origin of antibiotic residues in non-compliant samples taken within the frame of official veterinary drug residue control. Furthermore, biomarkers for changes in the metabolism of drug-treated, infected animals should be identified. Preferably, these biomarkers should be easily applicable in routine veterinary drug monitoring with minimal additional analytical effort. Non-compliant results may be caused by an improper drug administration to an infected animal or by sample contamination due to, e.g., cross-over during

sampling or analysis, polluted feed or drug-residues in animal farm housing. The feasibility of metabolic fingerprinting to distinguish between samples from infected, drug-treated animals (positive samples) and samples from healthy animals (control samples) was tested and successfully demonstrated. Two HRMS platforms (LC-Q-Orbitrap and LC-Q-TOF) were used to investigate the applicability of the approach on different instrument settings. The experiment was performed with muscle and kidney samples, taken from pigs ($n = 50$) and cattle ($n = 30$). The main focus laid on the examination of *porcine* muscle tissues, since the number of samples from pigs in German veterinary drug residue control clearly exceeds the number of samples taken from cattle. Principal components analysis (PCA) with subsequent quadratic discriminant analysis (QDA) was successfully applied in order to differentiate between *porcine* muscle samples from infected, drug-treated animals and samples from healthy animals. Sensitivity (0.826 (Q-Orbitrap) / 0.810 (Q-TOF)) and specificity (0.976 (Q-Orbitrap) / 0.923 (Q-TOF)) were high for both instruments. Interestingly, quality control samples of healthy animals spiked with 70 veterinary drugs were located in the same cluster as control samples in both PCA scores plots. These results indicate that the main difference between positive and control *porcine* muscle samples originated from changes in the metabolome of animals rather than from the presence of veterinary drugs. Univariate statistical tests, manual biomarker annotation and pathway mapping revealed changes in prostaglandin formation and the metabolism of unsaturated fatty acids as established markers of infection and inflammation. A Receiver Operating Characteristic (ROC) curve analysis was further used for biomarker selection and assessment. Two new biomarker candidates were identified as tripeptide prolylphenylalanylglycine and a lysophosphatidylcholine derivative. In order to increase the predictive power of these two biomarkers, a bivariate data analysis procedure was performed. A 0.9-prediction ellipse accepting an α error probability of 0.1 (specificity of 0.9) was calculated from the control samples and used as decision border. Applying this reliable approach on real world samples, contaminated samples could be unequivocally identified. Thus, this procedure is a valuable tool to gain legal certainty for non-compliant results in official veterinary drug residue control.

Further investigations showed that chemometric analysis of *bovine* muscle samples yielded comparable results. However, it is crucial to evaluate muscle samples from different species separately. The analysis of control muscle samples from pigs and cattle showed substantial differences in multivariate statistical analysis. The main difference was attributed to mass spectrometric features which were identified as $[M+H]^+$, $[M+2H]^{2+}$ and $[M+3H]^{3+}$ of the C-

terminal octapeptide of *bovine* myoglobin.

PCA with subsequent QDA of kidney samples of both species did not enable a sufficient separation of samples taken from drug-treated, infected animals and healthy ones. The sensitivity of the models was considered insufficient with a maximum of 0.572 (Q-Orbitrap data of *porcine* kidney samples) and 0.795 (Q-TOF data of *bovine* kidney samples), while the respective specificity was 0.903 and 0.899. However, a deviation of the expectation vectors of the two test groups was observed in the scores plots. In conclusion, the multivariate analysis of kidney samples was considered to be less useful for the differentiation of healthy and infected animals.

In the second practical application of the new procedure, the environmental release of altrenogest (ALT), a synthetic progestogen which is used in modern pig production for estrus synchronization, was investigated. For certain geographical areas it cannot be excluded that the zootechnical use of ALT in gilts may present a risk for fish and other aquatic organisms, since manure of ALT-treated gilts may be used as fertilizer for agricultural soils. According to current knowledge, no published scientific data on urinary excretion of ALT in gilts is available. In this work, a pilot study was conducted to collect urine samples from ALT-treated gilts as well as from non-medicated control animals. LC-HRMS analysis was applied to perform targeted analysis of ALT and known metabolites, i.e., conjugates with glucuronic acid and ALT sulfate. Additionally, non-targeted metabolomics analyses were conducted in order to investigate further ALT metabolites. A sample preparation and LC-HRMS protocol was developed and validated. Sample preparation was performed with and without a hydrolysis step in order to quantify ALT glucuronides after enzymatic cleavage using β-glucuronidase (*helix pomatia* Type H-2). The targeted investigation showed that glucuronide conjugates of ALT and its photo-isomerization product (ALT-CAP) are main urinary metabolites of ALT in gilts. The concentrations in urine ranged from 1900 ng/mL to 4140 ng/mL, which is in a similar order of magnitude to data reported for horses. A normalization to urinary creatinine showed that it is crucial to adjust for the fluid balance of animals in order to determine a concentration profile over time. Furthermore, a possibly hitherto unknown isomerization product of ALT was observed at trace level. ALT and ALT sulfate were not detected in the urine samples of the gilts from the study presented herein. The chemometric analysis of non-targeted data revealed a clear difference between ALT-treated gilts and control animals. ALT glucuronide, ALT-CAP glucuronide and their in-source fragmentation products were in the top 25 significant variables in a univariate test setting distinguishing the

two test groups. Furthermore, a hydroxylated ALT glucuronide was tentatively identified. Surprisingly, in the control group of the gilts, several mass spectrometric hits were significantly increased. Database and mass spectrometric library comparisons suggested that isoflavonoid derivatives and stilbene derivatives were increased in the control group. This may indicate that the metabolism of phytoestrogens from feed is altered under ALT treatment. However, further studies are needed to support this hypothesis. Pathway mapping based on database hits indicated differences in the tryptophan degradation as well as C21-steroid biosynthesis pathways in the ALT-treated group compared to the control group. These changes can be explained by the mechanism of action of the hormonally active substance ALT.

Taken together, the results of this thesis illustrate the exceptional potential of non-targeted LC-HRMS analyses and metabolomics workflows. This approach provides more legal certainty for non-compliant samples in veterinary drug residue control and reveales in-depth information about the sample under investigation for environmental analysis.

Zusammenfassung und Ausblick

Flüssigkeitschromatographie gekoppelt mit hochauflösender Massenspektrometrie (LC-HRMS) ermöglicht die Etablierung von Non-Targeted-Screening-Analysen und ergänzt somit die Triple-Quadrupol-Massenspektrometrie in der Tierarzneimittelrückstandsanalytik und Umweltanalytik. Metabolomics-Workflows bieten vielversprechende Möglichkeiten zur Untersuchung und Identifizierung von Metaboliten und Biomarkern für metabolische Veränderungen in verschiedenen biologischen Matrizes. Die Herausforderung bei der Auswertung von Non-Targeted-HRMS-Daten besteht im Umfang der erhaltenen Datenmenge, da diese mehrere Tausend massenspektrometrische Variablen enthalten kann.

In Rahmen dieser Dissertation wurde eine Strategie zur Prozessierung und Auswertung von Non-Targeted-LC-HRMS-Daten etabliert und der vielversprechende Einsatz von Non-Targeted-Metabolomics-Analysen („Metabolic-Fingerprinting“) in unterschiedlichen analytischen Arbeitsgebieten — der amtlichen Tierarzneimittelrückstandsüberwachung und der Umweltanalytik — aufgezeigt.
Der etablierte Metabolomics-Workflow besteht aus (1) einer Datenauswerteprozedur, welche eine Peak-Liste mit allen Analyten aller Proben erzeugt, (2) chemometrischen Analysen zur Unterscheidung von unterschiedlichen Testgruppen, (3) univariaten und bivariaten statistischen Methoden zur Biomarkerkandidatenauswahl und (4) Stoffwechselwegsanalysen als Ergänzung zur manuellen Biomarkeridentifizierung. Eine robuste Datenauswerteprozedur wurde mit Hilfe von statistischer Versuchsplanung durch Anwendung eines teilfaktoriellen Plans entwickelt und optimiert. Die minimale Peakbreite, die minimale Signalhöhe und die gewählte Amplitude des Hintergrundrauschens bei der Chromatogramm-Dekonvolution wurden als signifikante Faktoren für eine verlässliche Peakdetektion, Integration und das Chromatogramm-Alignment identifiziert. Die Ergebnisse der statistischen Versuchsplanung wurden weiterhin genutzt um die finalen Prozessierungsparameter zu optimieren. Dieser experimentelle Aufbau zeigt die Vorteile einer statistischen Versuchsplanung auf, da eine überschaubare Anzahl an Experimenten wertvolle Informationen für die Entwicklung und Optimierung der Datenauswerteprozedur lieferte.

Ziel der ersten praktischen Anwendung des neu etablierten Metabolomics-Workflows war es, weitere Informationen über den Ursprung von Antibiotikarückständen in beanstandungswürdigen Proben der amtlichen Tierarzneimittelrückstandsüberwachung zu erhalten und Biomarker für Veränderungen im Metabolismus von erkrankten, medikierten Tieren zu

identifizieren. Die neuen Biomarker sollten mit minimalem analytischem Aufwand in der Routineüberwachung von Tierarzneimittelrückstanden angewandt werden können. Beanstandungswürdige Befunde können durch eine unsachgemäße Medikamentenanwendung bei infizierten Tieren aber auch durch eine Kontamination der Probe begründet sein. Eine Kontamination der Probe kann während der Probennahme oder Untersuchung der Probe zustande kommen oder durch kontaminiertes Futter oder Rückstände im Stall verursacht werden. Die Anwendbarkeit von Metabolic-Fingerprinting zur Unterscheidung von Proben von infizierten, medikierten Tieren (positive Proben) von denen gesunder Tiere (Kontrollproben) wurde getestet und erfolgreich demonstriert. Es wurden zwei HRMS-Plattformen (LC-Q-Orbitrap und LC-Q-TOF) verwendet, um die Anwendbarkeit des Ansatzes auf unterschiedlichen Geräten zu testen. Das Experiment wurde mit Muskel- und Nierenproben von Schweinen ($n = 50$) und Rindern ($n = 30$) durchgeführt. Der Hauptfokus lag auf der Untersuchung von Schweinemuskulatur, da in Deutschland in der Tierarzneimittelrückstandsüberwachung deutlich mehr Proben von Schweinen als von Rindern untersucht werden müssen. Eine Hauptkomponentenanalyse (PCA) mit einer anschließenden quadratischen Diskriminanzanalyse (QDA) wurde erfolgreich angewandt, um Schweinemuskulatur von infizierten, medikierten Tieren von der gesunder Tiere zu unterscheiden. Die Modelle zeigten eine hohe Sensitivität (0,826 (Q-Orbitrap) / 0,810 (Q-TOF)) und Spezifizität (0,976 (Q-Orbitrap) / 0,923 (Q-TOF)) für beide Geräteplattformen. Auffällig war, dass in beiden PCA-Scores-Plots Qualitätskontrollproben von gesunden Tieren, die mit 70 Tierarzneimitteln dotiert wurden, im selben Cluster lokalisiert waren wie Kontrollproben von gesunden Tieren. Diese Ergebnisse deuten darauf hin, dass der Hauptunterschied zwischen positiven Schweine- und Kontrollmuskulaturproben durch Veränderungen im Metabolom der Tiere begründet ist und nicht auf die Anwesenheit von Tierarzneimitteln zurückzuführen ist. Univariate statistische Tests, manuelle Biomarkerannotation und Stoffwechselwegsanalysen offenbarten Veränderungen in der Prostaglandinbildung sowie dem Metabolismus von ungesättigten Fettsäuren, welche als etablierte Marker für Infektionen und Entzündungen beschrieben sind. Zusätzlich wurden ROC (Receiver Operating Characteristic)-Kurven zur Auswahl und Bewertung von Biomarkern genutzt. Zwei Biomarker wurden identifiziert: das Tripeptid Prolylphenylalanylglycin und ein Lysophosphatidylcholinderivat. Eine bivariate Datenanalyse wurde durchgeführt, um die Vorhersagekraft dieser zwei Biomarker zu erhöhen. Eine 0,9-Prognoseellipse mit einem α-Fehler von 0,1 wurde anhand der Kontrollproben berechnet (Spezifizität von 0,9) und als Entscheidungsgrenze verwendet. Unter Anwendung dieses Ansatzes auf offizielle Proben der Lebensmittelüberwachung wurden kontaminierte Proben

eindeutig identifiziert. Dies zeigt, dass diese Vorgehensweise ein wertvoller Ansatz ist, um Rechtssicherheit bei beanstandungswürdigen Proben in der amtlichen Tierarzneimittelrückstandsüberwachung zu gewähren.

Chemometrische Analysen von Rindermuskulaturproben lieferten vergleichbare Ergebnisse. Es ist allerdings zwingend erforderlich Muskulaturproben von unterschiedlichen Tierarten getrennt auszuwerten. Die multivariate statistische Analyse von Kontrollproben von Schweine- und Rindermuskulaturproben ergab erhebliche Unterschiede. Der Hauptunterschied konnte auf massenspektrometrische Variablen zurückgeführt werden, welche als $[M+H]^+$, $[M+2H]^{2+}$ und $[M+3H]^{3+}$ des C-terminalen Oktapeptides von Rindermyoglobin identifiziert wurden.
PCA mit anschließender QDA von Nierenproben beider Tierarten ermöglichte keine ausreichende Differenzierung von Proben infizierter, medikierter Tiere von denen gesunder Tiere. Die Sensitivität der Modelle war mit 0,572 (Q-Orbitrap-Daten von Schweinenierenproben) und 0,795 (Q-TOF-Daten von Rindernierenproben) unzureichend, während die Spezifizität bei 0,903 bzw. 0,899 lag. Allerdings wurde eine Abweichung der Erwartungsvektoren beider Testgruppen in den Scores-Plots beobachtet. Zusammenfassend kann festgehalten werden, dass die multivariate statistische Analyse von Nierenproben weniger zur Unterscheidung von gesunden und infizierten Tieren geeignet ist.

Die zweite praktische Anwendung des neu etablierten Metabolomics-Workflows beschäftigte sich mit möglichen Umwelteinträgen von Altrenogest (ALT), einem in der intensiven Schweinezucht zur Brunstsynchronisation eingesetzten synthetischen Progesteron-Derivat. Es kann nicht ausgeschlossen werden, dass für gewisse geographische Regionen der zootechnische Einsatz von ALT in Jungsauen ein Risiko für Fische und andere Wasserorganismen darstellt, falls Gülle von mit ALT behandelten Jungsauen als Wirtschaftsdünger für landwirtschaftliche Flächen verwendet wird. Nach aktuellem Wissensstand sind keine publizierten Daten zur renalen Ausscheidung von ALT in Jungsauen verfügbar. Im Rahmen dieser Arbeit wurde daher eine Pilotstudie zur Probennahme von Urinproben von mit ALT behandelten Jungsauen und unbehandelten Kontrolltieren durchgeführt. LC-HRMS-Analysen wurden zur zielgerichteten Untersuchung von ALT und seiner bekannten Metabolite (Konjugate mit Glucuronsäure und ALT-Sulfat) durchgeführt. Zusätzlich wurden Non-Targeted-Metabolomics-Analysen zum Auffinden weiterer Metabolite von ALT angewandt. Eine spezifische Probenvorbereitung und das LC-HRMS-Protokoll wurden entwickelt und validiert. Die Probenvorbereitung wurde mit und ohne Hydrolyseschritt durchgeführt, um

Glucuronid-Konjugate von ALT nach einer enzymatischen Spaltung mittels β-Glucuronidase (*helix pomatia* Type H-2) zu quantifizieren. Die zielgerichtete Untersuchung zeigte, dass Glucuronidkonjugate von ALT und seinem Photoisomerisationsprodukt (ALT-CAP) die Hauptmetabolite im Urin von Jungsauen darstellen. Die Konzentration der Glucuronidkonjugate im Urin variierte von 1900 ng/mL bis 4140 ng/mL. Dies ist vergleichbar mit veröffentlichten Daten für die Tierart Pferd. Eine Normalisierung auf Urinkreatinin zeigte auf, wie wichtig es ist, den Flüssigkeitshaushalt von Tieren miteinzubeziehen, um ein Konzentrationsprofil über die Zeit zu bestimmen. Es wurde auch ein möglicherweise bislang unbekanntes Isomer von ALT im Spurenbereich nachgewiesen. ALT und ALT-Sulfat wurden in den Urinproben der Jungsauen nicht detektiert. Die chemometrische Analyse der Non-Targeted-Daten zeigte einen eindeutigen Unterschied zwischen mit ALT behandelten Jungsauen und der Kontrollgruppe. Glucuronidkonjugate von ALT und ALT-CAP sowie deren In-Source-Fragmentierungs-Produkte waren unter den 25 signifikantesten Variablen in univariaten statistischen Tests zur Unterscheidung der beiden Testgruppen. Weiterhin wurde ein Hydroxy-Derviat des Glucuronidkonjugats von ALT anhand der massenspektrometrischen Daten identifiziert. Auffällig war, dass in der Kontrollgruppe der Jungsauen verschiedene massenspektrometrische Variablen signifikant erhöht waren. Vergleiche mit Datenbanken und Spektrenbibliotheken wiesen auf eine Erhöhung von Isoflavonoidderivaten und Stilbenderivaten in der Kontrollgruppe hin. Dies kann möglicherweise darauf hindeuten, dass der Metabolismus von Phytoestrogenen aus der Nahrung unter ALT-Behandlung beeinflusst wird. Allerdings sind weitere Studien erforderlich um diese Hypothese zu unterstützen. Stoffwechselwegsanalysen basierend auf Datenbankergebnissen wiesen weiterhin auf Unterschiede im Tryptophanabbau sowie in der C-21 Steroidbiosynthese in der mit ALT behandelten Gruppe hin. Diese Unterschiede können durch den Wirkmechanismus der hormonell aktiven Substanz ALT bedingt sein.

Abschließend kann festgehalten werden, dass die Ergebnisse dieser Doktorarbeit das außerordentliche Potential von Non-Targeted-LC-HRMS-Analysen sowie Metabolomics-Analysen aufzeigen. Für die Tierarzneimittelrückstandsüberwachung bietet dieses Verfahren eine erhöhte Rechtssicherheit bei beanstandungswürdigen Proben und für die Umweltanalytik liefert es tiefgehende Informationen über die zu untersuchende Probe.

1 Introduction

1.1 Veterinary drugs in modern livestock farming

Parts of the following passage have previously been published in [1].

In the European Union, rules for the manufacturing, distribution, control and use of veterinary medicinal products are laid down in the Regulation (EU) 2019/6, which shall apply in January 2022, replacing the Directive 2001/82/EC [2, 3]. According to Article 4 (1) of Regulation (EU) 2019/6, veterinary medicinal products are defined as "*any substance or combination of substances which fulfils at least one of the following conditions:*

(a) it is presented as having properties for treating or preventing disease in animals;

(b) its purpose is to be used in, or administered to, animals with a view to restoring, correcting or modifying physiological functions by exerting a pharmacological, immunological or metabolic action;

(c) its purpose is to be used in animals with a view to making a medical diagnosis;

(d) its purpose is to be used for euthanasia of animals;" [2].

Veterinary medicinal products may only be placed on the market in the European Union, if they have been authorized by the competent federal authority or in the course of a European authorization procedure. The applicant for the marketing authorization needs to provide documents describing the pharmaceutical quality, efficacy and safety of the new product, including but not limited to data about its identity, purity and stability, intended field of use, recommended dosage and target animal species as well as data on pharmacodynamics, pharmacokinetics, toxicology, microbiology, ecotoxicology and clinical benefit. Furthermore, a withdrawal period must be specified for veterinary drugs intended for administration to food producing animals [4].

In the course of the authorization procedure for a veterinary drug for food producing animals the Committee for Medicinal Products for Veterinary Use (CVMP) of the European Medicines Agency (EMA) issues an opinion including a scientific risk assessment and risk management recommendations [5]. If necessary the CVMP recommends maximum residue limits (MRL) of the pharmacologically active substance in food gained from previously medicated animals [5]. The MRL is calculated with regard to the daily intake of the target tissues

(muscle, kidney, liver, etc.) and estimated exposure data. It assures that the worst case theoretical daily intake of the new substance does not exceed the acceptable daily intake (ADI) [5]. The ADI is an estimate of the amount of the pharmacologically active substance that can be ingested daily over a lifetime without appreciable health risks [5]. It can be derived from pharmacological, toxicological and microbiological data, which ever has the lowest no observed adverse effect level (NOAEL) [5]. Hence, food which contains residues below the MRL does not present any health risks to humans.

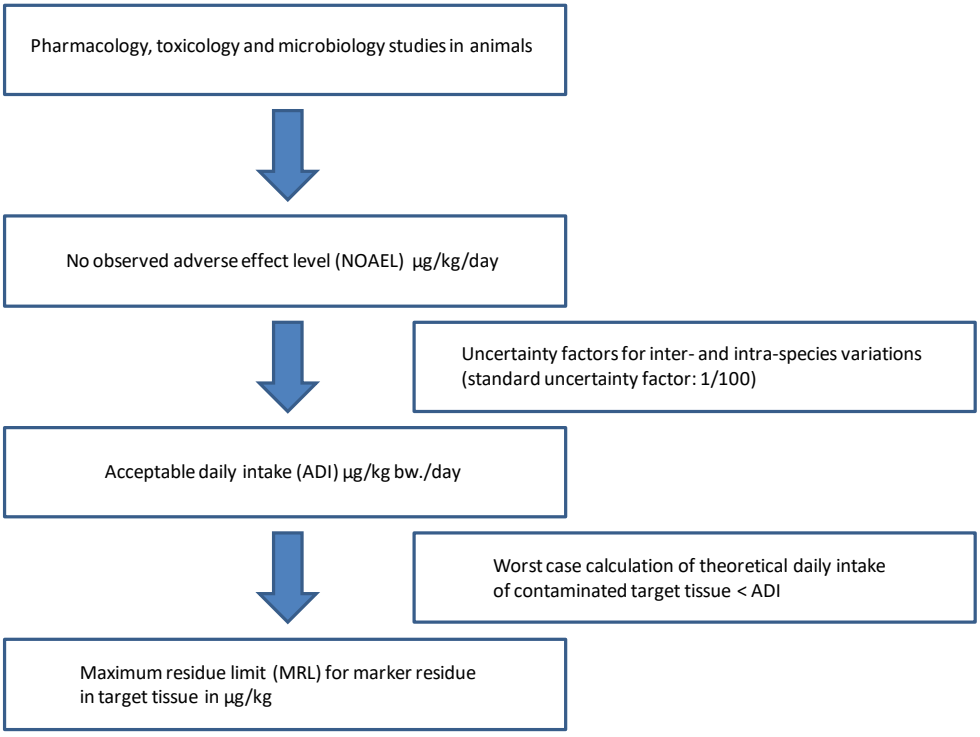


Figure 1. Derivation of maximum residue limits (MRL) for pharmacologically active substances under consideration of the acceptable daily intake (ADI) and daily consumption of target food tissue; bw.: bodyweight.

In consideration of the scientific opinion of the CVMP, the Commission classifies the pharmacologically active substance according to Article 14 (1) and (2) of Regulation (EC) No 470/2009 and sets a MRL, if necessary [5]. The classifications of the pharmacologically active substances are set out in the annex of Commission Regulation (EU) No 37/2010 [6]. Food from animal origin, which contains residues exceeding the MRL are considered to be not compliant with Community legislation according to Article 23 a) of Regulation (EC) No 470/2009 [5]. This food is not to be placed on the market according to Section 10 (1) of the German Food and Feed Code [7].

In the framework of preventive consumer health protection, residues of pharmacologically active substances in food are monitored in the national residue control plan according to Council Directive 96/23/EC[1] [8]. The national residue control plan (NRCP) is a food monitoring programme for pharmacologically active substances as well as contaminants like heavy metals or dioxins. It is implemented in every EU member state according to uniform criteria and is updated yearly. The NRCP lists the required kind and number of samples, the respective analyte spectrum, the analytical methods to be applied and the sampling method [9]. The German NRCP is coordinated by the Federal Office of Consumer Protection and Food Safety while the residue analysis is performed in each federal state independently. The number of samples for each state is determined based on the annual slaughter and production figures. Sampling is performed target-oriented at slaughterhouses and farming level, since the aims of the NRCP are the monitoring of the proper use of permitted veterinary drugs as well as the disclosure of illegal use of unauthorized or banned substances. Summarized data is reported to the European Commission [9]. Besides the European provisions in the NRCP, the German Animal Food Monitoring Regulation [10] states that two percent of all commercially slaughter calves and at least 0.5 percent of all other commercially slaughtered hoofed animals need to be tested for veterinary drug residues. In order to handle this large number of samples, a microbiological inhibiting test is used as an initial screening method. A positive result of this test needs to be confirmed with LC-MS analysis.

In Germany in 2018, 59,807 samples from animal origin were tested according to the NRCP and 277,308 samples were tested according to the German Animal Food Monitoring Regulation with a very low non-compliant rate of 0.42% and 0.09%, respectively [11]. These results are in line with the results of previous years [11].

The application of veterinary drugs such as antibiotics is essential for animal health maintenance and animal welfare. Thus, a prudent use of veterinary drugs is crucial in modern livestock farming. Several aspects need to be considered for their use. In general, veterinary drugs are only to be applied to treat or alleviate diseases. After the treatment of food producing animals it is important to respect withdrawal periods in order to ensure food safety. In the context of a one-health concept it is necessary to reduce antibiotic consumption as much as possible in order to prevent the selection and transmission of multi-drug resistant bacteria

[1] repealed by Regulation (EU) 2017/625 of the European Parliament and of the Council of 15 March 2017; translational measures according to Article 150 No. 1 of Regulation (EU) 2017/625 apply

[12]. Furthermore, environmental effects of veterinary drug residues in manure need to be considered [13].

A public summary of the consumption of veterinary drugs in Germany does not exist. However, in the light of the development of antimicrobial resistance, German pharmaceutical manufacturers are obligated to report their sold volume of antibiotics to an official register since 2011 [14]. In Germany, the sold volume of antimicrobials in veterinary medicine decreased constantly since 2011 from 1,706 metric tons (t) to a total of 670 t in 2019 [15]. The sold amount of antibiotics which are classified as "highest priority critically important antimicrobials for human medicine" by the World Health Organization (WHO) [16] such as fluorquinolones and 3rd and 4th generation cephalosporins also decreased by 1.7 t and 0.5 t, respectively [15]. However, regional differences in sales of antibiotics can be observed (see Figure 2).

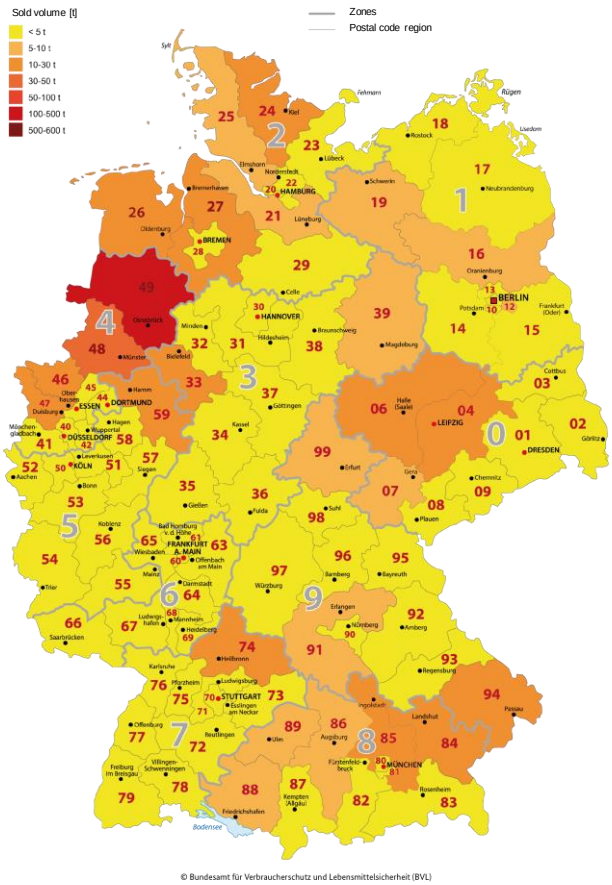


Figure 2. Map of sold volume of antibiotics in Germany for 2019, modified from [15].

This development may be attributed to an increased public awareness for the risks of the development of antibacterial resistance and initiatives from the European Union as well as changes in national laws [17]. In 2014, a bench-marking system was established in order to evaluate antibiotic use at individual farm level in correlation to nationwide antibiotic consumption. For this purpose, the semi-annual treatment frequency is determined according to § 58c of the German Medicinal Products Act [4]. The treatment frequency is calculated by multiplying the number of treated animals with the treatment days, divided by the mean number of animals of the farm in the time interval of six months. The median is set as a benchmark threshold to consult with a veterinarian and the 75% percentile requests a plan to reduce antibiotic use [4]. In general, the treatment frequency for broilers, turkeys and pigs are at a higher level than that for cattle, but a comparison of treatment frequencies of different species is difficult [18].

Up to now, veterinary residue analysis is mainly performed with conventional targeted multi-analyte methods using liquid chromatography coupled with mass spectrometry. This approach can reliably detect the presence of antibiotic residues, but it is not possible to determine if the residue is caused by improper drug administration to an infected animal or by sample contamination. Possible causes for sample contamination include analyte cross-over during sampling, transport or analysis as well as polluted feed, environmental contamination, drug residues in farm-animal houses [13, 19] or animal-to-animal transfer via ingestion of feces or urine [19]. Residues of veterinary drugs which are administered orally via feed can easily be transferred to non-medicated feed when milling tools or other tools like barrows are not properly cleaned. Naturally occurring toxins like mycotoxins such as α- and β-zearelenol in mouldy silage can be metabolized to zeranol, a banned estrogen [19]. Studies have shown that non-compliant residues of sulfamethazine can be found in unmedicated pigs which were housed in a stable which prior housed medicated pigs [20, 21]. Furthermore, it was shown that antibiotic residues can be detected in dust of pig confinement buildings. Thus, dust needs to be considered as cause for contamination and route for entry of veterinary drugs in the environment [22]. The presence of metabolites proofs that the parent drug was metabolized by a living animal and a cross-over during sampling, transport or analysis can be excluded. However, some antibiotics undergo no or only poor metabolic transformation, e.g., tetracyclines [23]. In such cases, biomarkers indicating that the non-compliant sample originated from an infected, drug-treated animal would be a promising approach in terms of legal certainty.

1.2 Hormonal preparations as veterinary medicinal products in modern livestock farming

Parts of the following passage have previously been published in [192].

Hormones are messenger substances which are produced by glands and are emitted in the blood circulation system reaching distant target organs and affecting various biochemical processes in the body, e.g., regulation of electrolytes, energy metabolism, regulation of blood pressure, growth, development and reproduction as well as behavior, stress and mood [24]. Hormonal effects are triggered by hormones binding to high affinity receptors at very low concentration levels (ppt level) [25]. Sex steroids, a subclass of steroid hormones, are crucial for reproduction and development. They are synthesized from cholesterol within the gonads, the adrenal glands and the placenta [26]. Important groups thereof are androgens, estrogens and progestogens (for details see Section 1.2.1).

In the past, substances with hormonal action were used for fattening purposes in livestock farming. Recognizing that pharmacologically active substances with hormonal action act at very low levels and their residues in foodstuff of animal origin may pose a risk for consumer health and affect the quality of meat, Council Directive 96/22/EC prohibits the use of a number of substances having a hormonal action in stock farming [27]. However, the cited Council Directive also lays down conditions for an exceptional use. One reason for an exceptional use is estrus synchronization as a zootechnical treatment. These regulations are the legal basis for the authorization of altrenogest (ALT), a synthetic progestogen, for the synchronization of estrus in gilts and mares. In 2012, the MRL of ALT was set to 2 µg/kg and 4 µg/kg in *porcine* liver and fat tissue, respectively [6]. This represents an increase from the previously set MRLs of 0.4 µg/kg and 1 µg/kg for *porcine* liver and fat tissue, respectively. The CVMP reasoned the recommendation for the new MRLs with the modification of the ADI for ALT. The most sensitive effects in pharmacology and toxicology studies were related to the hormonal activity of ALT resulting in a overall no observed adverse effect level (NOAEL) of 4 µg/kg bw/day. Using the uncertainty factor 20 (inter- and intraspecies uncertainty factors 2 and 10, respectively), a refined ADI of 0.2 µg/kg bw/day was established on the basis of new pharmacodynamic and pharmakinetic data submitted to the CVMP [28]. The selected marker residue was the parent compound, although this substance makes up only a small percentage of total residues in the target tissue liver. However, a ratio between the marker residue and the total hormonally active residue could be established and was considered in the deri-

vation of the MRLs [28, 29].

The application of ALT in pig breeding is important to ensure a smooth production rhythm. In 2019, pig meat production accounts for 60.4% of the total German commercial meat production followed by poultry meat and beef with 18.7% and 13.2%, respectively [30]. The German commercial pig meat production increased in the years of 2000 to 2019 from 3,923,453 t to 5,227,588 t [31]. In order to achieve these outputs, there were 21,200 agricultural holdings for pig keeping with an average of 1,229 pigs per holding in 2019 [30, 32]. The agricultural holdings with the highest number of animals were located in Lower Saxony followed by North Rhine-Westphalia with 1,600 and 1,065 livestock per holding, respectively [33, 34]. Since it is essential to uphold the production chain, 209,900 pregnant gilts and 1,081,700 pregnant sows were farmed in 7,200 holdings for pig breeding in Germany in 2019 [31, 32]. In order to efficiently integrate gilts in the production rhythm, synchronization of estrus and ovulation is of particular importance in pig breeding [35-37]. A simultaneous estrous cycle of gilts and sows facilitates strict batch farrowing (all-in-all-out system) [38]. Thus, it allows for more efficient use of labor as well as facilities and guarantees a steady production number of piglets for fattening. It further facilitates a better hygiene management with a decreased use of antibiotics in piglet breeding, since a disinfection of the entire housing complex after every piglet group can be performed [39, 40]. The optimization of the production rhythm has substantial economic value, since farrowing by gilts represents about 20% of the total farrowing rate [41].

The control of the reproduction cycle in pig breeding is mostly organized in a 3-week or in a 7-day production cycle depending on the size of the breeding farm [42] and is based on the synchronization of estrus in weaned sows and gilts. Artificial synchronization of estrus in pluriparous sows is not necessary, since the simultaneous weaning leads to a synchronized estrous cycle [38]. 4–6 days after weaning the estrus of sows can be expected [43]. ALT is the only authorized active ingredient for estrus synchronization in gilts. It is administered at 20 mg/d for 15–18 days to induce a blockage of cycle. Up to 5 days after the last ALT administration estrus can be observed [39]. An extrapolation from the sales numbers in Lower Saxony to Germany resulted in an estimated consumption of about 250 kg ALT per year in Germany [44].

1.2.1 Pharmacodynamics of ALT

In order to understand the pharmacodynamics of ALT, it is imperative to have a basic knowledge about the neuroendocrine control system of the estrous cycle. A brief overview is given in the following.

Neuroendocrine control of estrus in female pigs

The hypothalamus and the pituitary gland act as superior control systems for hormonal processes (see Figure 3). During the estrous cycle, they are subjected to a negative and positive feedback mechanism by sex steroids. Gonadotropin-releasing hormone (GnRH)—a decapeptide hormone—is synthesized in the hypothalamus and controls the release of sex steroids. It is secreted in the hypothalamic-hypophyseal portal circulation reaching the anterior pituitary. GnRH pulses stimulate the synthesis and secretion of the gonadotropins follicle-stimulating hormone (FSH) and luteinizing hormone (LH) in the anterior pituitary. FSH and LH stimulate the synthesis of estrogens (mostly 17-β-estradiol), progestogens (progesterone) and to a minor extent androgens (testosterone) in the follicular theca and granulosa cells. Progesterone and small amounts of 17-β-estradiol inhibit the GnRH-secretion in the hypothalamus and reduce the sensitivity of the GnRH-receptors of the anterior pituitary via a negative feedback mechanism. The glycoprotein inhibin, which is mostly synthesized in the granulosa cells, inhibits the secretion of FSH from the anterior pituitary. High estrogen levels exhibit a positive feedback mechanism on the secretion of gonadotropins [24, 39, 45, 46].

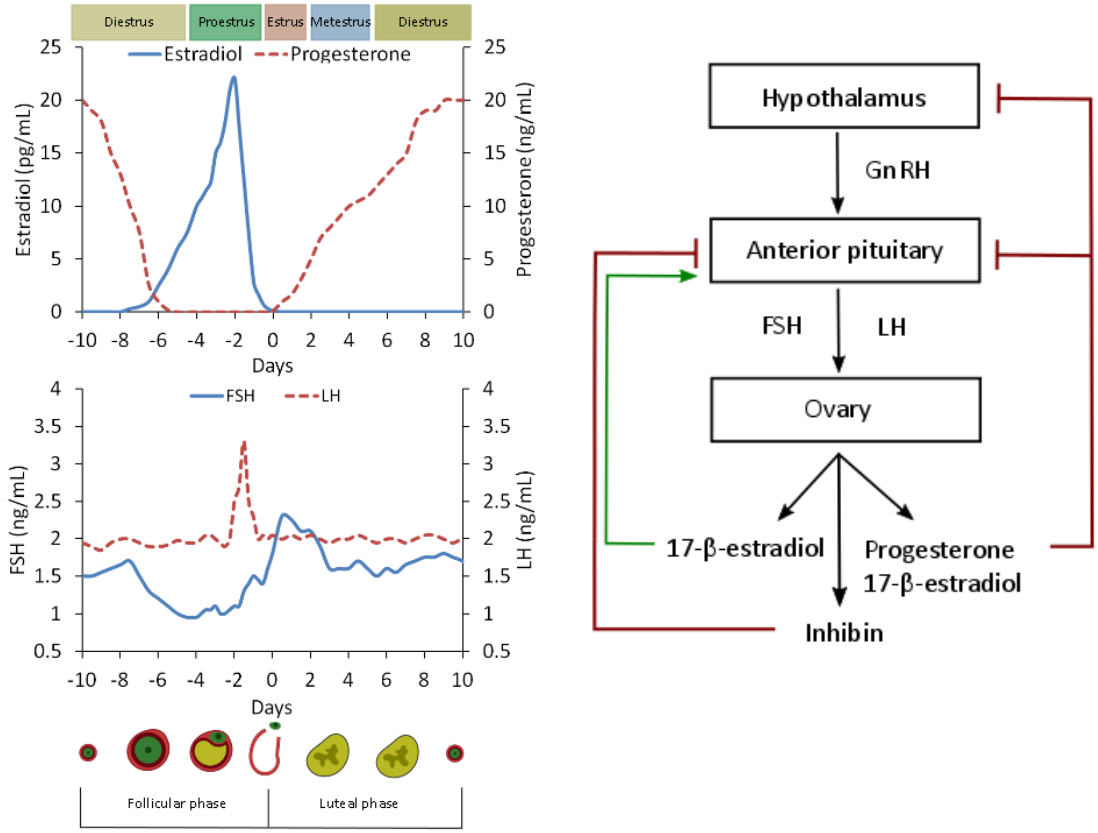


Figure 3. Left: plasma concentrations of estradiol, progesterone, LH and FSH in the course of estrous cycle (modified according to data from [47]). Right: superior controlling system of the synthesis of sex steroids during female cycle (modified according to [24]).

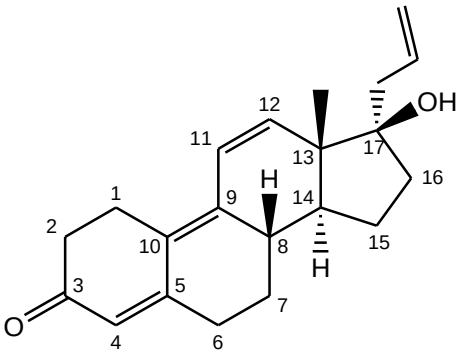
Table 1. Overview of peptide and sex hormones in estrous cycle, their most important physiological effects in female and male body as well as site of biosynthesis [24, 39].

Peptide hormones	
Gonadotropin releasing hormone (GnRH)	- synthesized and released from hypothalamus - stimulates biosynthesis of FSH and LH
Follicle stimulating hormone (FSH)	- synthesized and secreted from anterior pituitary - initiates follicle proliferation and maturation in ovaries - is in synergy with LH responsible for estrogen synthesis in follicles
Luteinizing hormone (LH)	- synthesized and secreted from anterior pituitary - triggers ovulation and development of corpus luteum - is in synergy with FSH responsible for biosynthesis of androgens in follicles
Sex steroids	
Estrogens, e.g., 17- β-estradiol	- mainly synthesized in ovaries and placenta - promotes growth of female sex organs and female secondary sex characteristics - regulates in synergy with progesterone estrous cycle - exerts anabolic effect to female reproductive system
Progestogens, e.g., progesterone	- synthesized mainly in corpus luteum, placenta and the adrenal glands - regulates estrous cycle and maintains pregnancy
Androgens, e.g., testosterone	- synthesized in testes, ovaries and adrenal glands - mainly responsible for the development of male primary sex organs and male secondary sex characteristics, spermatogenesis - precursor of estrogens in biosynthesis - exerts an anabolic effect on muscle growth

The estrous cycle of pigs is on average 21 days long, but can vary individually from 18–24 days. It can be divided in four phases: proestrus (days 18–21), estrus (days 1–2), metestrus (days 3–6) and diestrus (days 7–18) [39, 43, 48]. Luteolysis of corpus luteum starts at the end of diestrus while progesterone concentration sinks to its lowest level during the estrous cycle [47]. Consequently, the negative feedback mechanism of progesterone on the secretion of the gonadotrophic hormones stops and FSH as well as LH are secreted from the anterior pituitary, at the beginning of proestrus [39]. High estrogen concentrations at the end of the late follicular phase induce a pre-ovulatory LH surge, indicating the beginning of estrus. Estrus is the period around ovulation, which is characterized by the "standing response" of the sows in presence of a boar indicating their receptive behavior for mating [36, 39, 43]. Ovulation starts around 24–42 hours after the LH surge and lasts about 1–7 hours [37]. It is characterized by increasing progesterone levels [47]. During metestrus, theca and granulosa cells of the ovulated follicles transform to luteal cells. Progesterone synthesis and secretion rises with increasing rate of lutenization [47]. At the end of metestrus lutenization is complete and in case of a pregnancy the corpus luteum secrets progesterone inhibiting follicle proliferation [37]. Diestrus lasts from day 6–17. After the maximum progesterone levels are reached at day 12, luteolysis occurs at day 15 [39, 43]. Low levels of progesterone lead to the start of proestrus, while once again gonadotropins stimulate follicle proliferation [43].

Pharmacodynamics of ALT—estrus synchronization of gilts for zootechnical purpose

ALT, also called allyltrenbolone, belongs to the 19-nor-testosterone substance class and is a synthetic progesterone agonist. Its main pharmacodynamic actions are based on progestomimetic and anti-gonadotrophic effects. However, weak estrogenic, anabolic and androgenic effects are also described [28, 49, 50].



($C_{21}H_{26}O_2$)
Altrenogest, allyltrenbolone
CAS No: 850-52-2
17-a-allyl-17-b-hydroxyestra-4,9,11-trien-3-one

Figure 4. Chemical structure of altrenogest (ALT).

For a duration of 15–18 days, 20 mg ALT are administered once a day to cycling gilts to synchronize estrus [37, 51]. ALT suppresses the secretion of the gonadotropins FSH and LH from the anterior pituitary via a negative feedback mechanism, resulting in a blockage of cycle and a stop of follicle maturation at a medium sized stage (antral follicle, 5 mm) [51-53]. The function of the corpus luteum is not compromised, leaving it resistant to the luteolytic action of uterine prostaglandin $PGF_{2\alpha}$ secretion before day 12 of the estrous cycle [54]. After day 12 of the estrous cycle, the sensitivity of the corpus luteum to prostaglandins increases and luteolysis begins [37, 55]. Thus, at the end of the treatment a regression of corpora lutea of all gilts can be observed [36]. After cessation of ALT treatment, gonadotropins are secreted from the pituitary and LH induced follicle maturation of antral follicles to pre-ovulatory follicles starts. Peripheral estrogen concentrations increase about 3–4 days after the last administration of ALT and about 5 days after treatment the pre-ovulatory LH surge begins and estrus is observed [53, 56, 57]. These progestomimetic effects of ALT enable a synchronization of estrus, regardless of estrous cycle stage at the first ALT administration [56, 58]. Furthermore, studies have shown an increase in ovulation rate, size of antral follicle and litter size [55, 58, 59]. Soede *et al.* concluded, that ALT exhibits a lesser inhibiting effect on the development of antral follicles than endogenous progesterone [55].

The administration of ALT is recommended after the first estrus of gilts, since it can cause reduced fertility in pubertal gilts due to their not fully developed uteri [35, 37, 58]. Furthermore, ALT can be used to prolong gestation in order to synchronize farrowing and to postpone post-weaning estrus in order to improve reproductive performance by allowing the sows to recover a better body condition [54, 60-63].

1.2.2 Pharmacokinetics of ALT

Up until recently, scientifically published data about the pharmacokinetic profile and residue depletion of orally administered ALT in gilts was limited and fragmented. Pharmacokinetic data presented by the marketing authorization holders in the course of the European and American drug administration is not publicly available. The only data available was accessible through product information sheets or was taken from scientific opinions from the CVMP and a review of registration data of an ALT containing drug provided by the American Food and Drug Administration (FDA) according to the amendments to the Freedom of Information Act [28, 51, 64, 65]. The 2019 study from Xiao *et al.* was the first scientific and

publicly available plasma-pharmacokinetic study in gilts [66]. Urinary excretion data is still scarce.

In general, the results of Xiao and colleagues were coherent with the fragmented data from the marketing authorization holders. ALT is readily absorbed after the first oral administration, reaching a plasma peak concentration after 1.96 ± 1.45 hours with a mean plasma concentration of 66.16 ± 19.94 ng/mL [66]. After 24 hours the mean plasma concentration decreased to 5 ng/mL [66]. On day 18 of the treatment, maximum plasma levels of 71.32 ± 19.96 ng/mL were reached after about 3 hours [66]. Data presented to the FDA from marketing authorization holders showed a plasma peak concentration of 36.5–55.5 ng/mL at 1–4 hours after repeated dosing [65]. Plasma concentration decreased biphasically with an elimination half life of 9.7 ± 2.83 hours after repeated treatment [66]. The mean residence time of ALT was prolonged after repeated treatment and the apparent body clearance decreased. A significant accumulation could not be observed [66]. Similar pharmacokinetic parameters were observed when two different ALT preparations were investigated [67].

ALT is distributed mainly to the liver and to a minor extent to fat, kidney and muscle tissues [28, 68]. Data provided by the marketing authorization holders in the EU showed that ALT undergoes an extensive metabolism via oxidation and conjugation, which is in line with all steroids. Dealkylation was not observed [28]. In liver extracts, conjugation products with glutathione and glucuronic acid as well as an isobaric form and hydroxylated products were observed. These metabolites showed a reduced hormonal activity of about 14–21% compared to ALT in *in vitro* assays [28]. The reduced hormonal activity of conjugation products is in line with the findings for other steroid metabolites and can be explained by the higher polarity of the metabolites. Hence, they are suspected to have a reduced affinity to sex steroid receptors [28]. However, it can be expected that glucuronic acid conjugates regain their hormonal activity after enzymatic hydrolization in the gut [28].

Residue depletion studies cited in the authorization process showed that the highest tissue residues of radioactivity after dosing of radio-labeled ALT were detected in liver followed by kidney and to a lesser extent in fat and muscle (see Appendix Table S1) [28, 64, 65]. In contrast, the highest concentration of the parent compound ALT was found in liver followed by fat tissue and to a lesser extent in kidney and muscle (see Appendix Table S1) [28, 68]. The difference between the total residue levels (determined as radioactive residues) and the residue levels of the parent compound indicates that free ALT makes up only a small part of the

total residues in liver. In pig liver 68–79% of the administrated ALT dose is bound to macromolecules such as protein [65, 68]. The specified withdrawal time of registered veterinary drugs with ALT as active ingredient is 9 days [51, 66] and the depletion studies showed that after 7 days ALT residues were below MRL (see Appendix Table S1) [28, 68].

Excretion data is limited and inconsistent for the species pig. The only data available is given by the marketing authorization holders in the course of the authorization process. Two studies are cited in the scientific opinion of the CVMP, which report contradictory results. One study reported that elimination via bile is the major excretion pathway with only 20% of renal excretion, whereas in the second study primarily renal excretion (60%) was observed [28]. In the scientific opinion of the CVMP in 2016 concerning the potential environmental risk of ALT, studies with radio-labeled ALT were cited that showed that after prolonged treatment with 20 mg/d of ALT only 25% of the total dose was excreted in urine (sum of ALT and ALT conjugates) and 75% was excreted via bile [29]. In urine, 2% of the radioactivity was attributed to ALT and 24% to its conjugates, corresponding to 0.5% and 6% of the total dose. In bile, 6% of the radioactivity was caused by ALT and 14% by its conjugates, corresponding of 4.5% and 10.5% of the total dose, respectively. A total excretion of about 5% of ALT and 16.5% of its conjugates was reported [29]. Furthermore, a renal excretion of ALT of 44% and 53% excretion via bile within 24 hours in horses was reported [28]. Glucuronides and sulfate conjugates are described as renal metabolites of ALT in horses [69]. Twelve days after repeated treatment ALT levels in *equine* urine were below the limit of detection (2 ng/mL) [50].

1.2.3 Environmental impact of veterinary drugs with hormonal action

Synthetic substances with hormonal actions are subject to public discussion since they are environmental contaminants and belong to the group of endocrine disrupting chemicals [25]. According to the definition by the WHO and the United Nations Environment Programme (UNEP) an endocrine disruptor is "*an exogenous substance or mixture that alters function(s) of the endocrine system and consequently causes adverse health effects in an intact organism, or its progeny, or (sub) populations*" [25]. The increase in incidences of endocrine-related disorders of the reproduction system and immune system as well as of cardiovascular, thyroid or metabolic diseases cannot be explained solely by genetic predispositions, but epigenetic influences, e.g., nutrition, or environmental contamination, are likely to play a role in this development [25]. Studies showed an increase in feminized male fish living in waters contaminated with estrogens, such as, 17-β-estradiol, estrone and ethinylestradiol [25, 70-73]. Even low concentra-

tions of sex steroids (ppt-level) are environmentally relevant and lead to negative effects on the reproduction system of fish and other aquatic organisms [74]. Of particular concern is that sex steroids produce a non-linear dose-response curve and can lead to irreversible damages in the reproduction system, if the organism is exposed at critical time points, such as during fetal and postnatal development [25]. Thus, exogenous sex steroids exhibit a potent effect on the sensitive endocrine controlled reproduction system. Studies from Jenkins *et al.* and Orlando *et al.* describe a negative effect of sex steroids originated from effluent from paper mills or cattle feedlot on the reproduction systems of fish. Androgens, such as, androstendione, are the cause of a masculinization of female fish and mixtures of sex steroids alter the hypothalamic–pituitary–gonadal axis inhibiting the release of GnRH or gonadotropins [75, 76]. Furthermore, trenbolone, a growth promoting substance used in cattle feedlot, has an anabolic and androgenic effect and leads to masculinization of female fish [77].

Veterinary medicinal products containing ALT are also discussed to pose a risk to fish and other aquatic organisms [78]. Very low concentrations of ALT (< 0.4 ng/L) showed relevant effects on the fertilization and survival rate in a two-generation study in fish [78]. Furthermore, a shift in sex ratio toward males caused by the androgenic effect of ALT is described and is in line with findings of other synthetic progestogens used in medicinal products for human use [78, 79]. According to the "guideline on environmental impact assessment for veterinary medicinal products in support of the VICH guidelines GL6 and GL 38" the CVMP established a predicted environmental concentration (PEC) in soil for ALT and its conjugates of 0.013 µg/kg [29, 80]. For surface water, PEC ranged from 16–219 pg/L, depending on different soil types for run-off calculations [29, 80]. A variety of factors are considered in this modeling, i.e, ecotoxicological data, degradation in manure, different soil conditions in Europe, metabolism and application data. In consideration of the predicted no effect concentration (PNEC) for surface water of 0.04 ng/L ALT, a risk to fish and other aquatic organism associated with the zootechnical use of ALT could not be excluded for certain geographical areas [29, 81]. However, in the absence of viable alternatives to ALT for the synchronization of estrus in gilts, the CVMP considered the benefit-risk balance positive but recommended the addition of risk mitigation measures to the product information [78]. The hormonal activity of ALT is further underlined by a publication from the FDA, which highlights the potential health risks for people exposed to ALT products for horses or pigs [82]. Reported reproductive adverse effects include abnormal or absent menstrual cycles in women, and in men

decreased libido. Further adverse effects reported were headaches, fever, abdominal pain, nausea, diarrhea, vomiting and rashes [82, 83].

Scientific data about environmental residues of ALT and its behavior in environmental samples are scarce. In an environmental residue monitoring program in the U.S., ALT was found in 4% of 50 tested lakes in Minnesota [84]. In Czech Republic ALT was detected in influent and effluent of a wastewater treatment plant (WWTP) at a concentration of 0.35 ng/L and 0.15 ng/L, respectively [85]. The same study analyzed 17 progestogens in waste water and surface water. Progesterone was detected in all samples, while synthetic progestogens like ALT were only detected in influent and effluent waste water samples. Hence, WWTPs were not capable of removing progestogens entirely during treatment process [85]. A study from Sauer *et. al* showed that ALT exhibits androgenic activities of up to 14% of the reference substance dihydrotestosterone in an androgen receptor-specific chemically activated luciferase gene expression bioassay, but did not exhibit anti-androgenic activity [86]. This study also reported that progestins mostly occurred in influents of WWTPs and accounted for up to 29% of androgenic activity, but their contribution to an androgenic activity in effluents was negligible [86].

Furthermore, primary and secondary photo-products of ALT, generated in water under environmentally relevant conditions, showed significant androgenic activity in *in vitro* cell assays at picomolar to nanomolar concentrations [87]. These concentrations are in the range of the above mentioned PEC of ALT. Pflug *et al.* showed that the photo-isomer of ALT exerts 45–100% of the activity of ALT in activating estrogen receptors, progesterone receptor and pregnane X receptor [88]. This emphasizes the importance of further studies on the fate of ALT and its metabolites in environmental samples as well as of studies on the pharmacokinetics of ALT to determine the elimination rate and identify potential environmentally relevant metabolites.

Sorption studies at laboratory scale showed further that an accumulation of ALT in different soils cannot be excluded and biotic as well as abiotic elimination needs to be considered [44, 89]. Hence, a potential risk for the environment based on the administration of manure from ALT-treated gilts to agricultural soils cannot be ruled out completely, especially in regions were manure from gilts is used [44]. However, the estimated consumption of ALT of 250 kg in Germany does not suggest the need for the recommendation of a national environmental monitoring program [44].

1.3 Non-targeted LC-HRMS screening methods and metabolomics workflows

Parts of the following passage have previously been published in [1].

To this day, the majority of analytical methods for veterinary drug residue monitoring and environmental analysis are for good reasons based on targeted multi-residue approaches using liquid chromatography coupled to quadrupole-based tandem mass spectrometry. Triple quadrupole mass spectrometers (QqQ) provide high sensitivity and accuracy as well as robustness combined with a large dynamic range. Hence, QqQ are still considered to be the "gold standard" in veterinary drug residue analysis [90]. However, this technique is intrinsically limited to a predefined number of known analytes. For instance, it is not possible to perform a retrospective analysis of metabolites of veterinary drugs or biomarkers associated with a drug administration. Liquid chromatography hyphenated to high resolution mass spectrometry (LC-HRMS) offers a complementary approach. High resolution full scan methods can be used to quantify veterinary drugs at trace levels [90]. Moreover, it is possible to perform a non-targeted screening for analytes of interest such as transformation products or metabolites in parallel and retrospectively [91-93]. LC-HRMS provides important information for the structural elucidation of unknown analytes, such as chromatographic retention time, accurate mass, isotope distribution and fragmentation pattern. Thus, it is possible to identify analytes by library spectrum matches, although a reference standard still renders the highest level of analyte identification [94].

Besides non-targeted screening methods, metabolomics workflows are versatile tools to obtain in-depth information about the sample under investigation. Metabolomics approaches are widely used for biomarker discovery in clinical or pharmaceutical applications, cancer research or drug discovery [95], but they have also been successfully applied in food safety, food quality and traceability [96]. The terms "metabolomics" and "metabonomics" describe the exhaustive analysis of the metabolic profile of a biological system and can be used interchangeable [97, 98]. Metabolites are intermediates of metabolic pathways, typically with a molecular weight below 1.5 kDa [96]. Thus, metabolomics represents best the molecular phenotype, i.e., the current state of the biological system influenced by the genomic predisposition as well as environmental exposures. In general, there are two strategies in metabolomics approaches, i.e., metabolic fingerprinting and metabolic profiling. Fingerprinting refers to the analysis of as many compounds as possible in order to submit the obtained data to

statistical analysis for pattern recognition, whereas identification and quantification of the analytes are not mandatory steps. In contrast, profiling refers to the analysis of a set of related metabolites including their quantification [96].

The use of metabolomics for veterinary drug residue monitoring is a young field of research. Only a few studies on this subject have been published until now [99-104]. Two examples include the analysis of tissue samples from chickens, which were previously treated with amoxicillin [99] or enrofloxacin [100], respectively. Another study investigated metabolic changes in cattle urine after steroid treatments [101]. Regal *et al.* was able to highlight significant metabolic modifications in serum after to the administration of estradiol and/or progesterone [102]. Metabolomics-based strategies were used to determine diagnostic markers for the illegal use of low-dosed "cocktails" of β_2-agonists in *porcine* urine samples [103] and to propose potential biomarkers of illicit 5-nitroimidazole abuse in pigs [104]. These initial studies show that metabolomics approaches can provide insights into the metabolome of drug-treated animals. Hence, metabolic fingerprinting may be a valuable tool to evaluate the origin of residues of veterinary drugs, i.e., have the residues been caused by improper drug administration to an infected animal or simply by contamination.

Furthermore, the number of applications of non-targeted LC-HRMS analysis or metabolomics is also increasing in the field of environmental analysis [105, 106]. Anthropogenic substances, e.g., veterinary drugs, are released into the environment through various pathways such as effluents from WWTP, surface run-off or sorption to soil. However, not only the pharmaceutical itself is discharged into the environment but also its transformation products which can be formed due to biological or physicochemical processes. It is crucial to identify transformation products of these pharmaceuticals in order to assess the risk for exposed organisms since the toxicity of these mixtures may be altered compared to the toxicity of the original pharmaceutical product. Moreover, under real-world conditions environmental organisms are exposed to mixtures of pollutants at sub-lethal levels. Non-targeted LC-HRMS analysis can be applied to investigate transformation products of chemicals such as pharmaceuticals or other pollutants [107, 108]. Metabolomics analysis can be further used to study the organism's response to environmental stressors, e.g., after exposure to WWTP effluent [108]. These applications show the benefits of non-targeted LC-HRMS in order to investigate environmental exposure of pharmaceuticals including veterinary drugs.

1.3.1 High resolution mass spectrometry: Q-Orbitrap vs. Q-TOF

High resolution mass analyzers are characterized by a resolving power of > 10,000 [109]. Mass resolving power for an isolated peak is defined as $m/m_{\Delta 50\%}$, referencing the mass of the spectral peak to its full width at half maximum (FWHM) [110]. Mass resolution is defined as the minimum mass difference between two mass spectral peaks, $m_2 - m_1$, at a given percentage of the valley between (e.g., 10% valley method). It determines the ability to distinguish ions of different elemental composition but with the same nominal mass. Mass accuracy is defined as the difference between a measured mass and the calculated exact mass [109]. Two types of high resolution mass analyzer were used for this thesis and are described in the following.

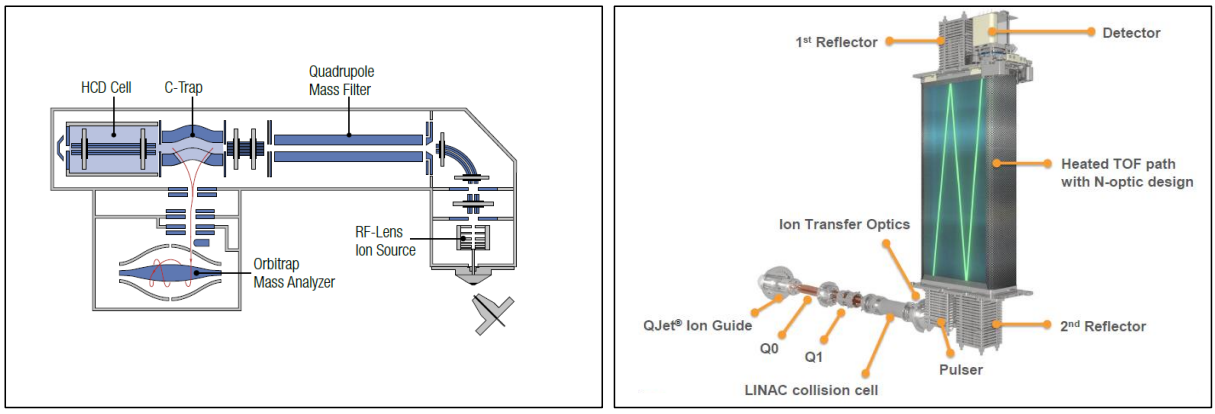


Figure 5. Left: schematic assembly of Q Exactive Focus (photo is provided courtesy of Thermo Fisher [111]). Right: schematic assembly of X500R QTOF (photo is provided courtesy of SCIEX (www.sciex.com) [112]).

Orbitrap™ mass spectrometers

Orbitrap™ mass spectrometers are high resolution mass analyzers which operate by trapping ions in electrostatic fields [113]. The first Orbitrap was described by Alexander Makarov in 2000 [114], but the principal concept of orbital trapping was yet described in 1923 [115]. There is a series of instrument models based on this mass analyzer technique with different spectrometer architectures. In this thesis, a Q Exactive Focus model was used (Thermo Fisher Scientific). Hence, the following description is focused on the operating principle of this particular model.

After ionization of the molecules via atmospheric pressure ionization (e.g., heated electro spray ionization (H-ESI)), charged ions are focused in the ion source optics by the means of direct current (DC) and radiofrequency (RF) voltage and a vacuum gradient. The ions pass a hyperbolic quadrupole mass filter, which can be used for precursor ion selection for parallel-

reaction monitoring (PRM), selected ion monitoring (SIM) or data-dependent fragmentation. After the quadrupole mass filter, the ions are accumulated in the RF-only bent C-Trap, where their energy is dampened in a bath gas (nitrogen). The C-Trap focuses the ions in a tight package before ejecting them either into the higher-energy collisional dissociation (HCD) cell to produce product ions or into the Orbitrap analyzer. All ions are released simultaneously to the Orbitrap analyzer by applying DC voltage, but light ions travel faster to the analyzer than heavy ones, getting squeezed closer to the central electrode [111]. The Orbitrap consists of an axial spindle shaped central electrode, which is surrounded by a pair of bell-shaped outer electrodes. It is an ion trap, capturing the moving ions in ion packets around the central electrode by electric fields, since the electrostatic attraction is compensated by the centrifugal force from the initial tangential velocity. The ions form stable rotations around the axial central electrode while oscillating harmonically along it. The frequency of this harmonic oscillation depends on the mass-to-charge (m/z) ratio and the device-specific field curvature. An image current produced by the oscillating ions is detected by the outer electrodes and via Fast Fourier Transformation (FFT) the frequency of the axial oscillation and hence, the m/z of the ions can be determined. Thus, multiple masses can be detected at the same time and with a high resolving power. Automatic gain control (AGC) is applied in order to avoid trap overfilling, which can cause mass shifts or detector saturation [111]. The Q Exactive Focus can be operated at a resolving power of up to 70,000 (full width half maximum (FWHM) at m/z 200) with a scan rate of 3 Hz. This high resolving power results in a mass accuracy below 5 ppm for most ions. Faster scan rates of 12 Hz can be achieved with a resolving power of 17,500. Furthermore, the wide linear dynamic range of up to 6 orders of magnitude is a powerful tool in pharmacokinetic experiments [111, 113, 116].

Time-of-flight mass spectrometer

The first time-of-flight (TOF) analyzer was called “velocitron” and was described in 1948 [117]. TOF analyzers measure the m/z ratio based on the time of flight through a field-free flight path. The measurement principle is based on the conversion of the potential energy of a charged particle into kinetic energy after acceleration [118].

$$\frac{1}{2}mv^2 = qU = zeV$$

$$\frac{m}{z} = \frac{2eV}{d^2} \times t^2$$

d: effective distance e: electron charge m: mass q: particle charge t: flight time U: voltage V: acceleration voltage v: velocity z: number of charges

Ions get accelerated by an electric field with the result that ions with the same charge have the same kinetic energy. The velocity of the ions depends on the *m*/*z* ratio. Hence, heavier ions with the same charge reach the detector later than lighter ones [118].

In the particular setting of the X500R QTOF (ABSciex), molecules get ionized in the ion source at atmospheric pressure. Then, a series of quadrupoles (Q-Jet) focuses the ion flux before they enter the first quadrupole (Q1) which can be used as a mass filter for a specified mass range or for specified *m*/*z* for MS/MS experiments. After mass filtering in Q1, mass fragmentation can be performed in Q2—a linear acceleration collision cell. In order to achieve the same starting position as well as starting velocity of the ions, the ions are accelerated in the direction orthogonal to the ion beam. This instrumental set-up leads to an enhanced resolving power [109]. Two reflectors increase the ion flight path and aid to focus ions, which also results in enhanced resolving power. At the end of the flight tube a microchannel plate detector is used for detection of the arrival of the ions [112, 119].

The resolving power of the X500R QTOF is specified as ≥ 42,000 (full width half maximum (FWHM) at *m*/*z* 956) but its scan rate is higher (25 Hz) than that of the Q Exactive Focus. QTOF systems typically have a smaller dynamic range of up to 5 orders based on the detector setup.

1.3.2 Data processing and statistical analysis

Metabolomics approaches render a huge amount of data with up to several thousand scans per sample. Data processing methods need to be developed which include peak picking, integration and peak alignment [120]. This is a crucial step and can also lead to data loss. Hence, it is important to identify important parameters which significantly influence peak detection. In order to do this in a time- and effort efficient way, a statistical design of experiment (DoE) with a fractional factorial design can be used [121]. The mathematical background of DoE is briefly described in the following. The theory of experimental design con-

siders the outcome of an experiment, Y, as a random variable which depends on several factors $(x_1, x_2, \dots x_k)$ [121]. In general, the model function f is assumed to be polynomial in several variables. f is often called response surface:

$$Y = f(x_1, x_2, \dots, x_k) + \epsilon \qquad x_h\text{: factors } (h = 1, \dots, k) \tag{1}$$

On the basis of the idea that a response function can be developed as a Taylor series about the point of interest, one can approximate the response function in a constrained space by a quadratic polynomial. In the case of two factors, for example, one can write

$$Y = a_0 + a_1x_1 + a_2x_2 + a_3x_1^2 + a_4x_2^2 + a_5x_1x_2 + \epsilon \tag{2}$$

A more simple approximation to the response function at the point of interest is provided by a linear model function, which is the idea that underlies a screening factorial design. In this case, when a full factorial design is used, the number of experiments can be calculated as

$$n = 2^k \qquad k\text{: number of factors to be investigated} \tag{3}$$

Screening designs are an efficient way to identify significant main effects. In a fractional factorial design, interactions between the effects are omitted and only main effects are considered. Each factor of a screening design can take two levels that are coded by -1 or $+1$.

In order to identify the relevant factors in the data processing setting, a 2^{m-p} fractional factorial design (i.e., $n = 2^{m-p}$, where n is the number of experiments, m is the number of factors and $p \in \mathbb{N}$, with $n \geq m + 1$) can be useful, since the number of experiments can be reduced significantly [121].

If the influence of five factors shall be investigated by a 2^{m-p} fractional factorial design and one only wants to consider main effects and omits all interactions, this leads to a model with 6 coefficients. Since one wants to estimate the standard deviation, a number of experiments that exceeds the number of the coefficients is needed.

$$n = 2^{5-2} = 8 \text{ (with } n \geq 6) \; m \dots \text{ sum of factors, } p \in \mathbb{N}\text{, with } n \geq m + 1. \tag{4}$$

$$Y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + \epsilon \tag{5}$$

Table 2. Example of fractional factorial design of experiment plan.

Experiment	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
1	+1	+1	+1	+1	+1
2	+1	+1	-1	+1	-1
3	+1	-1	+1	-1	-1
4	+1	-1	-1	-1	+1
5	-1	+1	+1	-1	+1
6	-1	+1	-1	-1	-1
7	-1	-1	+1	+1	-1
8	-1	-1	-1	+1	+1

The DoE is drawn up in such a way that the mean of every column is 0 and each column is orthogonal to any other column. The estimates of the coefficients $\hat{a}_{0\ldots5}$ are stochastically independent from each other. Furthermore, the sample variance of all factors is identical.

In case of the optimization of a peak detection workflow, the percentage of peaks detected by the data processing workflow in relation to peaks detected using vendor specific software with additional manual integration can be defined as the response Y (identification ratio (ID ratio)).

Ordinary least-squares regression of equation (5) provides the estimates $\hat{a}_{0\ldots5}$ and $\hat{s}$. The significance of the estimates can be evaluated by a lower (in case of $\hat{a} > 0$) or an upper one-sided confidence limit (in case of $\hat{a} < 0$), respectively. For example, if a lower one-sided 0.9-confidence limit is > 0 (in case of $\hat{a} > 0$), one can conclude that the expectation of a is higher than 0 and the factor is of significance and cannot be eliminated [121]. Only significant factors need to be considered for optimization of peak detection.
Hence, with the help of this procedure the best settings for peak detection can be determined in a time- and effort efficient way.

Furthermore, chemometric tools are essential to reduce data complexity and to identify statistical significant differences. Ultimately, this can lead to the identification of biomarker candidates. Principal components analysis (PCA) is a widely used unsupervised exploratory technique for data reduction and pattern recognition [122, 123]. PCA consists of an orthogonal transformation of a set of correlated variables into a set of uncorrelated ones, the so-called principal components (PC). The principal components are constructed in such a way that they explain in a descending order of magnitude the variation of the data. First, the data are mean-centered, and then the covariance matrix $A_{mxm} = \frac{1}{n}X'X$ is calculated. The eigenvectors of A are the so called loadings. If the first principal components capture most of the

variance, the others can be neglected [124]. Thus, an efficient data reduction is achieved. Commonly, the first two or three PCs are used to span a two- or three-dimensional space referred to as scores plot. Samples can be represented by their scores which are calculated from the original data by multiplying them by the loadings of the PCs. Hence, one can imagine a PCA in the following way. First, a new axis is drawn through the center of the data points in such a way that the empirical variance of the new coordinates is maximized. Subsequently, a second axis is drawn through the center of the data points, which is orthogonal to the first axis, in such a way that the empirical variance of the coordinates referenced to the second axis is also maximized and so forth up to k number of axes ($k \leq \min(n, m)$) [121].

In order to perform classification on the basis of bi- or multivariate data from two (or more) classes (either original data or scores) a statistically defined decision criterion is necessary. One approach is to use so called prediction ellipsoids [124]. A prediction ellipsoid is an ellipsoid which encloses, with a certain error probability, a single future observation from a given class. Another approach is the use of discriminant analysis. This supervised discrimination technique can be performed as a linear discriminant analysis (LDA) or a quadratic discriminant analysis (QDA). LDA implies that the covariance matrices of both classes are equal. QDA assumes that the covariance matrices of both classes differ from each other. Both approaches provide a border which consists of data points at which the probability density of both classes is equal. In a bivariate case, LDA renders a straight line and QDA provides a quadratic model function [125].

Monte Carlo simulations can be conducted to assess the performance of discriminant analysis via cross-validation. The number of misclassifications (NMC) can be used as a diagnostic statistic. The predicted class membership of samples is compared to the true class membership and the samples are classified as True Positive (TP), True Negative (TN), False Negative (FN) or False Positive (FP) [126]. The NMC is defined as the sum of FP and FN. Further performance characteristics can be calculated as follows:

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

The sensitivity describes the ability of the model to correctly classify positive samples, while the specificity is a measure of how well the model performs to identify control samples [127].

In contrast to PCA, partial least squares-discriminant analysis (PLS-DA) is a supervised classification technique which can be used to classify samples in predefined categorical classes [128]. In a first step, partial least squares regression is performed to obtain a relationship between the observations and a predefined categorical vector by maximizing the square of the covariance matrix. In a second step, a class membership is assigned to the prediction by applying a threshold above which the sample is classified as one of the predefined classes [129]. Hence, one can imagine a PLS-DA in the following way. First, a new axis is drawn through the center of the data points in such a way that the square of the covariance between the new coordinates and the predefined categorical vector is maximized. Subsequently, a second axis is drawn through the center of the data points, which is orthogonal to the first axis, in such a way that the square of the covariance between the new coordinates and the predefined categorical vector is maximized and so forth up to k number of axes ($k \leq \min(n, m)$). Nowadays, there are different PLS-algorithms. Some algorithms implement the original approach described above, in which case the loadings are orthogonal to each other while other algorithms provide scores which are orthogonal to each other. In any case the classification outcome is the same.

Supervised methods are prone to over-fit data. Thus, Monte Carlo simulations have to be conducted to assess the PRESS statistics in order to check for over-fitting. The predicted residual sum of squares (PRESS) is divided by the total sum of predictions to calculate the mean square error of prediction (MSEP). MSEP is used to evaluate the error between the predicted categorical variable and the known category [128]. If the MSEP of the cross-validation via Monte Carlo simulations is clearly higher than the MSEP obtained from the residuals of the regression, one can conclude that the model over-fitted the data [121]. This part of the introduction provides some theoretical background information, detailed information on the data processing and statistical analysis is given in Section 3.1.4 and 3.2.3.3.

2 Aims of study

Non-targeted LC-HRMS screening and metabolomics workflows can be applied to explore changes in the metabolome of tissues or body fluids caused by various diseases or after the application of drugs. The use of metabolic fingerprinting in veterinary drug residue monitoring and environmental analysis has recently been described as an emerging field of science.

The overall objectives of this study were to develop non-targeted LC-HRMS metabolomics approaches and to assess their feasibility to reveal differences in biological matrices and to apply these approaches in order to carry out two separate scientific projects, one in the field of veterinary drug monitoring, the other in the field of environmental analysis.

For this purpose, a general workflow needed to be developed. Primary aims for this workflow development were:

- to develop or to improve sample preparation procedures and LC-HRMS protocols,
- to establish a workflow for LC-HRMS data processing,
- to implement an evaluation procedure subjecting non-targeted LC-HRMS data to chemometric analysis,
- to assess the feasibility of different chemometric methods to differentiate between positive samples and control samples and
- to identify unknown metabolites or biomarkers, which could be used as diagnostic markers.

Project 1: Non-targeted LC-HRMS analysis of samples taken within the frame of official veterinary drug residue control

In the first project of this thesis, a workflow should be developed for the investigation of LC-HRMS data obtained in the frame of official veterinary drug residue control. Up to now, official residue control is based on targeted multi-analyte methods in order to detect non-compliant residues of veterinary drugs. However, with the current methodology it is not possible to evaluate whether the presence of antibiotic residues points to a therapeutic application or was the result of a contamination or cross-over. Hence, the aim of this project was to apply metabolic fingerprinting to analyze samples from drug-treated, infected animals and samples from untreated, healthy animals and to highlight potential biomarkers for changes in the metabolism of drug-treated, infected animals. The identified biomarkers

should be easily included in routine veterinary drug monitoring with minimal additional analytical effort. A routine protocol for multi-residue analysis of antibiotics in muscle tissue should be chosen as starting point. The study aimed to provide analytical tools that can be applied for more legal certainty in order to decide whether a non-compliant result was caused by improper medication of infected animals or if further investigations are needed and sample contamination cannot be excluded. Furthermore, the applicability of the approach was to be investigated on two different HRMS systems (Q-Orbitrap and Q-TOF).

Project 2: LC-HRMS analysis of urine samples from ALT-treated gilts

In the second project of this thesis, the aim was to develop a non-targeted LC-HRMS method in order to investigate urine samples from ALT-treated gilts and from non-medicated gilts. ALT, a synthetic progestogen, is used in modern pig production for zootechnical purposes, i.e., for estrus synchronization in gilts. For certain geographical areas it cannot be excluded that the zootechnical use of ALT in gilts may present an environmental risk. Environmental residues of veterinary drugs with hormonal action are of rising concern, because even low level residues can pose a risk to the environment. There is only limited and inconsistent data on the urinary excretion of ALT and its metabolites. In the course of this study a pilot study should be conducted in order to collect urine samples from gilts which were treated with ALT as well as urine samples from non-medicated gilts which served as a control group. The aims were to determine the amount of ALT and its metabolites excreted in urine and to investigate the potential of metabolomics workflows (1) to distinguish between ALT-treated gilts and non-medicated gilts and (2) to identify new metabolites of ALT, which may be released in the environment after the application of manure from ALT-treated gilts as fertilizer to agricultural soils.

3 Material and methods

3.1 Non-targeted LC-HRMS analysis of samples taken within the frame of official veterinary drug residue control

Parts of the following passage have previously been published in [1].

3.1.1 Samples

***Porcine* samples**

Twenty-five non-compliant *porcine* samples from infected, drug-treated animals were collected within the framework of the European residue control plan [8] and used as positive samples. These samples exhibited residues of veterinary drugs which exceeded the European MRL in muscle tissue, kidney tissue or both (see Supplementary data Table S2 for details on analyte spectrum). Twenty-five compliant *porcine* samples from untreated, healthy animals served as controls. All samples were stored at -20°C prior to sample preparation. For reasons of quality control (QC), spiked QC samples as well as pooled samples of all positive (positive QC) and control samples (control QC) were prepared. Spiked QC samples contained 70 veterinary drugs (including analytes for marker residues) which were added prior to sample preparation at MRL level (see Supplementary data Table S4).

In order to establish a bivariate biomarker model for *porcine* muscle samples, 50 positive and 50 control *porcine* muscle samples were used. These samples were analyzed in several analytical batches over the course of one year and were retrospectively evaluated for the two biomarker candidates. Detailed sample information is given in Table 3.

Table 3. Sample information (gender, age, storage time) on the data set used for PCA and the data set used for the bivariate model.

	Data set for PCA		Data set for bivariate model	
	Positive	**Control**	**Positive**	**Control**
Gender				
Male, castrated		14		26
Female	4	11	9	20
n.a.	21		41	4
Age				
Piglet (3 months)	10		21	
Fattened pig (4-6 months)	15	25	29	50
Storage time				
< 12 months	8	19	19	45
12–24 months	17	6	31	5

n.a.: not available.

***Bovine* samples**

Fifteen non-compliant *bovine* samples from infected, drug-treated animals were collected within the framework of the European residue control plan [8] and used as positive samples. These samples exhibited residues of veterinary drugs which exceeded the European MRL in muscle tissue, kidney tissue or both (see Supplementary data Table S3 for details on analyte spectrum). Fifteen compliant *bovine* samples from untreated, healthy animals served as controls. All samples were stored at -20°C prior to sample preparation. For reasons of quality control (QC), spiked QC samples as well as pooled samples of all positive (positive QC) and control samples (control QC) were prepared. Spiked QC samples contained 70 veterinary drugs (including analytes for marker residues) which were added prior to sample preparation at MRL level (see Supplementary data Table S4).

Table 4. Sample information (gender, age, storage time) on the *bovine* data set used for PCA.

	***Bovine* samples**	
	Positive	**Control**
Gender		
Male	4	5
Female	11	10
Age		
1–2 years	2	2
3–5 years	5	4
6–10 years	4	5
n.a.	4	4
Storage time		
< 12 months	9	12
12–24 months	6	3

n.a.: not available.

3.1.2 Sample preparation

An in-house validated protocol for multi-residue analysis of veterinary drugs was used for sample preparation [130]. In brief, 1 g homogenized muscle tissue was weighed into a centrifuge tube, fortified with 100 µL of a mixture of internal standards (amoxicillin-d4, benzylpenicillin-d7, demeclocycline, enrofloxacin-d5, oxolinic acid-d5, roxithromycin, sarafloxacin-d8, sulfadimethoxine-d6 and sulfadoxine-d3 each at a concentration of 1000 ng/mL) and left at room temperature for 15 min. Ten mL of EDTA-McIlvaine buffer [130] was added. The samples were vortexed, shaken for 10 min and then sonicated for 5 min. After centrifugation for 10 min at 2,500 × g, the supernatants were collected and filtered through folded filters (> 0.4 µm, Macherey-Nagel, Düren, Germany). The extraction was repeated with an addi-

tional 5 mL buffer and the supernatants were combined. After rinsing the filters with 2 mL buffer, the entire solution was applied to a 6 mL/200 mg OASIS HLB® solid phase extraction (SPE) cartridge (Waters, Eschborn, Germany), pre-activated with 6 mL of methanol and 6 mL of H_2O. After a washing step with 6 mL of 5% methanol and 10 min drying *in vacuo*, the samples were eluted with 6 mL of methanol. The extracts were evaporated to near dryness under nitrogen (40°C) and reconstituted in 1 mL of H_2O/acetonitrile (90/10 (v/v)). After centrifugation at 21,500 × g for 5 min, the supernatants were transferred into HPLC vials for LC-HRMS analysis.

3.1.3 LC-HRMS methods (Q-Orbitrap and Q-TOF)

The LC-MS method was adopted from Bohm *et al.* and slightly modified [130]. In brief, the chromatographic separation was carried out using a Dionex Ultimate 3000 RS UHPLC (Thermo Scientific, Idstein, Germany) equipped with a Phenomenex Luna Omega C18 polar column (1.6 µm, 100 Å, 100 mm x 2.1 mm; Phenomenex, Aschaffenburg, Germany). Twenty µL of extract was injected for each analytical run. The mobile phase consisted of H_2O with 0.1% formic acid (mobile phase A) and acetonitrile with 0.1% formic acid (mobile phase B). The chromatographic method started with an initial phase of 10% B for 1 min, followed by a linear gradient to 60% B over 12 min and keeping these conditions for 2 min. A washing step increased mobile phase B to 95% within 1 min and conditions were held constant for an additional 5 min. B was decreased to 10% thereafter and held for 5 min to equilibrate the column to starting conditions. The flow rate was 0.4 mL/min. The column operated at 30°C. The temperature of the autosampler was set to 15°C. The UHPLC-system was coupled to a Q-Exactive Focus Orbitrap mass spectrometer (Thermo Scientific, Idfstein, Germany) equipped with an heated electrospray (HESI-II) source operating using the following conditions: capillary temperature 300°C, sheath gas flow rate 30 a.u., aux gas flow rate 10 a.u., spray voltage 3.5 kV, S-lens RF level 60 and aux gas heater temperature 400°C. Data was acquired in positive ion mode at a resolving power of 70,000 (FWHM at *m*/*z* of 200) over a mass range of *m*/*z* 75–1050. The AGC target was set to 1.0×10^6, with a maximum injection time of 220 ms. MS/MS-spectra were triggered in fullscan data-dependent MS^2 discovery mode by a data-dependent fragmentation of the most intense ion per scan at 17,500 FWHM covering a mass range from *m*/*z* 50 to the precursor mass. The isolation window for the precursor was set to 1 *m*/*z* and fragmentation was performed with stepped collision energy of 20, 40 and 60 eV. The AGC target was set at 2.0×10^5 with an injection time of 50 ms. To avoid triggering MS/MS

data from background ions an exclusion list with the top 50 background ions which were observed in reagent blanks was used. Data acquisition and qualitative analysis were carried out using Trace finder 4.0 (Thermo Scientific, Idstein Germany).

Furthermore, the samples were analyzed using a LC-Q-TOF system. The chromatographic separation was carried out using an Agilent 1290 Infinity UHPLC system (Agilent Technologies, Waldbronn, Germany) with the chromatographic conditions as mentioned above. The UHPLC-system was coupled to a X500R Q-TOF (Sciex, Darmstadt, Germany) mass spectrometer equipped with a Turbo V™ source operating in positive electrospray ionization (ESI) mode using the following conditions: Atomizing gas 50 psi, auxiliary gas 50 psi, curtain gas 35 psi, collision gas 7 a.u., source temperature 550°C and spray voltage 5.5 kV. Data was acquired using the information dependent acquisition (IDA) mode. For the IDA experiment, TOF MS mass range spanned from *m*/*z* 75–1055 with an accumulation time of 0.1 s, a declustering potential (DP) of 80 V with DP spread of 0 V and a collision energy (CE) of 10 V with CE spread of 0 V were applied. MS/MS-spectra were triggered for the ten most intense ions per scan (intensity threshold of 10 cps) with a collision energy of 35 V (CE spread of 15 V), declustering potential of 80 V (DP spread of 0 V) in an accumulation time of 0.05 s. Candidate ions were excluded for 5 seconds after 3 occurrences. For TOF-MS/MS the start mass was set to *m*/*z* 75 and the TOF stop mass was set to *m*/*z* 1055. Data acquisition and qualitative analysis were carried out using SciexOS 1.3 (Sciex, Darmstadt, Germany).

3.1.4 Data processing and statistical analysis

Raw files were converted to mzXML format [131] and imported into MZmine 2 [120]. In order to optimize the MZmine 2 workflow a statistical DoE with a fractional factorial design was applied.

3.1.4.1 DoE with fractional factorial design

Initial experiments indicated that different settings for noise level, peak time span, minimum peak height, smoothing and amplitude of noise used for chromatogram deconvolution had relevant effects on peak detection performance in a MZmine 2 workflow. Hence, a screening design with these five factors was used in order to identify significant main effects (see Figure 6).

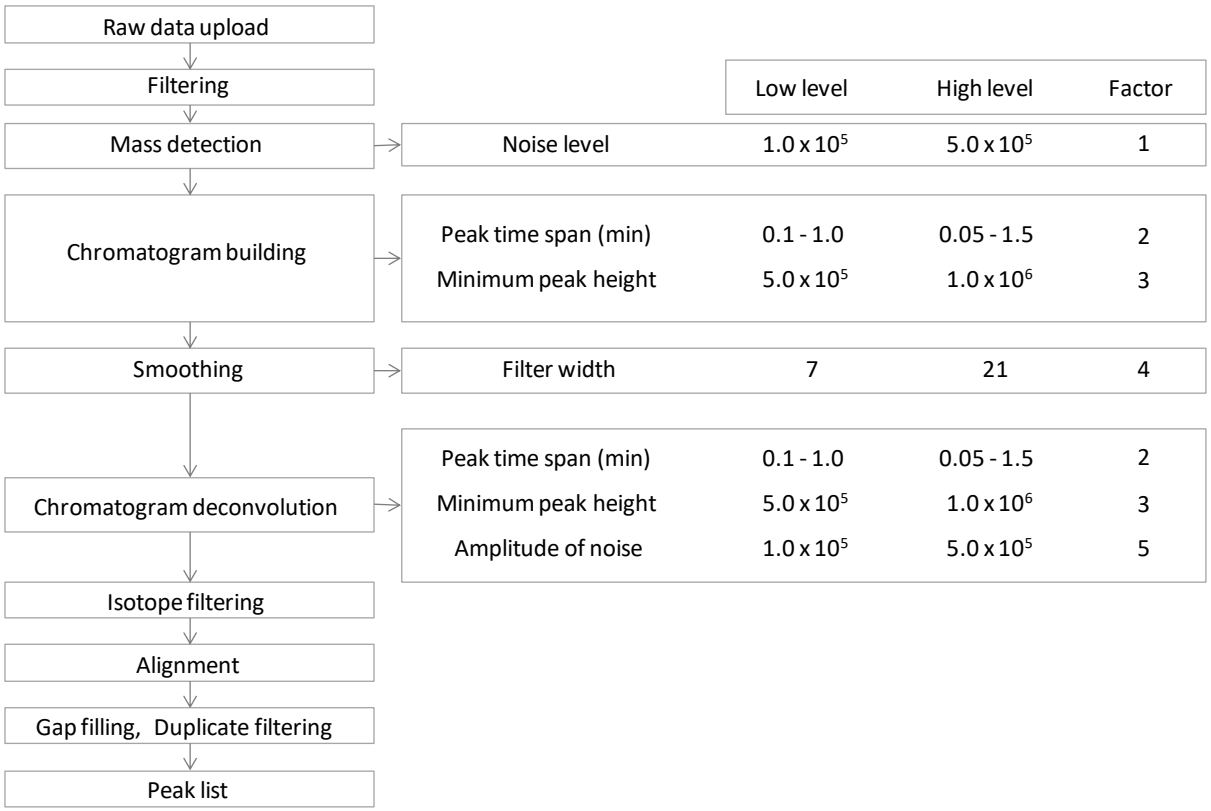


Figure 6. Workflow of MZmine 2 for LC-Q-Orbitrap data processing with low and high level values of factors for fractional factorial design of experiment.

The 2^{m-p} fractional factorial DoE consisted of 8 experiments ($n = 2^{5-2}$). For each factor a high and a low level was set (see Figure 6) and the DoE was performed with a test data set according to Table 5.

Table 5. Fractional factorial design of experiment plan.

Experiment	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
1	+1	+1	+1	+1	+1
2	+1	+1	-1	+1	-1
3	+1	-1	+1	-1	-1
4	+1	-1	-1	-1	+1
5	-1	+1	+1	-1	+1
6	-1	+1	-1	-1	-1
7	-1	-1	+1	+1	-1
8	-1	-1	-1	+1	+1

The dataset used consisted of three sub datasets of *porcine* muscle samples each acquired at different time points on the LC-Q-Orbitrap system over the time span of two months. One sub dataset consisted of one blank and four spiked samples (n = 15). The spiked samples were fortified with 70 different antibiotics prior to sample preparation at 10%, 50%, 100% and 150% of MRL level. The percentage of peaks detected by the MZmine 2 workflow in re-

lation to peaks detected using vendor specific software (Tracefinder) with additional manual integration was defined as the response *Y* (identification ratio (ID ratio)).

Validation of MZmine 2 peak detection

In order to evaluate the performance of the data processing method 15 non-compliant *porcine* muscle samples from drug-treated, infected animals were analyzed as described above (see Section 3.1.2 and 3.1.3) with the LC-Q-Orbitrap method and processed with the chosen MZmine 2 procedure. These 15 muscle samples contained 16 antibiotics at or above MRL-level.

3.1.4.2 Data processing chosen for study

The final data processing workflow was determined as per the fractional factorial design experiments described above (see Section 3.1.4.1 and 4.1.1.1). The following data processing workflow was then used for all further experiments.

***Porcine* muscle samples**

Raw data from Q-Orbitrap analysis was filtered (0.5–12.5 min) and masses were detected with centroid algorithm at a noise level of 1.0×10^5. Chromatogram building was carried out with a minimum time span of 0.05 min, a minimum height of 5.0×10^5 and a mass accuracy of 5 ppm or 0.002 Da. Chromatograms were smoothed (filter width 15). For chromatogram deconvolution, the noise amplitude algorithm with a minimum peak height of 5.0×10^5, peak duration of 0.05–1.5 min and amplitude of noise of 1.0×10^5 was applied. After deisotoping with an *m*/*z* tolerance of 5 ppm or 0.002 Da, a retention time tolerance of 0.01 min, a maximum charge of 3, peaks were aligned using the join aligner algorithm. The *m*/*z* tolerance was set to 10 ppm or 0.005 Da weighted with 10%, the retention time tolerance was 0.2 min weighted with 10%, only using ions with the same charge. Gap filling was performed with 20% intensity tolerance at a mass accuracy of 10 ppm or 0.005 Da and 0.05 min retention time tolerance after retention time correction. The final steps were duplicate filtering with a mass accuracy of 10 ppm or 0.005 Da and 0.2 min absolute retention time tolerance as well as row filtering eliminating rows with less than 5 peaks. Data post-processing consisted of eliminating all mass spectrometric features (*m*/*z* at retention time) which were present in the reagent blank sample at an intensity of $> 1.0 \times 10^7$.

Raw data from Q-TOF analysis was treated similarly, except the noise level was adjusted to 1.0×10^3 detecting peaks with a minimum height of 5.0×10^3 and *m*/*z* tolerance for peak

alignment was set to 0.02 Da or 50 ppm. Data post-processing consisted of eliminating all mass spectrometric features (*m/z* at retention time) which were present in the reagent blank sample at an intensity of > 1.0 x 10^6.

After data processing, a data integrity check was performed. Over the time of two analytical batches, 8 spiked QC samples were measured and the identification ratio (ID ratio) of the spiked analytes was calculated as well as the relative standard deviation (RSD). The ID ratio was defined as the percentage of peaks detected by the MZmine 2 workflow in relation to peaks detected using vendor specific software with additional manual integration. Furthermore, the data was evaluated for batch effects, which would have indicated the need for QC based signal correction [132].

Further investigations

Raw data of *bovine* muscle samples was processed in exact the same way as *porcine* muscle samples (see Supplementary data Table S5) except that a data integrity check was performed on 4 spiked QC samples in the course of one analytical batch.

Porcine and *bovine* kidney samples were processed in a similar way with minor changes (see Supplementary data Table S6). Data integrity check was performed with 8 and 4 spiked QC samples for *porcine* and *bovine* kidney samples, respectively. In general, the analysis of *porcine* samples was performed in two analytical batches due to the number of samples, while the analysis of *bovine* samples was completed in one analytical batch.

3.1.4.3 Statistical analysis

***Porcine* muscle samples**

The data was subjected to statistical analysis using an evaluation procedure that was implemented in MATLAB R2019b (The MathWorks Inc., Natick, Massachusetts). In a first step, all zero values were replaced by the half of the minimum detected peak area. The data set was normalized by the sum of all peaks, since no batch effect was observed, that would have indicated the need for more complex normalization strategies. Subsequently, the data was transformed by taking the natural logarithm. After transposing and standardizing (mean-centering and division by standard deviation) multivariate statistical analyses were performed, i.e., PCA and PLS-DA.

Subsequently to PCA, the scores of the 3D-PCA model were submitted to a QDA. Monte Carlo cross-validation was used in order to assess the ability of the PCA coupled with QDA

to discriminate between the two classes (control vs. positive samples). A test set of 10 samples was randomly drawn from the data set and PCA with subsequent QDA was performed with the remaining training set. This procedure was repeated 5,000 times. The number of misclassifications can be used as a diagnostic statistic. The predicted class membership of samples is compared to the true class membership and the samples are classified as true positive, true negative, false negative or false positive. Sensitivity of the PCA with subsequent QDA was calculated as the ratio of the number of true positive samples to the sum of true positive and false negative samples, specificity was obtained by dividing the number of true negative samples by the sum of true negative and false positive samples. The number of misclassifications was calculated as the sum of false positive and false negative samples.

Furthermore, Monte Carlo simulations were conducted to assess the quality of the supervised PLS-DA model by calculating the mean square error of prediction (MSEP). The MSEP based on the residuals of the PLS-DA regression and the MSEP derived from the cross-validation procedure were compared and used as diagnostic parameters for over-fitting [133]. If the MSEP derived from the cross-validation is higher than MSEP from the residuals of the PLS-DA regression, the model over-fitted the data. For exemplary MATLAB codes see Supplementary data Section 10.

Further investigations

The statistical analysis of *bovine* muscle samples was performed in the exact same way except the test set for Monte Carlo simulations consisted of 5 samples.

In the case of the kidney samples, further normalization procedures were necessary, since batch effects and analytical drifts were observed. In the case of *porcine* kidney samples, a pronounced batch effect between the two analytical batches was noticed. In order to compensate for batch-effects, normalization with QC samples was performed. The mean values of all variables of the QC samples from day two of measurement were divided by the mean values of all variables of the QC samples from day one of measurement. Subsequently, all variables (*m*/*z* at retention time) from day two of measurement were divided by this coefficient. In the case of the data set from *bovine* kidney samples measured at the Q-TOF system, an analytical drift as intra-batch variation was observed. Thus, normalization by applying a correction curve fitted to the QC samples in respect to the order of injection was performed. A quadratic least-squares regression curve was fitted to the mean peak area of each variable of the QC samples as dependent variable Y with the run number as independent variable x. Then, a correction curve for the whole analytical batch was interpolated for each variable

(*m/z* at retention time) and applied to the data set. For exemplary MATLAB codes see Supplementary data Section 10.

3.1.5 Study design

In the process of study design, it was necessary to determine if different factors have an influence on the multivariate statistical analysis. The factors investigated were species, gender and storage time.

In order to evaluate the possibility to build a species-independent model, 10 blank *porcine* and *bovine* muscle samples were analyzed as described in Section 3.1.2 and 3.1.3. Furthermore, mixtures of *porcine* and *bovine* muscle samples with a ratio of *porcine:bovine* muscle tissue of 20:80, 50:50 and 80:20 (w/w) were analyzed. The data was processed as described in Section 3.1.4.2. For multivariate statistical analysis, all zero values were replaced by the half of the minimum detected peak area. The data set was normalized by the sum of all peaks. Subsequently, the data was transformed by taking the natural logarithm. After transposing and standardizing (mean-centering and division by standard deviation) multivariate statistical analyses were performed, i.e., PCA and PLS-DA.

Since multivariate statistical analysis showed substantial differences between the two species, the data was additionally submitted to MetaboAnalyst 4.0 for further statistical analysis [134]. For structural elucidation of important features discriminating *bovine* from *porcine* muscle samples the open access software "seven golden rules" [135] was used. Input parameters were the accurate mass of the neutralized analyte with a mass accuracy of 20 ppm, the isotopic pattern with an isotopic abundance error of 10% and an elemental restriction to carbon, hydrogen, nitrogen, sulfur and oxygen.

The molecular formulas obtained from the "seven golden rules" were searched against available chemical databases, such as ChemSpider [136], PubChem [137], Metlin [138] and the Human Metabolome Database (HMBD) [139]. Product ion spectra were searched in MasCot against MS/MS files in NCBIprot and SWISSprot [140-142]. Parameters were set as follows:

1) Tryptic digestion with the maximum number of missed cleavages was chosen, since trypsin is the most common enzyme used, thus, resulting in many database entries. However, it must be kept in mind that enzymatic digestion was not used in the study presented in this thesis.

2.) Peptide and MS/MS tolerance was set to 20 ppm.

In a final step, the proposed peptide sequence in FASTA format was searched against nonredundant protein sequences using the protein Basic Local Alignment Search Tool (BLASTp) of the NCBI database [141].

The influence of the variables gender and storage time was examined in 25 control *porcine* samples. The samples were analyzed and processed as described in Section 3.1.2, 3.1.3 and 3.1.4.2 and PCA was performed in order to check for variable dependent clustering. In control *bovine* samples the influence of the variable age was tested additionally, since for this species data about the age of the animal was provided in the course of sampling.

3.1.6 Selection of biomarker candidates and pathway mapping in *porcine* muscle samples

For biomarker selection in *porcine* muscle samples, the data was submitted to MetaboAnalyst 4.0 [8, 134]. Data pretreatment consisted of (1) replacing zero values by the half of the minimum detected peak, (2) normalization by the sum of all peaks, (3) log-transformation and (4) standardizing (mean-centering and division by standard deviation). Univariate tests (*Student's t*-test and a fold change test) were applied for biomarker selection. In order to account for multiplicity, all *p*-values were adjusted using the procedure of Benjamini and Hochberg [143]. Biomarker identification was attempted for the top 25 analytes based on data obtained from the Q-Orbitrap method. In order to classify the confidence in candidate identification, the levels of confidence for unknown annotation described by Schymanski *et al.* were applied [94]. The authors distinguish 5 levels of confidence.

- Level 5 is chosen for the identification of a unique variable with its accurate mass at a specific retention time.
- Level 4 is determined by the postulation of a molecular formula via isotope abundance distribution, charge state and adduct determination.
- Level 3 and 2 include structural identification via library spectrum matching or literature hits as well as via experimental information of diagnostic evidence. Level 3 corresponds to a tentative structure with more than one possible structure, e.g., positional isomers and level 2 refers to a probable structure with one exact proposed structure based on diagnostic fragments or an unambiguous library match.
- In level 1, the biomarker identity is validated by a reference standard [94].

For biomarker annotation, several vendor specific as well as open access software packages were used in addition to online databases, i.e., Compound Discoverer 3.1 (Thermo Scientific, Idstein, Germany), "seven golden rules" software [135], Metlin [138] and the HMDB [139].

Furthermore, the data was submitted to pathway mapping by "MS Peaks to Pathway (Version 2)" in MetaboAnalyst 4.0 [134]. This functional module is based on the *mummichog* algorithm, which is able to predict the functional activity from mass spectrometric data without *a priori* identification, since metabolite identification can be derived from local enrichments in metabolic pathways [144]. As input parameters, a peak list (m/z at retention time) with the corresponding p-values was provided. The mass accuracy was set to 5 ppm and the top 10% significant mass spectrometric features (p-value cut-off: 1.0×10^{-5}) were submitted to the *mummichog* algorithm. The manually curated pathway library for *Homo sapiens* based on KEGG, BiGG and Edinburgh Model was selected as knowledge-base for pathway mapping [144].

ROC curves were plotted for two biomarker candidates in MATLAB. ROC curves visualize the specificity and sensitivity of the biomarker [145]. In this graph, the sensitivity is plotted against 1-specificity. ROC curve analysis can be used to illustrate the compromise between sensitivity and specificity regarding the biomarker performance at different discrimination thresholds [134]. The area under the ROC (AUROC) is often used as a diagnostic tool for biomarker selection [146].

For final biomarker evaluation, bivariate data analysis and construction of prediction ellipses were performed using Excel 2010 (Microsoft Office, Redmond, U.S.). According to [124], given n pairs $(x_i, y_i)_{i=1,\dots,n}$ randomly taken from a bivariate normal population, a $(1-\alpha)$ −prediction ellipse for a single further observation is formed by all points (ξ, ψ) that satisfy

$$\psi = \frac{\left(s_x^2\bar{y} - s_{xy}\bar{x} + s_{xy}\xi\right)}{s_x^2} \pm \frac{\sqrt{\left(s_{xy}^2 - s_x^2 s_y^2\right)\xi^2 - 2\left(s_{xy}^2 - s_x^2 s_y^2\right)\bar{x}\xi + \left(s_x^2\bar{y} - s_{xy}\bar{x}\right)^2 - s_x^2 s_y^2\bar{x}^2 - s_x^4\bar{y}^2 + 2s_x^2 s_{xy}\bar{x}\bar{y} - s_x^2\left(s_{xy}^2 - s_x^2 s_y^2\right)\left(1+\frac{1}{n}\right)T_{1-\alpha}^2}}{s_x^2}$$

where

$$s_x^2 = \frac{1}{n-1}\sum_i (x_i - \bar{x})^2 \quad , \quad s_y^2 = \frac{1}{n-1}\sum_i (y_i - \bar{y})^2 \quad , \quad s_{xy} = \frac{1}{n-1}\sum_i (x_i - \bar{x})(y_i - \bar{y})$$

and $T_{1-\alpha}^2$ is the quantile of Hotelling's distribution with 2 and $n-2$ degrees of freedom. Note that in the bivariate case $T_{1-\alpha}^2 = F_{1-\alpha,2,n-2}\frac{2(n-1)}{n-2}$ where $F_{1-\alpha,2,n-2}$ is the quantile of the F-distribution with 2 and $n-2$ degrees of freedom. To graph an ellipse with Excel, the lower and upper limit of the ellipse's domain are needed. These values can be obtained by setting the term under the root equal to zero and solving the equation for ξ:

$$\xi_{lower} = \bar{x} - \sqrt{\bar{x}^2 - \frac{\left(s_x^2\bar{y} - s_{xy}\bar{x}\right)^2 - s_x^2 s_y^2 \bar{x}^2 - s_x^4 \bar{y}^2 + 2s_x^2 s_{xy}\bar{x}\bar{y}}{\left(s_{xy}^2 - s_x^2 s_y^2\right)} + s_x^2\left(1+\frac{1}{n}\right)T_{1-\alpha}^2}$$

and

$$\xi_{upper} = \bar{x} + \sqrt{\bar{x}^2 - \frac{\left(s_x^2\bar{y} - s_{xy}\bar{x}\right)^2 - s_x^2 s_y^2 \bar{x}^2 - s_x^4 \bar{y}^2 + 2s_x^2 s_{xy}\bar{x}\bar{y}}{\left(s_{xy}^2 - s_x^2 s_y^2\right)} + s_x^2\left(1+\frac{1}{n}\right)T_{1-\alpha}^2}$$

3.2 LC-HRMS analysis of urine samples from ALT-treated gilts

Parts of the following passage have previously been published in [192].

3.2.1 Samples

In order to investigate the urinary excretion of ALT and its metabolites, a pilot study with conventionally housed gilts treated with ALT as part of routine breeding procedures was conducted at the "Landesanstalt für Schweinezucht" in Boxberg (Baden-Wuerttemberg, Germany) in December 2019. Spontaneous urine samples collected from 5 non-medicated gilts (Baden-Wuerttemberg hybrid, age 6 month) and 3 ALT-treated gilts (Baden-Wuerttemberg hybrid, age 6 month) were studied. Both groups were given the same feed stuff consisting of wheat (418 g/kg), barley (398 g/kg), soy extract grist (80 g/kg), fibre mix (40 g/kg), mineral feed (30 g/kg), rapeseed oil (20 g/kg) and feed acidifiers (14 g/kg). In the treatment group, 20 mg ALT were administered daily at 08:00 a.m. over the course of eighteen days. The samples from medicated gilts were taken at the end of the eighteen-day treatment period, i.e., in the steady-state phase. The sampling time of the urine samples is given in Table 6.

Table 6. Sampling times of spontaneous urine samples from ALT-treated gilts.

Gilt	Day 17 of treatment	Day 18 of treatment
A	3:30 p.m.	11:15 a.m.
B	4:00 p.m.	11:15 a.m.
C	3:00 p.m.	07:00 a.m., 10:15 a.m.

For non-targeted analysis, the following QC samples were prepared additionally: QC control by pooling all samples from non-medicated gilts, QC positive by pooling all samples from medicated gilts and QC total by pooling all samples. All samples were stored at -20°C prior to samples preparation.

3.2.2 Targeted LC-HRMS analysis

3.2.2.1 Sample preparation

Development of sample preparation protocol

Three different SPE cartridges were tested during method development: OASIS HLB® SPE cartridges (200 mg, 6 cc) acquired from waters (Eschborn, Germany), Bond Elut C18 SPE car-

tridges (500 mg, 6 cc) purchased from Agilent Technologies (Santa Clara, USA) and STRATA XL SPE cartridges (100 mg, 3 cc) purchased from Phenomenex (Aschaffenburg Germany).

The sample preparation was adopted from Domínguez-Romero *et al.* and was slightly modified [147]. Prior to sample preparation ALT-free urine samples were centrifuged (10 min, 2,500 x g) and 2 mL of the supernatants were spiked with ALT at the concentration levels 5 and 500 ng/mL, respectively. The samples were incubated at room temperature for 10 min. Two mL of ammonium formate buffer (50 mM, pH 2.6) was added and the extracts were used for SPE. OASIS HLB® and Bond Elut C18 SPE cartridges were conditioned using 6 mL of methanol and 6 mL of H_2O. After the conditioning step the extracts were loaded to the cartridges followed by a washing step with 6 mL of 5% methanol and 10 min drying *in vacuo* to remove excess water. Elution of the analyte was performed with 6 mL of methanol. Using the STRATA XL SPE cartridges the analytical procedure was the same except using 3 mL for the conditioning and washing step, respectively. The extracts were evaporated to near dryness at a temperature of 50°C under nitrogen gas flow. The samples were reconstituted in 0.5 mL of methanol/water (60:40 (v/v)) resulting in a pre-concentration factor of 4:1. Prior to LC-HRMS analysis, the samples were filtered (0.45 µm filter) and transferred to a HPLC vial. In order to assess losses during the extraction procedure, blank urine extracts were spiked after the extraction procedure with 20 and 2000 ng/mL, respectively. Recovery was determined via external calibration. Furthermore, samples of each loading and washing step were collected and analyzed to determine possible losses during the extraction procedure of ALT spiked urine samples.

In order to investigate if an additional defattening step would reduce the effect of matrix components on the recovery, the reconstituted samples (0.5 mL) were extracted with hexane (0.5 mL) (n = 10). After the extracts were shaken for 10 min and centrifuged for 5 min at 21,500 x g, the upper hexane layer was removed. The extracts were transferred to a HPLC vial prior to analysis.

Furthermore, stabile isotope dilution analysis was performed to compensate for matrix effects. Prior to sample preparation, 2 mL of the supernatant of ALT-free urine was spiked with a fixed concentration of ALT-d5 and the sample treatment was performed as described above. For quantification the response ratios of ALT and ALT-d5 were plotted against the corresponding concentration levels.

Final sample preparation method

After method development the OASIS HLB® (200 mg, 6 cc) SPE cartridges were chosen for

sample preparation. Furthermore, a hydrolysis step was incorporated into the protocol in order to quantify ALT after β-glucuronidase treatment. Hence, the method validation and quantification of ALT in urine samples was performed twice, i.e., with and without hydrolysis step.

a) Sample preparation without hydrolysis step

Prior to sample treatment, urine samples were centrifuged (10 min, 2,500 x g) and 2 mL of the supernatants were spiked with internal standard ALT-d5 at a fixed concentration depending on the operating range (see Section 3.2.2.3 and 3.2.2.4). The samples were incubated at room temperature for 10 min. After the addition of 5 mL of ammonium formate buffer (50 mM, pH 2.6), the samples were used for SPE. OASIS HLB® (200 mg, 6 cc) SPE cartridges were conditioned using 6 mL of methanol and 4 mL of H_2O. After the conditioning step, the samples were loaded to the cartridges followed by a washing step with 6 mL of 5% methanol and 10 min drying *in vacuo* to remove excess water. Elution of the analyte was performed with 6 mL of methanol. The extracts were evaporated to near dryness at a temperature of 50°C under nitrogen gas flow. The samples were reconstituted in 0.5 mL of methanol/water (60:40 (v/v)). Prior to LC-HRMS analysis, the samples were filtered (0.45 μm filter) and transferred to a HPLC vial.

b) Sample preparation with hydrolysis step

2 mL of urine were spiked with internal standard ALT-d5 at a fixed concentration depending on the operating range (see Section 3.2.2.3 and 3.2.2.4). The samples were incubated at room temperature for 10 min. Subsequently, 5 mL of ammonium acetate (50 mM, pH 4.8) was added and pH was adjusted to pH 5.0 with aqueous KOH solution (6 M). The samples were incubated with 100 μL β-glucuronidase (*helix pomatia* Type H-2, ≥85,000 units/mL) for 24 hours. After centrifugation (10 min, 2,500 x g), the addition of 5 mL of ammonium formate buffer (50 mM, pH 2.6) and pH adjustment to 2.6 with formic acid, the samples were used for SPE as described above. Quantification was performed via matrix calibration.

3.2.2.2 LC-HRMS method

The chromatographic separation was carried out using a Dionex Ultimate 3000 RS UHPLC (Thermo Scientific, Idstein Germany) equipped with a Phenomenex Prodigy C18 column (5 μ, ODS (3), 100 Å, 150 mm x 3 mm; Phenomenex, Aschaffenburg Germany). 20 μl of extract was injected for each analytical run. The mobile phase consisted of water with 0.1%

formic acid (mobile phase A) and methanol with 0.1% formic acid (mobile phase B). The chromatographic method started with an initial phase of 70% B for 3 min, followed by a linear gradient to 100% B in 2 min and keeping these conditions for 5 min. In 1 min the proportion of B was decreased to 70% and held for 10 min to equilibrate the column to starting conditions. The flow rate was set to 0.3 mL/min. The column temperature was held at 30°C whereas the temperature of the autosampler was set to 15°C.

The UHPLC-system was coupled to a Q-Exactive Focus Orbitrap mass spectrometer (Thermo Scientific, Idstein Germany) equipped with an heated electrospray (H-ESI) source using the following conditions: Capillary temperature 300°C, sheath gas flow rate 30 a.u., aux gas flow rate 9 a.u., spray voltage 3.5 kV, S-lens RF level 60 and aux gas heater temperature 350°C. Data was acquired in positive and negative ion mode at a resolving power of 70,000 (FWHM at *m*/*z* of 200) over a mass range of *m*/*z* 75–1055. The AGC target was set to 1.0 x 10^6, with a maximum injection time of 220 ms. MS/MS-spectra were triggered in fullscan data-dependent MS^2 confirmation mode at 17,500 FWHM covering a mass range from *m*/*z* 50 to the precursor mass. The isolation window for the precursor was set to 1 *m*/*z* and fragmentation was performed with stepped collision energy of 20, 40 and 60 eV. The AGC target was set at 2.0 x 10^5 with an injection time of 50 ms. To avoid triggering MS/MS data from background ions an exclusion list with the top 50 background ions which were observed in reagent blanks was used. Mass tolerance was set to 5 ppm. Data acquisition and quantitative analysis was carried out using Trace finder 4.0 (Thermo Scientific, Idstein Germany).

Table 7. Parameters (retention time, ionization mode, precursor ion and fragment ions) of the LC-HRMS analysis.

	Retention time (min)	Ionization mode	Precursor ion (*m*/*z*) used as Quantifier	Characteristic fragment ions *m*/*z*
Altrenogest	7.67	Positive	311.2006	227.1425, 251.1427, 269.1533, 293.1905, 199.1118, 159.0802
Altrenogest-CAP	7.31	Positive	311.2006	227.1425, 251.1427, 269.1533, 293.1905, 199.1118, 159.0802
Altrenogest-d5	7.65	Positive	316.2319	227.1425, 251.1427, 269.1533, 298.2207, 199.1118, 159.0802
Altrenogest glucuronides	6.20	Positive	487.2327	311.2006, 293.1905, 269.1533, 227.1425, 159.0802
Altrenogest-CAP glucuronides	5.30	Positive	487.2327	311.2006, 293.1905, 269.1533, 227.1425, 159.0802
Altrenogest sulfate	3.83	Negative	389.1433	309.1855, 197.0966, 211.1123, 213.1279, 225.1280, 253.1230

CAP: cycloaddition product

3.2.2.3 Method validation

Matrix calibration curves were analyzed following the respective urinary extraction methods as described in Section 3.2.2.1 and 3.2.2.2. Two concentration ranges were covered to assess the operating range of the method. The spiking levels ranged from 5–3000 ng/mL ALT and 2.5–50 ng/mL ALT, respectively, using ALT-d5 as internal standard at a concentration of 50 ng/mL. Matrix calibration curves were analyzed on consecutive working days using four different *porcine* urine samples ($n = 4$).

Performance characteristics

Analytical methods used in testing of official samples for veterinary drug residues have to be validated according to the EU Commission Decision 2002/657/EC [148]. This Decision makes demands on the analytical proficiency and provides rules and recommendations for the determination of performance characteristics.

In this work, performance characteristics as defined in Commission Decision 2002/657/EC as well as performance characteristics as defined in DIN 32645 [149] were established. For stability testing, also the "Guidance for Industry on Bioanalytical Method Validation" of the Food and Drug Administration (FDA) was taken into account [150].

Performance characteristics according to CD 2002/657/EC

Selectivity

Selectivity describes the ability of a method to distinguish structurally similar analytes from one another. The EU Commission Decision 2002/657/EU states the following requirements for selective substance identification in confirmatory methods: The relative retention time of the analyte in the sample shall correspond to the calibration solution in the range of ± 2.5%. Besides the chromatographic match, mass spectrometry requirements also need to be met, since they provide information of the chemical structure of the target analyte. Using triple quadrupole mass spectrometry, at least two product ions need to be detected and their ion ratio must correspond to standard solutions or spiked matrix samples with predefined tolerances depending on the relative intensity of the detected ion expressed as a percentage of the base peak [148]. Using fullscan HRMS methods, the detection of one fullscan ion (i.e., molecular ion or characteristic adducts) [148] and one product ion with a mass deviation lower than 5 ppm is sufficient to identify the analyte unambiguously.

Specifity

A method can be considered to be specific if the detection and identification of the analyte is not impaired by other substances or matrix components. In order to evaluate the specificity of the method, blank matrix samples were analyzed and checked for peaks co-eluting at the expected retention time.

Stability

The chemical stability of the analyte in solvent is assessed over a period of one year to determine the storability of the stock and spiking solutions at 4°C. A deviation of 20% is considered acceptable. The analysis of the chemical stability of the analyte in matrix is important in order to determine the best storage conditions and maximum storage time during the study. The stability of the analyte in matrix is analyzed at bench top conditions (room temperature with and without light exposure) and storage conditions (4°C and -20°C) over the course of 28 days as well as after three freeze and thaw cycles.

Uncertainty-related performance characteristics

The uncertainty-related performance characteristics that have to be determined are the standard deviation of repeatability, the standard deviation of reproducibility, the decision limit CC_α, the detection capability CC_β and the trueness. The decision limit CC_α is defined as: *"[...] the limit at and above which it can be concluded with an error probability α that a sample is non-compliant"* [148]. Put another way, CC_α represents the measured concentration from which on a certain threshold value x_0 is exceeded significantly. For substances with an established maximum residue limit (MRL) the threshold x_0 equates to this MRL value. The error probability than can be set to 0.05. For banned substances x_0 is zero and α = 0.01. The official definition of the detection capability CC_β reads: "*Detection capability means the smallest content of the substance that may be detected, identified and/or quantified with an error probability of β. In the case of substances for which no permitted limit has been established, the detection capability is the lowest concentration at which a method is able to detect truly contaminated samples with a statistical certainty of $1-\beta$. In the case of substances with an established permitted limit, this means that the detection capability is the concentration at which the method is able to detect permitted limit concentrations with a statistical certainty of $1-\beta$."* [148] Practically, the detection capability is the true concentration value whose measurement with a probability of $1-\beta$ brings forth a result that exceeds CC_α.

To determine the standard deviation of repeatability, the standard deviation of reproducibility, the decision limit CC_α and the detection capability CC_β, a variance components model was fitted to the experimental data of the validation study that was performed as described above and included several analytical runs. The data analysis approach assumes empirical data to obey a linear, first order regression model. The deviation between the observed measurement and the "true" response is regarded as the sum of two error terms. The one error term reflects the measurement dispersion within an arbitrarily selected calibration series, while the other expresses the run-related variability that leads to fluctuations among different calibration series [151, 152].

Trueness or accuracy of the mean describes the closeness of agreement between the average value obtained from a large series of test results and an accepted reference value [148] and can be assessed by the recovery of the analytical method. Recovery represents the percentage of the true concentration recovered during the analytical procedure [148]. In the concrete case, the analytical method uses matrix-calibration to accomplish quantification. That is, the solutions used to calibrate the analytical instrument are obtained by spiking blank samples at different levels and submitting those samples to the whole sample preparation procedure. In the case of matrix calibration, the expectation of the recovery within the calibration range is approximately 100%, provided that sample matrix and calibration matrix do not differ substantially from each other [153]. Since matrix calibration allows to quantify accurately without the necessity of correcting results for the bias error, no recovery has to be determined. Nevertheless, recovery is a valuable tool in the course of method development to assess systematic errors that may arise due to the sample preparation procedure (e.g. loss of analyte) or are caused by matrix effects (e.g. ion suppression or enhancement). Hence, recovery was investigated for the case of external calibration.

Performance characteristics according to DIN 32645

Performance characteristics according to DIN 32645 are determined under repeatability conditions. The decision limit according to DIN 32465 (in German called “Nachweisgrenze”) is the concentration, which according to the calibration curve, corresponds to the critical response. The critical response is the response at and above which it can be concluded with an error probability of $\alpha = 0.05$ that the analyte is present in the sample. The detection limit (in German called “Erfassungsgrenze”) is the true concentration value whose measurement with a probability of $1 - \beta$ brings forth a result that exceeds the de-

cision limit. The determination limit (in German called "Bestimmungsgrenze") is the concentration from which on, the uncertainty is adequately low in comparison to the measured value. In residue analysis a relative uncertainty of 50% is considered to be acceptable [149].

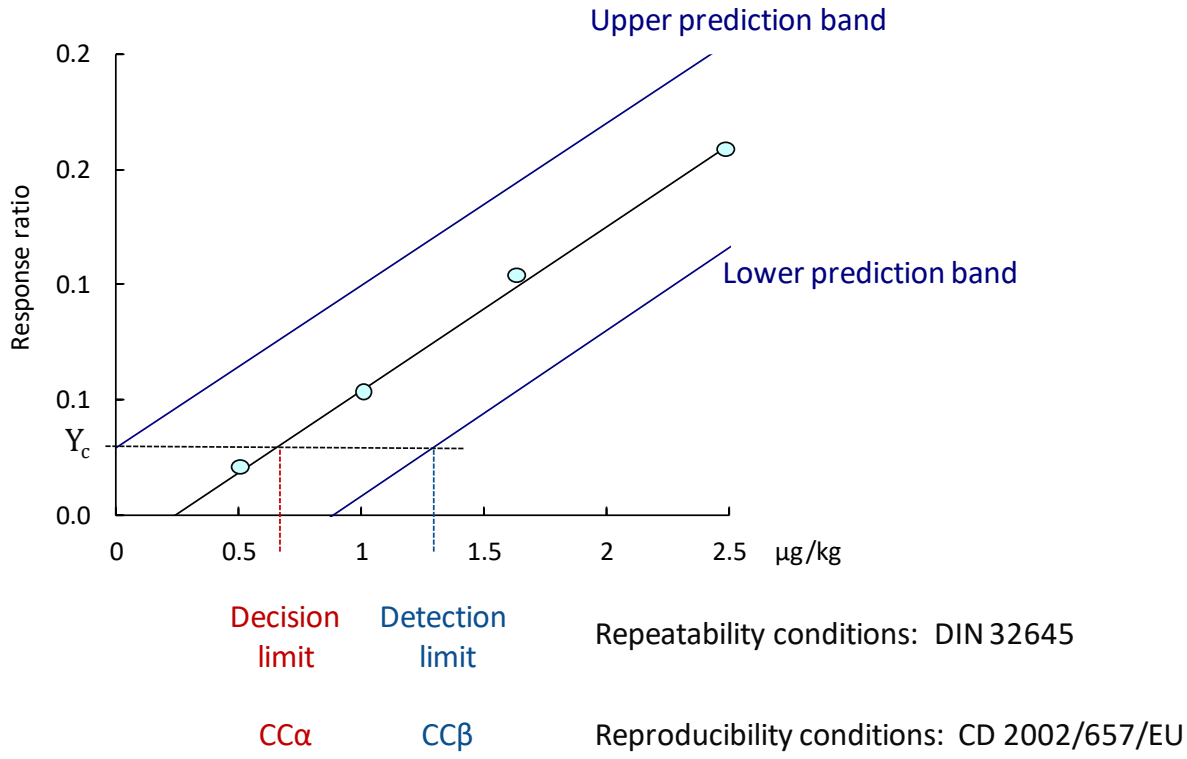


Figure 7. Schematic illustration of the performance characteristics decision limit ($CC\alpha$) and detection capability ($CC\beta$) under reproducibility conditions at a threshold value x_0 = 0 compared to the decision limit and detection limit according to DIN 32645 under repeatability conditions; Yc: critical response.

3.2.2.4 Analysis of urine samples from ALT-treated gilts

First, sample preparation was performed without a hydrolysis step in order to check for the presence of known metabolites such as ALT-glucuronide, sulfate conjugate, isomers and the parent drug (ALT) as described above (see Section 3.2.2.1) with internal standard ALT-d5 at a concentration of 500 ng/mL. In a second step, quantification of ALT after β-glucuronidase treatment was performed. For this purpose, 400 µL of urine were spiked with internal standard ALT-d5 at a concentration of 500 ng/mL and sample preparation was done as described above (see Section 3.2.2.1). Quantification was performed via matrix calibration using pooled samples of the non-medicated gilts. Calibration levels ranged from 2.5–100 ng/mL (for the

quantification of ALT and its isomer in non-hydrolyzed samples) and 50–2500 ng/mL (for the quantification of deconjugated ALT after hydrolysis).

3.2.2.5 NMR analysis of creatinine for normalization purposes

^{1}H NMR data acquisition was performed by Svenja Wenig (CVUA Karlsruhe) and Eva Gottstein (CVUA Karlsruhe). Text section regarding NMR data acquisition was provided by Svenja Wenig. Data evaluation was performed by myself.

^{1}H NMR data acquisition

^{1}H NMR spectra were acquired using a Bruker Avance III 400 Ultrashield spectrometer (BrukerBiospin, Rheinstetten, Germany) equipped with a 5-mm selective inverse probe (SEI) with Z-gradient coils, using a Bruker Automatic Sample Changer (Sample Xpress). ^{1}H NMR spectra were acquired using the standard pulse program noesygppr1d (1D NMR spectra) with presaturation of water signal and a prolonged d7. With a time domain of 65 k, 64 scans and 4 dummy scans, a spectral width of 20.55 ppm (8223 Hz), an acquisition time of 3.98 s and receiver gain of 64 for the ^{1}H 1D measurement. The temperature was set at 300 K (+/- 0.2 K). All spectra were automatically phased and baseline corrected. Data were acquired under the control of Sample Track Client (BrukerBiospin, Rheinstetten, Germany), requiring about 15 min per sample. NMR spectra were analyzed using TopSpin version 3.2 (Bruker Biospin, Rheinstetten, Germany). Creatinine concentration was determined via external calibration in the concentration range from 0.05 to 10.0 mg/mL.

Urine sample preparation for ^{1}H NMR analysis

For ^{1}H NMR, 500 µL urine or creatinine calibration standard, respectively, were spiked with 100 µL K_2HPO_4 buffer in D_2O (pH 7.4, 1.5 M) containing 5 mM TSP (trimethylsilylpropanoic acid-2,2,3,3-d4). Thus, a final D_2O concentration of 16.6% (v/v) for lock and shim purposes and TSP concentration of 0.87 mM was achieved.

3.2.3 Non-targeted LC-HRMS analysis

3.2.3.1 Sample preparation

For non-targeted screening the sample preparation was performed as described above (see Section 3.2.2.1) with 2 mL of urine and without the hydrolysis step. Furthermore, a reagent blank sample was also submitted to sample preparation.

3.2.3.2 LC-HRMS method

For untargeted analysis the same chromatographic and mass spectrometric parameters as listed under targeted analysis were used except MS/MS-spectra were triggered in fullscan data-dependent MS^2 discovery mode by a data-dependent fragmentation of the most intense ion per scan (see Section 3.2.2.2). Subsequently, instead of analyzing predefined target ions (ALT and ALT metabolites) a non-targeted data processing and statistical analysis workflow was utilized (see below).

3.2.3.3 Data processing and statistical analysis of non-targeted data

Raw files were converted to mzXML format [131] and imported into MZmine 2 [120]. Raw data were filtered (0.5–13 min) and masses were detected with centroid algorithm at a noise level of 1.0×10^6. Chromatogram building was carried out with a minimum time span of 0.05 min, a minimum height of 1.0×10^6 and a mass accuracy of 5 ppm or 0.002 Da. Chromatograms were smoothed (filter width 15). For chromatogram deconvolution, the noise amplitude algorithm with a minimum peak height of 3.0×10^6, peak duration of 0.05–1.5 min and amplitude of noise of 1.0×10^6 was applied. After deisotoping with an m/z tolerance of or 5 ppm or 0.002 Da, a retention time tolerance of 0.01 min and a maximum charge of 3, peaks were aligned using the join aligner algorithm. The m/z tolerance was set to 10 ppm or 0.005 Da weighted with 10%, the retention time tolerance was 0.2 min weighted with 10%, only using ions with the same charge. Gap filling was performed with 20% intensity tolerance at a mass accuracy of 10 ppm or 0.005 Da and 0.05 min retention time tolerance after retention time correction. The final steps were duplicate filtering with a mass accuracy of 10 ppm or 0.005 Da and 0.2 min absolute retention time tolerance as well as row filtering eliminating rows with less than 5 peaks. Data post-processing consisted of eliminating all mass spectrometric features (m/z at retention time) with relative standard deviation > 25% in QC samples (QC total: all samples pooled) and all mass spectrometric features which were present in the reagent blank sample at an intensity of $> 1.0 \times 10^7$.

The data was subjected to statistical analysis using an evaluation procedure that was implemented in MATLAB R2019b (The MathWorks Inc., Natick, Massachusetts). In a first step, all zero values were replaced by the half of the minimum detected peak area. The data set was normalized by the sum of all peaks. Subsequently, the data was transformed by taking the natural logarithm. After transposing and standardizing (mean-centering and division by standard deviation) multivariate statistical analysis was performed, i.e., PCA. The calculation

of a 0.9-prediction ellipse for the PCA data was performed using EXCEL 2010 (Microsoft Office, Redmond, U.S.).

3.2.3.4 Metabolite and biomarker identification

For metabolite identification, the data was submitted to MetaboAnalyst 4.0 [134]. Data preprocessing was performed as described above (see Section 3.2.3.3). Univariate tests (*Student's t*-test and a fold change test) were applied to identify significant features distinguishing the ALT-medicated from the non-medicated group. All *p*-values were adjusted using the procedure of Benjamini and Hochberg to account for multiplicity [143]. In order to visualize the results a 2D heatmap of the top 25 significant features of the *Student's t*-test was constructed in MetaboAnalyst 4.0. The samples were hierarchically clustered based on Euclidean distance as distance metric and Ward's linkage (algorithm to perform clustering to minimize the sum of squares of any two clusters) [134].

Structural elucidation was attempted for the top 25 features. In order to classify the confidence in candidate identification, the levels of confidence for unknown annotation described by Schymanski *et al.* were used [94]. Several vendor specific as well as open access software packages in addition to online databases for biomarker annotation, i.e., Compound Discoverer 3.1 (Thermo Scientific, Idstein, Germany), "seven golden rules" software [135], Metlin [138] and the HMDB [139] were used.

Furthermore, the data was submitted to pathway analysis by "MS Peaks to Pathway (Version 2)" in MetaboAnalyst 4.0 [134]. This functional module is based on the *mummichog* algorithm, which is able to predict the functional activity from mass spectrometric data without *a priori* identification, since metabolite identification can be derived from local enrichments in metabolic pathways [144]. As input parameters, a peak list (*m/z* at retention time) with the corresponding *p*-values was provided. The mass accuracy was set to 5 ppm and the top 10% significant mass spectrometric features were submitted to *mummichog* algorithm. The manually curated pathway library for *Homo sapiens* based on KEGG, BIGG and Edinburgh Model was selected as knowledge-base for pathway mapping as no model for the genus *sus* was available [144]. Pathways with the most mass spectrometric hits were further checked using the "Metabolika Pathways" function in Compound Discoverer 3.1.

4 Results and discussion

4.1 Non-targeted LC-HRMS analysis of samples taken within the frame of official veterinary drug residue control

Parts of the following passage have previously been published in [1].

4.1.1 Workflow development

The aim of this project was to develop and apply a data analysis procedure to data obtained by the application of a routine LC-HRMS method in order to gain further information about the origin of veterinary drug residues in non-compliant samples taken within the frame of official veterinary drug residue control. The LC-HRMS method used was an already existing in-house validated method for the quantitative determination of veterinary drugs in muscle and kidney samples. The validation was previously performed according to Commission Decision 2002/657/EC [148], but was not part of this thesis. The focus of this study laid on the development of a reliable data processing procedure for peak detection as well as on the application and assessment of chemometric methods in order to evaluate HRMS data.

4.1.1.1 Development of a data processing procedure using MZmine 2

For an assessment of the impact of different settings for peak detection and their optimization a statistical design of experiment with a fractional factorial approach was employed. For this purpose, 70 veterinary drugs were spiked in blank *porcine* muscle samples prior to sample preparation at different spiking levels ranging from 10%–150% of MRL of the respective veterinary drugs. Three data sets were prepared in three analytical batches over the course of two months in order to account for intensity variations caused by instrument related drifts. Figure 8 shows the identification (ID) ratios of the 8 different experiments at the spiking levels ranging from 10% MRL to 150% of MRL. The settings of experiment 6 led to the best ID ratios with 90% at a spiking level of 10% of MRL and 96% at spiking level of 150% of MRL, respectively. The settings of experiment 4 and 8 resulted in the lowest ID ratios.

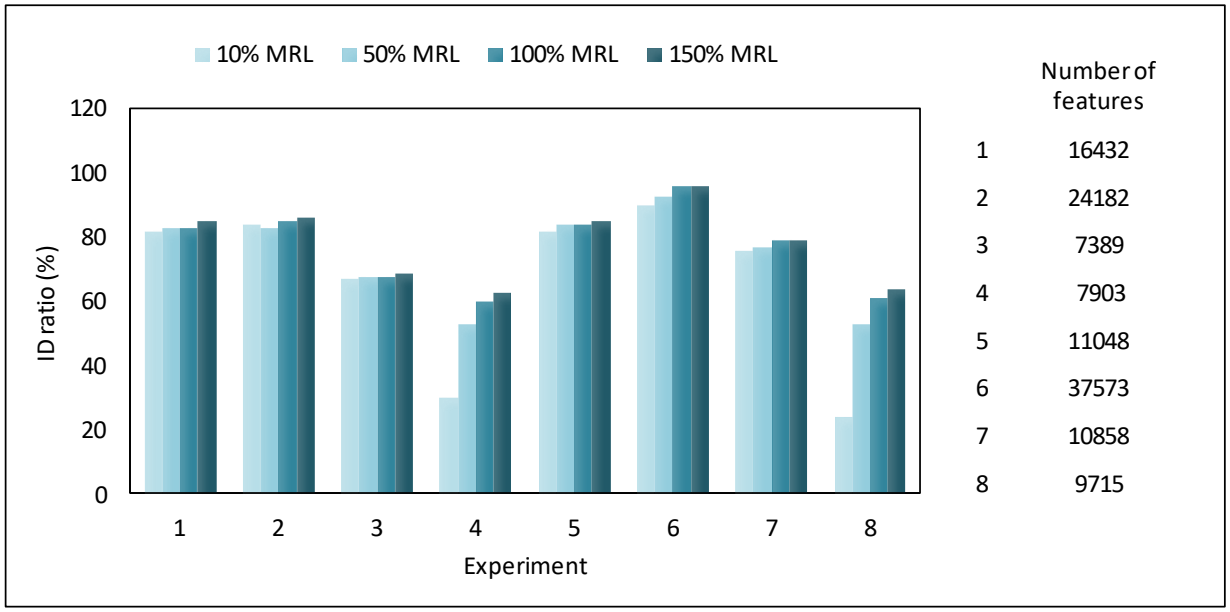


Figure 8. ID (identification) ratio of spiked antibiotics in *porcine* muscle samples at 10% MRL to 150% MRL using the fractional factorial DoE; mean of three datasets (*n* = 3, for reasons of clarity the measurement uncertainty is not given since it was very low (range of 0–0.12%)) and number of mass spectrometric features detected with respective experiment conditions.

The response function of the applied fractional factorial DoE is given by

$$Y = a_0 + a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 + a_5x_5 + \epsilon$$

Ordinary least-squares regression provided the estimates $\hat{a}_{0,\ldots,5}$ and $\hat{s}$, which are given in Table 8. The significance of the estimates was evaluated by the lower one-sided (in case of $\hat{a} > 0$) or the upper one-sided 0.9-confidence limit (in case of $\hat{a} < 0$), respectively. The lower 0.9-confidence limit of $\hat{a}_2$ and $\hat{a}_3$ is > 0 as well as the upper 0.9-confidence limit of $\hat{a}_5$ is < 0, respectively. Hence, these factors are of significance and cannot be eliminated.

Table 8. Estimates of coefficients obtained by ordinary least-squares regression and 0.9-confidence limits.

	$\hat{a}_0$	$\hat{a}_1$	$\hat{a}_2$	$\hat{a}_3$	$\hat{a}_4$	$\hat{a}_5$	$\hat{s}_a$
Estimated value	66.875	-1.125	17.625	9.875	-0.375	-12.375	2.019
Lower 0.9-confidence limit	63.067	-4.932	**13.817**	**6.067**	-4.183	-16.183	
Upper 0.9-confidence limit	70.683	2.683	21.433	13.683	3.433	**-8.567**	

$\hat{a}_1$: noise level, $\hat{a}_2$: peak time span (min), $\hat{a}_3$: minimum peak height, $\hat{a}_4$: smoothing $\hat{a}_5$: amplitude of noise (chromatogram deconvolution); bold values: upper/lower confidence border is less/greater than zero and thus the factor is of significance.

The factors of significance were minimum time span, minimum height and the amplitude of noise for chromatogram deconvolution, while smoothing and the initial noise level showed no significant effect. This approach shows the benefits of DoE, since a manageable number of experiments rendered valuable information for workflow development, i.e., to focus particular on the optimization of these 3 parameters.

Furthermore, the ID ratios of experiment 6 were very satisfactory. Hence, no further optimization using central composite design was needed. In experiment 6 the amplitude of noise and minimum height of peaks were set to the low level of 1.0×10^5 and 5.0×10^5, respectively, while the time span was set to the maximum level 0.05–1.5 min. Thus, the highest number of features (*m*/*z* at retention time) was detected, and consequently the highest ID ratio was achieved, detecting also analytes with lower intensities and analytes with a narrow peak shape as well as analytes which showed a broader peak shape.

The final data processing procedure was further validated with real world data of 15 non-compliant *porcine* muscle samples analyzed with the Q-Orbitrap method (as described in Section 3.1.4.1). In total, 16 antibiotics were identified with the vendor specific data acquisition software at or above MRL level (see Figure 9). Using the chosen Mzmine 2 data processing procedure, 15 out of the 16 antibiotics were detected resulting in an identification rate of 96%. Only one sample was false-negative for benzylpenicillin. This can be explained by its very narrow peak shape of < 0.1 min.

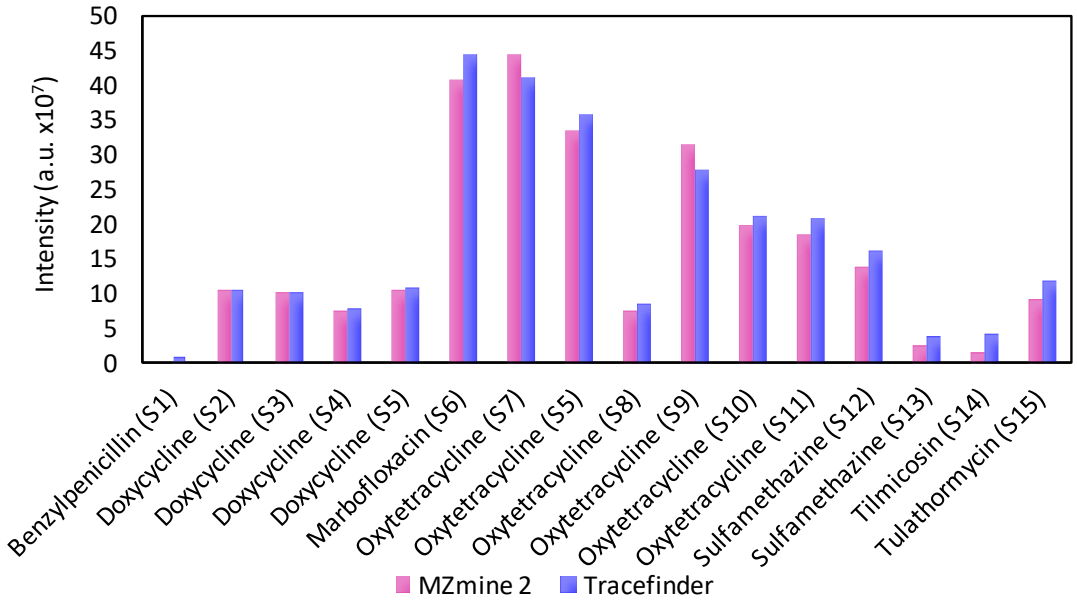


Figure 9. Comparison of peak integration via MZmine 2 and vendor specific software (Tracefinder) using veterinary drug residues in 15 non-compliant *porcine* muscle samples (S1–S15).

As a result, the settings of experiment 6 were maintained for all further data processing of Q-Orbitrap data for muscle samples, with the exception of the change of the smoothing width to the mean value of the two tested levels. The final data processing procedure was transferred to the Q-TOF data with minor changes which were due the instrument specification.

The *m*/*z* tolerance was set to 50 ppm or 0.02 Da and the intensity values for noise level and minimum peak height were adjusted according to the instrument (see also Section 3.1.4.2). Furthermore, the data processing procedure was adapted for the analysis of kidney samples only by decreasing the intensity threshold for the minimum peak height. In general, ion suppression in kidney samples was observed to be higher than in muscle samples. Hence, this adaption was rendered necessary.

4.1.1.2 Evaluation of a species-independent model

In order to evaluate whether samples from different species can be used in one multivariate statistical model, blank *porcine* and *bovine* muscle samples were examined. It was observed that the extracts of the *bovine* muscle samples were darker than that of *porcine* muscle samples and that they were slightly red colored (see Figure 10).

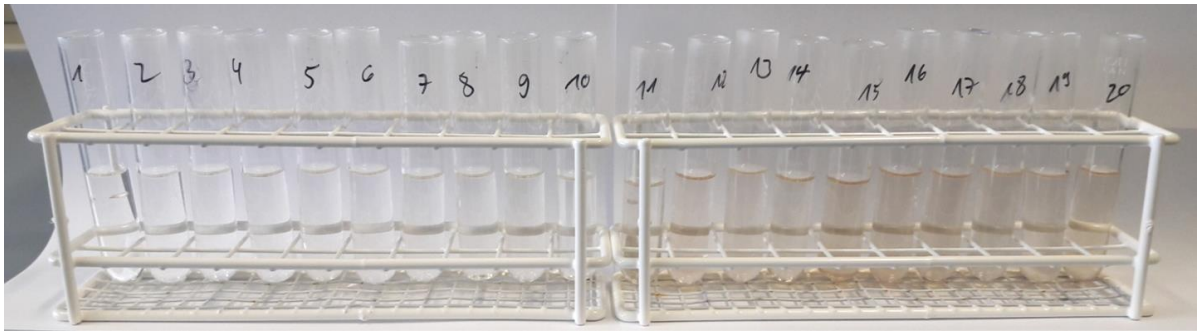


Figure 10. Extracts after SPE from *porcine* muscle samples (left) and *bovine* muscle samples (right).

The experimental data of *porcine* and *bovine* muscle samples was subjected to a PCA and PLS-DA. The scores plot of the PCA showed that the two species were well separated. The main variance was explained by PC1 (40.6%) which also was the main principal component responsible for the separation of the two species. PLS-DA also showed a good separation of the two species, but one must keep in mind that PLS-DA belongs to the supervised methods and that these methods are prone to over-fit data. Especially, when the data set used is small (n = 20).

Furthermore, mixtures of *porcine* and *bovine* muscle samples were analyzed and the data was projected in the PCA scores plot. For this purpose, the experimental data of the mixtures was normalized by the sum of all peaks, log-transformed and standardized (mean-centered and divided by standard deviation). The standardized data was multiplied by the loadings of the PCA and the resulting scores were graphed in the scores plot. The projection of the mixtures showed a gradual difference considering their composition (see Figure 11).

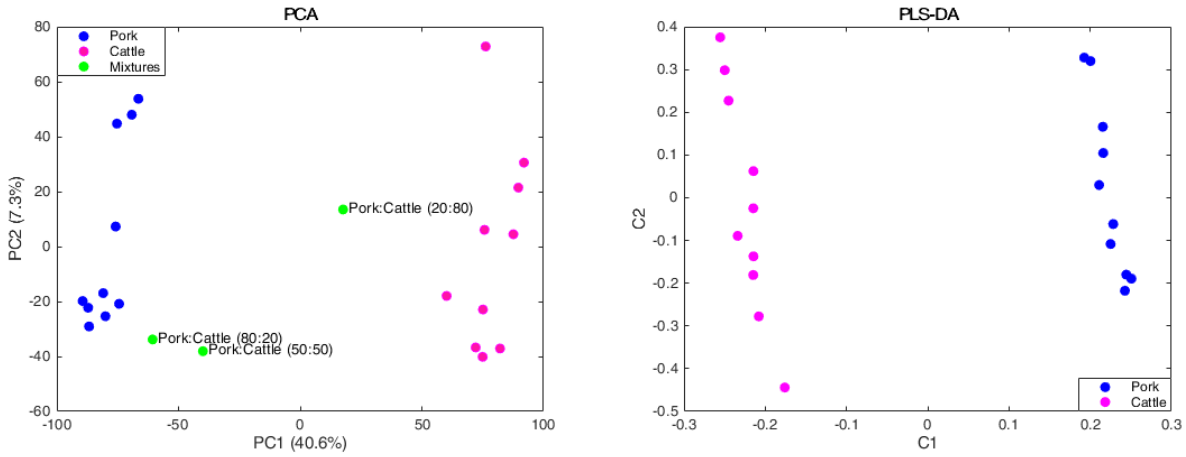


Figure 11. Left: scores plot of PCA of LC-HRMS (Q-Orbitrap) data of *porcine* and *bovine* muscle samples; blue: *porcine* muscle samples, red: *bovine* muscle samples, green: mixtures of *porcine* and *bovine* muscle samples (w/w); explained variance is shown in brackets. Right: scores plot of PLS-DA of LC-HRMS data of *porcine* and *bovine* muscle samples.

In order to identify significant analytes, univariate statistical analysis was performed, i.e., two-sided *Student's t*-test, fold change test and volcano plot [134]. A volcano plot illustrates the combined results from a two-sided *Student's t*-test and a fold change (FC) test, plotting the significance ($-\log10(p)$) versus the ratio between the area of a feature in *bovine* samples divided by the area of the same feature in *porcine* samples ($\log2(FC)$). The result is shown in Figure 12.

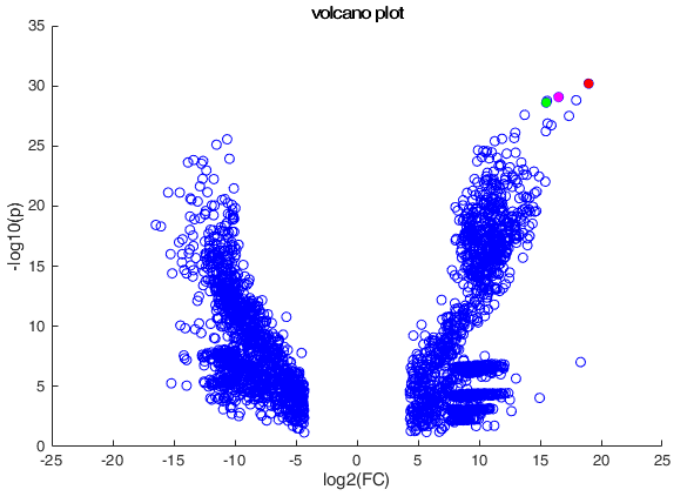


Figure 12. Volcano plot illustrating fold change (threshold 20) and statistical significance (*t*-test threshold 0.01); fold changes and *p*-values are log transformed; blue circles represent features (*m/z* at retention time) above the thresholds; red filled circle: *m/z* 460.7528 (@ 4.054 min), magenta filled circle: *m/z* 307.5042 (@ 4.053 min) and green filled circle: *m/z* 920.4978 (@ 4.053 min).

The five most significant variables in a two-sided *Student's t*-test are listed in Table 9 with the respective fold change indicating higher levels in *bovine* muscle samples. Four of the five top significant variables elute at the same retention time. Hence, the raw data was checked for these variables. The isotope patterns of four of the five variables revealed that they all belong to one analyte, which elutes at 4.05 min. *M*/*z* 460.7528 is a double-charged ion ($[M+2H]^{2+}$), *m*/*z* 307.5042 is a triple-charged ion ($[M+3H]^{3+}$) and *m*/*z* 920.4978 corresponds to the product ion ($[M+H]^{+}$). *M*/*z* 307.8386 is the isotope M+1 of *m*/*z* 307.5042 (see Figure 13).

Table 9. Top five significant variables in two-sided *Student's t*-test at a significance level of 5% with the respective fold change (FC).

m/*z* (at retention time in min)	FC	*p*-adj.	$-\log10(p)$
460.7528 (@ 4.054)	512630	6.41×10^{-31}	30.193
307.5042 (@ 4.053)	93294	8.45×10^{-30}	29.073
397.5486 (@ 4.212)	253860	1.55×10^{-29}	28.810
307.8386 (@ 4.053)	48454	1.74×10^{-29}	28.759
920.4978 (@ 4.053)	46170	2.41×10^{-29}	28.618

FC: fold change, *p*-adj.: multiple test corrected *p*-value based on the Benjamini Hochberg false discovery rate.

Thus, the focus was set on the identification of this analyte with the accurate mass of 919.4905, denoted as analyte X in the following. For structural elucidation, a validated, freely accessible software called "the seven golden rules" was used [135]. The name of the algorithm of the software originates from seven heuristic rules, which are used for filtering the possible molecular formulas, i.e., (1) restriction for the number of elements, (2) LEWIS and SENIOR chemical rules, (3) isotopic patterns, (4) hydrogen/carbon ratios, (5) element ratio of nitrogen, oxygen, phosphor and sulphur versus carbon, (6) element ratio probabilities and (7) presence of trimethylsilylated compounds[2] [135]. On the basis of the accurate mass and the corresponding isotope pattern, the software computes proposals for elemental compositions with an assigned score. Subsequently, it performs a search against the database of PubChem and an integrated peptide database. Table 10 shows the top ten proposals for molecular formulas for analyte X.

[2] One should note that rule number 7 (i.e., trimethylsilylated products being considered) is only relevant for structural elucidations of GC-MS data where derivatization procedures such as silylation are employed.

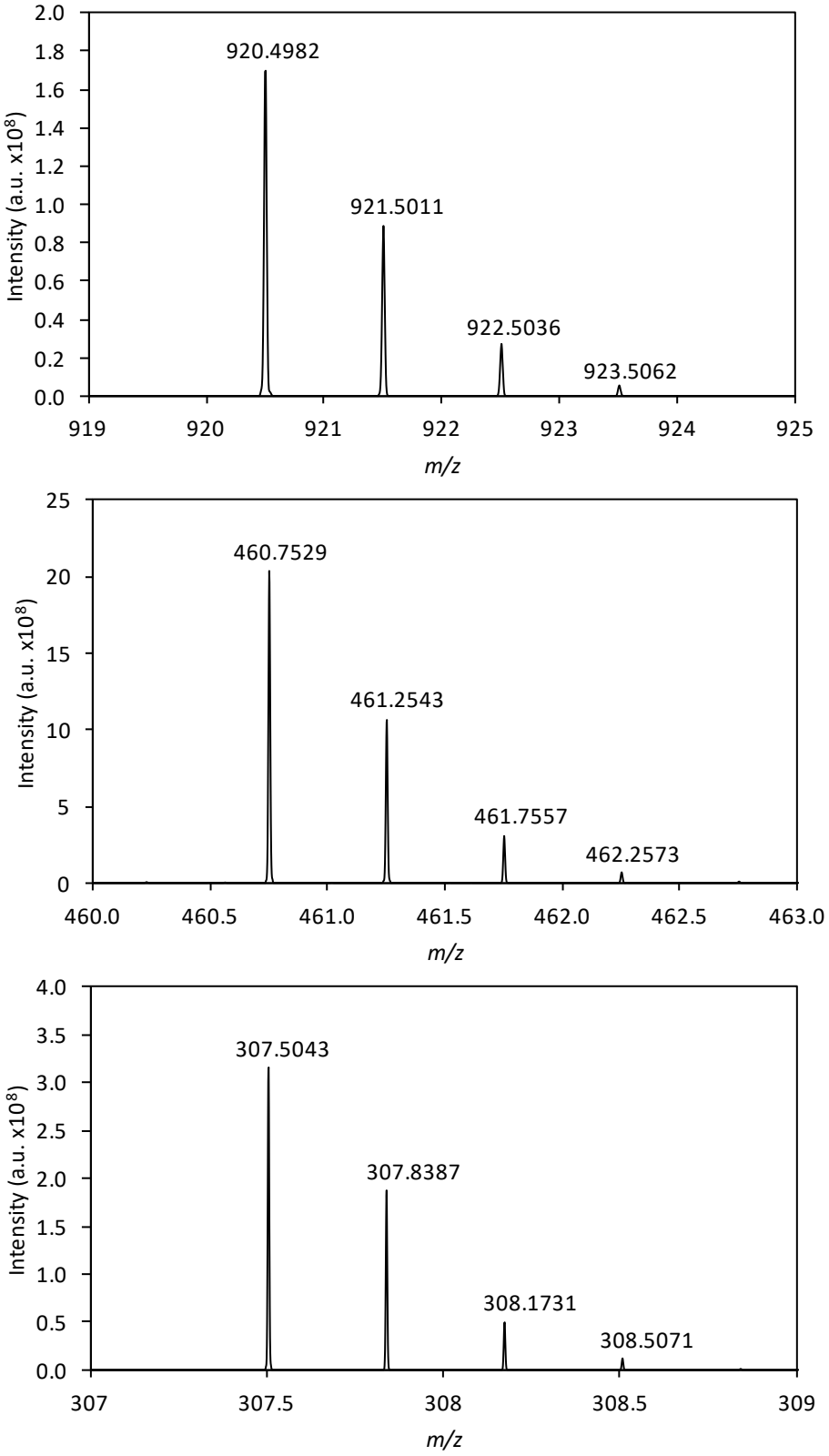


Figure 13. Isotopic pattern of *m/z* 920.4982, 460.7529 and 307.5043 at retention time 4.05 min. The mass differences between M, M+1, M+2 and M+3 indicate the charge state of the ions as single-charged, double charged and triple charged.

Table 10. Top ten proposals for molecular formulas obtained with seven golden rules; score: isotopic pattern score for respective molecular formula; peptide: match in peptide database of software.

Molecular formula	Mass accuracy (ppm)	Score (%)	Pubchem result	Peptide
$C_{42}H_{61}N_{15}O_9$	14.01	98.34		Found
$C_{41}H_{57}N_{23}O_3$	11.88	98.33		-
$C_{41}H_{61}N_{17}O_8$	1.79	98.20		Found
$C_{44}H_{65}N_{13}O_9$	13.35	97.90		Found
$C_{39}H_{53}N_{25}O_3$	15.47	97.76		-
$C_{42}H_{57}N_{21}O_4$	0.33	97.71		-
$C_{40}H_{61}N_{19}O_7$	10.43	97.57		-
$C_{43}H_{69}N_9O_{13}$	11.89	97.40	octapeptide FGLQLELT	Found
$C_{44}H_{69}N_7O_{14}$	0.32	97.24		-
$C_{45}H_{65}N_{11}O_{10}$	1.13	97.10		Found

It needs to be noted that five out of the top ten results were found in the peptide database of the software and in consideration of the multiple charge states observed for the analyte X, it can be hypothesized that analyte X could be a peptide.

Additionally, an MS/MS-spectra search against the Metlin database and MasCot database was performed in order to identify characteristic fragments in the product ion spectra of analyte X. MasCot is particularly useful since it offers an MS/MS spectra search against various protein databases such as SWISSprot or NCBIprot. Although, the sample preparation was not designed for peptide analysis, it is possible, that the acidic buffering conditions led to protein digestion into smaller peptides. With the help of the database search, various fragments in the product ion sprectra were identified (see Figure 14).

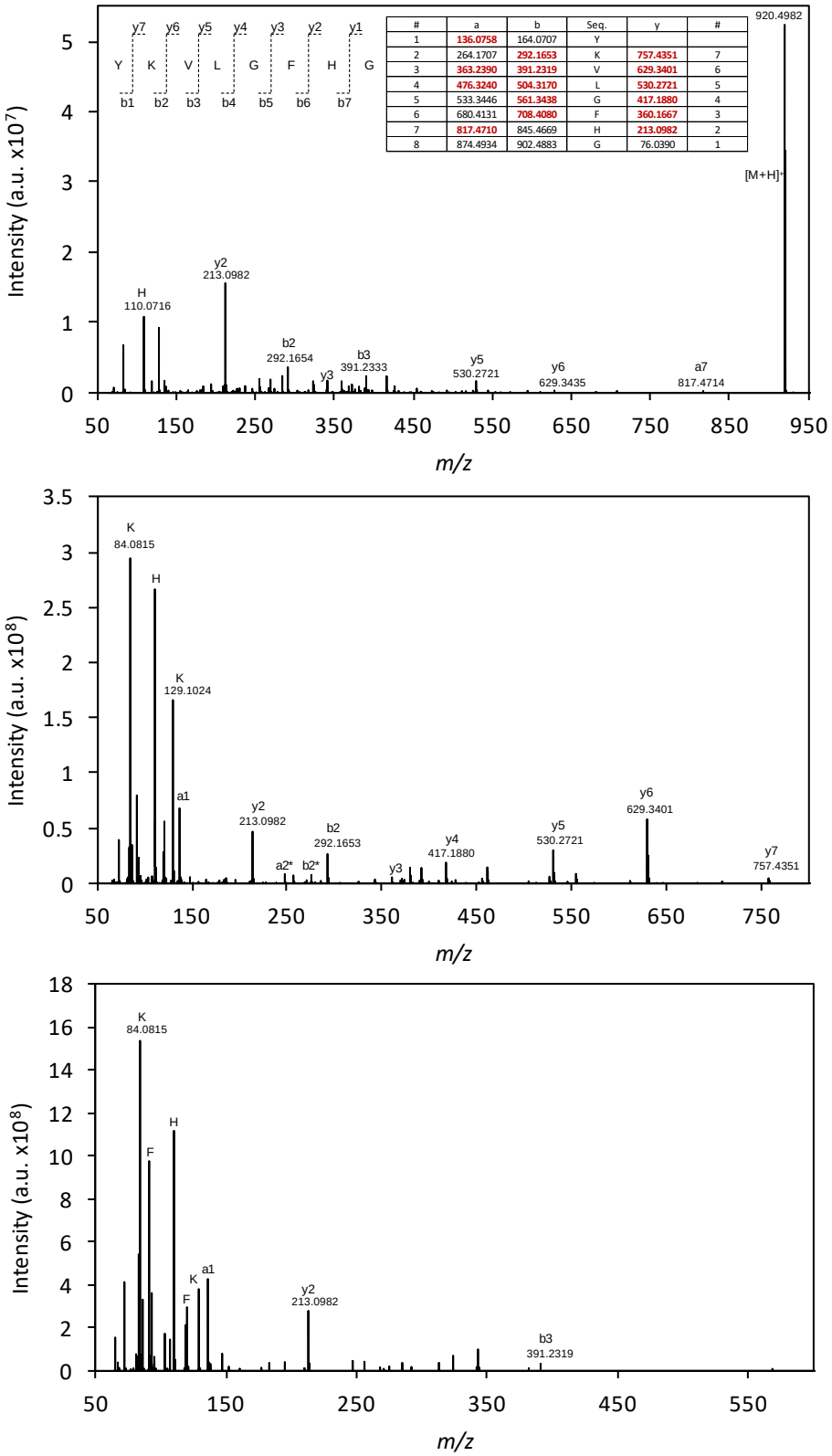


Figure 14. Representative product ion spectra of *m/z* 920.4982 $[M+H]^+$ (top), 460.7529 $[M+2H]^{2+}$ (middle) and 307.5044 $[M+3H]^{3+}$ (bottom) at 4.05 min. The table illustrates the fragment ions of the peptide YKVLGFHG as predicted by [140]. a-, b- and y-fragments which match are illustrated in red; a2*, b2* represents the respective a- and b-fragment after loss of ammonia (-17 Da); K: immonium related ions of K; H: immonium ion of H; F: immonium and immonium related ions of F (for details for peptide fragment identification see Supplementary data Figure S1).

These fragments led to the proposal of the peptide sequence YKVIGFHG or YKVLGFHG. However, I (isoleucine) and L (leucine) could not be differentiated with the available data as they represent molecular ions with the same exact mass. In evaluation of all three MS/MS-spectra, all y-fragments of the proposed peptide sequence except y1 were observed as well as some b- and a-fragments. a2* and b2* are diagnostic fragments for YK, since they are formed after a loss of ammonia. Characteristic immonium ions and immonium related ions of H, F and K were also observed. Furthermore, the molecular formula of the proposed peptide sequence is $C_{45}H_{65}N_{11}O_{10}$, which was hit number 10 of the seven golden rules with a mass accuracy of 1.13 ppm and an isotopic pattern score of 97.10%.

In a final step, both peptide sequences YKVIGFHG and YKVLGFHG were compared against the NCBI protein database using BLASTp, resulting in a match with the sequence string 147-154 of myoglobin in cattle (*bos taurus*, UniProt accession number P02192) for YKVLGFHG [154]. A comparison of the myoglobin in pork (*sus scrofa,* UniProt accession number P02189) showed that this specific sequence is not present in *porcine* myoglobin (see Figure 15). In the literature, myoglobin has already been described as a protein with species-specific peptide sequences and can be used as biomarker for species differentiation [155, 156]. Furthermore, the areas of the corresponding peaks of YKVLGFHG in the extracted ion chromatograms of *m/z* 460.7529 ($[M+2H]^{2+}$) are gradually descending with decreasing amount of cattle tissue in extracts of mixtures of cattle and pork muscle tissues (see Figure 16). These observations are line with the literature, since this marker peptide can also be applied quantitatively [155]. Even 1 % of cattle muscle tissue spiked in pork muscle tissue was clearly detected via *m/z* 460.7529 ($[M+2H]^{2+}$) with high sensitivity.

```
                10         20         30         40         50
P02192      MGLSDGEWQL VLNAWGKVEA DVAGHGQEVL IRLFTGHPET LEKFDKFKHL
P02189      MGLSDGEWQL VLNVWGKVEA DVAGHGQEVL IRLFKGHPET LEKFDKFKHL
                60         70         80         90        100
P02192      KTEAEMKASE DLKKHGNTVL TALGGILKKK GHHEAEVKHL AESHANKHKI
P02189      KSEDEMKASE DLKKHGNTVL TALGGILKKK GHHEAELTPL AQSHATKHKI
               110        120        130        140        150
P02192      PVKYLEFISD AIIHVLHAKH PSDFGADAQA AMSKALELFR NDMAAQYKVL      GFHG
P02189      PVKYLEFISE AIIQVLQSKH PGDFGADAQG AMSKALELFR NDMAAKYKEL      GFQG
```

Figure 15. Protein sequence of myoglobin in cattle (top, P02192) and pig (bottom, P02189). The red sequence string indicates the marker peptide for cattle.

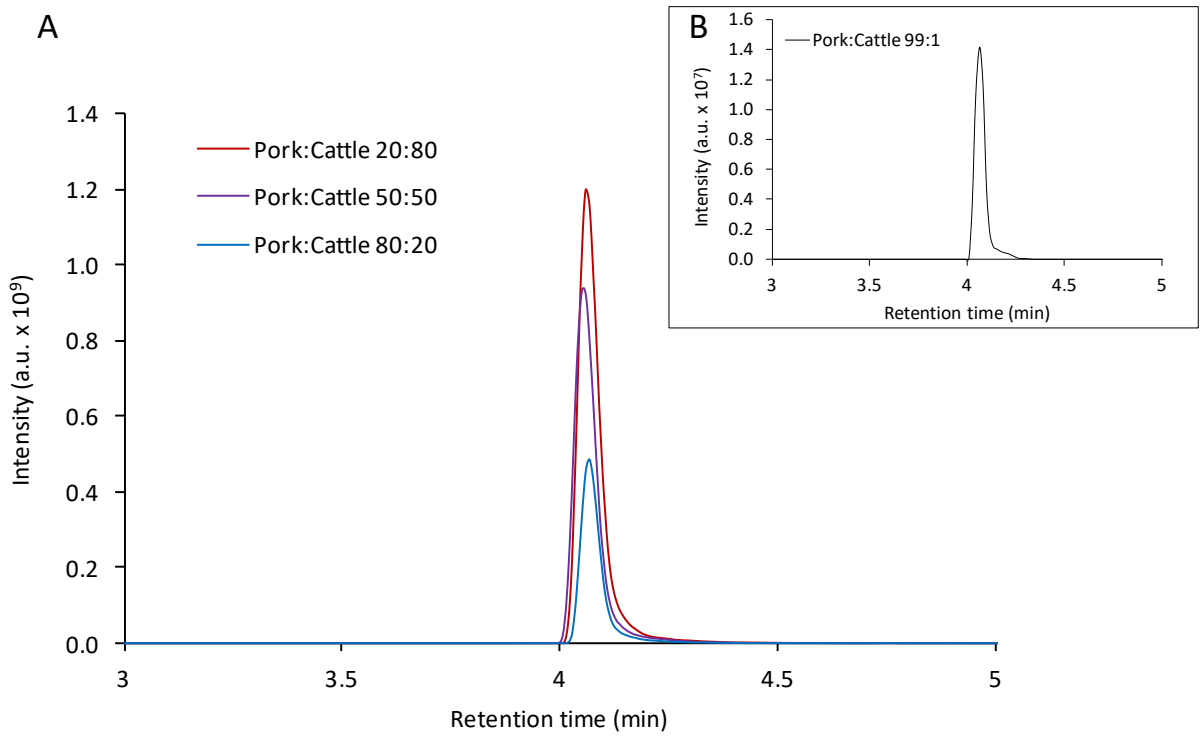


Figure 16. Extracted ion chromatograms of *m/z* 460.7529 ($[M+2H]^{2+}$) of extracts from mixtures of pork and cattle muscle tissue (w/w). A: YKVLGFHG is decreasing with descending order of *bovine* tissue in the mixture. B: One percent of *bovine* muscle tissue spiked in *porcine* muscle tissue is shown in enlarged section.

The robustness of the peptide detection of the used method which originally was developed for veterinary drug residue detection was investigated with regard to changes in pH of buffer solutions. The sample preparation was performed as described above, with the exception that the extraction was carried out at various pH levels (pH = 4.0, 4.3, 4.7). A significant decrease in peak area was observed with a buffer at pH 4.7 in comparison with a buffer at pH 4.0 (see Figure 17). The decrease in peak area of the marker peptide showed that out of the three pH levels examined, a McIlvaine buffer at pH 4.0 represents the best condition for an efficient formation or extraction of the marker peptide. The results showed further that the sample preparation is not robust with regard to an increase in pH.

However, these results do not provide a complete explanation for the decrease in peak area of the marker peptide with increasing pH. Several possible reasons can be hypothesized. One reason could be a decreased extraction efficiency with higher pH. Other explanations could be a poor formation rate via acid cleavage of the marker peptide at higher pH or that pH 4.0 represents a pH optimum for an unspecific peptidase which cleaves the marker peptide sequence. An online search in the MEROPS database for peptidases with a preference of glu-

tamine in P1 and tyrosine P1' of the cleaved bond did not lead to conclusive results [157]. Thus, it is not possible to decide whether the octapeptide is cleaved from the C-terminal of myoglobin due to acidic conditions of the buffer or due to an unspecific endopeptidase after acidic denaturation of the protein.

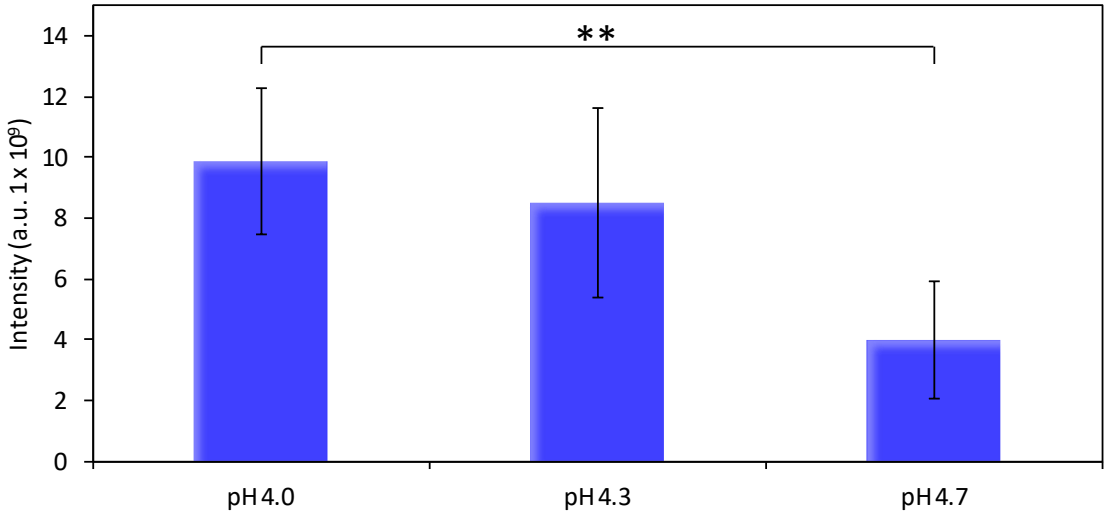


Figure 17. Peak areas of *m/z* 460.7529 ($[M+2H]^{2+}$) at 4.05 min in *bovine* muscle tissue extracts prepared with McIlvaine buffer at 3 different pH levels (n = 3); mean ± measurement uncertainty (MU, half-width of the 0.95-confidence interval); ** p-value from *Student's t*-test: $p < 0.01$.

In summary, these results showed that samples from different species, i.e., pigs and cattle, needed to be evaluated separately. The most significant analyte for discriminating these two species was identified as the octapeptide YKVLGFHG cleaved from the C-terminal of *bovine* myoglobin. These observations served further as proof of concept for the chosen workflow and the approach to identify differences in two test groups based a non-targeted approach with subsequent multivariate statistical analysis and biomarker identification.

4.1.1.3 Evaluation of the variables sample storage time and gender

In the process of study design, the influence of sample storage time and gender of the animal were also investigated in order to determine whether these factors have an effect. For *porcine* samples, 25 samples from healthy animals of both genders (female vs. male, castrated) with different storage times (three, four and twelve months) were analyzed and the experimental data was submitted to PCA. The scores plots showed no sample clustering according to gender or storage time in muscle samples (see Figure 18).

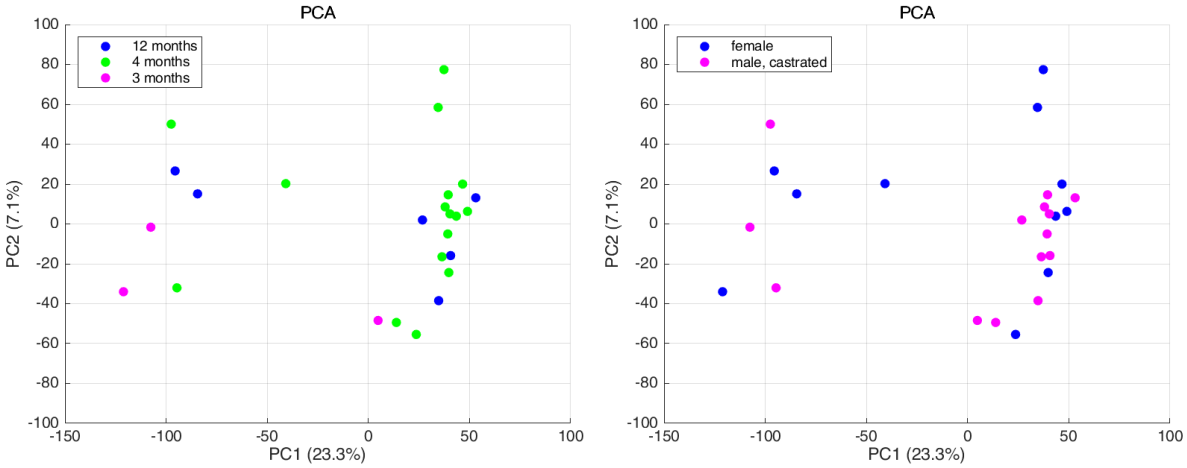


Figure 18. 2D-scores plot of PCA of LC-HRMS (Q-Orbitrap) data obtained from the examination of control *porcine* muscle samples. Left: samples were colored according to sample storage time; right: samples were colored according to gender of the animal from which the sample was taken; PC: principal component with explained variance shown in brackets.

Hence, storage times as well as gender of animals were not deemed confounding variables for LC-HRMS. The influence of age of the animals could not be checked. Normally, pigs are slaughtered around 6 months of age after fattening. Hence, all control samples consisted of samples from fattening pigs. Yet, the positive samples also consisted of samples from piglets which were approximately 3 months of age. However, the age difference is small with maximum 3 months and consequently, it was decided to move forward despite this limitation of the sample population.

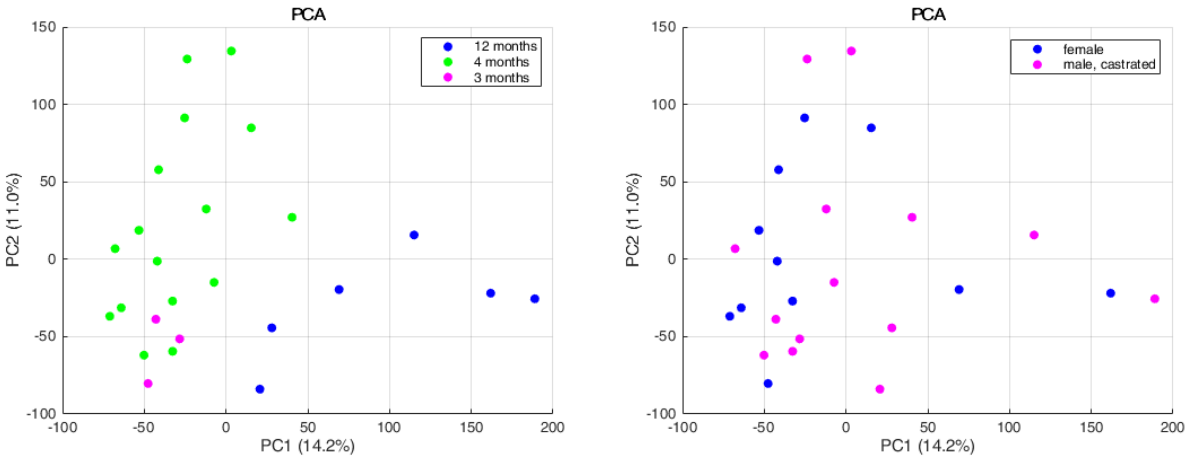


Figure 19. 2D-scores plot of PCA of LC-HRMS (Q-Orbitrap) data obtained from the examination of control *porcine* kidney samples. Left: samples were colored according to sample storage time; right: samples were colored according to gender of the animal from which the sample was taken; PC: principal component with explained variance shown in brackets.

In *porcine* kidney samples a trend towards clustering of samples stored for 12 months was observed in the PCA scores plot. The limitation of a change of the metabolome due to prolonged storage can thus not be excluded for *porcine* kidney samples. The gender of the animals showed no effect in the PCA scores plot of *porcine* kidney samples (see Figure 19).

Comparable results were observed in the PCA scores plots of *bovine* samples (see Figure 20). In muscle samples no apparent clustering according to storage time or gender was visible. However, in kidney samples an effect of storage time cannot be excluded, since a clustering tendency of samples stored for 12 months was observed. Furthermore, the effect of age of the control *bovine* samples was evaluated. Cattle farming differs substantially from pig farming. Heifers are slaughtered at an age of 18–24 months, while dairy cows can live for up to 10 years [158]. However, the scores plots of *bovine* muscle and kidney samples showed no clustering according to age. In general, one needs to keep in mind that the sample size for the examination of *bovine* samples was low with 15 control samples.

Considering the available sample size, the observations obtained during study design and the economic importance of the two species, the focus of the following study was on the examination of *porcine* muscle sample. In Germany commercial slaughtering of pigs (2018: 56,825,400 animals corresponding to 5,363,300 t) significantly exceeds commercial slaughtering of cattle (2018: 3,456,000 animals corresponding to 1,115,600 t) [159]. Hence, pig slaughtering has a higher economic value. Furthermore, the treatment frequency of pigs with antibiotics is significantly higher than that of cattle [18], although a direct comparison is difficult.

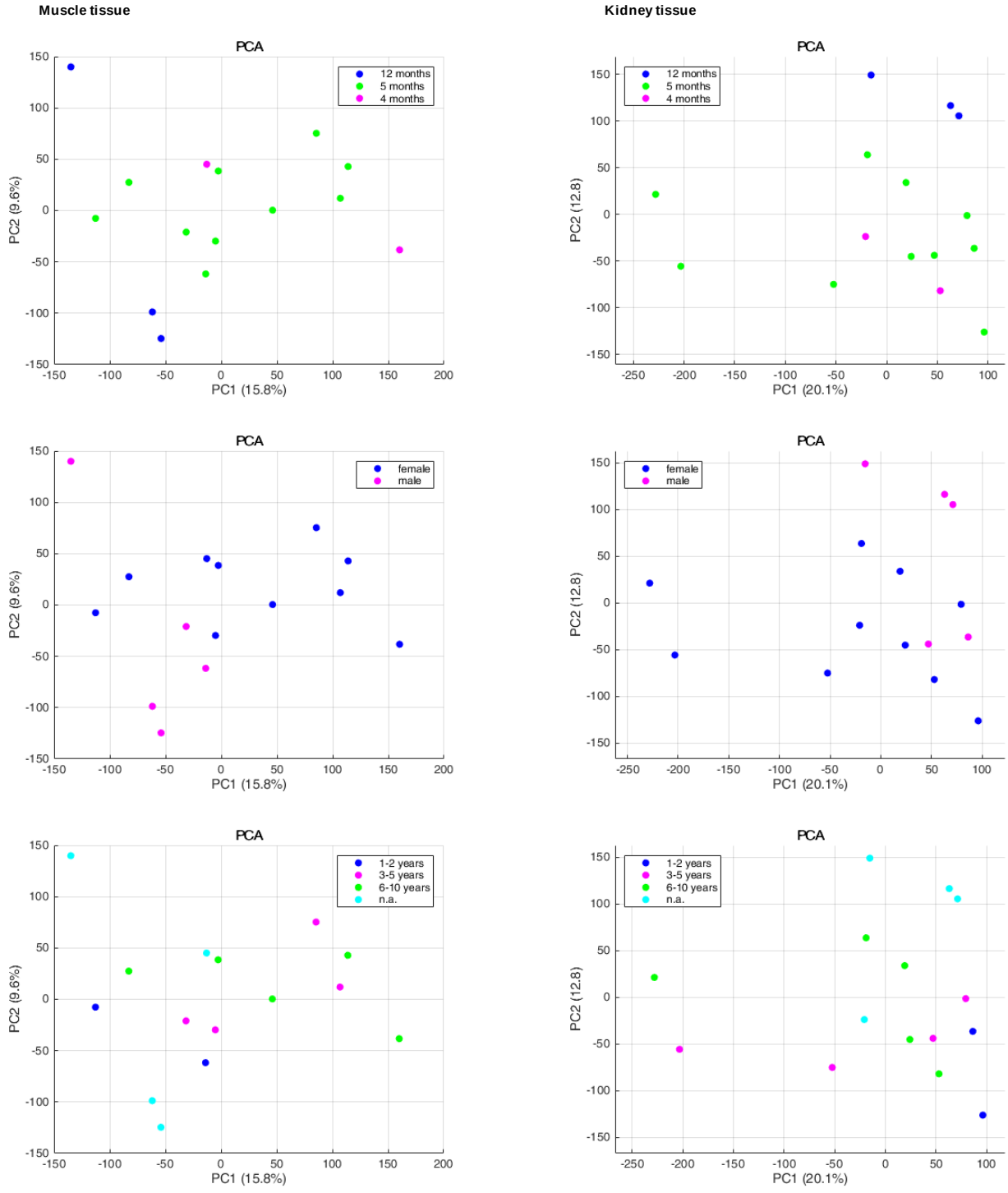


Figure 20. 2D-scores plot of PCA of LC-HRMS (Q-Orbitrap) data obtained from the examination of control *bovine* samples. Left: *bovine* muscle samples; right: *bovine* kidney samples; top: samples were colored according to sample storage time; middle: samples were colored according to gender of the animal from which the sample was taken; bottom: samples were colored according to age; PC: principal component with explained variance shown in brackets.

4.1.2 Non-targeted LC-HRMS analysis of *porcine* muscle samples

Data integrity

In order to verify that data processing via MZmine 2 performed a reliable peak detection as well as peak integration and to check for analytical drifts, a data integrity check was performed. For this purpose, the number of spiked analytes in the spiked QC samples was determined and compared against the expected detection rate. Eight spiked QC samples were scattered throughout two analytical batches. The data integrity check revealed that 87% and 90% of all spiked analytes were detected in the Q-Orbitrap data and the Q-TOF data, respectively (see Table 11). The relative standard deviation (RSD) of the peak areas of the spiked analytes were mostly below 20%, which indicated that the analytical method was robust. Nine analytes in the Q-Orbitrap data and 7 analytes in the Q-TOF data were not detected with the chosen data processing method. A check of the raw data showed that the peak intensities of these analytes were below the chosen minimum peak height for mass detection. This compromise was necessary since otherwise too many artifact peaks would have been detected as chromatographic peaks.

Table 11. Mean peak areas of spiked analytes at MRL level in QC samples with measurement uncertainty (MU, half-width of the 0.95-confidence interval) and relative standard deviation (RSD; *n* = 8).

Analyte	Q-Orbitrap			Q-TOF		
	Mean (a.u.x10^7)	MU (a.u.x10^7)	RSD (%)	Mean (a.u.x10^4)	MU (a.u.x10^4)	RSD (%)
Cephalosporins						
Cefalexin	6.85	0.19	3.35	20.46	1.21	7.25
Cefapirin	1.59	0.07	5.41	7.27	0.31	5.31
Ceftiofur	19.36	1.13	7.16	196.05	19.33	12.09
Cefquinome	0.65	0.06	11.42	4.20	0.29	8.37
Desacetylcefapirin	1.61	0.09	6.96	8.35	0.46	6.80
Desfuroylceftiofur	5.94	0.72	14.87	33.02	7.78	28.91
Coccidiostats						
Halofuginone	0.24	0.04	20.36	n.d.		
Corticosteroids						
Betamethasone	n.d.			n.d.		
Methylprednisolone	0.35	0.04	14.35	1.66	0.81	60.17
Prednisolone	n.d.			0.45	0.38	103.79
Diaminopyrimidines						
Trimethoprim	28.6	1.37	5.87	59.49	3.67	7.56
Lincosamides						
Lincomycin	16.3	0.73	5.47	95.79	6.11	7.82
Pirlimycin	7.28	1.37	23.05	85.18	5.85	8.43

Analyte	Q-Orbitrap			Q-TOF		
	Mean (a.u.x10^7)	MU (a.u.x10^7)	RSD (%)	Mean (a.u.x10^4)	MU (a.u.x10^4)	RSD (%)
Macrolides						
3-O-Acetyltylosin[a]	1.90	0.19	11.91	34.79	4.21	14.84
Azithromycin	0.36	0.07	24.53	18.14	1.55	10.50
Erythromycin	2.39	1.28	65.55	11.74	2.32	24.23
Gamithromycin	3.60	0.4	13.44	58.03	3.73	7.89
Josamycin	2.33	0.1	5.5	43.63	4.16	11.68
Oleandomycin	0.1	0.02	21.9	13.25	1.55	14.36
Spiramycin	1.98	0.5	31.09	34.08	4.22	15.17
Tildipirosin	27.92	2.02	8.88	356.98	21.14	7.26
Tilmicosin	1.35	0.37	33.24	32.29	4.95	18.79
Tulathromycin	4.16	0.6	17.7	117.52	10.76	11.24
Tylosin	7.30	0.53	8.86	121.80	13.53	13.62
Tylvalosin	2.20	0.15	8.57	39.46	2.61	8.12
Pleuromutilins						
8-A-Hydroxymutilin[b]	n.d.			n.d.		
Tiamulin	28.87	0.93	3.93	2.81	0.41	17.94
Valnemulin	7.69	0.33	5.30	130.92	8.32	7.79
Quinolones						
Ciprofloxacin	13.73	1.45	12.94	36.52	1.70	5.72
Danofloxacin	32.34	1.31	4.95	103.94	7.81	9.22
Difloxacin	56.86	6.77	14.60	223.39	18.22	10.00
Enrofloxacin	12.55	0.55	5.34	31.08	14.37	56.69
Flumequine	46.73	2.57	6.75	69.60	6.67	11.75
Marbofloxacin	22.74	2.87	15.48	82.31	3.88	5.79
Nalidixic Acid	16.57	0.92	6.82	17.94	1.40	9.59
Oxolinic Acid	32.83	3.22	12.04	33.74	4.18	15.21
Sarafloxacin	1.76	0.19	12.91	5.89	0.44	9.13
Sulfonamides						
Dapsone	0.97	0.11	13.44	n.d.		
Sulfachlorpyrazine	n.d.			8.96	0.81	11.11
Sulfachlorpyridazine	3.59	0.26	8.76	8.36	0.49	7.24
Sulfadiazine	6.97	0.31	5.36	13.63	0.72	6.48
Sulfadimethoxine	12.24	0.89	8.91	33.78	3.63	13.18
Sulfadoxine	12.31	0.43	4.30	28.80	3.20	13.62
Sulfamerazine	1.33	0.12	11.08	20.63	1.20	7.13
Sulfamethazine	13.97	1.15	10.08	22.19	2.23	12.35
Sulfamethoxazole	3.63	0.19	6.43	18.47	1.15	7.61
Sulfamethoxypyridazine	10.42	0.95	11.15	18.83	2.06	13.43
Sulfanilamide	n.d.			n.d.		
Sulfaquinoxaline	5.05	0.49	11.91	15.83	1.33	10.28
Sulfathiazole	4.91	0.25	6.14	10.64	0.44	5.09
Tetracyclines						
Chlortetracycline	2.97	0.78	32.12	22.77	2.96	15.93

Analyte	Q-Orbitrap			Q-TOF		
	Mean (a.u.x10^7)	MU (a.u.x10^7)	RSD (%)	Mean (a.u.x10^4)	MU (a.u.x10^4)	RSD (%)
Doxycycline	4.16	0.21	6.04	6.18	0.88	17.38
Oxytetracycline	3.66	0.51	17.04	0.19	0.04	26.70
Tetracycline	5.87	0.73	15.28	36.66	1.04	3.49
β-Lactam antibiotics						
Ampicillin	2.61	0.09	4.19	18.64	1.23	8.10
Benzylpenicillin	0.08	0.02	28.26	n.d.		
Cloxacillin	n.d.			39.96	3.36	10.30
Dicloxacillin	n.d.			21.93	1.85	10.34
Nafcillin	8.34	0.84	12.41	26.36	4.36	20.26
Oxacillin	n.d.			48.23	4.92	12.50
Phenoxymethylpenicillin	n.d.			n.d.		
Internal standards						
Amoxicillin-D4	0.02	0.01	96.19	0.09	0.05	72.12
Benzylpenicillin-D7	0.38	0.11	34.16	1.53	0.72	57.24
Demeclocycline	2.79	0.14	6.27	25.01	1.06	5.19
Enrofloxacin-D5	11.1	0.72	7.87	33.30	2.71	10.00
Oxolinic Acid-D5	33.67	3.90	14.20	39.09	5.37	16.86
Roxithromycin	0.70	0.13	23.60	142.54	13.89	11.95
Sarafloxacin-D8	8.04	0.43	6.56	46.60	1.75	4.60
Sulfadimethoxine-D6	12.62	1.23	11.98	37.05	4.71	15.58
Sulfadoxine-D3	15.05	0.62	5.08	36.77	5.08	16.95

n.d.: not detected; [a]: marker residue for tylvalosin, [b]: marker residue for tiamulin.

Multivariate data analysis

The data pretreatment consisted of data normalization, transformation and scaling. Data normalization was performed by dividing each variable by the sum of all peaks. This is a common normalization technique in order to adjust for analytical drifts and matrix effects. In the case of *porcine* muscle samples, no further QC based normalization was performed since the data integrity check showed that there was no observable batch effect. Log-transformation was applied to account for skewness of the original data. Furthermore, the data was standardized (mean-centering and division by standard deviation), since peak areas spanned over different orders of magnitude. This scaling method helps to make all variables equally important for further analysis [160].

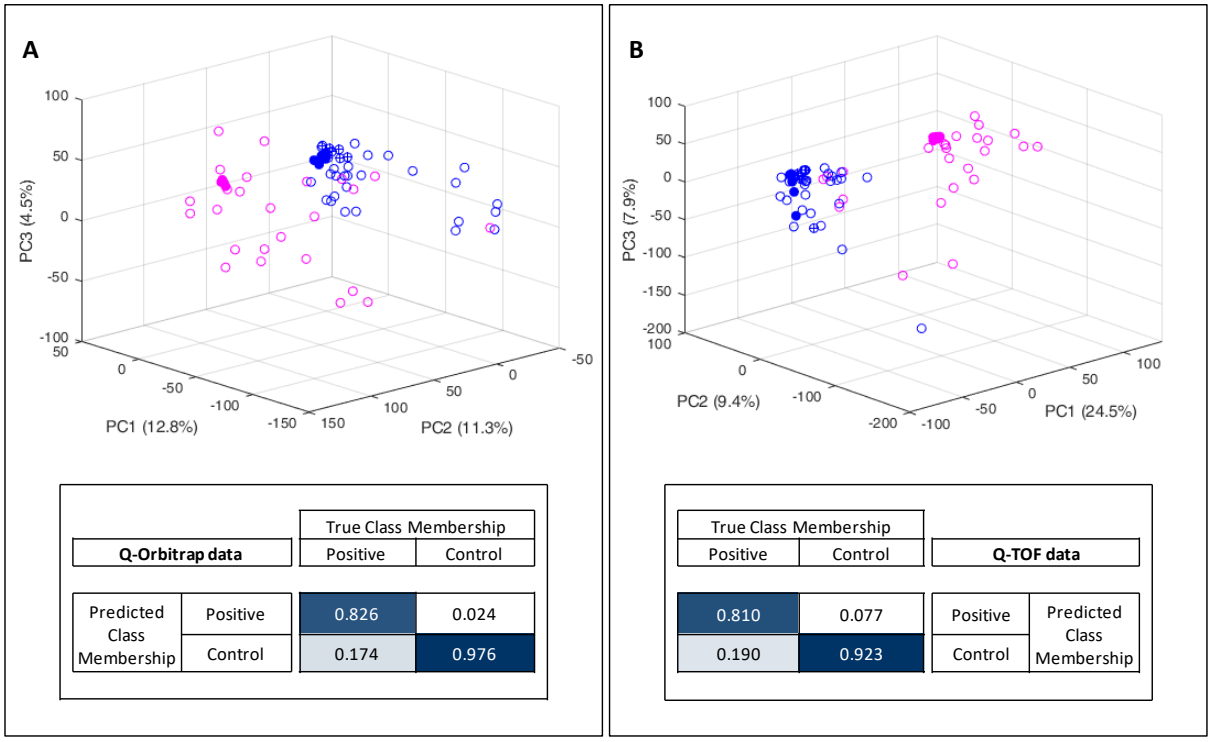
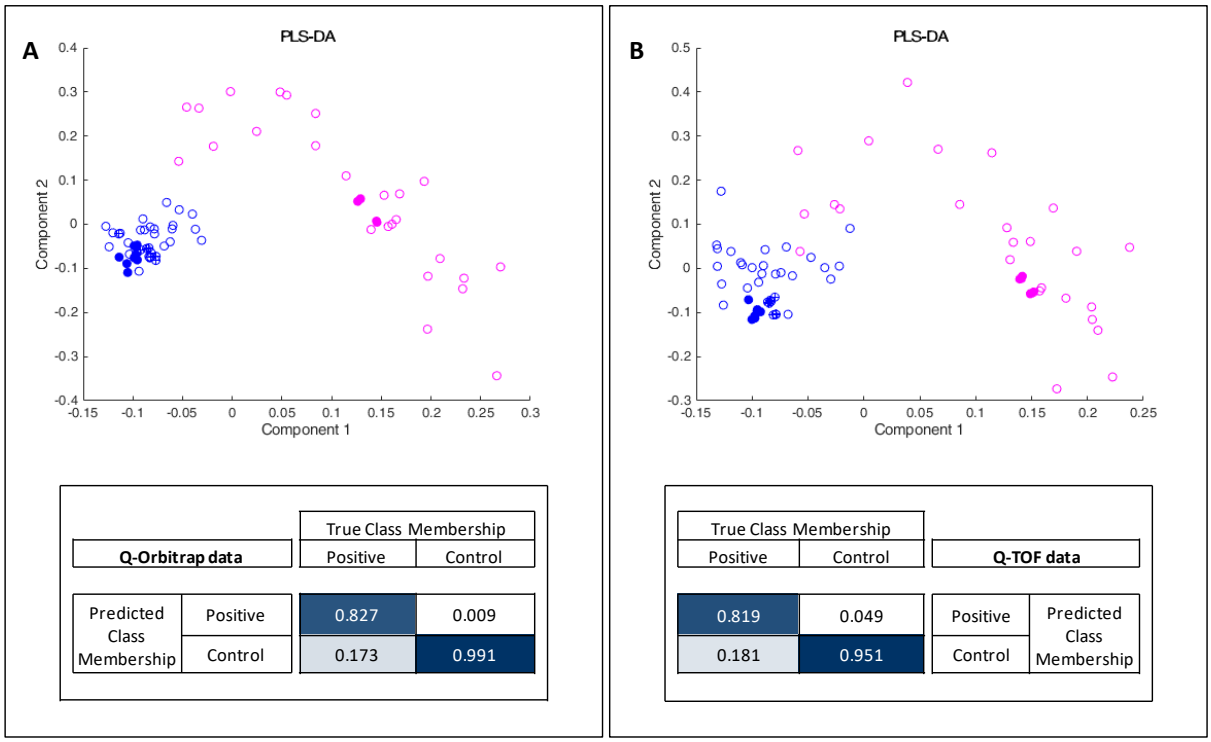


Figure 21. PCA of HRMS data obtained from the examination of *porcine* muscle samples with Q-Orbitrap mass spectrometer (A) and Q-TOF mass spectrometer (B). Top: 3D-scores plot, blue: control samples (*porcine* muscle samples of untreated, healthy animals), magenta: positive samples (*porcine* muscle samples of infected, drug-treated animals); blue filled: control QC samples, blue filled with cross: QC samples spiked with 70 veterinary drugs, magenta filled: positive QC samples; PC: principal component with explained variance shown in brackets. Bottom: confusion matrix of *porcine* muscle samples classified as positive samples or control samples via PCA with subsequent QDA using a Monte Carlo cross-validation (*n* = 5,000; see Figure 1 in [1]).

After data pretreatment, the experimental data of *porcine* muscle samples was subjected to PCA. As illustrated in Figure 21, both PCA models revealed differences between samples from untreated, healthy animals (control) and samples from drug-treated, infected animals (positive samples). Since all QC samples coincided in the plot, sample preparation and data processing workflows were considered robust. Monte Carlo cross-validation was used in order to assess the ability of the PCA coupled with QDA to discriminate between the two classes (control vs. positive samples). The number of misclassifications based on the PCA with subsequent QDA was 19.8% (Q-Orbitrap) and 26.7% (Q-TOF), respectively (see Figure 21). The models reached a high specificity of 0.976 (Q-Orbitrap) and 0.923 (Q-TOF), respectively. The sensitivity of 0.826 (Q-Orbitrap) and 0.810 (Q-TOF) was also satisfactory. Hence, both HRMS systems showed comparable results, although the false-positive rate (falsely classified as infected) of the Q-TOF method was slightly higher. Interestingly, the QC samples spiked with 70 veterinary drugs located in the same cluster as control samples in both models. These results suggest that the main difference between positive and control *porcine*

muscle samples originated from changes in the metabolome of animals rather than from the presence of veterinary drugs.

Furthermore, the experimental data was subjected to PLS-DA and Monte Carlo cross-validation was conducted to check for over-fitting. The number of misclassifications based on the 2D PLS-DA was 18.2% (Q-Orbitrap) and 23.0% (Q-TOF), respectively (see Figure 22). The models reached a high specificity of 0.991 (Q-Orbitrap) and 0.951 (Q-TOF), respectively. The sensitivity was also high with 0.827 (Q-Orbitrap) and 0.819 (Q-TOF), respectively.



Q-Orbitrap data		True Class Membership: Positive	True Class Membership: Control
Predicted Class Membership	Positive	0.827	0.009
	Control	0.173	0.991

True Class Membership: Positive	True Class Membership: Control	Q-TOF data	
0.819	0.049	Positive	Predicted Class Membership
0.181	0.951	Control	

Figure 22. PLS-DA of HRMS data obtained from the examination of *porcine* muscle samples with Q-Orbitrap mass spectrometer (A) and Q-TOF mass spectrometer (B). Top: 2D-scores plot, blue: control samples *porcine* muscle samples of untreated, healthy animals), magenta: positive samples *porcine* muscle samples of infected, drug-treated animals); blue filled: control QC samples, blue filled with cross: QC samples spiked with 70 veterinary drugs, magenta filled: positive QC samples; PC: principal component with explained variance shown in brackets. Bottom: confusion matrix of *porcine* muscle samples classified as positive samples or control samples via 2D PLS-DA using a Monte Carlo cross-validation (n = 5,000).

However, the evaluation of the MSEP indicated that the PLS-DA models over-fitted the data. The MSEP based on the residuals of the PLS-DA regression (Q-Orbitrap data: 0.022; Q-TOF: 0.041) was significantly lower than the MSEP obtained from the cross-validation (Q-Orbitrap data: 0.388; Q-TOF: 0.406). Models which incorporated too many parameters tend to show low deviations from observations that were used to build the model, but higher deviations to

new observations. This is common in the case of a data set which consists of a limited number of samples and a very high number of variables. The data set consisted of 50 samples and several thousand mass spectrometric features (m/z at retention time). Therefore, merely unsupervised models, PCA with subsequent QDA, were used for the performance assessment of sample classification.

4.1.3 Biomarker candidate selection and pathway mapping

In order to further select biomarker candidates for the differentiation of control *porcine* muscle samples from muscle samples taken from drug-treated, infected animals, univariate statistical tests were applied (see Section 3.1.6). The results of the top 25 mass spectrometric features of the Q-Orbitrap data are given in Table 12. All adjusted p-values of the top 25 features were considered highly significant (p-adj. $< 3.9 \times 10^{-11}$) and the mean fold changes ranged from 0.3 for analyte m/z 1004.4665 (@6.13 min) to 404.8 for analyte m/z 552.3288 (@8.87 min). The mass spectrometric data obtained with the Q-Orbitrap method was used for structural elucidation and biomarker annotation due to the higher mass accuracy and reproducibility of MS/MS-spectra. Although, 20 out of the 25 top mass spectrometric features from the Q-Orbitrap data were also highly significant in the Q-TOF data (p-adj. $< 1.1 \times 10^{-6}$, see Supplementary data Table S7).

Table 12. Top 25 significant analytes (*m/z* at retention time in min) responsible for the difference between positive and control *porcine* muscle samples with proposed elemental composition and biomarker annotation (level of identification confidence (ID) according to [94]); positive fold changes indicate higher levels in positive samples compared to control samples (see Table 1 in [1]).

m/z (@min)	Adduct	Proposed elemental composition	δ ppm	ID level	Annotation	*p*-adj.	FC
393.2239 (@8.55)	$[M+Na]^+$	$C_{20}H_{34}O_6$	-2.2	3	similar to 6-keto-prostaglandin $F_{1\alpha}$	5.8×10^{-15}	76.7
320.1599 (@2.09)	$[M+H]^+$	$C_{16}H_{21}N_3O_4$	-3.5	1	PFG	2.1×10^{-14}	3.7
375.2133 (@8.93)	$[M+Na]^+$	$C_{20}H_{32}O_5$	-2.4	3	similar to thromboxane B_2	2.7×10^{-14}	62.3
570.3394 (@6.75)	$[M+H]^+$	$C_{26}H_{52}NO_{10}P$	-2.3	3	oxidation product of lyso-PC	4.0×10^{-14}	37.5
348.2737 (@8.81)	$[M+NH_4]^+$	$C_{18}H_{34}O_5$	-2.2	3	similar to trihydroxy-derivative of octadecenoic acid	4.0×10^{-14}	12.0
404.2637 (@7.10)	$[M+H]^+$	$C_{16}H_{33}N_7O_5$	4.4	4		5.2×10^{-14}	68.1
386.2529 (@8.23)	$[M+H]^+$	$C_{16}H_{31}N_7O_4$	4.4	4		1.4×10^{-13}	76.3
480.1642 (@1.19)	$[M+H]^+$	$C_{19}H_{25}N_7O_6S$	-3.5	4		1.4×10^{-13}	56.3
335.2209 (@8.54)	$[M+H]^+$	$C_{20}H_{30}O_4$	-2.5	4		2.4×10^{-13}	97.7
552.3286 (@7.64)	$[M+H]^+$	$C_{26}H_{50}NO_9P$	-1.8	3	similar to hydroperoxy derivative of lyso-PC C18:2	2.4×10^{-13}	66.6
375.2136 (@11.14)	$[M+Na]^+$	$C_{20}H_{32}O_5$	0.8	3	similar to thromboxane B_2	4.9×10^{-13}	117.2
391.2083 (@8.18)	$[M+H]^+$	$C_{18}H_{26}N_6O_4$	-1.7	4		9.9×10^{-13}	17.4
372.2738 (@9.64)	$[M+H]^+$	$C_{20}H_{37}NO_5$	-2.1	4		2.2×10^{-12}	41.6
534.3184 (@10.20)	$[M+H]^+$	$C_{26}H_{48}NO_8P$	-2.5	3	similar to oxo derivative of lyso-PC C18:2	2.2×10^{-12}	109.9
506.1800 (@3.94)	$[M+H]^+$	$C_{21}H_{27}N_7O_6S$	-3.3	4		2.2×10^{-12}	36.5
388.2686 (@8.19)	$[M+H]^+$	$C_{16}H_{33}N_7O_4$	4.5	4		2.5×10^{-12}	163.6
173.117 (@8.87)	$[M+H]^+$	$C_9H_{16}O_3$	-1.6	4		2.7×10^{-12}	22.1
170.081 (@1.57)	$[M+H]^+$	$C_8H_{11}NO_3$	-1.1	4		2.9×10^{-12}	14.4
204.1050 (@4.95)	$[M+H]^+$	$C_9H_{17}NO_2S$	-1.4	3	similar to L-Cystein	2.9×10^{-12}	13.1
536.3337 (@9.99)	$[M+H]^+$	$C_{26}H_{50}NO_8P$	-1.6	3	similar to lyso-PC derivative	5.0×10^{-12}	58.2
301.2157 (@9.63)	$[M+H]^+$	$C_{20}H_{28}O_2$	-2.0	4		5.0×10^{-12}	61.8
552.3288 (@8.87)	$[M+H]^+$	$C_{26}H_{50}NO_9P$	-1.7	3	similar to hydroperoxy derivative of lyso-PC C18:2	1.4×10^{-11}	404.8
394.2275 (@8.55)	$[M+Na]^+$	$C_{20}H_{34}O_6$	-2.2	3	^{13}C isotope of 393.2239 (@8.55)	2.0×10^{-11}	6.8
1004.4665 (@6.13)	$[M+H]^+$	$C_{47}H_{71}N_7O_{13}P_2$	0.5	4		2.2×10^{-11}	0.3
335.2209 (@9.69)	$[M+H]^+$	$C_{20}H_{30}O_4$	-2.5	3	similar to 15(S)-HpEPE	3.9×10^{-11}	21.6

FC: fold change, *p*-adj.: multiple test corrected *p*-value based on the Benjamini Hochberg false discovery rate, PFG: prolylphenylalanylglycine, lyso-PC: lysophosphatidylcholine, 15(S)-HpEPE: 15S-hydroperoxy-5Z,8Z,10E,14Z,17Z-eicosapentaenoic acid.

Metabolite annotation is time-consuming and challenging. It is the bottleneck of biomarker discovery in data-driven, non-targeted analysis. Therefore, the levels of identification confidence as described by Schymanski *et al.* [94] were used as means to communicate the confidence of biomarker annotation. It was possible to verify the identity of analyte *m/z* 320.1599 (@2.09 min) as tripeptide prolylphenylalanylglycine (PFG) with a reference standard (level 1, see Figure 23).

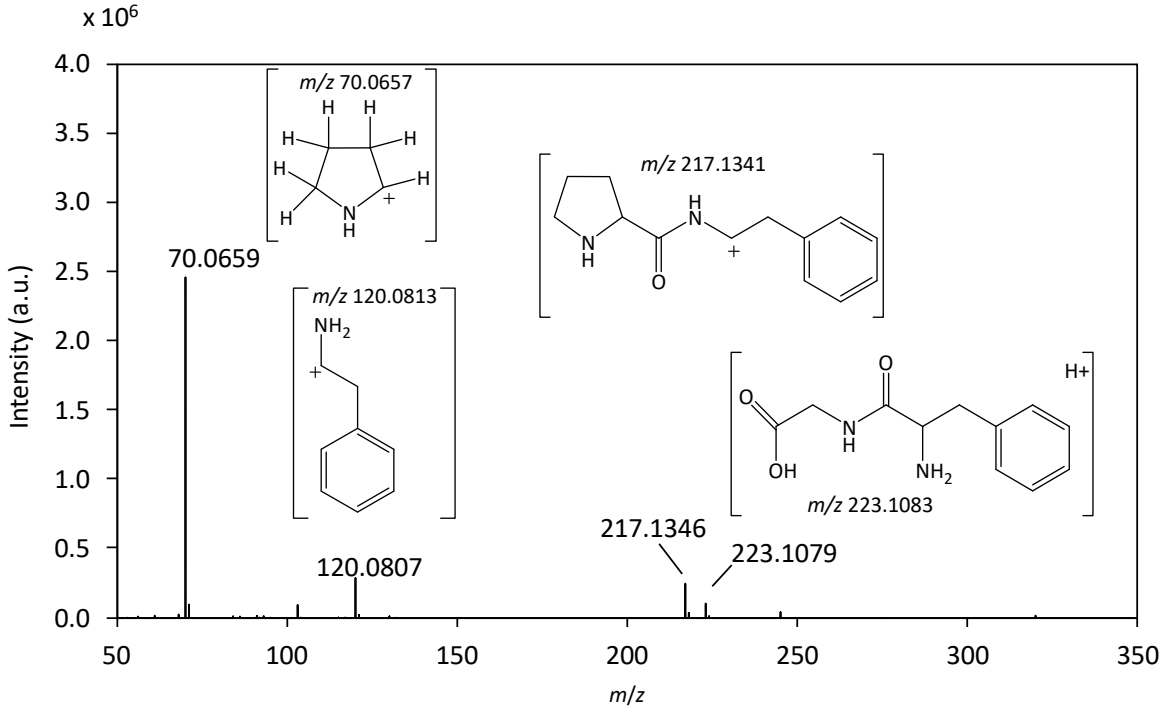


Figure 23. Representative product ion spectrum of tripeptide PFG (*m/z* 320.1599 $[M+H]^+$ @ 2.09 min). Characteristic fragments were confirmed via reference standard; *m/z* ratio next to structure represents exact calculated mass compared to measured accurate mass of fragments (see Figure 2 (top) in [1]).

Furthermore, the analytes *m/z* 570.3394 (@6.75 min), 552.3286 (@7.64 min), 534.3184 (@10.20 min), 536.3337 (@9.99 min) and 552.3288 (@8.87 min) were tentatively identified as lysophosphatidylcholine (lyso-PC) derivatives (level 3), since their product ion scans showed the characteristic fragments *m/z* 184.0736 and *m/z* 104.1073 [161]. These characteristic fragments were confirmed with the reference standard lyso-PC 17:0. Of particular note is the loss of three H_2O molecules from the molecular ion in the product ion scan of *m/z* 570.3394 (@6.75 min). A similar MS/MS-pattern has most recently been described for hydroperoxides and hydroxides of lipid mediators derived from ω-3 eicosapentaenoic acids [162]. Thus, it is proposed that the analyte *m/z* 570.3394 (@6.75 min) is a trihydroxide of lyso-PC C18:1 (level 3, see Figure 24). A proposal for the fragmentation mechanism is given in Figure 25. This analyte will be referred to as "lyso-PC derivative 570" hereafter.

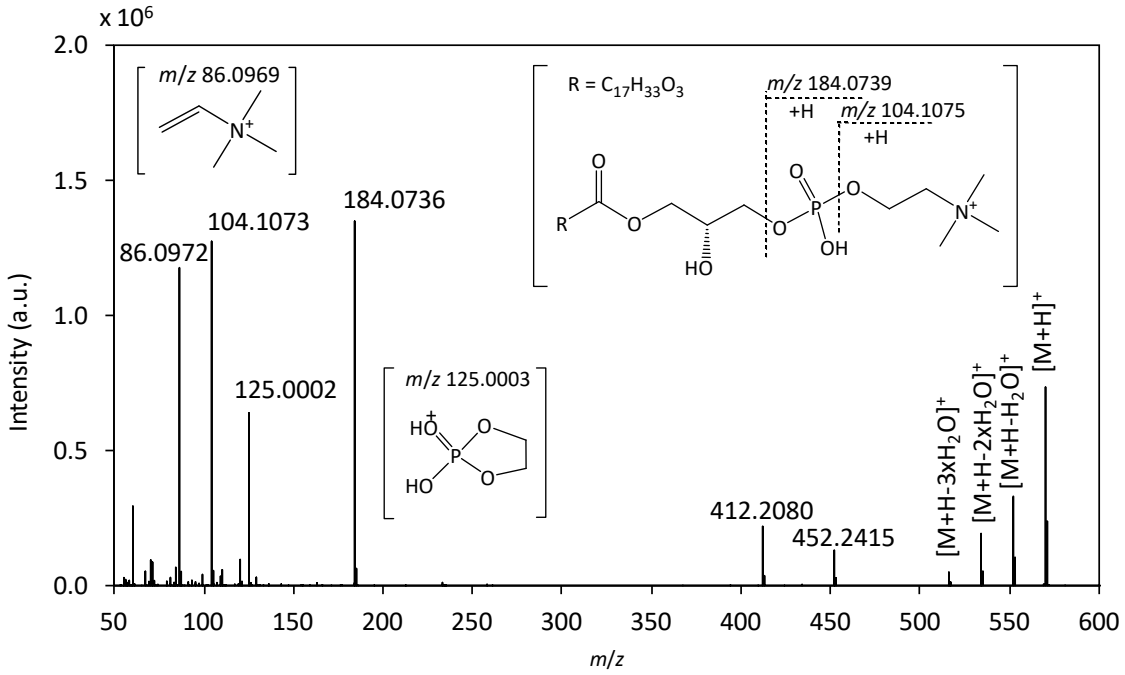


Figure 24. Representative product ion spectrum of lyso-PC derivative 570 (*m/z* 570.3394 $[M+H]^+$ @ 6.75 min). *M/z* 86, 184, 125 and 104 are characteristic fragments for lysophosphatidylcholines (lyso-PC) and were confirmed via reference standard; *m/z* ratio next to structure represents exact calculated mass compared to measured accurate mass of fragments (see Figure 2 (bottom) in [1]).

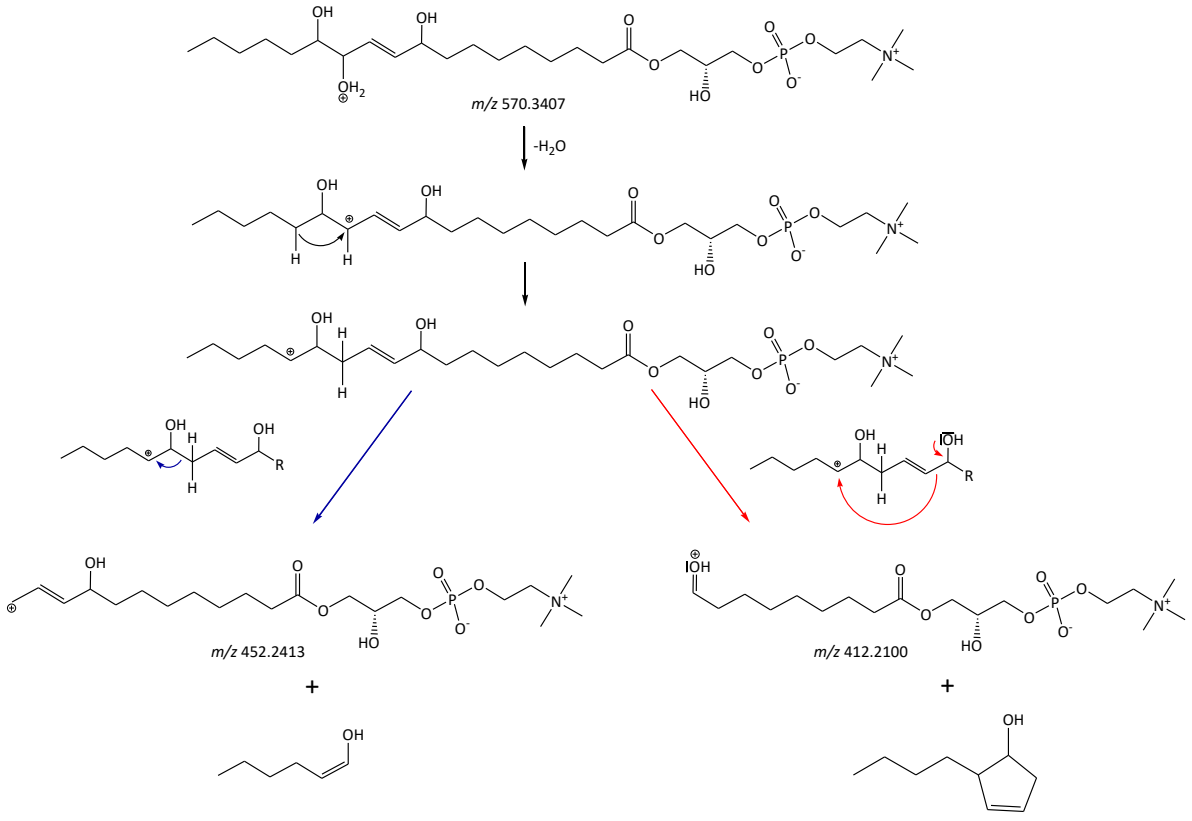


Figure 25. Proposal of fragmentation mechanism of tentatively identified trihydroxy derivative of lyso-PC (C18:1), referred to as "lyso-PC derivative 570" (see Supporting information Figure S3 in [1]).

Database hits suggested that the analytes *m/z* 393.2239 (@8.55 min), 375.2133 (@8.93 min), 375.2136 (@11.14 min) and 394.2275 (@8.55 min; ^{13}C isotope of m/z 393.2239) belong to the substance class of prostanoids (level 3). The product ion spectra of the analytes *m/z* 348.2737 (@8.81 min) and 335.2209 (@9.69 min) showed similarities to oxidation products of unsaturated fatty acids (level 3). For some analytes, only level 4 of identification was achieved, i.e., proposal of chemical composition based on isotope pattern.

Additionally an "MS peak to pathway" mapping was performed in MetaboAnalyst 4.0 to complement manual biomarker identification. This procedure applies a computational algorithm called *mummichog* to map mass spectrometric features to functional activities by leveraging the collective insights of metabolic pathways [144]. 1523 significant mass spectrometric features (*p*-value from *t*-test < 1.0×10^{-5}) were used for pathway analysis and 105 empirical compounds were annotated by "MS peaks to pathway". The significant hits per pathway were checked manually for plausibility (chromatographic retention and adduct identification). The pathway enrichment factor is calculated as the ratio of the number of significant pathway hits and the expected number of randomly matched compound hits [134]. The significance of the pathway enrichment is determined by the *p*-value of Fisher's exact test (FET) [144]. In order to take the size of each pathway into account, an adjusted *p*-value per pathway is calculated based on permutation-testing [144]. Prostaglandin formation from arachidonate (FET *p*-value: 0.01, adjusted *p*-value: 0.04, enrichment factor: 1.76) and arachidonic acid metabolism (FET *p*-value: 0.04, adjusted *p*-value: 0.04, enrichment factor: 1.95) followed by linoleate metabolism (FET *p*-value: 0.10, adjusted *p*-value: 0.04, enrichment factor: 1.50) were the three most stimulated pathways differentiating positive samples from control samples (see Figure 26). These results were conclusive with the manually performed biomarker annotations. However, it is to be kept in mind that isobaric mass spectrometric features cannot be distinguished by this method. Hence, this tool was mainly used for an unbiased identification of patterns, i.e., pathway enrichments in order to complement manual biomarker identification.

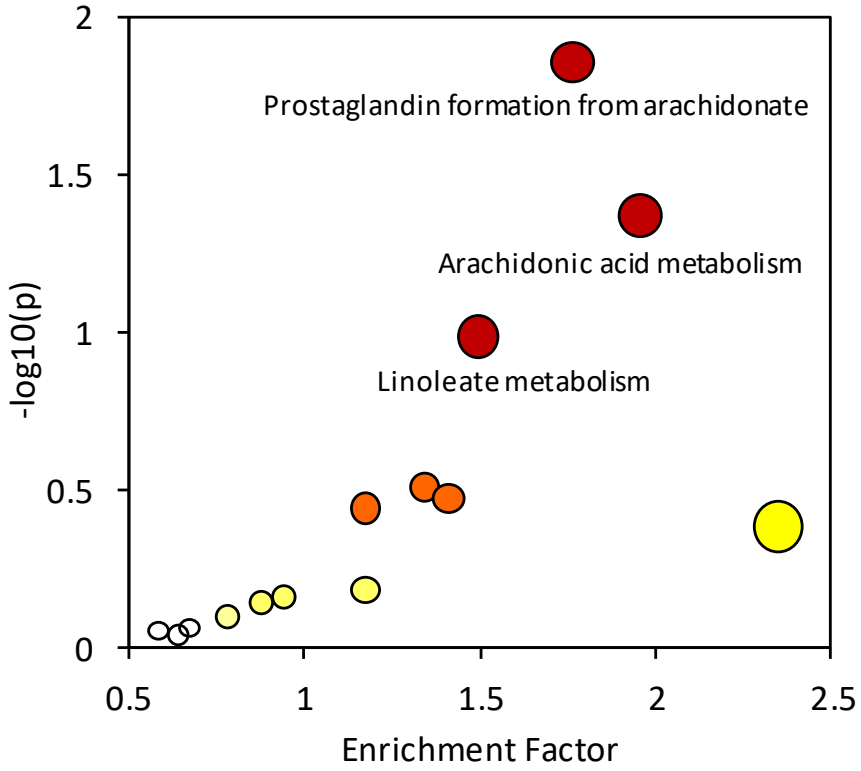


Figure 26. Pathway mapping summary: all matched pathways are displayed as circles. The color and size of each circle corresponds to its Fisher's exact test (FET) *p*-value and enrichment factor, respectively. Small white circle represent non-significant hits, while larger, red circles represent significant pathway enrichments (see Figure 3 in [1]).

Prostanoids, i.e., prostaglandins and thromboxane A_2, are lipid mediators derived from arachidonic acid. Their biosynthesis is catalyzed by cyclooxgenase (COX) isoenzymes (COX-1 and COX-2) which are targets for medication by non-steroidal anti-inflammatory drugs (NSAIDs) [163]. Prostaglandins play a key role in inflammatory response with partly counteracting effects [163]. For example, prostaglandin E_2, one of the most abundant prostaglandins *in vivo* is involved in typical signs of infection, such as swelling, redness and pain [164]. Thromboxane A_2 increases platelet-aggregation and activates endothelial inflammatory responses [163]. Increased concentrations of prostaglandins in inflamed tissues have been described [165]. Stimulated linoleic acid metabolism is reasonable, since elevated levels of oxidation products of linoleic acid have recently been reported during inflammation processes due to oxidative stress [166]. Lysophosphatidylcholines (lyso-PCs) are lipid mediators, which are formed by cleavage of phosphatidylcholine via phospholipase A_2 or by the activity of lecithin-cholesterol acyltransferase [167]. Lyso-PCs exert pro-inflammatory effects and are positively associated with inflammatory diseases, e.g., atherosclerosis [168]. The application

of lyso-PCs as biomarkers for different pathophysiological changes has been described [167]. Oxidation products similar to the tentative proposals, e.g., monohydroxyperoxy derivatives, trihydroxides and oxo derivatives have been found in human plasma [169].

To the best of my knowledge, the tripeptide PFG has not yet been described as biomarker for inflammation or infection. However, in the present study elevated levels of PFG were determined reproducibly in *porcine* muscle tissue from infected animals and PFG was unequivocally identified using a reference standard. Furthermore, the biological importance of tripeptides has been recently summarized [170]. The review of Ung *et al.* shows that endogenous tripeptides as well as contiguous tripeptide sequences which are incorporated in larger proteins are capable to function as efficient ligands for protein receptors. The review also gives examples for tripeptides which exhibit anti-inflammatory activity [170]. Recently, extensive research into the development of anti-inflammatory drugs based on tripeptide motifs has been reported [170].

The tripeptide PFG and lyso-PC derivative 570 were used as specific markers in the further investigations. The rationale behind this decision was that both analytes were among the top five most significant analytes and chromatographic baseline separation was achieved. Furthermore, the identity of both analytes at least at substance class level was verified.

4.1.4 Bivariate biomarker model

In a first step, ROC curves were constructed for the two chosen biomarker candidates based on the log-transformed, standardized data used for PCA in order to evaluate their performance as biomarkers for the differentiation of *porcine* muscle samples from infected, drug-treated animals and healthy, untreated animals. The area under the ROC curve (AUROC) of lyso-PC derivative 570 and tripeptide PFG was 0.9472 and 0.9552, respectively. This indicated an excellent discrimination between the two test groups based on the selected biomarker candidates.

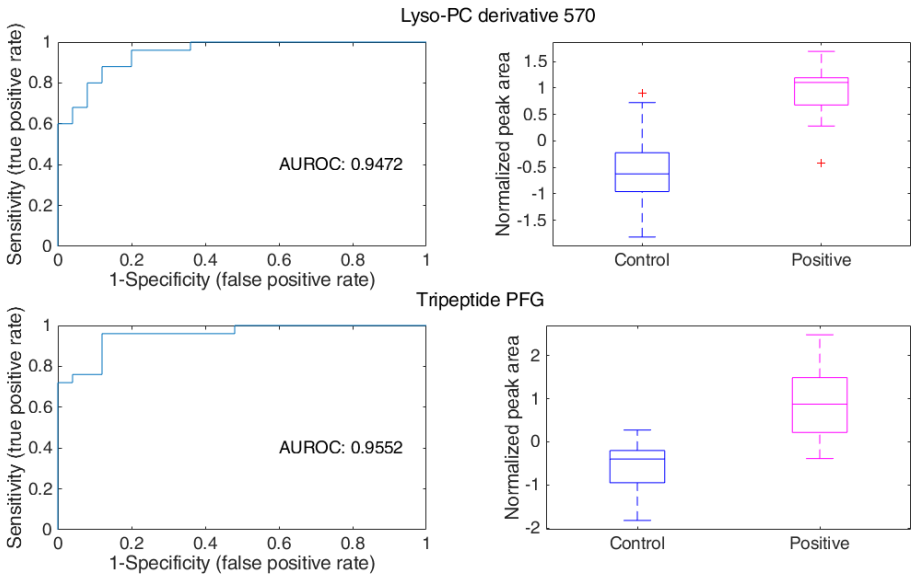


Figure 27. Receiver operating characteristic (ROC) curves (left) and boxplots (right) of normalized, log-transformed and standardized peak areas of biomarker candidates lyso-PC derivative 570 (top) and tripeptide PFG (bottom) in control and positive *porcine* muscle samples; AUROC: area under the curve; boxplot: the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers (+).

AUROC values close to 0.5 indicate that the discrimination between the test groups is simply by chance. In contrast, an AUROC value of 1 indicates that a biomarker enables perfect separation between the two test groups [145]. The sensitivity at a fixed specificity of 1.0 (i.e., no false positives allowed) was 0.60 for lyso-PC derivative 570 and 0.72 for tripeptide PFG. The specificity at a sensitivity of 1.0 (i.e., no false negatives allowed) was 0.64 for lyso-PC derivative 570 and 0.52 for tripeptide PFG, respectively. These results indicate that both biomarkers were suitable to distinguish between positive and control *porcine* muscle samples in the underlying data set. The drawback of this evaluation strategy is that it is hard to incorporate new samples in this model, since in LC-HRMS analysis the intensity of peak areas in differ-

ent analytical batches can vary based on the instrument status, i.e., instrument related drifts over time. However, the aim of this study was to identify biomarker candidates and develop a model which can be used in routine veterinary drug analysis to determine whether a residue is caused by medication or contamination. Hence, if one would want to use the described procedure in order to classify a sample in question as originated from a drug-treated, infected animal or a healthy one, one would have to reanalyze the data set of positive and control *porcine* muscle samples to construct a model which is generated with data measured in a short time period. This is not feasible for routine veterinary drug analysis.

Since a correlation between the two biomarkers was observed in healthy animals, a bivariate data model for the control group was developed. The peak areas of the two biomarker candidates were divided by the peak areas of the internal standard oxolinic acid-d5. This referencing procedure led to run-independent values, which can be compared to values from other measurements.

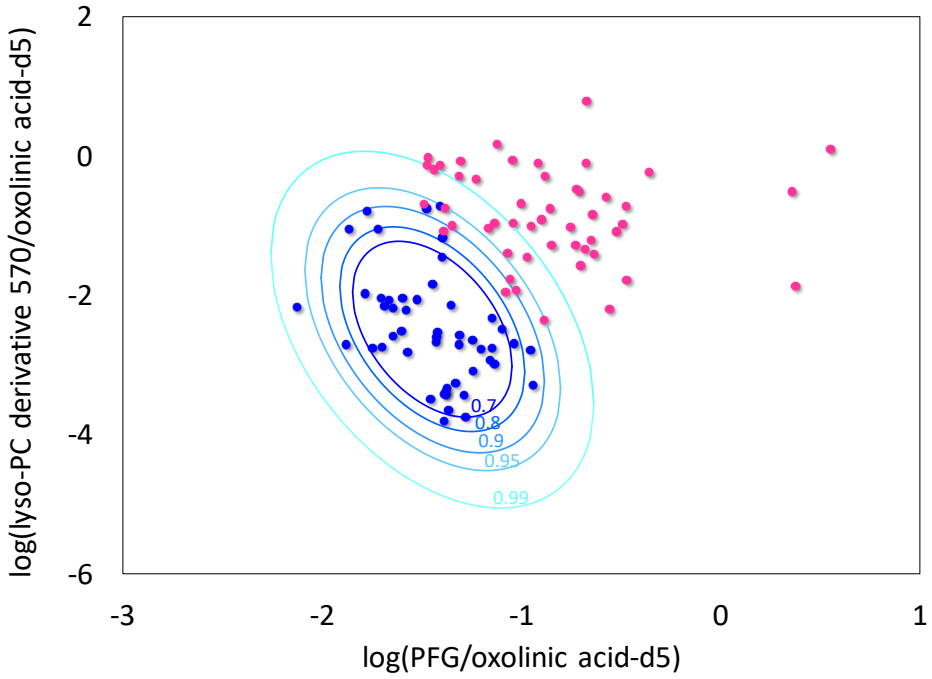


Figure 28. Bivariate data analysis of the log-transformed response ratios of the two biomarker candidates tripeptide PFG and lyso-PC derivative 570 (internal standard: oxolinic acid-d5). Control *porcine* muscle samples from healthy animals (blue) with prediction ellipses and positive *porcine* muscle samples from drug-treated, infected animals (magenta); blue shaded lines illustrate prediction ellipses of control samples constructed with α error probability of 0.3–0.01.

Hence, 50 positive and 50 control *porcine* muscle samples which were analyzed during the course of one year in several analytical batches were retrospectively evaluated for the two biomarker candidates. The ratios were transformed by taking their natural logarithm and visualized in a scatter plot. As shown in Figure 28, the log-transformed values of control samples appear to follow a bivariate normal distribution. In contrast, this cannot be assumed for positive samples, which is not surprising, since this group consisted of a heterogeneous muscle sample population of pigs having different diseases. For practical purposes, it is necessary to apply a statistically defined criterion by which one can decide whether a *porcine* muscle sample can be classified as taken from a drug-treated, infected animal or from a healthy one. In this case it is reasonable to use a prediction ellipse calculated from the control samples as decision border. A prediction ellipse for a single future observation is an ellipse that will, with a specified degree of confidence, enclose a region that contains the next (or some other prespecified) randomly selected observation of a bivariate population [124]. Thus, a 0.9-prediction ellipse accepting an α error probability of 0.1 (specificity of 0.9) was employed. As shown in Figure 28, only 2 out of 50 positive samples were classified as false negative applying this approach, which is equivalent to a sensitivity of 0.96.

The referencing procedure with the internal standard oxolinic acid-d5 was applied, since it was not possible to select *a priori* an optimal internal standard, because the outcome of the investigations was unknown. In the routine method 9 internal standards are included into the analyses. Oxolinic acid-d5 was chosen because it was the internal standard with the lowest difference in retention time to lyso-PC derivative 570. Thus, one can anticipate similar chromatographic properties to lyso-PC derivative 570, keeping in mind that sample preparation also includes a chromatographic method, i.e., a SPE step. This procedure was transferred to the tripeptide PFG because of the lack of a suitable alternative internal standard. One additional aim was to retrospectively evaluate data for further validation purposes. This would not have been possible by including another internal standard. However, the effect of using internal standards was investigated on the overall performance of the bivariate model and it was found that a clear separation between the test groups was observed, even without using any internal standards (see Figure 29). However, especially in LC-MS, it is important to perform a referencing procedure in order to compare different measurement series.

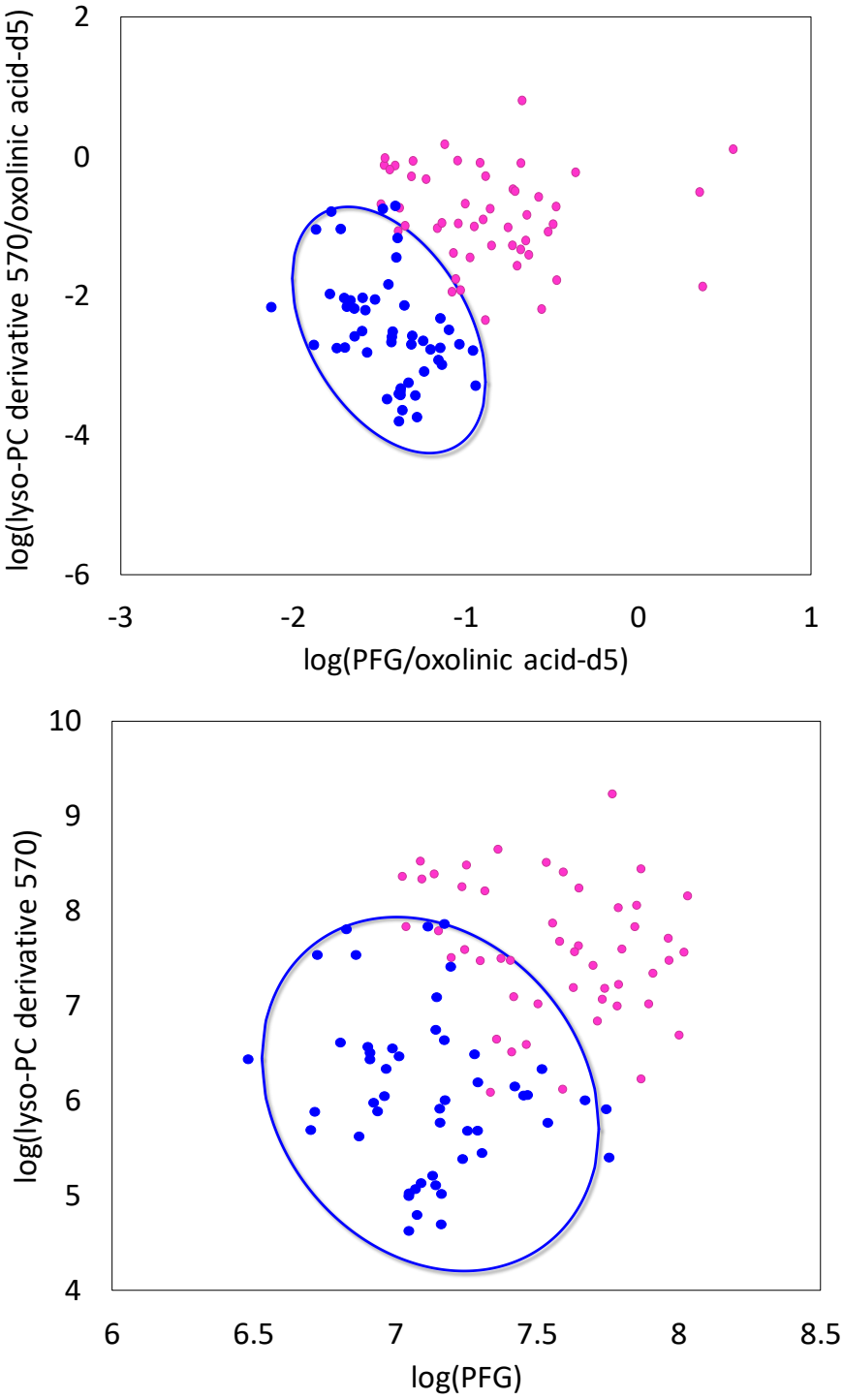


Figure 29. Bivariate data analysis of the two biomarker candidates tripeptide PFG and lyso-PC derivative 570; control *porcine* muscle samples from healthy animals (blue) with 0.9-prediction ellipse, positive *porcine* muscle samples from drug-treated, infected animals (magenta); top: data was log-transformed and referenced to internal standard oxolinic acid-d5; bottom: data was merely log-transformed.

In order to illustrate the practical value of this sophisticated approach, two samples unambiguously contaminated during sampling and transport were analyzed as described above. The areas of PFG and lyso-PC derivative 570 were divided by the area of the internal standard oxolinic acid-d5 and after log-transformation the results were plotted (see Figure 30).

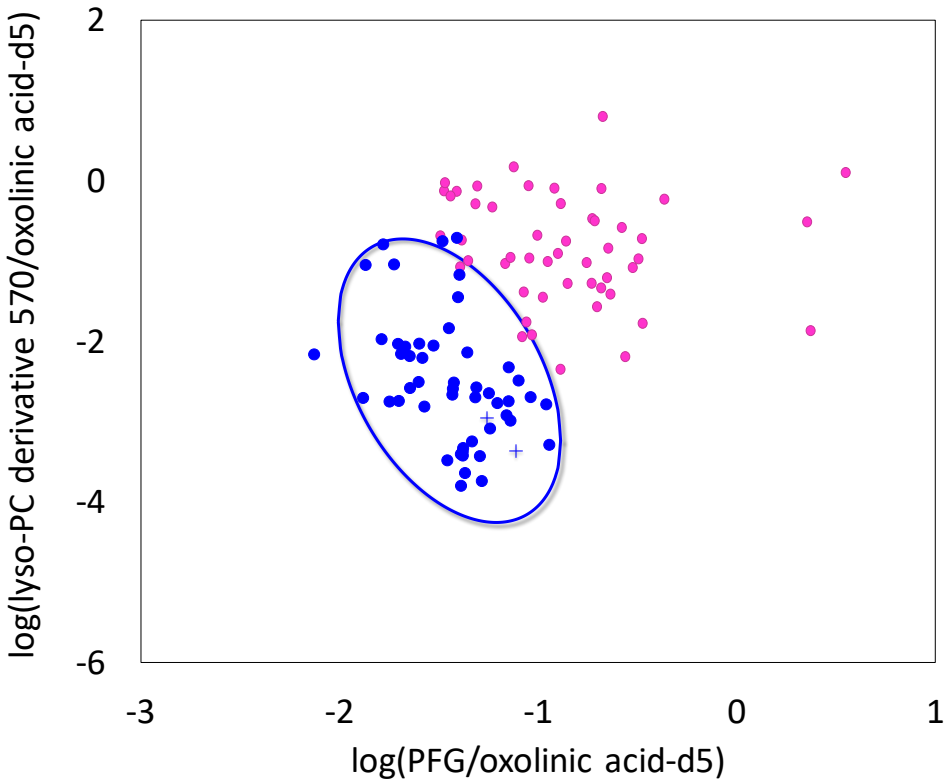


Figure 30. Bivariate data analysis of the log-transformed response ratios of the two biomarker candidates tripeptide PFG and lyso-PC derivative 570 (internal standard: oxolinic acid-d5). Control *porcine* muscle samples from healthy animals (blue) with 0.9-prediction ellipse and positive *porcine* muscle samples from drug-treated, infected animals (magenta); p*orcine* muscle samples known to be contaminated with residues of veterinary drugs (blue cross) illustrate the practical relevance of 0.9-prediction ellipse in routine analysis (see Figure 4 in [1]).

Both points lie within the 0.9-prediction ellipse of the control samples. Hence, they were classified as samples taken from healthy animals but contaminated with veterinary drugs. If a *porcine* muscle sample with veterinary drug residues is classified as negative (taken from a healthy animal), further investigations at the farm level are needed. Possible reasons for this observation could be a contamination during sampling, transport or analysis or a cross-contamination based on contaminated feed. The latter can only be excluded by the absence of metabolites. However, it is to be kept in mind that many metabolites of veterinary drugs are not detectable in muscle tissue. Nonetheless, this newly developed approach can be of great benefit in routine laboratory practice, since a non-targeted LC-HRMS setting offers the ad-

vantage, that these two biomarkers can easily be analyzed retrospectively. Hence, this procedure can ensure to rule out a suspected cross-contamination of a sample under investigation.

4.1.5 Further investigations

The main focus of this project was the multivariate statistical analysis of *porcine* muscle samples and the identification of biomarker candidates in order to discriminate samples from drug-treated, infected pigs from untreated, healthy ones. The rationale behind the decision to focus on *porcine* muscle samples was based on the fact that in Germany the commercial slaughtering numbers of pigs are higher compared to cattle [159]. Thus, the number of official samples for veterinary drug residue control of *porcine* samples is higher than of *bovine* samples. Furthermore, in kidney samples a clustering trend according to the storage time of samples was observed in control samples. However, multivariate statistical analysis was performed for both species (pig and cattle) in both sample types in order to test the methodological approach. The results are shown in the following further investigations.

***Bovine* muscle samples**

Data integrity

After data processing using Mzmine 2, a data integrity check was performed. In the course of one analytical batch 4 spiked QC samples were measured. The QC samples were scattered throughout the batch in order to check for instrument related drifts. Eigthy-three percent of the spiked analytes were detected in the Q-Orbitrap data as well as in the Q-TOF data. The RSDs of the peak areas were high for the majority of the spiked analytes (on average about 30%, see Supplementary data Table S8) on both instruments. A closer look at the raw data showed an intensity loss over time, i.e., an instrument related drift. Therefore, normalization by the sum of all peaks for each sample was performed in order to evaluate the suitability of the data after this data pretreatment step. Normalization by the sum of all peaks is a common way to account for intensity loss over the course of the analysis and it is part of the data pretreatment prior to PCA. After normalization the RSDs of the peak areas were acceptable with a mean of 20% (Q-Orbitrap) and 11% (Q-TOF, see Supplementary data Table S8), respectively. Twelve analytes were not detected with the chosen data processing method for both data sets. A check of the raw data showed that the peaks were low in intensity and

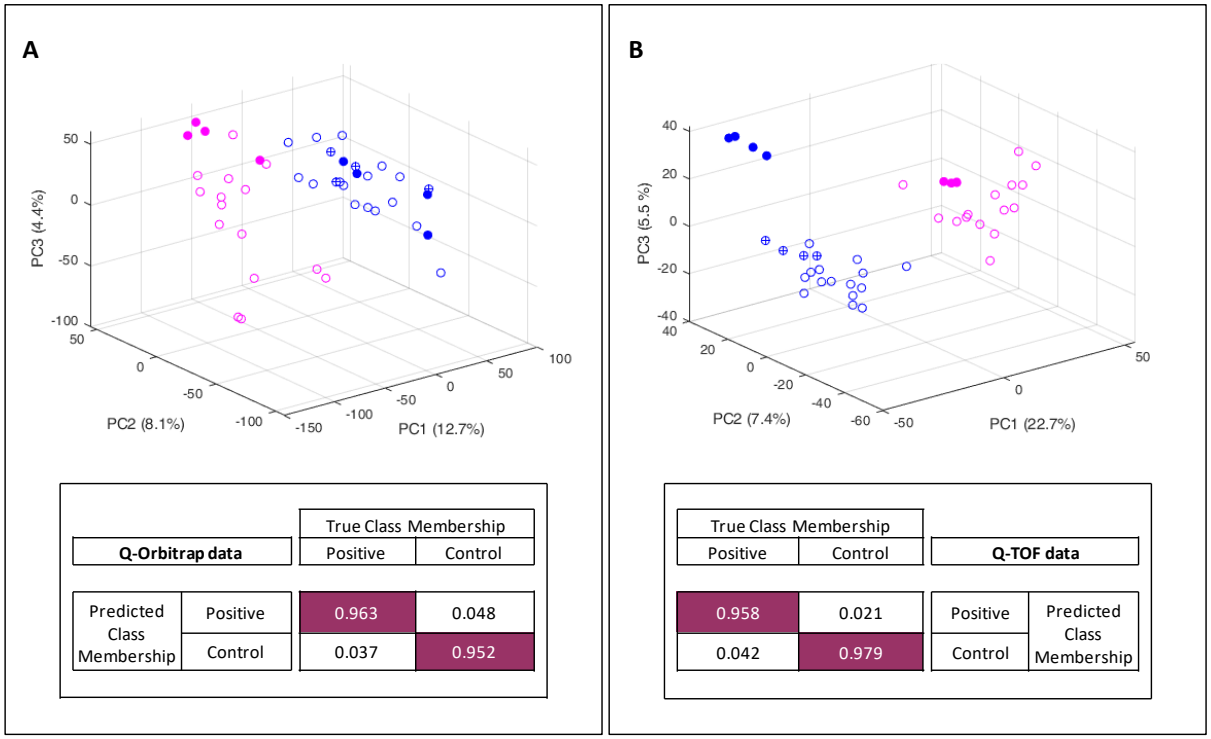
showed a narrow peak shape. Subsequently, the number of data points across the peak was too low for the chosen data processing method.

Multivariate statistical analysis

After normalization, transformation and scaling, the experimental data of *bovine* muscle samples was subjected to PCA. As illustrated in Figure 31 differences between the two test groups (drug-treated, infected cattle and untreated, healthy cattle) were observed. In the case of the Q-Orbitrap data, the main variance contributing to the separation was observed in PC1, which accounted for 12.7% of the total variance in the data set. In evaluation of the QC samples, a drift during the course of the analysis was noted. However, the drift was mainly observed in PC2 and did not impede the discrimination of the two groups. The observations in the Q-TOF data were similar. The main variance contributing to the separation was observed in PC1, which accounted for 22.2% of the variance in the data set. In evaluation of the QC samples, a slight drift in PC2 was observed. However, this analytical drift was less pronounced than in the analysis using the LC-Q-Orbitrap method. Hence, sample analysis using the LC-Q-TOF and the data processing workflow was considered to be robust.

In the analysis of both HRMS instruments, the spiked QC samples from untreated, healthy animals were located in the same cluster as the negative samples from untreated animals. This observation indicated that the main difference in the data did not originate from the presence of antibiotic residues but from a change in the metabolism in infected animals, similar to the observations made in *porcine* muscle samples (see Section 4.1.2).

In order to assess the performance of the PCA with subsequent QDA, Monte Carlo simulations were conducted. The results are shown in Figure 31 as a confusion matrix. Both models showed a high specificity of 0.952 (Q-Orbitrap) and 0.979 (Q-TOF), respectively. Thus, in both cases the data analysis procedure was suitable for the correct identification of control samples as samples taken from untreated, healthy cattle. The sensitivity of both methods was also very high with 0.963 (Q-Orbitrap) and 0.958 (Q-TOF), respectively. The number of misclassifications was 8.5% (Q-Orbitrap) and 6.3% (Q-TOF), respectively. These data indicate that in both cases PCA with subsequent QDA performed well in distinguishing samples from the two test groups.



Q-Orbitrap data		True Class Membership: Positive	True Class Membership: Control
Predicted Class Membership	Positive	0.963	0.048
	Control	0.037	0.952

True Class Membership: Positive	True Class Membership: Control	Q-TOF data	
0.958	0.021	Positive	Predicted Class Membership
0.042	0.979	Control	

Figure 31. PCA of HRMS data obtained from the examination of *bovine* muscle samples with Q-Orbitrap mass spectrometer (A) and Q-TOF mass spectrometer (B). Top: 3D-scores plot, blue: control samples (*bovine* muscle samples of untreated, healthy animals), magenta: positive samples (*bovine* muscle samples of infected, drug-treated animals), blue filled: control QC samples, blue filled with cross: QC samples spiked with 70 veterinary drugs, magenta filled: positive QC samples; PC: principal component with explained variance shown in brackets. Bottom: confusion matrix of *bovine* muscle samples classified as positive samples or control samples via PCA with subsequent QDA using a Monte Carlo cross-validation (n = 5,000).

Furthermore, the experimental data was subjected to PLS-DA and Monte Carlo cross-validation was conducted to check for the performance and quality of the PLS-DA. The performance was very good, but the MSEP of the cross-validation indicated that the PLS-DA models over-fitted the data (data not shown). The same observation was made for the *porcine* muscle data. In the case of the *bovine* data set, the sample number was even smaller with a total of 30 samples. Hence, merely unsupervised methods were used for the evaluation of sample classification performance.

Porcine kidney samples

Data integrity

In the course of two analytical batches 8 spiked QC samples were measured and used for a data integrity check after data processing using Mzmine 2. The QC samples were scattered throughout the batch in order to check for instrument drifts. Eighty percent (Q-Orbitrap) and 87% (Q-TOF) of the spiked analytes were found, respectively (see Supplementary data Table S9). The RSDs of the peak areas were mostly below 20% in the Q-Orbitrap data. The RSDs of the Q-TOF data was generally higher than the RSDs of the Q-Orbitrap data, but still mostly below 30%, which was considered satisfactory. Fourteen analytes (Q-Orbitrap) and 9 analytes (Q-TOF) were not detected with the chosen data processing method. A check of the raw data showed that the peak areas of these analytes were below the chosen minimum peak height. This is in line with the general observation that ion suppression was high in extracts from kidney tissue leading to lower peak intensities.

Multivariate statistical analysis

After normalization, transformation and scaling, the experimental data of *porcine* kidney samples was subjected to PCA. The two scores plots (Figure 32, top) illustrate that the PCA did not reveal clear differences between samples from untreated, healthy animals and samples from infected, drug-treated animals, since the distribution of the scores of the two groups overlapped considerably. However, the expectation vectors of the two test groups appear to deviate from each other. In evaluation of the QC samples, an analytical batch effect was observed, i.e., QC samples from different analytical batches did not coincide. This indicates that the applied data normalization process was still susceptible to variations based on the instrument performance or variations based on the sample preparation across analytical batches. In order to compensate for batch effects, a normalization process was performed using the QC samples which were measured regularly within each batch (Figure 32, bottom). This normalization effectively compensated the batch effect in the data set. The variation explained by PC1 and PC2 was reduced by the normalization procedure, since part of the variation explained by these two principal components was due to fluctuations based on the batch effect. This normalization step was implemented in the workflow of data processing for every multivariate analysis of the experimental data of *porcine* kidney samples. Yet, the outcome of the PCA was not significantly improved. A distinct class separation was not observed; merely a deviation in the expectation vectors was observed. It is noticeable, however, that the spiked QC samples coincided with the control samples.

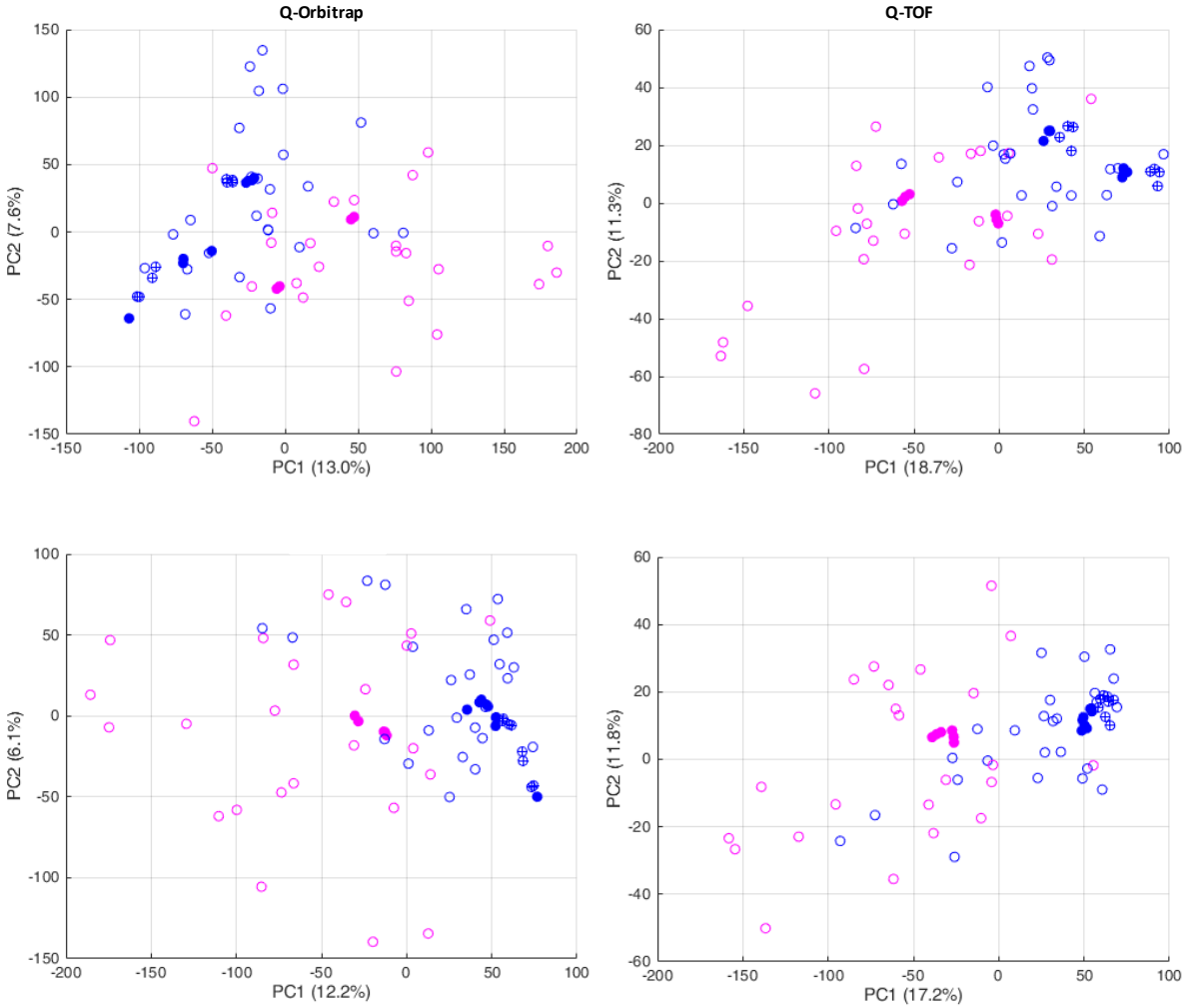


Figure 32. 2D-scores plot of PCA of LC-HRMS (left: Q-Orbitrap, right: Q-TOF) data of *porcine* kidney samples; blue: control samples (*porcine* kidney samples of untreated animals), magenta: positive samples (*porcine* kidney samples of infected, drug-treated animals); blue filled: negative QC samples, blue filled with cross: spiked QC samples, magenta filled: positive QC samples; PC: principal component with explained variance shown in brackets. Top: data prior QC based normalization; bottom: data after QC based normalization.

Monte Carlo cross-validation was applied in order to assess the ability of the PCA coupled with QDA to discriminate between the two test groups (control samples vs. positive samples). The results are shown in Figure 33. The models show a specificity of 0.903 for the Q-Orbitrap data and 0.884 for the Q-TOF data, respectively. Thus, in both cases the data analysis procedure was suitable for the correct identification of control samples as samples taken from untreated, healthy pigs with a false positive rate of maximum 11.6%. However, the sensitivity of the models was unsatisfactory with 0.572 (Q-Orbitrap data) and 0.561 (Q-TOF data), respectively. Hence, kidney samples were not a suitable matrix in order to determine if the sample was taken from an infected, drug-treated pig.

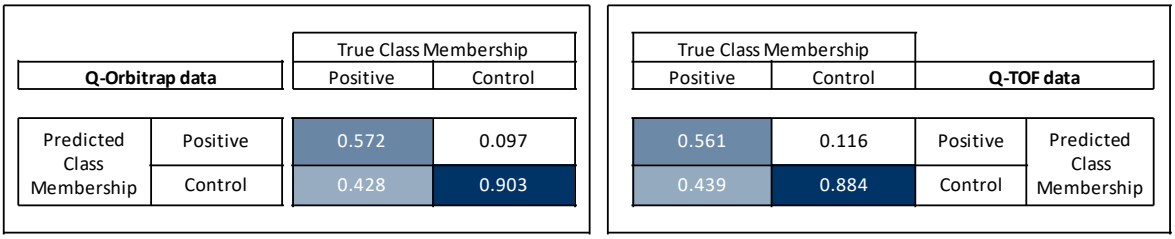


Figure 33. Confusion matrix of *porcine* kidney samples classified as positive samples (samples of infected, drug-treated animals) or control samples (samples of untreated animals) via 3D-PCA with subsequent QDA using a Monte Carlo cross-validation (n = 5,000). Left: Q-Orbitrap data, right: Q-TOF data.

Supervised multivariate statistical analyses, i.e., PLS-DA, were performed, but the MSEP of the Monte Carlo cross-validation indicated that the model over-fitted the data (data not shown). These observations were in line with the observations for *porcine* muscle samples, since in the case of a small sample size these methods are prone to over-fit data.

Bovine kidney samples

Data integrity

A data integrity check using 4 spiked QC samples scattered throughout the analytical batch was performed. Eighty-six% (Q-Orbitrap) and 91% (Q-TOF) of the spiked analytes were found. The RSDs of the peak areas of the Q-Orbitrap data were mostly below 20% (see Supplementary data Table S10). Hence, the analytical method in the case of the Q-Orbitrap data was considered robust. In the case of the Q-TOF data, the RSD of the peak areas of the spiked QC samples were overall high with a maximum of 62%. This indicated the need for a normalization procedure, since analytical drifts could not be excluded. The data pretreatment prior PCA consisted of normalization by the sum of all peaks. This normalization was performed and subsequently, the data was reevaluated. This normalization process lead to RSDs of spiked QC samples which were mostly below 20%. Thus, it was concluded in a first data integrity check that after a normalization step the data was fit for further multivariate analysis (see Supplementary data Table S10). Ten analytes (Q-Orbitrap) and 6 analytes (Q-TOF) were not detected with the chosen data processing method. A check of the raw data showed that the peak areas of these analytes were below the chosen minimum peak height, which was due to pronounced ion suppression. This observation was comparable to the observations made during the analysis of *porcine* kidney samples.

Multivariate statistical analysis

After data pretreatment, the experimental data of *bovine* kidney samples was subjected to PCA. In the data set obtained from Q-TOF analysis, an analytical drift was observed over time (Figure 34, left). A distinct analytical drift in the direction of the first principal component was critical for the further evaluation of the data, since differences between the two test groups were mainly revealed by a distinct distribution of the samples along the first principal component. Thus, a correction based on the number of injection was implemented to minimize any influence due to order of injection. A quadratic least-squares regression curve was fitted to the mean peak areas of each variable of the QC samples in respect to the order of injection. Then the correction curve was applied to the whole data set. The clustering of the QC samples showed that the analytical drift was effectively removed by the normalization procedure (Figure 34, right). Hence, this step was implemented in the workflow of data processing for every multivariate analysis of the experimental data of *bovine* kidney samples obtained by Q-TOF analysis. Furthermore, it was noticeable that this normalization procedure resulted in a reduced variance explained by PC1, but the general outcome remained the same.

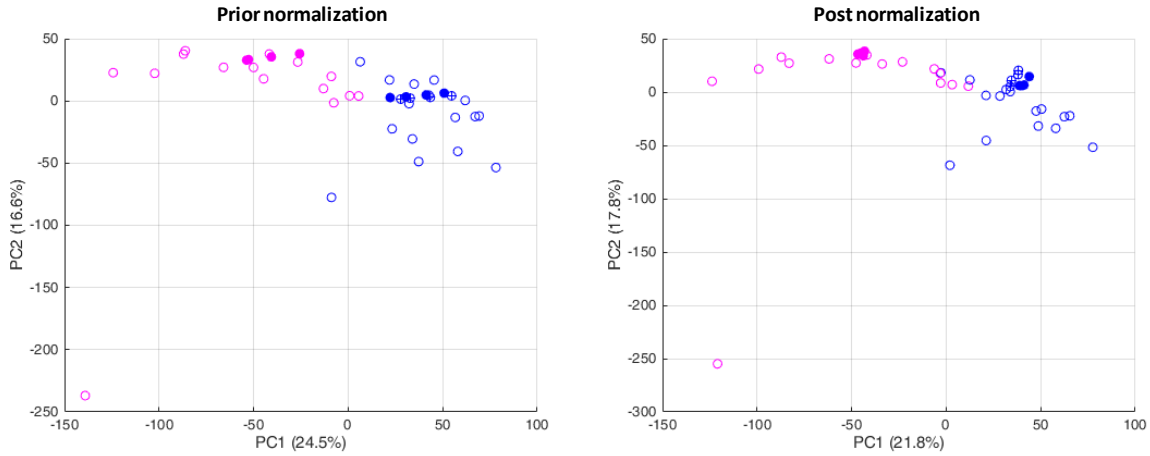


Figure 34. 2D-scores plot of PCA of LC-HRMS (Q-TOF) data of *bovine* kidney samples before (left) and after (right) normalization by applying a correction curve fitted to the QC samples with respect to the order of injection; blue: control samples (*bovine* kidney samples of untreated, healthy animals), magenta: positive samples (*bovine* kidney samples of infected, drug-treated animals); blue filled: negative QC samples, blue filled with cross: spiked QC samples, magenta filled: positive QC samples; PC: principal component with explained variance shown in brackets.

The scores plots of the Q-Orbitrap and Q-TOF data (see Figure 35) illustrate that the PCA revealed differences between samples from untreated, healthy animals and samples from infected, drug-treated animals, since the samples of the two groups showed a distinct dispersion along PC1, which accounted for 12.5% (Q-Orbitrap) and 21.8% (Q-TOF) of the total vari-

ance. However, a slight overlap of the scores of the sample groups was observed. The QC samples in the Q-Orbitrap data showed a variation in PC2. However, further normalization attempts were not performed since the main differences between the two groups were observed in PC1. Again, the spiked QC samples coincided with the control samples.

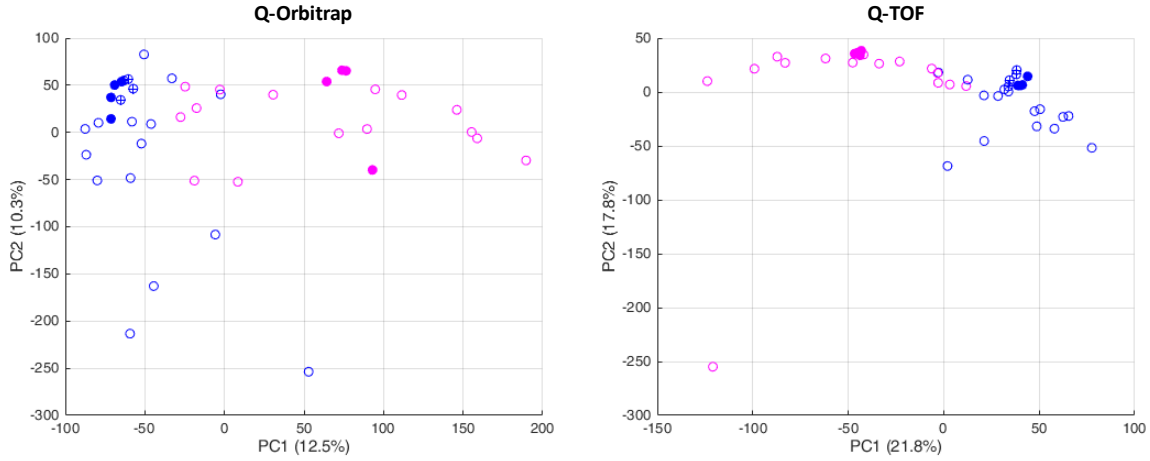


Figure 35. 2D-scores plot of PCA of LC-HRMS (left: Q-Orbitrap, right: Q-TOF) data of *bovine* kidney samples; blue: control samples (*bovine* kidney samples of untreated animals), magenta: positive samples (*bovine* kidney samples of infected, drug-treated animals); blue filled: negative QC samples, blue filled with cross: spiked QC samples, magenta filled: positive QC samples; PC: principal component with explained variance shown in brackets.

Monte Carlo cross-validation was applied in order to assess the ability of the PCA coupled with QDA to discriminate between the two classes (control samples vs. positive samples). The results are shown in Figure 36 as a confusion matrix. The models showed a specificity of 0.947 for the Q-Orbitrap data and 0.899 for the Q-TOF data, respectively. Thus, in both cases the data analysis procedure was suitable for the correct identification of control samples as samples taken from untreated, healthy cattle with a false positive rate of maximum 10%. The sensitivity of the model obtained from the Q-Orbitrap data (0.772) was comparable to the sensitivity of the model obtained from the Q-TOF data (0.795) with a false negative rate of maximum 22.8%.

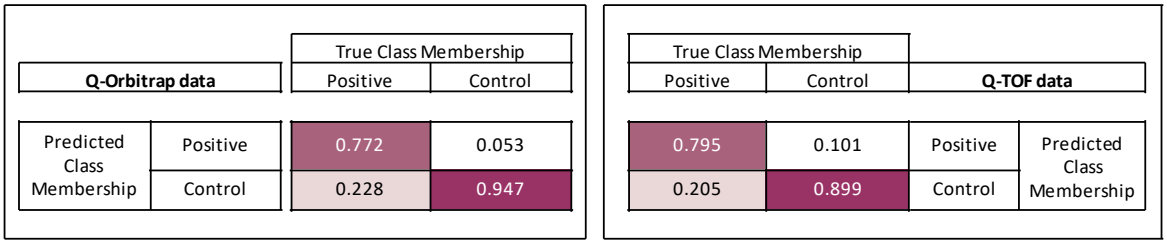


Figure 36. Confusion matrix of *bovine* kidney samples classified as positive samples (*bovine* kidney samples of infected, drug-treated animals) or control samples (*porcine* kidney samples of untreated animals) via 3D-PCA coupled with QDA using a Monte Carlo cross-validation (*n* = 5,000). Left: Q-Orbitrap data, right: Q-TOF data.

Supervised multivariate statistical analyses, i.e., PLS-DA, were performed, but the results are not shown since the model over-fitted the data.

In conclusion, the further investigations confirmed the overall applicability of the established approach to all matrices under investigation. Several methodological findings were recognized in the course of these investigations.

First, it is crucial to evaluate the quality of the raw data used for multivariate statistical analysis. A data integrity check of the mean peak areas and RSDs of known analytes in spiked QC samples provides a first impression of the data quality. High values of RSD indicate that the variation in the data set is high, which may be due to analytical drifts. This may negatively affect sample classification based on PCA with subsequent QDA, since the principle of PCA is that the principal components are constructed in such a way that they explain in a descending order of magnitude the variation of the data. Hence, a high variation in the data set originated from the instrument performance may superimpose variance based on different test groups. In general, a RSD below 20% is considered to be a good value for LC-MS data while a RSD of 30% is still considered to be sufficient [134]. In this study, in the case of higher variation, a normalization step was performed in order to check if the transformation by the division by the sum of all peaks rendered satisfactory values for RSD of analytes. This was done for the data obtained from the analysis of *bovine* samples. After the normalization the variation of the data was at an acceptable level. Normalization by the sum of all peaks is very common and was always the first step in the data pretreatment prior to PCA.

Besides an overall high variation in data sets, batch effects or a gradual intensity loss over time are also common phenomena in LC-MS analysis. These observations were also made in this study. In the case of the analysis of *porcine* kidney samples a pronounced batch effect was observed and in the case of the analysis of *bovine* kidney samples a gradual analytical

drift was noted. Two different QC based normalization procedures were applied to the respective data sets. In the first case, a correction factor based on the ratio of the mean peak areas of the QC samples from the two different batches was applied. In the second case, a correction curve based on the number of injection was used for normalization. Both procedures successfully reduced the variation based on analytical drifts. These examples show that it is crucial to evaluate the data integrity and, if necessary, apply normalization strategies in order to reduce analytical variance which may superimpose differences from different test groups.

Furthermore, the investigations via PCA showed that for both species and matrices a difference between positive (drug-treated, infected animals) and control (untreated, healthy animals) samples was observed. However, the difference was more pronounced in the data sets obtained from the examination of muscle samples. In all test scenarios, spiked QC samples coincided with control samples. Neither the control nor the positive samples contained all 70 veterinary drugs, but still the spiked QC samples coincided with the control samples. Assuming that the presence of one or multiple veterinary drugs at the level of MRL is the main difference between non-compliant samples and control samples, one would have expected that spiked QC samples coincided with positive (drug-treated, infected animals) samples. Since this was not the case, but QC samples spiked with veterinary drugs coincided with control samples, it can be concluded that in an unsupervised setting the main difference between these two test groups is not the presence of veterinary drug residues, but changes in the metabolome of the positive samples caused by the underlying infection of the animal.

In *bovine* muscle samples the sensitivity and specificity were high. Hence, PCA with subsequent QDA was considered to be a suitable approach in order to discern samples from the two test groups. This result was comparable to the outcome of the multivariate analysis of *porcine* muscle samples (see Section 4.1.2). However, one must keep in mind that the sample size used for the examination of cattle tissue was very small with only 15 samples per group. In future investigations, the sample size should be increased. Based on a larger data set it may then become feasible to attempt biomarker identification for this matrix.
In the multivariate statistical analyses of kidney samples of both species no clear separation of the two test groups was observed. Yet, it was evident that the expectation vectors of the two test groups deviated from each other. PCA coupled with QDA showed that the specificity of the models was high with 0.884 (Q-TOF data of *porcine* kidney samples)–0.947 (Q-Orbitrap data of *bovine* kidney samples), but the sensitivity was insufficient with a maximum

of 0.795 (Q-TOF data of *bovine* kidney samples). The false negative rate varied from 20.5% (Q-TOF data of *bovine* kidney samples) to 43.9% (Q-TOF data of *porcine* kidney samples), which indicated that this approach was not feasible in order to identify true positive samples, i.e., samples taken from drug-treated, infected animals. One possible reason could be that renal tissue is more metabolically and hormonally active than muscle tissue [171]. The kidneys exert diverse functions which include the maintenance of volume and composition of body fluids as well as the control of electrolyte and acid-base balance [172]. They participate in the regulation of blood pressure via the regulation of sodium balance as well as the synthesis and secretion of renin, erythropoietin and vasoactive eicosanoids [172]. Furthermore, polar chemicals as well as metabolites of xenobiotics and endogenous substances are excreted renally [24]. Hence, the metabolome of the renal tissue is influenced by various metabolic processes. This may explain the high variation in the data set of positive samples and thus, explain the high false negative rate. Besides, one must keep in mind that a trend according to storage time was observed in kidney samples and thus this parameter may present a confounding factor.
In consideration of all these observations, it can be concluded that the analysis of muscle tissue as described in Section 4.1.2 was the most suitable approach.

In general, the investigations showed further that supervised multivariate statistical analysis, i.e., PLS-DA, over-fitted the data and thus, PCA coupled with QDA was considered to be a better choice for sample classification.

4.2 LC-HRMS analysis of urine samples from ALT-treated gilts

Parts of the following passage have previously been published in [192].

4.2.1 Targeted LC-HRMS analysis

4.2.1.1 Method development and validation

Development of sample preparation protocol

Based on extraction methods for anti-doping drug testing [147] three different SPE cartridges were tested. Two of them, OASIS HLB® (200 mg, 6 cc) and STRATA XL (100 mg, 3 cc) SPE cartridges, contain a polymeric sorbent, which offers a reversed-phase retention mechanism of acidic, basic and neutral compounds [173]. The Bond Elut C18 (500 mg, 6 cc) SPE cartridge consists of a hydrophobic, silica-based sorbent allowing the retention of non-polar compounds. However, it is not recommended for the analysis of polar compounds [173].

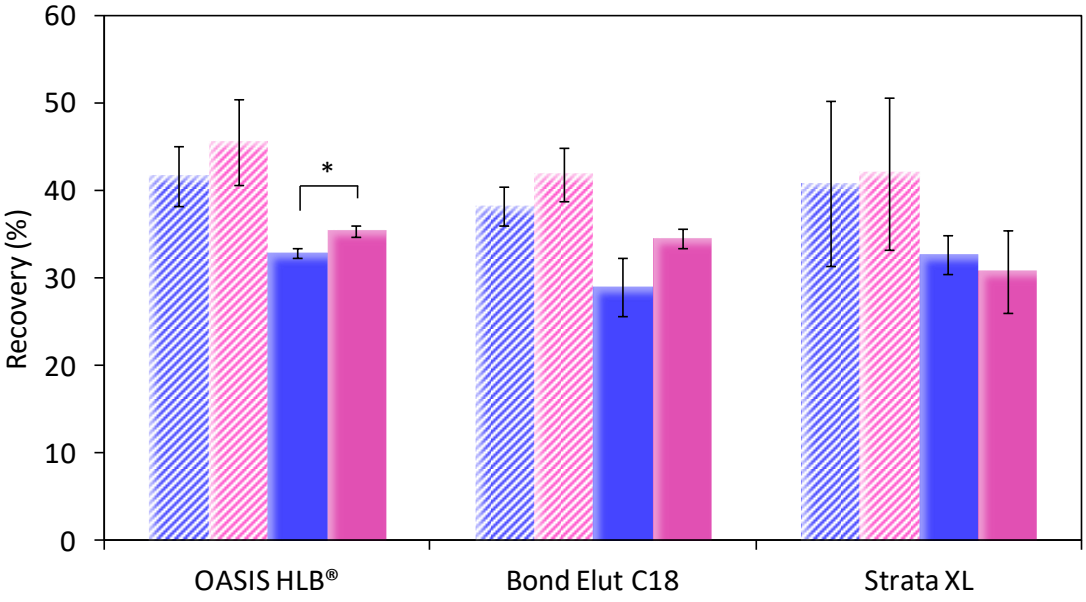


Figure 37. Recovery rates of ALT in urine samples using 3 different SPE cartridges (n = 5); mean ± measurement uncertainty (half-width of the 0.95-confidence interval); blue: spiking was performed before sample preparation; magenta: spiking was performed after sample preparation; hatched columns: spiking level: 5 ng/mL; filled columns: spiking level 500 ng/mL; *: p-value from *Student's* t-test: $p < 0.05$.

All three of the SPE cartridges showed comparable recovery rates of ALT in urine samples at the two tested concentration levels. The results varied from 29.0% to 45.6% (see Figure 37). The highest measurement uncertainty was observed for the results obtained by the Strata XL SPE procedure with a maximum of 9.4% while the variation of the measurement uncertainty

obtained by the OASIS HLB® and Bond Elut C18 procedure were comparable with a maximum of 4.9%.

The recovery rates at the spiking level 5 ng/mL were generally higher than the recovery at the spiking level 500 ng/mL. A comparison of means by a two-sided *Student's t*-test showed no significant difference between the recovery rates of the samples spiked before and after the SPE procedure for all three SPE cartridges at the concentration range of 5 ng/mL ($\alpha = 0.05$). The only difference was observed in the recovery rates of samples prepared with the OASIS HLB® cartridge (pre vs. post SPE procedure) at the concentration level of 500 ng/mL ($p < 0.05$). Note, that the OASIS HLB® cartridge's measurement uncertainty values at 500 ng/mL were distinctly lower than for all the other SPE cartridges. Overall, these results indicate that there was no notable loss of the analyte during the solid phase extraction procedure. The analysis of solutions recovered from the loading and washing steps also confirmed that no analyte was eluted during those steps (data not shown). Hence, the low recovery rates of 29.0% to 45.6% resulted from a reduced signal abundance of the analyte in matrix samples compared to standard solutions. This ion suppression was addressed with an additional extract cleaning step prior to LC-HRMS analysis. Hexane was used for defattening of the extracts. The hexane extraction of matrix components significantly increased signal abundance and recovery increased from 39.6% ± 12.9% to 52.5% ± 17.2% (mean ± measurement uncertainty, $p < 0.01$, see Figure 38). Thus, an additional cleaning step of the final extract prior to analysis is recommended if no internal standard is used.

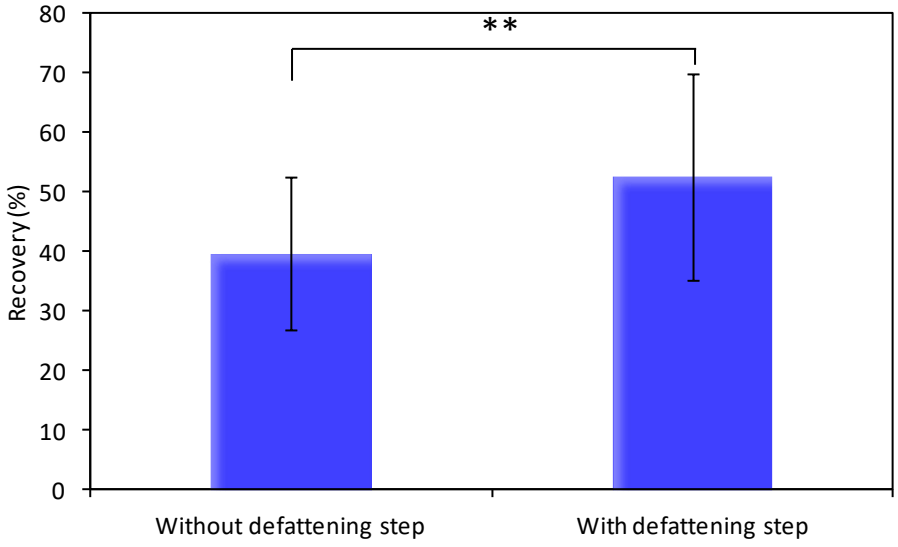


Figure 38. Recovery of ALT in urine extracts (5 ng/mL) prepared with OASIS HLB SPE procedure without and with defattening step (n = 10); mean ± measurement uncertainty (half-width of the 0.95-confidence interval); **: p-value from *Student's t*-test: $p < 0.01$.

However, in this study an isotopic-labeled internal standard (ALT-d5) was used to compensate for matrix induced ion suppression. The internal correction resulted in a trueness of 99.3% to 107.0% (see Table 13). These results show that quantification via internal standard is a valuable tool to compensate matrix suppression during LC-MS analysis.

Table 13. Trueness of ALT in urine sample using three different SPE cartridges, corrected with internal standard ALT-d5 at concentration level of 5 ng/mL (n = 5); mean ± measurement uncertainty (half-width of the 0.95-confidence interval).

	OASIS HLB®	Bond Elut C18	STRATA XL
Trueness (%)	106.7 ± 10.8	107.0 ± 14.6	99.3 ± 8.4

Conclusively, all three SPE cartridges showed comparable recovery rates for ALT. None of the extraction procedures led to a significant loss of the analyte during the solid phase extraction. The matrix effect which reduced the recovery can be corrected using an internal standard in the final method.

The results obtained with the STRATA XL SPE procedure showed the highest measurement uncertainty. The lowest measurement uncertainty was observed after the sample preparation with the OASIS HLB® SPE cartridges at spiking level 500 ng/mL. The precision achieved with the Bond Elut C18 procedure was also considered satisfying. However, considering the retention mechanism this cartridge is not ideal for analyzing more polar metabolites of ALT. Hence, the SPE cartridge OASIS HLB® was selected for further method validation, keeping in mind the scope of the method analyzing ALT and potentially more polar metabolites, such as its glucuronides and sulfates [69, 174].

LC-HRMS method development and evaluation of fitness for purpose

Chromatographic parameters were optimized in terms of chromatographic resolution and total time of analysis. When analyzing ALT standard solutions the main analyte (ALT) was retained at 7.67 min. Additionally, a small peak at 7.34 min was observed in standard solution as well as in spiked urine samples. The isotopic labeled internal standard ALT-d5 was retained at 7.65 min (see Figure 39).

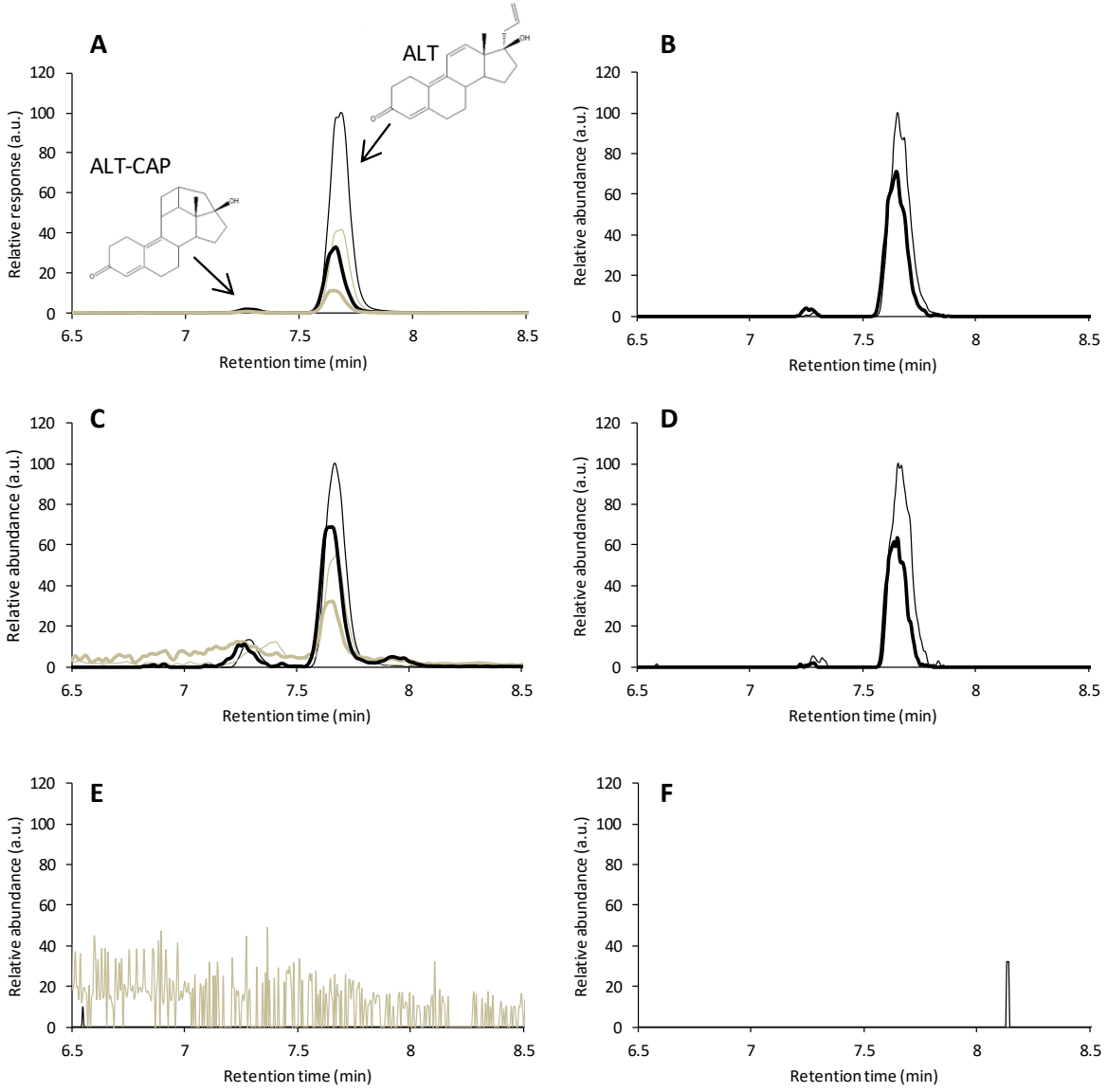


Figure 39. Comparison of extracted ion chromatograms obtained by PRM method (left) and full scan mode (right) for ALT (@ 7.67 min) and internal standard ALT-d5 (@ 7.65 min, bold line); in PRM mode *m/z* 227.1425 (black) was used as quantifier and *m/z* 159.0804 (grey) was used as qualifier; in full scan $[M+H]^+$ was used for quantification (ALT: *m/z* 311.2006; ALT-d5: *m/z* 316.2309); A/B: extracted ion chromatograms of standard solutions of ALT and ALT-d5; C/D: extracted ion chromatograms of ALT and ALT-d5 in urine samples at spiking level 5 ng/mL; E/F: extracted ion chromatograms of blank urine samples. The photo-isomer ALT-CAP (@ 7.34 min) can also be observed in standard and matrix samples.

As described in the literature, the isomerization of ALT has been observed and the photo-isomer has been described as a 2 + 2 cycloaddition product (ALT-CAP) [87]. Hence, the identity of the isomer detected at 7.34 min was further investigated and the isomer was identified as ALT-CAP via ^{1}H NMR (see Section 8.1.1). The gradient program was set to separate the isomers (chromatographic resolution R: 1.6) while reducing the total analysis time to 20 min per analytical run.

Initially, two different scan modes were tested: parallel reaction monitoring (PRM) and full scan mode with confirmatory fragmentation. PRM mode is comparable to the multiple reaction monitoring (MRM) mode used in triple quadrupole mass spectrometry. However, a full scan of all product ions is performed after inducing a collision in the HCD cell. Therefore, all product ions can be acquired simultaneously. In contrast, in triple quadrupole mass spectrometry using MRM mode, a second *m/z* filter in Q3 selects only some of the product ions for data acquisition. The quantification in PRM mode was performed using the product ion *m/z* 227.1425 as quantifier for ALT and the internal standard ALT-d5, while *m/z* 159.0804 was monitored as qualifier. The molecule ions $[M+H]^+$ were detected and used for quantification in full scan mode. Additionally, fragments were used for confirmatory purposes.

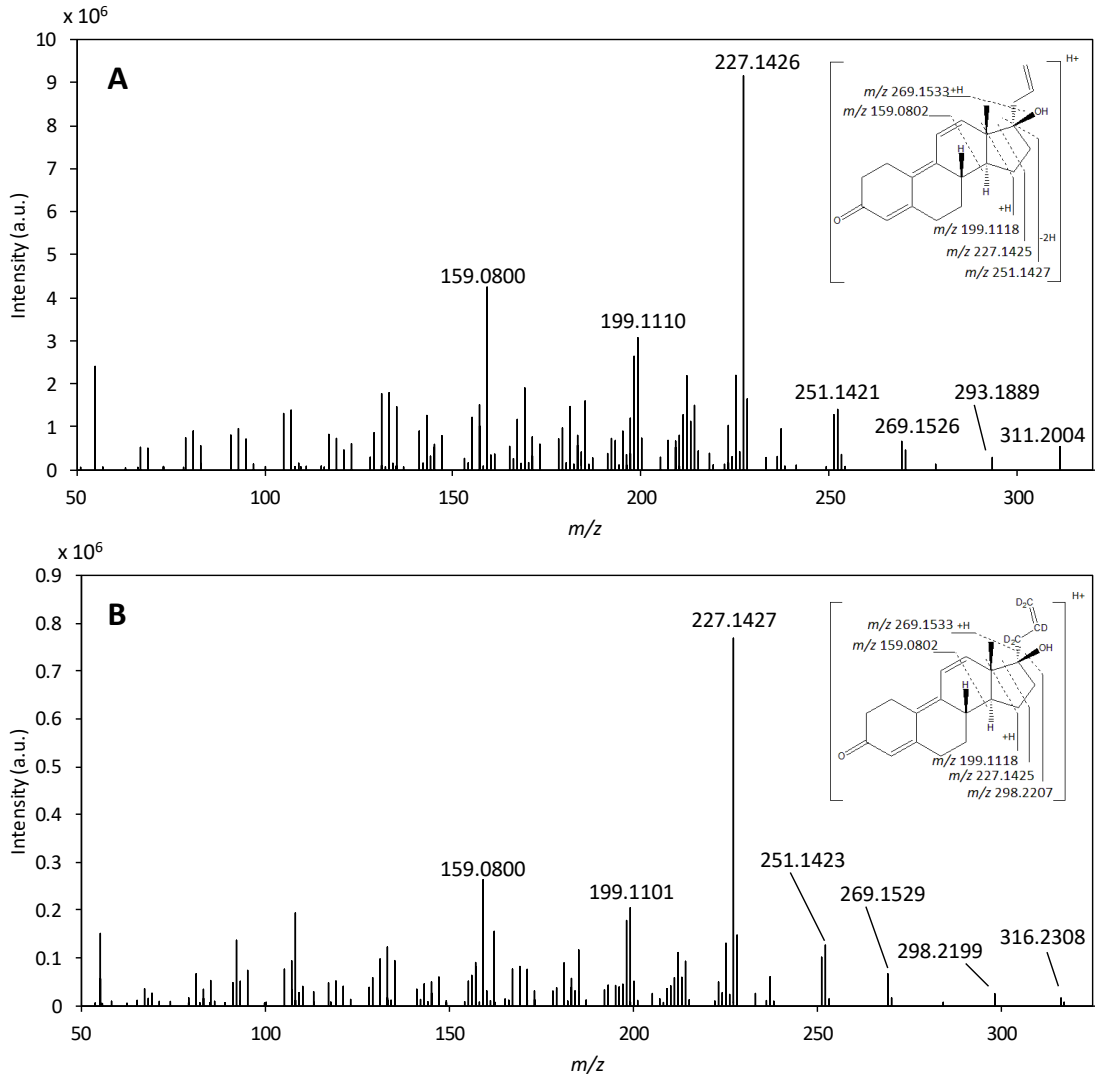


Figure 40. A: product ion spectrum of ALT obtained at Q-Orbitrap at 7.67 min (precursor ion *m/z* 311.1991; $[M+H]^+$) by the analysis of spiked urine samples; B: product ion spectrum of ALT-d5 obtained at Q-Exactive Focus at 7.65 min (precursor ion *m/z* 316.2309; $[M+H]^+$) by the analysis of spiked urine samples.

Commission Decision 2002/657/EC [148] defines requirements for the identification of an analyte (see Section 3.2.2.3). Besides matching retention times of a standard solution and the analyte in a test sample, confirmatory methods need to provide information on the chemical structure of the analyte. Using PRM this can be achieved by measuring at least 2 product ions and comparing their ion ratio to standard solutions and spiked matrix samples. Using high resolution full scan with confirmatory fragmentation, the accurate mass of the analyte with a mass deviation lower than 5 ppm to the calculated exact mass as well as isotopes of the analyte and characteristic fragments allow for unequivocal identification. Figure 39 shows that the ion ratio of the detected ions can be used as diagnostic parameter in order to assure the selectivity of a PRM method. The ion ratios of ALT in standard solutions and matrix samples were comparable. In a HRMS full scan method, the mass deviation compared to the calculated exact mass was -1.8 ppm in standard solutions and -2.5 ppm in matrix samples, respectively. Furthermore, more than 3 characteristic fragments were observed in the product ion spectrum of ALT in urine samples (see Figure 40). Thus, the results of both scan modes meet the demands for selective substance identification in confirmative methods set by the Commission Decision 2002/657/EC. The analysis of blank matrix samples showed no interfering signals of matrix components in both scan modes (see Figure 39). Hence, both methods were considered to be specific. With respect to the comparable results for selectivity and specificity for the two scan modes, full scan mode with data-dependent fragmentation was chosen for further method development, since it can be used to quantify analytes at trace level and offers valuable advantages for metabolic fingerprinting, i.e., the possibility to perform retrospective analyses of relevant biomarker candidates combined with MS/MS information for structural elucidation. Thus, the chosen mass spectrometric method is fit for targeted as well as for non-targeted analyses.

Validation data

Stability

A stability study demonstrated that after 28 days under storage conditions (-20°C) ALT showed a minor decrease of 10%. Storage at 4°C led to a decrease of 20%. ALT stored at room temperature decreased to 67% without light exposure and to 33% (sum of ALT and ALT-CAP) with light exposure, respectively (see Figure 41). Furthermore, storage at room temperature with light exposure led to an elimination of the analyte accompanied by the formation of the photo-isomer ALT-CAP as it has been reported before [87, 174]. Nevertheless, the formation of the photo-isomer resulted only in 33% of the initial concentration of the parent compound. Thus, different elimination pathways must be considered, but remain unrevealed.

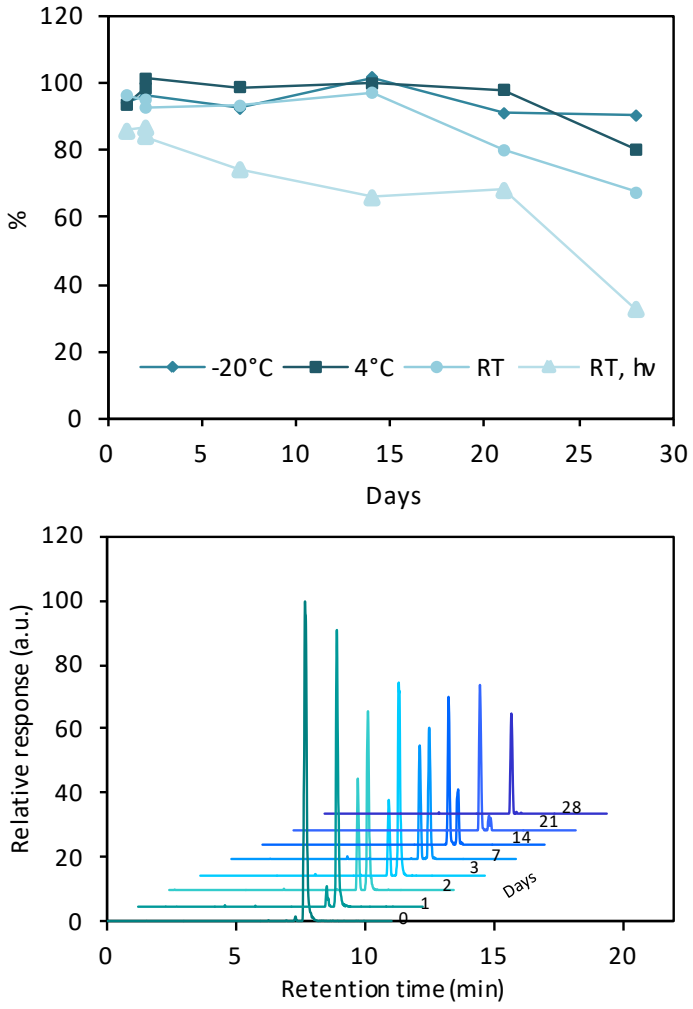


Figure 41. Top: concentration profile of ALT (including its isomer ALT-CAP) in urine extracts over time under different storage conditions (rhombus: -20°C; square: 4°C; dot: room temperature without light exposure; triangle: room temperature with light exposure). Bottom: isomerization of ALT in matrix extracts over time when stored at room temperature with light exposure.

Furthermore, three freeze-thaw cycles were performed with spiked matrix extracts at three different concentration levels (high, medium, low). No significant elimination was observed (see Table 14).

Table 14. Recovery of ALT in urine extracts after three freeze-thaw cycles (n = 3) at three spiking levels (low, medium, high); mean ± measurement uncertainty (MU; half-width of the 0.95-confidence interval).

	Recovery (%) @ spiking level (ng/mL)		
	5	**500**	**3000**
three freeze-thaw cycles	92.8 ± 13.9	109.6 ± 10.7	107.5 ± 4.5
no freeze-thaw cycles	96.9 ± 6.6	110.8 ± 10.3	105.5 ± 8.9
p-value (t-test)	N.S.	N.S.	N.S.

Recovery rates of matrix extracts after three freeze-thaw cycles and no freeze-thaw cycles were compared by *Student's t*-test (error probability of $\alpha = 0.05$); N.S.: not significant.

Uncertainty-related performance characteristics

Uncertainty-related performance characteristics were calculated for ALT in *porcine* urine applying the sample preparation procedure without and with an additional hydrolysis step in order to quantify ALT and the sum of ALT and the hydrolyzed conjugates of ALT, respectively.

No scientific data concerning the urinary excretion of ALT in gilts was available, only data on the plasma-pharmacokinetics of ALT have recently been published [66]. Therefore, pharmacokinetic data on urinary excretion of ALT in horse were used to estimate the concentration range for method validation, i.e, determination of uncertainty-related performance characteristics. In a pharmacokinetic study in horse the maximum urinary concentration values of hydrolyzed ALT varied between 1000–3000 ng/mL [174]. The calibration range was adjusted accordingly. A linear relationship between the response Y and the concentration x was observed for the operating range of both methods (see Figure 42).

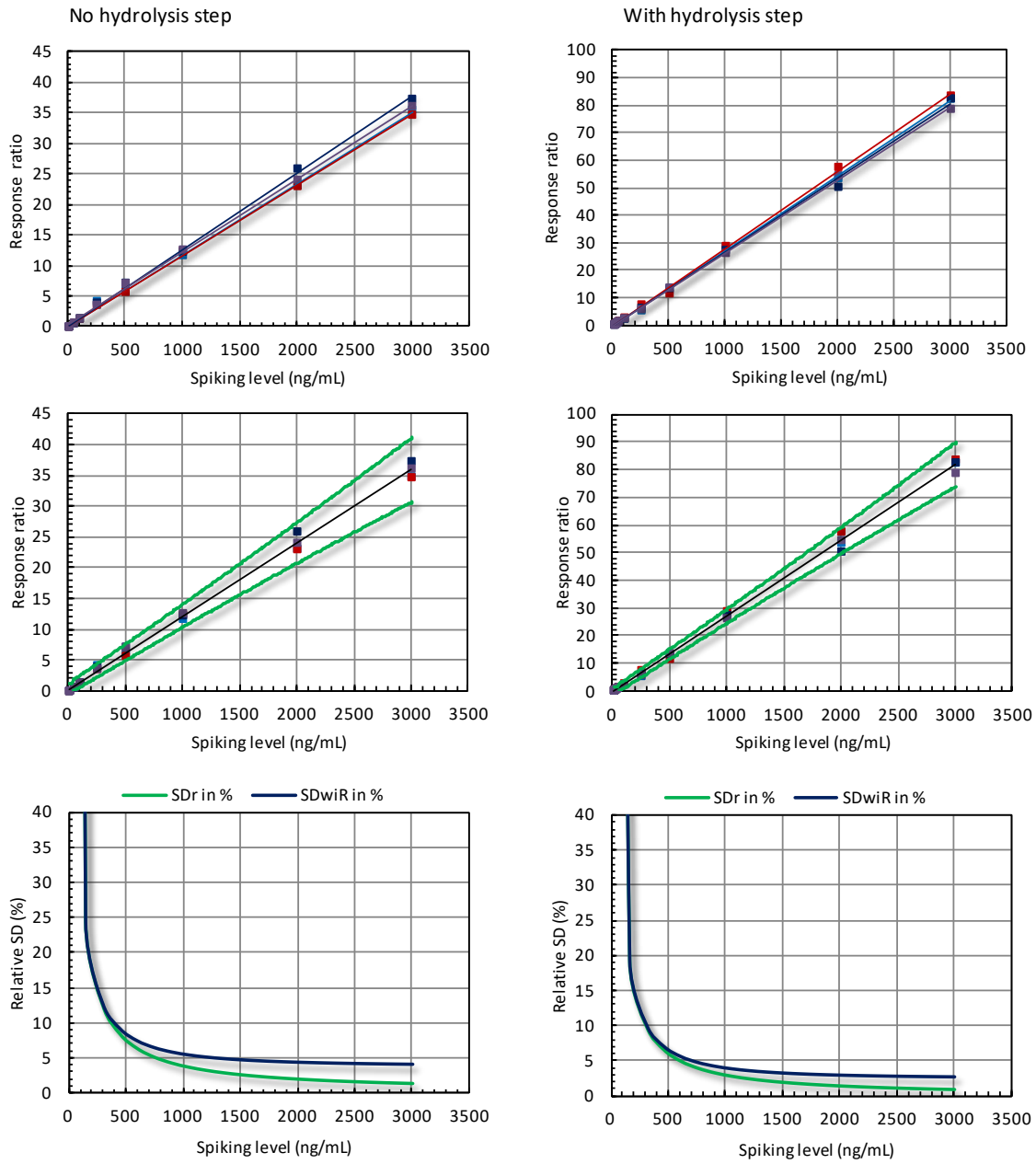


Figure 42. Top: four individual matrix calibration series at the concentration range of 5–3000 ng/mL; middle: variance components model fitted to data obtained from the analysis of four individual matrix calibration series; green curves: 0.95-prediction bands; bottom: relative standard deviation of repeatability (SD_r in%) and relative standard deviation of within-laboratory reproducibility (SD_{wiR} in%) as a function of concentration. Left: data obtained from the analysis without β-glucuronidase treatment; right: data obtained from the analysis with β-glucuronidase treatment.

The uncertainty-related performance characteristics as defined in Commission Decision 2002/657/EC as well as performance characteristics as defined in DIN 32645 [149] were established for both methods (with and without hydrolysis step). Decision limit ($CC\alpha$) and detection capability ($CC\beta$) were obtained under within-laboratory reproducibility conditions according to Commission Decision 2002/657/EC, while the decision limit (CL), detection limit

(LOD) and determination limit (LOQ) were calculated under repeatability conditions according to DIN 32645 (see Figure 43). The results showed that $CC\alpha$ and $CC\beta$ as well as the performance characteristics according to DIN 32645 cover the same concentration range (see Table 15). $CC\alpha$ was 1.3 ng/mL and 0.9 ng/mL using the method without hydrolysis step and with an additional β-glucuronidase step, respectively. $CC\beta$ was 2.5 ng/mL and 1.8 ng/mL using the method without any pre-treatment and with an additional β-glucuronidase step, respectively. Interestingly, the method with an additional pre-treatment step resulted in a slightly decreased $CC\alpha$ and $CC\beta$. CL and LOQ according to DIN 32645 were 1.6 ng/mL and 3.9 ng/mL (without hydrolysis step) as well as 1.2 ng/mL and 3.0 ng/mL (with hydrolysis step).

Table 15. Comparison of performance characteristics for the determination of ALT in urine extracts using different sample pre-treatment procedures: $CC\alpha$ and $CC\beta$ obtained via variance components model vs. decision limit (CL), detection limit (LOD) and determination limit (LOQ) according to DIN 32645.

Pre-treatment	DIN 32645*			CD 2002/657/EC	
	CL (ng/mL)	LOD (ng/mL)	LOQ (ng/mL)	$CC\alpha$ (@x_0 = 0) (ng/mL)	$CC\beta$ (@x_0 = 0) (ng/mL)
None	1.6	3.2	3.9	1.3	2.5
β-glucuronidase	1.2	2.4	3.0	0.9	1.8

* worst case values of 4 independent calibration curves.

The standard deviation of repeatability (SD_r) and the standard deviation of reproducibility (SD_{wiR}) were below 5.0% at the highest concentration level of ALT (3000 ng/mL) and maximum 13.9% at the lowest concentration level (5 ng/mL) (see Table 16). At higher concentration levels the deviation between the prediction bands increased, which is a result of a heteroscedastic variance. This demonstrates a phenomenon which is often observed in residue analysis: Experimental errors often are heteroscedastic under reproducibility conditions, since the variance of the measurement error in that case is related to the concentration [175].

Table 16. Relative standard deviation of repeatability (SD_r in %) and relative standard deviation of within-laboratory reproducibility (SD_{wiR} in %) at three spiking levels (low, medium, high) in urine extracts.

Pre-treatment	SD_r in % @ spiking level (ng/mL)			SD_{wiR} in % @ spiking level (ng/mL)		
	5	500	3000	5	500	3000
None	13.9	7.5	1.3	13.9	8.4	4.0
β-glucuronidase	10.3	6.1	1.0	11.1	6.4	2.6

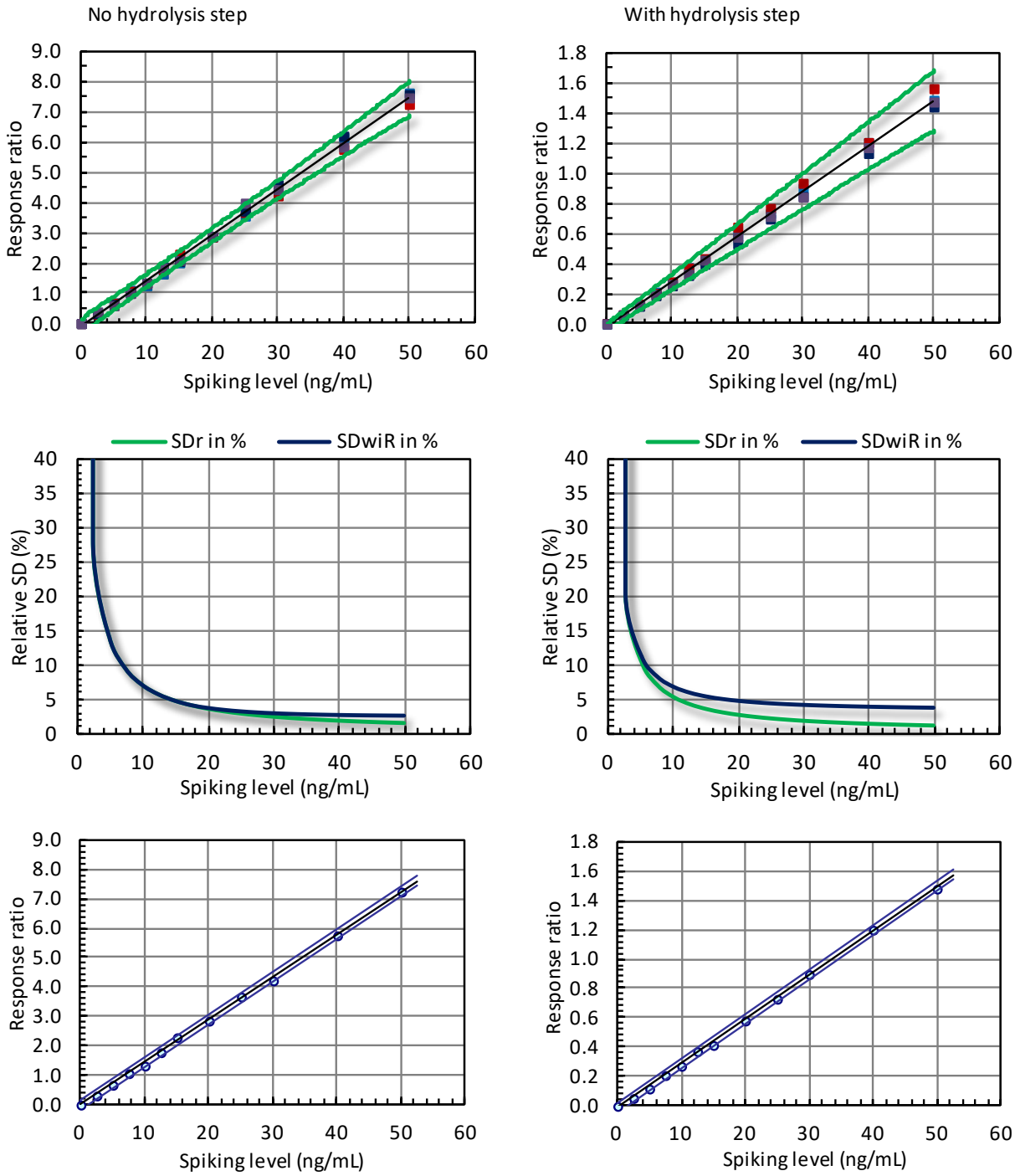


Figure 43. Top: variance components model fitted to data obtained from the analysis of four individual matrix calibration series to determine $CC\alpha$ and $CC\beta$ according to CD 2002/657/EC; green curves: 0.95-prediction bands; middle: relative standard deviation of repeatability (SD_r in%) and relative standard deviation of within-laboratory reproducibility (SD_{wiR} in%) as a function of concentration; bottom: exemplary data of ordinary least squares regression (OLS) to determine performance characteristics according to DIN 32645, blue curves: 0.95-prediction bands. Left: data obtained from the analysis without β-glucuronidase treatment; right: data obtained from the analysis with β-glucuronidase treatment.

Recovery was assessed by plotting measured values obtained via external calibration against actual values in the concentration range from 10–200 ng/mL in the final extract. The recovery was 101%–105% using the method without hydrolysis step and 81%–100% using the method with hydrolysis step, respectively. However, in this study quantification was done via matrix calibration. Provided that sample matrix and calibration matrix do not differ substantially from each other, the expectation of the recovery is nearly 100% [153].

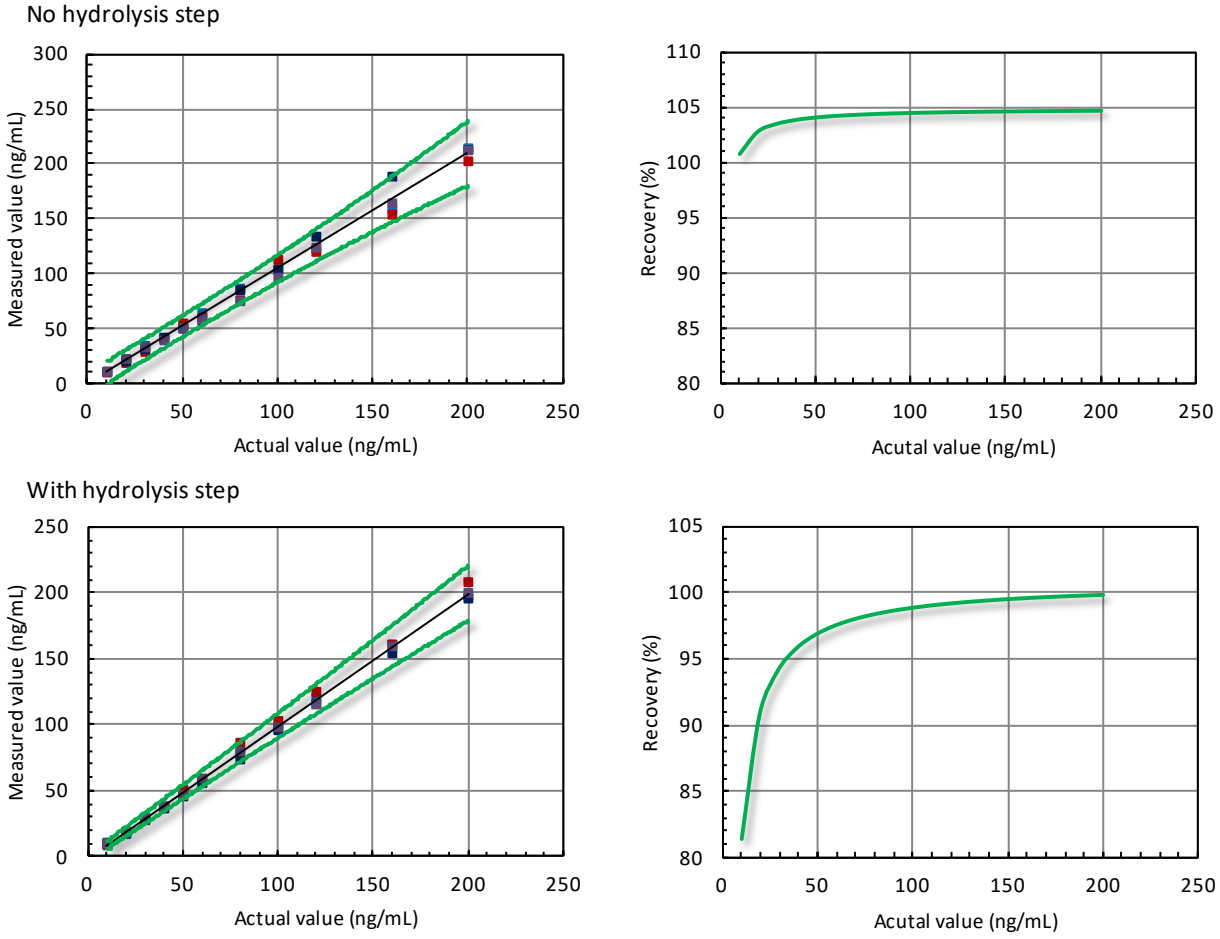


Figure 44. Left: measured value plotted relative to actual value in urine extracts in the concentration range from 10–200 ng/mL (corresponding to a spiking level from 2.5–50 ng/mL); right: recovery (%) as a function of the concentration. Top: data obtained from the analysis without β-glucuronidase treatment; bottom: data obtained from the analysis with β-glucuronidase treatment

Overall, method validation showed that matrix calibration and the use of an internal standard lead to accurate and precise results. The uncertainty-related performance criteria (CCα as well as CL) were in the range of 1 ng/mL ALT in urine samples (with and without hydrolysis) and thus, meet the purpose of the method analyzing urine samples to determine the urinary excretion rate of ALT and ALT metabolites, i.e., glucuronides.

4.2.1.2 Identification and quantitative determination of ALT glucuronides

In the targeted analysis of urine samples of ALT-treated gilts pronounced peaks at 5.3 and 6.2 min (*m*/*z* 487.2327 $[M+H]^+$) were detected (see Figure 45). These peaks were identified as glucuronides of the parent drug ALT (peak at 6.2 min) and of the primary photo-isomer ALT-CAP (peak at 5.3 min). Since the glucuronide of ALT was not commercially available, it was synthesized using an *in vitro* enzymatic reaction [176]. For details see Supplementary data Section 8.1.2. The *in vitro* synthesis in aqueous solution also rendered the glucuronide of ALT-CAP, which was consisted with the observation in urine samples. The product ion spectra of the two glucuronides showed characteristic fragments, i.e., *m*/*z* 311.2006 (ALT), 293.1895, 269.1531, 227.1426. The mass deviation of the product ions compared to the calculated mass of ALT glucuronides was very low with a maximum of -3.3 ppm and within the accuracy of the instrument (5 ppm). A peak with the mass-to-charge ratio of ALT was observed at the same retention time than of ALT glucuronide. Hence, this peak was contributed to an in-source fragmentation of ALT glucuronide.

ALT and its photo-isomer were not observed at the expected retention time (7.3 and 7.7 min, *m*/*z* 311.2006 $[M+H]^+$) in the extracted ion chromatogram of samples from ALT-treated gilts. However, a new peak at 7.9 min was detected with this mass-to-charge ratio. It can be hypothesized that this peak corresponds to an unknown isomer, since its product ion spectrum showed all characteristic fragments of ALT. However, the isomer was only detected at trace level and thus, preparative purification was not possible. Furthermore, in negative ion mode, no ALT sulfate was observed.

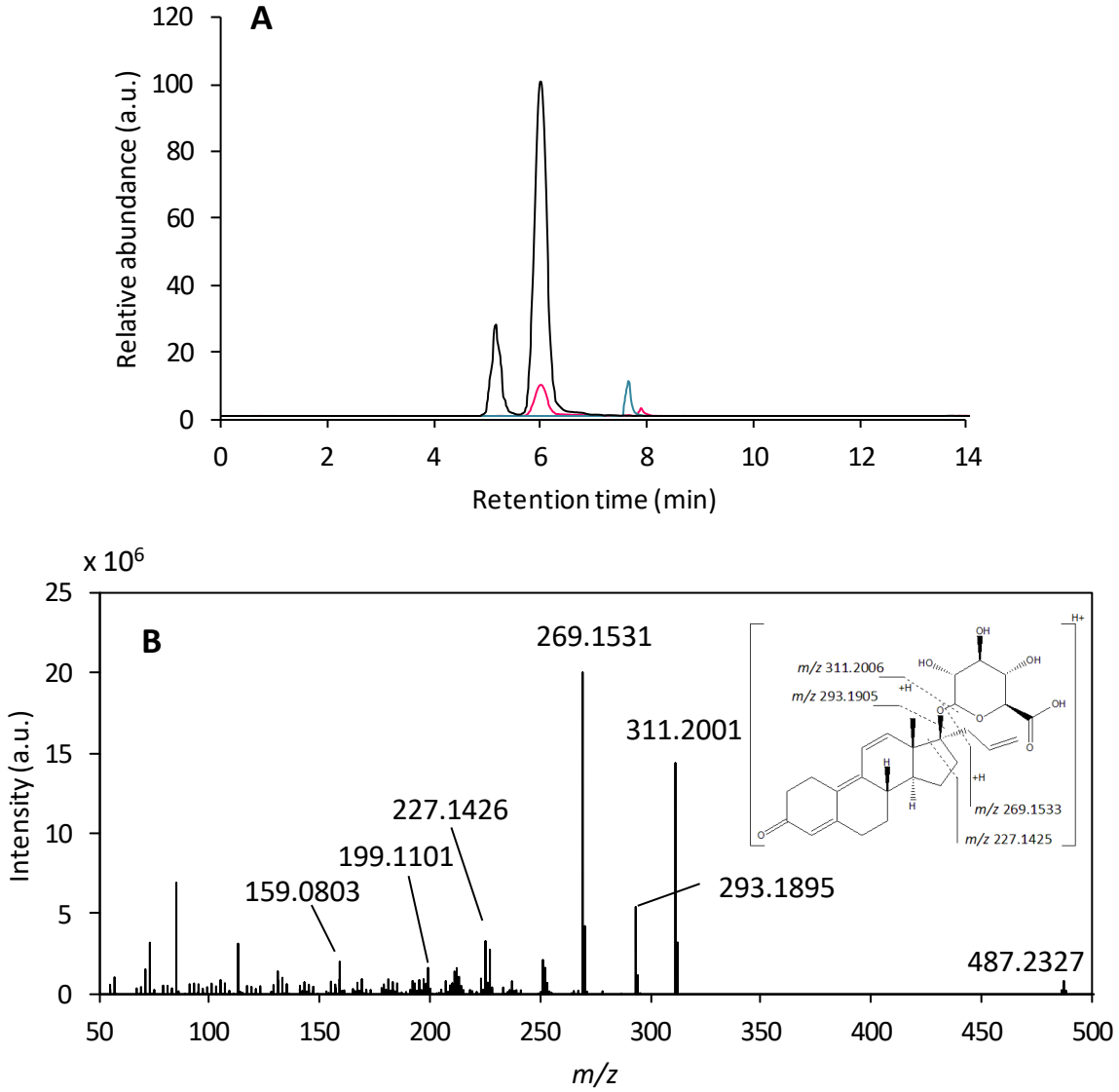


Figure 45. A: overlay of representative extracted ion chromatogram of urine sample from ALT-treated gilt and reference standard; black: ALT glucuronide (@ 6.2 min) and ALT-CAP glucuronide (@ 5.3 min) (*m/z* 487.2327); magenta: in-source fragmentation product of ALT glucuronide (@ 6.2 min) and unknown isomer of ALT (@ 7.9 min) (*m/z* 311.2006); blue: internal standard ALT-d5 (*m/z* 311.2006 @ 7.7 min); B: product ion spectrum of ALT glucuronide at 6.2 min (precursor ion *m/z* 487.2327; $[M+H]^+$).

For quantification of ALT glucuronides (i.e., ALT glucuronide and ALT-CAP glucuronide), enzymatic hydrolysis was performed for 24 hours with a sample volume of 400 µL in order to reduce the amount of ALT glucuronides which needed to be hydrolyzed. However, a complete hydrolysis could not be achieved since the response ratios of ALT glucuronides/ALT-d5 only decreased by 42–63% (see Figure 46). This phenomenon is not uncommon, since enzymatic hydrolysis often results in incomplete hydrolysis [177].

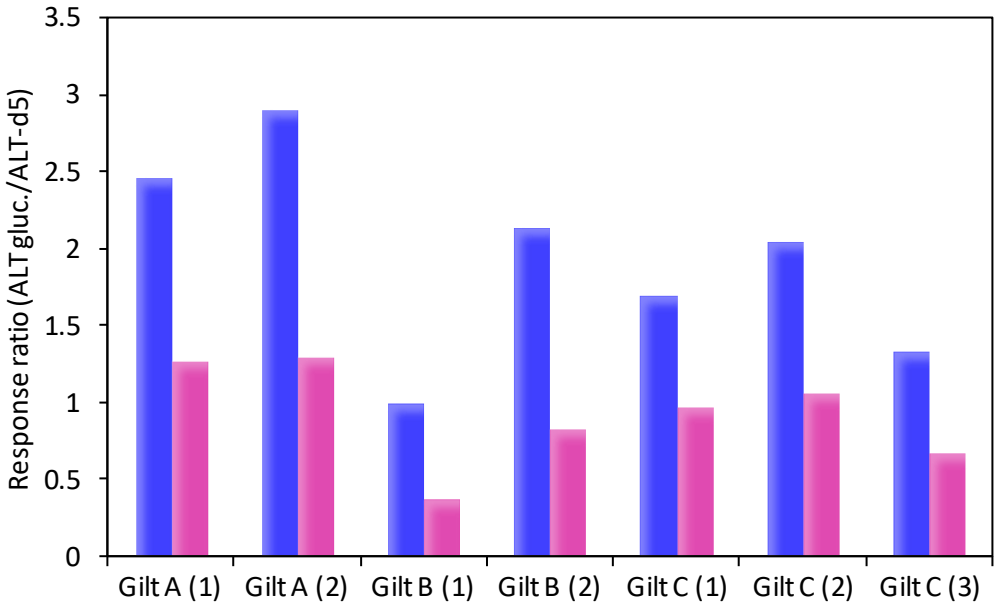


Figure 46. Response ratio of (sum of ALT glucuronide and ALT-CAP glucuronide)/ALT-d5 in urine samples from ALT-treated gilts before (blue) and after (magenta) hydrolysis step; numbers in brackets indicate sampling time (see Table 6).

In order to accomplish a complete hydrolysis, an acid hydrolysis was also tested (50% HCl for 1.5 hours at 100°C) [177]. However, ALT was not stable under these conditions (data not shown).

Hence, quantification of ALT glucuronides (sum of ALT glucuronide and ALT-CAP glucuronide) was performed as follows:

1. The concentration of ALT (sum of ALT and ALT-CAP) in mol/L after enzymatic hydrolysis, c^+, was quantified via matrix calibration.

2. Response ratios y^+ and y^- were calculated as

$$y^- = \left(\frac{Area\ (ALT\ \text{glucuronides})}{Area\ (\text{ALT-d5})}\right)_{\text{before hydrolysis}}$$

and

$$y^+ = \left(\frac{Area\ (ALT\ \text{glucuronides})}{Area\ (\text{ALT-d5})}\right)_{\text{after hydrolysis}}$$

3. The concentration of ALT glucuronides in mol/L, $c_{\text{ALT glucuronides}}$, was estimated as

$$c_{\text{ALT glucuronides}} = \frac{y^-}{y^- - y^+} \cdot c^+$$

The results are shown in Table 17. The concentrations varied from 1900 ng/mL to 4140 ng/mL, which is comparable to data reported for horses by Machnik *et al.* [50]. However, no pronounced decrease of ALT glucuronides according to sampling time was observed. Hence, a normalization procedure was attempted.

Table 17. Quantification of ALT, an unkown isomer of ALT and ALT glucuronides (sum of ALT glucuronide and ALT-CAP glucuronide) in urine samples of gilts after ALT treatment of 20 mg/day at different sampling times (*n* = 2; mean ± measurement uncertainty).

Sample	Hours after last treatment	ALT[1]	ALT isomer (@ 7.9 min)[2]	ALT glucuronides[3]
		ng/mL		
gilt A				
3:30 pm (day 17)	7.5	n.d. (<CL)	<LOQ	3900 ± 250
11:15 am (day 18)	3.25	n.d. (<CL)	<LOQ	3950 ± 250
gilt B				
4:00 pm (day 17)	8	n.d. (<CL)	<LOQ	1900 ± 250
11:15 am (day 18)	3.25	n.d. (<CL)	<LOQ	4140 ± 250
gilt C				
3:00 pm (day 17)	7	n.d. (<CL)	<LOQ	3780 ± 250
7:00 am (day 18)	23	n.d. (<CL)	<LOQ	3320 ± 250
10:15 am (day 18)	2.25	n.d. (<CL)	<LOQ	2780 ± 250

n.d.: not detected, decision limit CL: 1.6 ng/mL, determination limit LOQ: 3.9 ng/mL; [1]ALT analyzed by method without enzymatic hydrolysis; [2]unknown ALT isomer analyzed by method without enzymatic hydrolysis; [3]ALT glucuronides quantified as ALT after enzymatic hydrolysis.

Normalization of ALT glucuronide to urinary creatinine concentration

Concentrations of metabolites in urine are depending on the hydration status of the respective animal and thus, the total urine volume in which urinary metabolites are excreted [178]. In order to account for urinary volume, it is a common practice to reference urinary biomarkers to urinary creatinine concentration since creatinine—a product of muscle metabolism—is excreted at a fairly constant rate within healthy individuals [178]. Urinary creatinine concentration was determined via ^{1}H NMR. Method validation of quantitative ^{1}H NMR method is described in the Supplementary data Section 8.2. The results are shown in Table 18. The concentration of ALT glucuronides normalized to urinary creatinine decreased over time for gilt B and gilt C. The ALT glucuronides/U_{cr} of gilt A was nearly the same at 3.25 hours and 7.5 hours after the last treatment.

Table 18. Normalization of urinary ALT glucuronides (sum of ALT glucuronide and ALT-CAP glucuronide) to urinary creatinine concentration U_{cr} (*n* = 2; mean ± measurement uncertainty and relative measurement uncertainty based on error propagation calculation).

Sample	Hours after last treatment	ALT glucuronides[1]	Creatinine (U_{cr})	ALT glucuronides/U_{cr}
		ng/mL	mg/mL	µg/mg
gilt A				
3:30 pm (day 17)	7.5	3900 ± 250	1.04	3.76 ± 6%
11:15 am (day 18)	3.25	3950 ± 250	1.02	3.86 ± 6%
gilt B				
4:00 pm (day 17)	8	1900 ± 250	0.36	5.32 ± 13%
11:15 am (day 18)	3.25	4140 ± 250	0.29	14.06 ± 6%
gilt C				
3:00 pm (day 17)	7	3780 ± 250	0.80	4.74 ± 6%
7:00 am (day 18)	23	3320 ± 250	1.77	1.87 ± 7%
10:15 am (day 18)	2.25	2780 ± 250	0.36	7.69 ± 9%

[1]ALT glucuronides quantified as ALT equivalents after enzymatic hydrolysis.

According to the "summary of medicinal product characteristics"[51] a steady-state is described after several days of treatment. Since the sample collection was done during the last two days of treatment, a steady-state can be assumed. Consequently, the concentration of urinary ALT glucuronides was plotted against the hours after the last treatment in order to visualize the effect of normalization to urinary creatinine U_{cr}. A continuous decrease in the ratio of ALT glucuronides/U_{cr} can be observed, which confirms that a normalization of urinary biomarkers is crucial in order to determine a pharmacokinetic profile. This observation is especially evident in the concentration profile over time for gilt C. The non-normalized values were nearly constant, while after normalization a decrease was observed (see Figure 47). This can be explained by the fact that the spot urine taken at 23 hours was most likely morning urine, since it was highly concentrated and strongly colored.

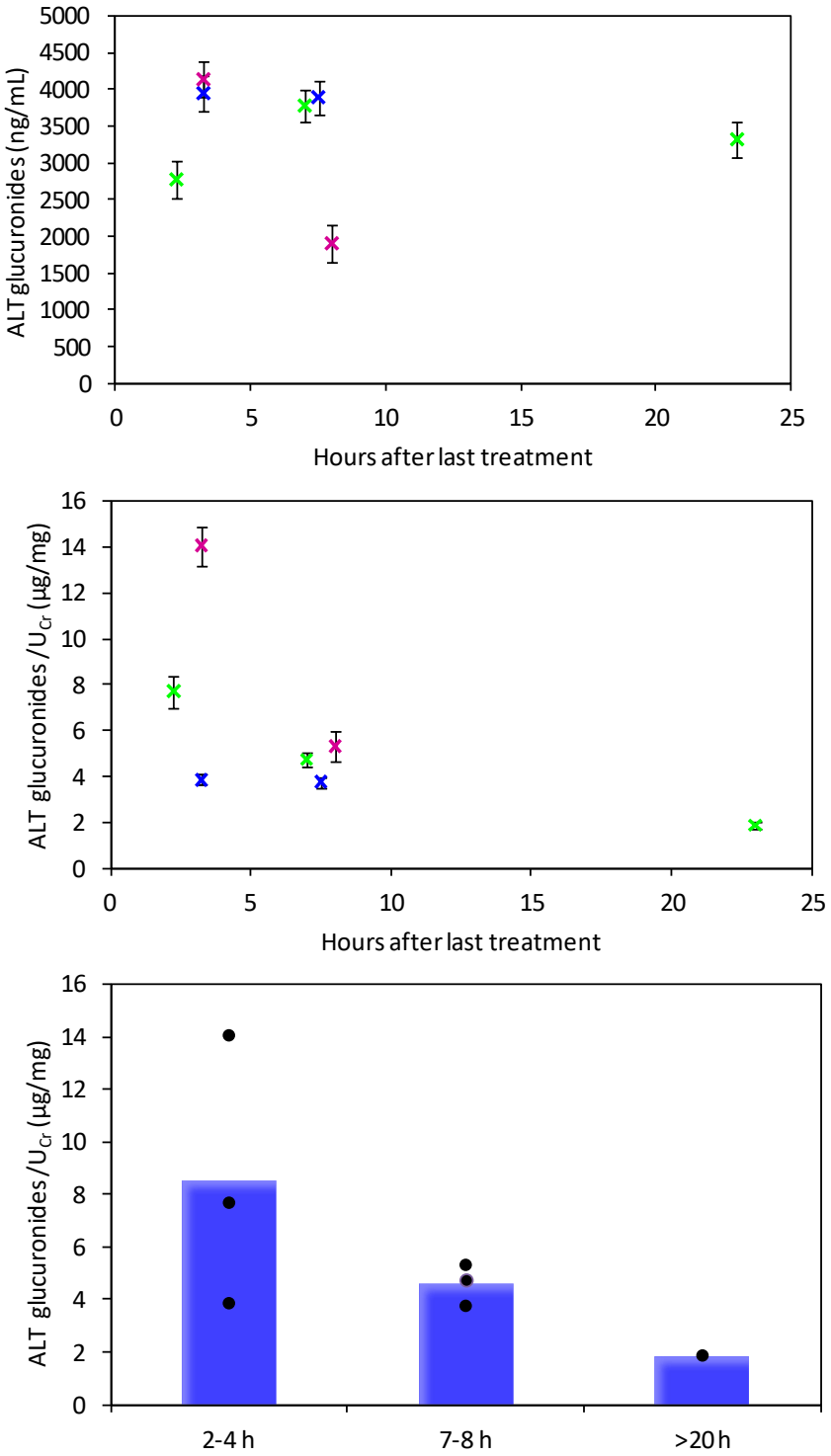


Figure 47. Top: urinary ALT glucuronides (sum of ALT glucuronide and ALT-CAP glucuronide) concentration (ng/mL) plotted against hours after last treatment; middle: ratio of ALT glucuronides to urinary creatinine (U_{Cr}) plotted against hours after last treatment; blue: gilt A, magenta: gilt B, green: gilt C, mean values (n = 2) ± measurement uncertainty; bottom: mean of ALT glucuronides normalized to urinary creatinine (U_{Cr}) of urine samples from ALT-treated gilts grouped by time intervals after last treatment with individual values plotted as dots.

Unfortunately, it was not possible to fit an excretion function to this data since the number of samples per individual was too small. This real world study took place under routine breeding conditions without any additional interventions. Hence, only spontaneous spot urine was sampled. It was not possible to sample urine at fixed time points or with a predefined number of samples per individual animal. Furthermore, inter-individual variations in the excretion rate appear to be high, since urinary ALT glucuronides/U_{cr} values at 2–4 hours after the last treatment range from 3.86 µg/mg to 14.06 µg/mg. In contrast, the ratios of ALT glucuronides/U_{Cr} at approximately 7–8 hours only varied from 3.76–5.32 µg/mg. However, a pronounced decrease over time was observed.

In general, the results are in line with the literature with glucuronide conjugation being described as a major metabolic pathway for ALT [28]. Sulfate conjugates which have been already described as renale metabolites in horses [69] were not detected in the present study.
No urinary excretion data of ALT in gilts are available in the published scientific literature. The only urinary excretion data for ALT in gilts available was taken from scientific opinions from the CVMP. However, these data are limited and range from 20% of the total dose excreted via urine to 60% renal excretion [28]. The CVMP used a release rate of 25% (ALT or ALT conjugates) of the total dose of ALT via urine for the refinement of the predicted environmental concentrations in soil (PEC_{soil}) in their phase II environmental risk assessment [29]. After the application of radio-labeled ALT, 2% of the total radioactivity in urine was identified as ALT and 24% as ALT conjugates, corresponding to 0.5% to 6% of the total dose [29]. Since only spot urine was available, it is not possible to calculate the total amount of ALT excreted as glucuronides and compare the obtained data to the literature. In order to determine the total urinary excretion an administration of radio-labeled ALT as well as the collection of urine over a longer time period would be necessary.
In addition, excretion via feces was not investigated in the present study. However, it also needs to be taken into account for the assessment of the environmental release of ALT. The CVMP stated that in bile 6% of the radioactivity was identified as ALT and 14% as ALT conjugates, corresponding to 4.5% and 10.5% of the total dose [29].

The environmental impact of ALT needs to be discussed, since manure is often used as fertilizer for agricultural soils. In order to estimate worst case predicted environmental concentrations for ALT in Germany, a worst case scenario of 5 mg/kg ALT in manure was calculated in a literature study [44]. The observed excretion data in this thesis was in the same concentration range of up to 4.14 mg/L ALT in urine samples. Hence, the assumption taken for a

worst case scenario of the cited literature study can be confirmed with the data generated as part of this thesis. However, it has to be taken into account that under realistic conditions manure consists of urine and feces from many animals and only a minority of which are likely to have been treated with ALT. Therefore, a dilution of synthetic hormonal residues in manure cannot be excluded. A study from Biswas *et al.* demonstrated that manure from cattle treated with different hormonal substance, i.e., α-zearalanol, trenbolone, 17-β-estradiol and melengesterol acetate contained total residues from 9.9–26.7 μg/kg [179]. Yet, in the swine industry production phases are segregated and therefore, regional hot spots can occur, when manure from sow farms is used for fertilization.

In this study, only ALT glucuronide and ALT-CAP glucuronide were detected in urine samples at higher concentration levels, ALT was not detected and only traces of an unknown and for the first time observed isomer of ALT were observed. ALT conjugates may exhibit a hormonal effect after microbial deconjugation and thus, need to be considered in the environmental risk assessment. However, according to CVMP summary report in order to specify a MRL for ALT, the hormonal activity of ALT conjugates after deconjugation was 14% of the activity of ALT [28]. The CVMP stated further that the hormonal activity of the isomeric form of ALT (ALT-CAP) was very low [28], which is in contrast to the findings of Wammer *et al.* and Pflug *et al.* [87, 88]. Pflug *et al.* showed that the photo-isomer ALT-CAP exhibited 45–100% of the activity of ALT in activating estrogen receptors, progesterone receptor and pregnane X receptor [88].

The environmental fate of ALT is a topic of recent research. The octanol/water distribution coefficient ($\log K_{OW}$) of ALT was estimated at 3.54 by the US environmental protection agency (EPA) [180] and experimental values range from 3.44–3.74 [89, 181]. This indicates that sorption processes to soil or sediments are likely to occur. Sorption studies for ALT at laboratory scale with an estimated worst case concentration in manure of 5 mg/kg showed a high affinity to soil and an accumulation in soil, while the total recovery decreased over the course of the study (28 days) [44]. Experiments with ALT spiked manure resulted in higher sorption to soil matrix compared to experiments with ALT spiked directly to the added water. After 28 days the concentration of ALT in water was still at a level above the lowest observed effect concentration of progesterone in fathead minnows [44]. Yang *et al.* showed in sorption studies with different soil types that sorption of ALT occurred concurrently with abiotic-transformation catalyzed by soil-minerals [89]. It is important to keep in mind that ALT undergoes rapid photo-isomerization, which may have an impact on the sorption behavior

when manure is spread on agricultural fields. Investigations of the sorption behavior of the photo-isomers of ALT (ALT-CAP and ALT-CAP-OH) revealed that they have a reduced sorption affinity to sand-soil mixtures relative to parent drug, which is in line with their lower $\log K_{OW}$ of 3.25 and 2.88 [181]. Hence, the more polar photo-products exhibit a higher transport potential in soil-water systems [181]. However, studies haven shown that photolysis on soil surfaces is not only determined by the photolysis itself but is a function of layer thickness and transport processes [182].

The excretion data described above showed that no parent drug and only minor amounts of an isomer of the parent drug with similar physicochemical properties are present in the urine of ALT-treated animals. But one can hypothesize that more polar metabolites, such as glucuronides, are more likely to exhibit a low sorption to soil matrix and may leach faster into groundwater. However, it cannot be excluded that bacterial β-glucuronidases in soil render the parent drug and the photo-isomer after cleavage of the glucuronic acid moiety.

In conclusion, the targeted investigation of real world urine samples from ALT-treated gilts showed that glucuronide conjugates of ALT and its photo-isomerization product (ALT-CAP) are main urinary metabolites of ALT in gilts. Furthermore, another isomerization product of ALT was observed at trace level, which to the best of my knowledge has not yet been described in literature. The quantification of the ALT glucuronides was performed after enzymatic hydrolysis. The concentrations ranged from 1900 ng/mL to 4140 ng/mL, which is comparable to data reported for horses and worst-case calculations used for laboratory-scale sorption studies. The results showed further that it is crucial to adjust for the fluid balance of the animals at different time points in order to determine a concentration profile over time.

4.2.2 Non-targeted LC-HRMS analysis

Non-targeted LC-HRMS analysis was performed in order to investigate the potential of metabolomics workflows to (1) distinguish between ALT-treated gilts and non-medicated gilts and (2) to identify unknown metabolites of ALT.

Data integrity

In the course of data post-processing, the experimental data was filtered by removing all mass spectrometric features which were present in the reagent blank sample at an intensity of $> 1.0 \times 10^7$ and features with RSD > 25% in QC samples (all samples pooled). This de-

creased the total number of features from 7504 to 3497. The purpose of this step was to remove analytical noise and non-informative variables in the data set [134]. Variables with low repeatability are considered to be of low quality, since they may impede the identification of true biologically meaningful differences between the test groups since important biologically meaningful features could be lost in an abundance of analytical noise [183].

Subsequently, data quality was checked by means of the RSD of the internal standard ALT-d5 of the QC samples (all samples pooled) which were scattered throughout the analytical batch as well as the RSD of the internal standard ALT-d5 of all samples of the analytical batch. The RSD of ALT-d5 in the QC samples and in all samples were 4% and 12%, respectively. Hence, the analytical method was considered robust.

Multivariate statistical analysis of non-targeted data

After data pretreatment, the experimental data was subjected to PCA. As illustrated in Figure 48, the PCA scores plot revealed differences between samples from non-medicated gilts (control) and ALT-treated gilts (positive samples). Sample preparation and data processing workflows were considered robust, since all QC samples coincided in the scores plot. For practical purposes, it is necessary to apply a statistically defined criterion by which one can distinguish the two test groups. 0.9-prediction ellipses with an α error probability of 0.1 were employed. As can be seen in Figure 48 the prediction ellipses only overlapped slightly.

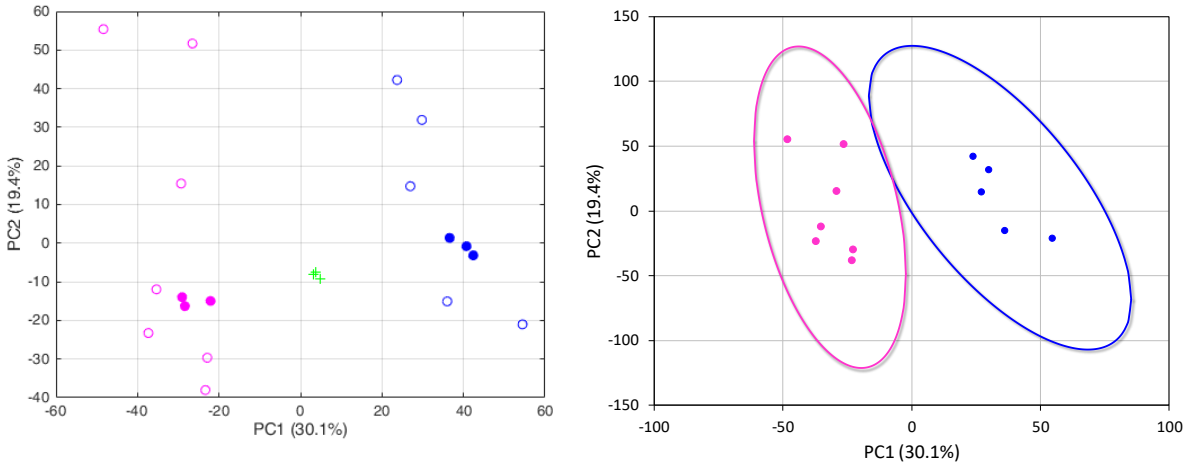


Figure 48. Left: 2D-scores plot of PCA of LC-HRMS data of urine samples; blue: control samples (non-medicated gilts), magenta: positive samples (ALT-treated gilts); green cross: pooled QC samples, blue filled: control QC samples, magenta filled: positive QC samples; right: 2D-scores plot of PCA of LC-HRMS data of urine samples; blue: control samples (non-medicated gilts) with 0.9-prediction ellipse, magenta: positive samples (ALT-treated gilts) with 0.9-prediction ellipse; PC: principal component with explained variance shown in brackets.

Metabolite and biomarker identification

Univariate statistical tests were used for metabolite and biomarker identification. The results of the top 25 mass spectrometric features based on their level of significance (adjusted *p*-value) are given in Table 19. All adjusted *p*-values of the top 25 features were highly significant (p-adj. < 3.1 x 10^{-9}) and the mean fold changes ranged from 0.001 for analyte *m/z* 827.2521 (@6.78 min) to 22,200 for analyte *m/z* 487.2328 (@6.16 min) which was identified at level 1 as ALT glucuronide. In order to visualize the results and aid pattern recognition a 2D heatmap was constructed (see Figure 49).

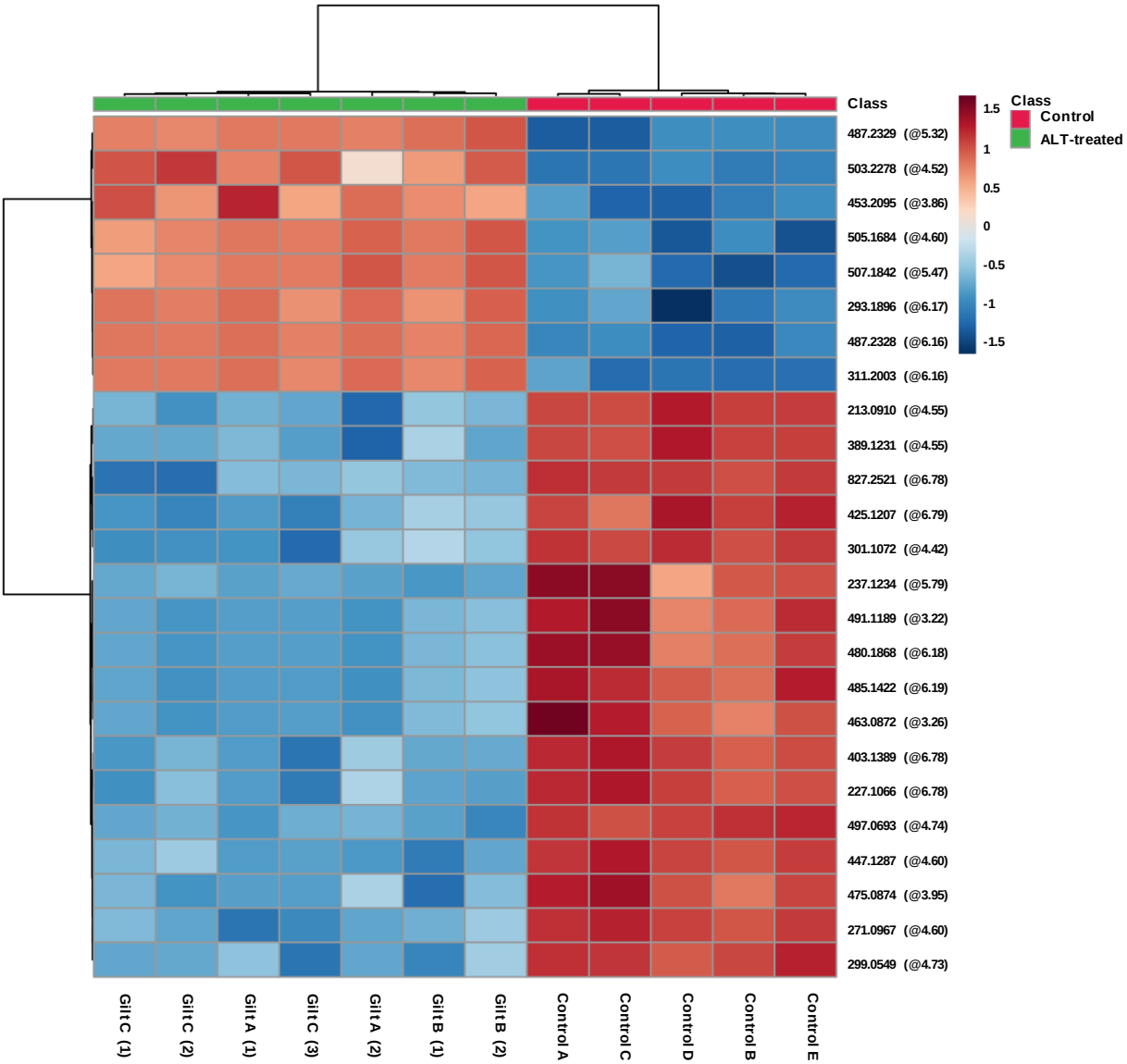


Figure 49. 2D heatmap in order to visualize the top 25 significant features of *Student's t*-test and to aid pattern recognition; 5 control gilts (A-E) and 3 ALT-treated gilts (A-C) at different sampling times (1–3, see Table 6) were clustered with dendograms at sample and analyte level.

Table 19. Top 25 significant features (*m/z* at retention time in min) with adduct determination, proposed molecular formulas and level of identification (ID) confidence for annotation according to [94]).

m/z at retention time (min)	Adduct	Proposed elemental composition	δ ppm	ID Level	Annotation	*p*-adj.	FC
311.2003 (@6.16)	$[M+H]^+$	$C_{21}H_{26}O_2$	0.3	1	ALT caused by in-source fragmentation	1.7×10^{-13}	1390
301.1072 (@4.42)	$[M+H]^+$	$C_{17}H_{16}O_5$	0.2	4	database hits for isoflavonone derivatives	2.3×10^{-13}	0.009
487.2328 (@6.16)	$[M+H]^+$	$C_{27}H_{34}O_8$	0.9	1	ALT glucuronide	2.6×10^{-13}	22200
447.1287 (@4.60)	$[M+H]^+$	$C_{22}H_{22}O_{10}$	0.3	3	glucuronide of $C_{16}H_{14}O_4$ (similar to isoflavone derivatives)	9.8×10^{-12}	0.020
487.2329 (@5.32)	$[M+H]^+$	$C_{27}H_{34}O_8$	0.9	1	ALT-CAP glucuronide	1.0×10^{-11}	4680
271.0967 (@4.60)	$[M+H]^+$	$C_{16}H_{14}O_4$	0.5	3	similar to isoflavone derivatives[1]	2.7×10^{-11}	0.012
485.1422 (@6.19)	$[M+H]^+$	$C_{18}H_{29}O_{13}P$	0.1	4		2.7×10^{-11}	0.131
497.0693 (@4.74)	$[M+H]^+$	$C_{24}H_{16}O_{12}$	-4.5	4		2.7×10^{-11}	0.014
403.1389 (@6.78)	$[M+H]^+$	$C_{21}H_{22}O_8$	0.1	3	glucuronide of $C_{15}H_{14}O_2$ (similar to stilbene derivative)	3.5×10^{-11}	0.055
227.1066 (@6.78)	$[M+H]^+$	$C_{15}H_{14}O_2$	0.1	3	similar to stilbene derivative[1]	7.2×10^{-11}	0.051
491.1189 (@3.22)	$[M+H]^+$	$C_{23}H_{22}O_{12}$	0.2	4	glucuronide of $C_{17}H_{14}O_6$	1.3×10^{-10}	0.095
213.0910 (@4.55)	$[M+H]^+$	$C_{14}H_{12}O_2$	0.2	4	database hits for stilbene derivative[1]	1.4×10^{-10}	0.014
507.1842 (@5.47)				5	unknown glucuronide	1.5×10^{-10}	339
827.2521 (@6.78)				5		2.4×10^{-10}	0.001
389.1231 (@4.55)	$[M+H]^+$	$C_{20}H_{20}O_8$	-0.1	4	glucuronide of $C_{14}H_{12}O_2$	3.7×10^{-10}	0.013
480.1868 (@6.18)				5		3.7×10^{-10}	0.098
505.1684 (@4.60)				5	unknown glucuronide	3.7×10^{-10}	554
425.1207 (@6.79)				5		3.7×10^{-10}	0.090
237.1234 (@5.79)	$[M+H]^+$	$C_{12}H_{16}N_2O_3$	0.3	4		4.1×10^{-10}	0.001
453.2095 (@3.86)	$[M+H]^+$	$C_{19}H_{28}N6O_7$	-0.7	4		7.6×10^{-10}	10.5
463.0872 (@3.26)	$[M+H]^+$	$C_{21}H_{18}O_{12}$	0.2	4	glucuronide of $C_{15}H_{10}O_6$	9.2×10^{-10}	0.112
503.2278 (@4.52)	$[M+H]^+$	$C_{27}H_{34}O_9$	0.7	3	ALT-OH glucuronide	1.7×10^{-9}	10.4
299.0549 (@4.73)	$[M+H]^+$	$C_{16}H_{10}O_6$	-0.2	4	database hits for isoflavonoid derivatives	2.5×10^{-9}	0.010
475.0874 (@3.95)	$[M+H]^+$	$C_{22}H_{18}O_{12}$	0.5	4	glucuronide of $C_{16}H_{10}O_6$	2.6×10^{-9}	0.012
293.1896 (@6.17)	$[M-H_2O+H]+$	$C_{21}H_{26}O_2$	-1.3	1	ALT caused by in-source fragmentation	3.1×10^{-9}	159

FC: fold change: positive fold changes indicate higher levels in ALT-treated gilts, *p*-adj.: multiple test corrected *p*-value based on the Benjamini Hochberg false discovery rate; [1]in-source fragmentation of the respective glucuronide; ALT-CAP: cycloaddition product of ALT.

Eight of the top 25 significant features were increased in the medicated group and 5 of these 8 increased features were identified as metabolites and in-source fragments of ALT metabolites. The metabolite with the highest change was the glucuronide of ALT (*m/z* 487.2328 @ 6.16 min), followed by the glucuronide of ALT-CAP (*m/z* 487.2329 @ 5.32 min). Based on the aligned retention time, the features *m/z* 311.2003 $[M+H]^+$ and *m/z* 293.1896 $[M-H_2O+H]^+$ were identified as in-source fragments of ALT glucuronide. The mass spectrometric feature *m/z* 503.2278 at 4.52 min was identified as glucuronide conjugate of a hydroxy-ALT at level 3 (see Figure 50). The characteristic fragments *m/z* 199.1101 and 159.0806 for ALT indicated that the hydroxyl group is positioned in ring D.

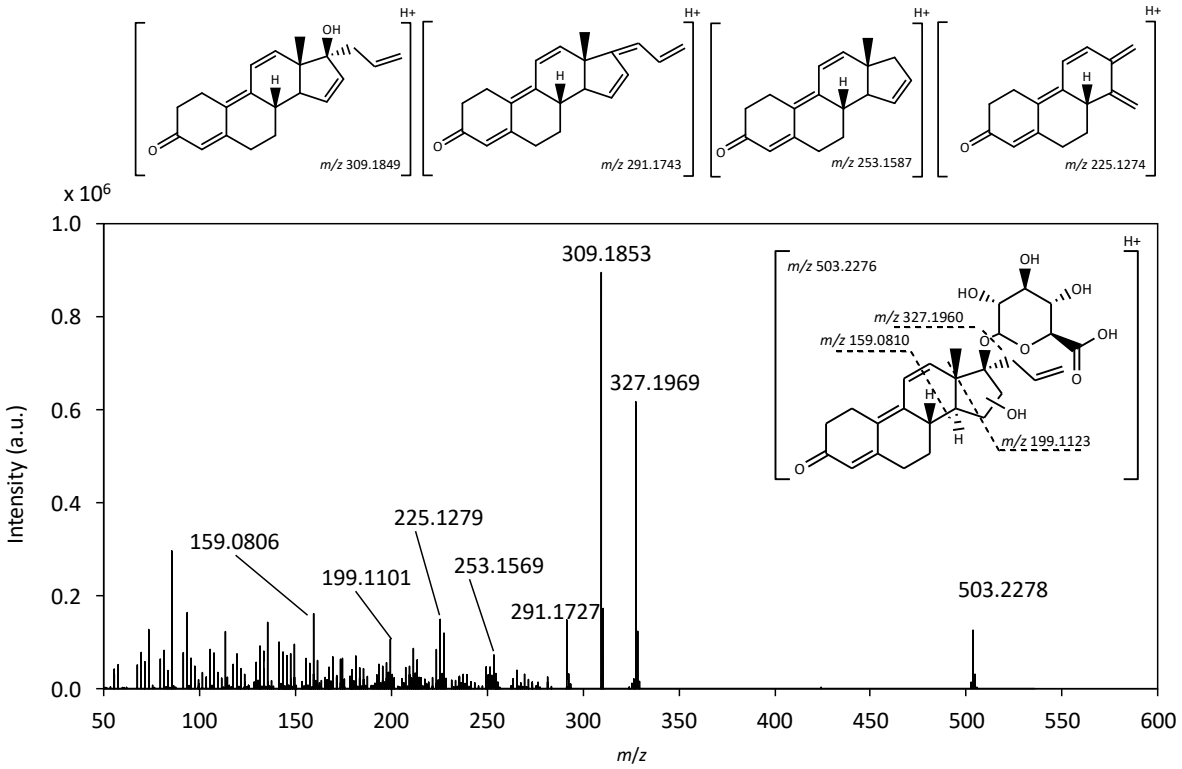


Figure 50. Representative product ion spectrum of glucuronide conjugate of hydroxy-ALT (*m/z* 503.2278 $[M+H]^+$ @ 4.52 min) with proposal of product ion fragments.

The mass spectrometric features *m/z* 507.1842 at 5.47 min *m/z* 505.1684 at 4.60 min could not be further identified except for a characteristic neutral loss of 176.0309 which corresponds to a loss of glucuronic acid. A sum formula of $C_{19}H_{28}N_6O_7$ was proposed for *m/z* 453.2095 (@3.86 min) which was also increased in the medicated group.

Interestingly, 17 of the 25 top significant features showed elevated intensities in the control group. Six features correspond to glucuronide conjugates given the characteristic neutral loss of 176.0309. The mass spectrometric features *m/z* 271.0967 at 4.60 min and *m/z* 227.1066 at

6.78 min produced mass spectral library hits for isoflavone derivatives (5,7-dihydroxy-3-(4-hydroxyphenyl)-4H-chromen-4-one) and stilbene derivatives (4,4-prop-1-ene-1,2-diyldiphenol), respectively, by the integrated library "mzCloud" of Compound Discoverer 3.1. Further database hits in ChemSpider also suggest isoflavonoid derivatives. This may lead to the assumption that the metabolism of nutrients, i.e., phytoestrogens, is altered under ALT treatment. However, further studies are needed to support this hypothesis.

Pathway mapping

We additionally performed an "MS peak to pathway" mapping in MetaboAnalyst 4.0 to complement manual biomarker identification. This procedure applies the *mummichog* algorithm to map mass spectrometric features to functional activities by leveraging the collective insights of metabolic pathways [144]. 1871 significant mass spectrometric features (p-value $< 1.0 \times 10^{-3}$) were used for pathway analysis and 310 empirical compounds were annotated by "MS peaks to pathway". The significant hits per pathway were checked manually for plausibility (chromatographic retention and adduct identification). The pathway enrichment factor is calculated as the ratio of the number of significant pathway hits and the expected number of randomly matched compound hits [134]. The significance of the pathway enrichment is determined by the p-value of Fisher's exact test (FET) [144]. In order to take the size of each pathway into account, an adjusted p-value per pathway is calculated based on permutation-testing [144].

Tryptophan metabolism (FET p-value: 0.003, adjusted p-value: 0.001, enrichment factor: 1.42) and leukotriene metabolism (FET p-value: 0.20, adjusted p-value: 0.003, enrichment factor: 1.52) followed by C21-steroid hormone biosynthesis and metabolism (FET p-value: 0.23, adjusted p-value: 0.003, enrichment factor: 1.22) were the three most enriched pathways (see Figure 51). A second algorithm was applied using "Metabolika Pathways" in Compound Discoverer 3.1. This algorithm also revealed several mass spectrometric hits for the tryptophan metabolism and the C21-steroid biosynthesis. Hence, these two pathways were investigated further. Significant hits that were independently found by both algorithms were highlighted in the pathways in Figure 52 and Figure 53. However, it is crucial to keep in mind that the pathway mapping is based on database hits only and no analytes were unambiguously identified via reference standards. Nonetheless, it is noteworthy that non-targeted explorative methods suggest metabolic changes that can be explained by the mechanism of action of ALT.

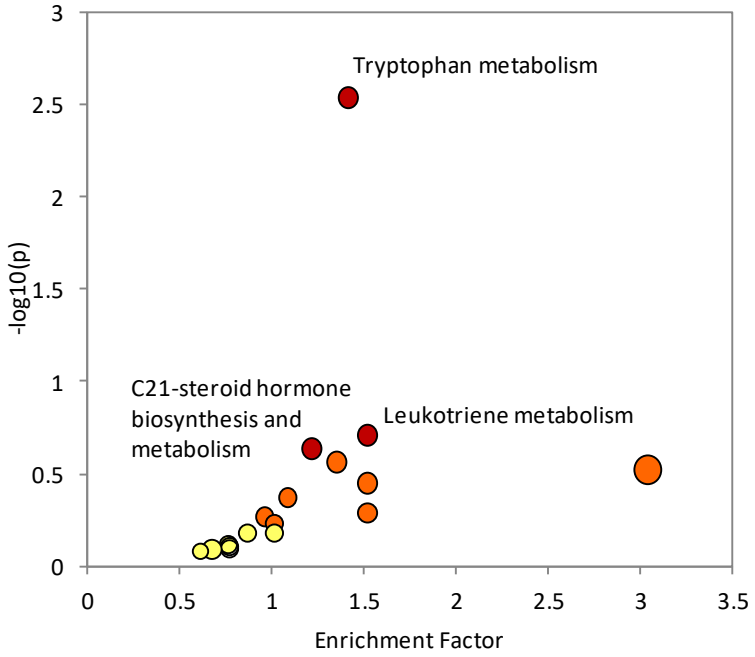


Figure 51. Pathway mapping summary: all matched pathways are displayed as circles. The color and size of each circle corresponds to its Fisher's exact test (FET) *p*-value and enrichment factor, respectively. Small yellow circle represent non-significant hits, while larger, red circles represent significant pathway enrichments.

Tryptophan is an essential amino acid for humans and pigs. It is involved in protein synthesis and the synthesis of the neurotransmitter serotonin. An effect of the menstrual cycle on tryptophan as well as kynurenine (a degradation product of tryptophan) is described in women [184, 185]. In the follicular phase lower plasma levels of tryptophan as well as lower urinary kynurenine levels have been observed [184, 185]. ALT exhibits progestomimetic actions and blocks the menstrual cycle in the follicular phase. Hence, lower levels of degradation products of tryptophan in the ALT-treated group can reasonably be explained by ALT's mode of action.

Decreased levels of progesterone and other sex steroids can be explained by the progestomimetic action of ALT, since it suppresses the secretion of the gonadotropins FSH and LH from the anterior pituitary via a negative feedback mechanism and thus inhibits the sex steroid biosynthesis [51-53]. Furthermore, the samples were taken at the end of the ALT-treatment, when luteolysis is progressed and endogenous progesterone levels are at their lowest levels [24].

These results illustrate the exceptional potential of non-targeted LC-HRMS analyses, since they render valuable information for metabolite and biomarker discovery. One major ad-

vantage of the described approach is that real world samples could be used which were taken non-invasively and under realistic conditions for ALT application. Thus, metabolomics can be a powerful tool in the context of the principle of the 3Rs (Replacement, Reduction and Refinement) for an ethical use of animal testing. The comprehensive analysis of metabolites may partly replace more invasive animal testing by giving insights in changes of the metabolome of the animal.

In conclusion, the non-targeted analysis of real world urine samples showed a clear separation of ALT-treated gilts and control animals. ALT glucuronides and their in-source fragmentation products were in the top 25 significant hits in order to distinguish the two test groups. Furthermore, a hydroxylated ALT glucuronide could be identified at level 3 via its product ion scan. In the control group, several mass spectrometric hits were significantly increased. Database and mass spectrometric library hits suggested that isoflavonoid derivatives and stilbene derivatives were increased in the control group. This may imply that the metabolism of nutrients, i.e., phytoestrogens, is altered under ALT treatment. However, further studies are needed to support this hypothesis. Pathway analysis showed differences in tryptophan degradation and C21-steroid biosynthesis. These changes can be explained by the mechanism of action of the hormonally active substance ALT.

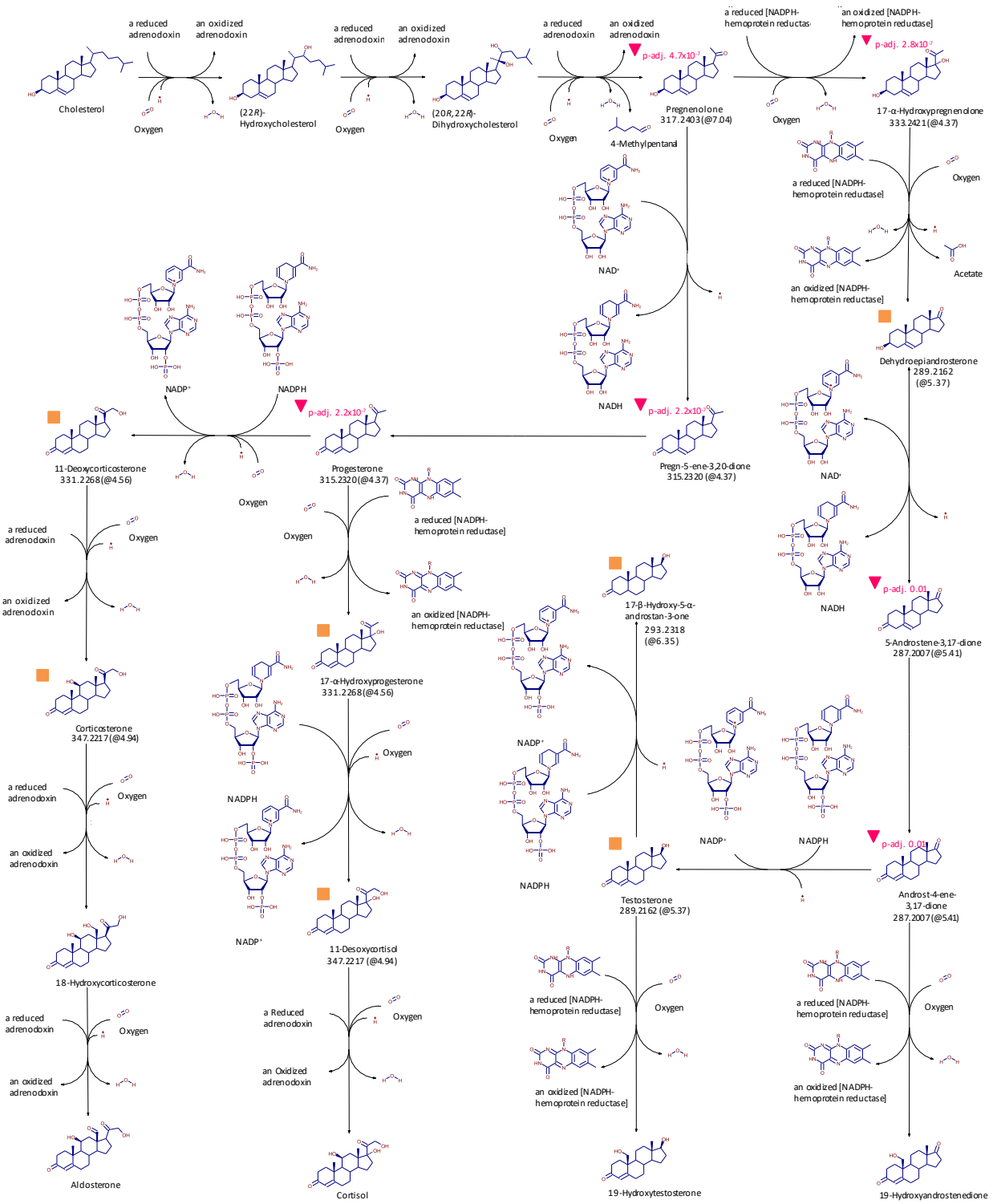


Figure 52. C21-steroid biosynthesis with assigned mass spectrometric hits based on database hits; triangle indicates significant decrease, square indicates hits without significant increase or decrease, *p*-adj.: multiple test corrected p-value based on the Benjamini Hochberg false discovery rate.

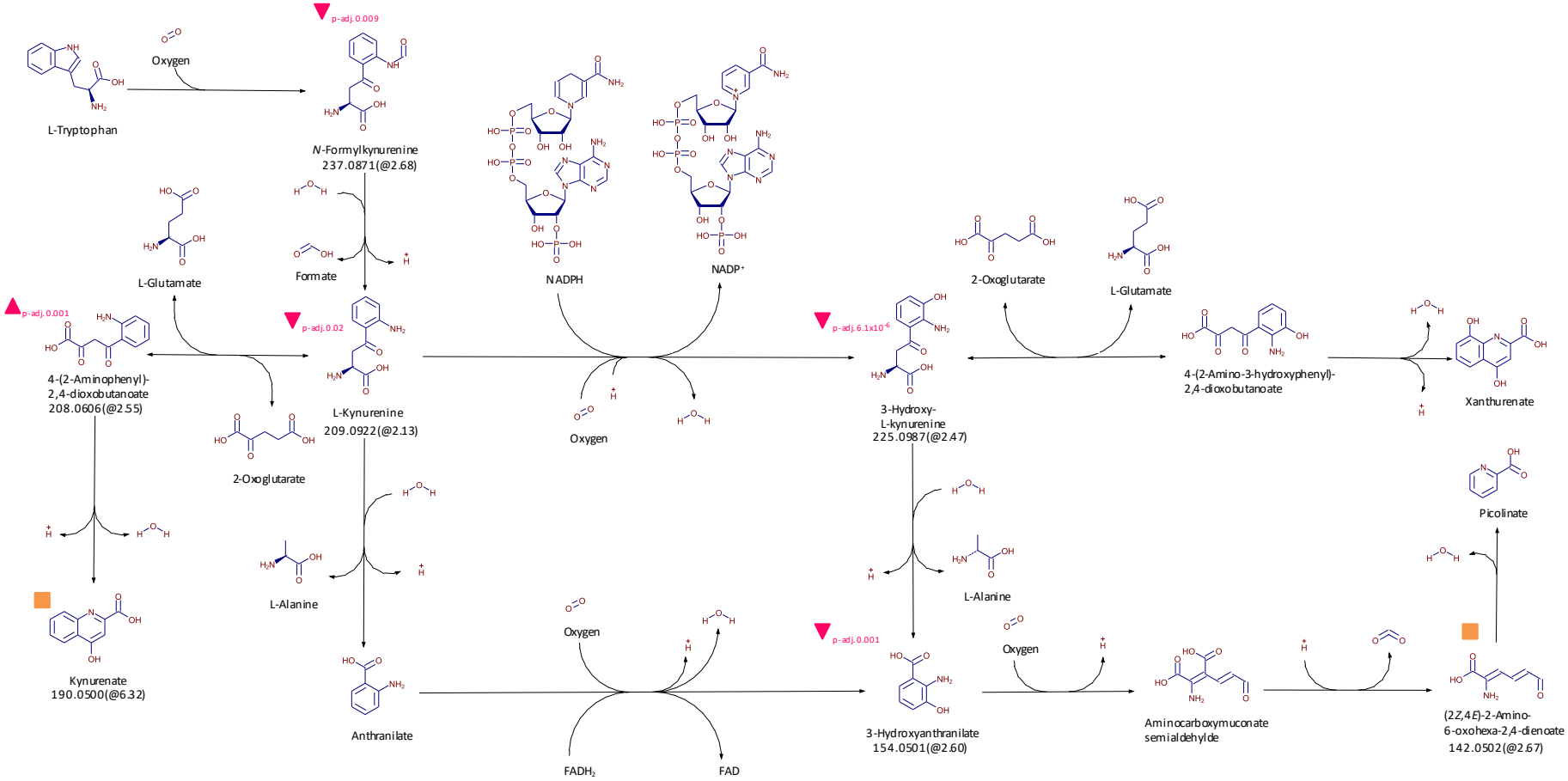


Figure 53. Tryptophan degradation with assigned mass spectrometric hits based on database hits; triangle indicates significant decrease or increase, square indicates hits without significant increase or decrease, *p*-adj.: multiple test corrected *p*-value based on the Benjamini Hochberg false discovery rate.

5 References

1. Liesenfeld, S.; Steliopoulos, P.; Hamscher, G., Comprehensive metabolomics analysis of nontargeted LC-HRMS data provides valuable insights regarding the origin of veterinary drug residues. *Journal of Agricultural and Food Chemistry*, **2020**, 68 (44), 12493-12502.

2. Regulation (EU) 2019/6 of the European Parliament and of the Council of 11 December 2018 on veterinary medicinal products and repealing Directive 2001/82/EC (OJ L 4, 7.1.2019, p. 43–167).

3. Directive 2001/82/EC of the European Parliament and of the Council of 6 November 2001 on the Community code relating to veterinary medicinal products (OJ L 311, 28.11.2001, p.1), as lastly amended by Regulation (EC) No. 596/2009 of the European Parliament and of the Council of 18 June 2009 (OJ L 188, 18.7.2009, p.14).

4. Arzneimittelgesetz in der Fassung der Bekanntmachung vom 12. Dezember 2005 (BGBl. I, S. 3394), das zuletzt durch Artikel 2 Absatz 1 des Gesetzes vom 25. Juni 2020 (BGBl. I S. 1474) geändert worden ist.

5. Regulation (EC) No 470/2009 of the european parliament and of the council of 6 May 2009 laying down Community procedures for the establishment of residue limits of pharmacologically active substances in foodstuffs of animal origin, repealing Council Regulation (EEC) No 2377/90 and amending Directive 2001/82/EC of the European Parliament and of the Council and Regulation (EC) No 726/2004 of the European Parliament and of the Council (OJ L 152, 16.6.2009, p.11-22).

6. Commission Regulation (EU) No 37/2010 of 22 December 2009 on pharmacologically active substances and their classification regarding maximum residue limits in foodstuffs of animal origin (OJ L 015, 20.1.2010, p. 1), as lastly amended by Commission Implementating Regulation (EU) 2020/1712 of 16 November 2020 (OJ L 384, 17.11.2020, p.3).

7. Lebensmittel- und Futtermittelgesetzbuch in der Fassung der Bekanntmachung vom 3. Juni 2013 (BGBl. I S.1426), das durch Artikel 97 der Verordnung vom 19. Juni 2020 (BGBl. I S. 1328) geändert worden ist.

8. Council Directive 96/23/EC of 29 April 1996 on measures to monitor certain substances and residues thereof in live animals and animal products and repealing Directives 85/358/EEC and 86/469/EEC and Decisions 89/187/EEC and 91/664/EEC, repealed by Regulation (EU) 2017/625 of the European Parliament and of the Council of 15 March 2017 (OJ L 125, 23.5.1996, p.10); translational measures according to Article 150 No. 1 of Regulation (EU) 2017/625 apply.

9. Federal Office of Consumer Protection and Food Safety, *National residue control plan (NRCP) and import control plan (ICP) for food of animal origin.* https://www.bvl.bund.de/EN/Tasks/01_Food/01_tasks/02_OfficialFoodControl/05_NRCP/nrcp_node.html (accessed Apr 25, 2021).

10. Tierische Lebensmittel-Überwachungsverordnung in der Fassung der Bekanntmachung vom 3. September 2018 (BGBl. I S. 1358), die durch Artikel 3 der Verordnung vom 19. Juni 2020 (BGBl. I S. 1480) geändert worden ist.

11. Bundesamt für Verbraucherschutz und Lebensmittelsicherheit, *BVL-Report · 14.5 Berichte zur Lebensmittelsicherheit.* 2020. https://www.bvl.bund.de/SharedDocs/Berichte/05_Weitere_Berichte_LM_Sicherheit/Berichte_zur_Lebensmittelsicherheit_2018.pdf;jsessionid=83DABC4ABD2E08F58179BF514D2781EF.1_cid360?__blob=publicationFile&v=3 (accessed Apr 25, 2021).

12. Merle, R.; Hajek, P.; Käsbohrer, A.; Hegger-Gravenhorst, C.; Mollenhauer, Y.; Robanus, M.; Ungemach, F. R.; Kreienbrock, L., Monitoring of antibiotic consumption in livestock: a German feasibility study. *Preventive Veterinary Medicine*, **2012**, 104 (1-2), 34-43.

13. Hamscher, G.; Bachour, G., Veterinary drugs in the environment: current knowledge and challenges for the future. *Journal of Agricultural and Food Chemistry*, **2018**, 66 (4), 751-752.

14. DIMDI-Arzneimittelverordnung vom 24. Februar 2010 (BGBl. I S. 140), die zuletzt durch Artikel 1 der Verordnung vom 13. Juli 2020 (BGBl. I S. 1692) geändert worden ist.

15. Bundesamt für Verbraucherschutz und Lebensmittelsicherheit, *Abgabe an Anitbiotika in der Tiermedizin sinkt weiter.* 2020. https://www.bvl.bund.de/SharedDocs/Pressemitteilungen/05_tierarzneimittel/2020/2020_07_29_PI_Antibiotikaabgabe.html (accessed Apr 25, 2021).

16. World Health Organization (WHO), *Critically important antimicrobials for human medicine, 6th revision*. 2018. https://www.who.int/foodsafety/publications/antimicrobials-sixth/en/ (accessed Apr 25, 2021).

17. German Federal Ministry of Health, *DART 2020 Fighting antibiotic resistance for the good of both humans and animals.* 2020. https://www.bundesgesundheitsministerium.de/fileadmin/Dateien/5_Publikationen/Gesundheit/Berichte/BMG_DART_2020_Bericht_en.pdf (accessed Apr 25, 2021).

18. Bundesamt für Verbraucherschutz und Lebensmittelsicherheit, Bekanntmachung des Medians und des dritten Quartils der vom 1. Juli 2019 bis 31. Dezember 2019 erfassten bundesweiten betrieblichen Therapiehäufigkeiten für Mastrinder, Mastschweine, Masthühner und Mastputen nach § 58c Absatz 4 des Arzneimittelgesetzes (BAnz AT 31.03.2020 B11).

19. Kennedy, D. G.; Cannavan, A.; McCracken, R. J., Regulatory problems caused by contamination, a frequently overlooked cause of veterinary drug residues. *Journal of Chromatography A*, **2000**, 882 (1-2), 37-52.

20. McCaughey, W. J.; Elliott, C. T.; Crooks, S. R., Carry-over of sulphadimidine in the faeces and urine of pigs fed medicated feed. *Veterinary Record*, **1990**, 126 (15), 351-4.

21. Elliott, C. T.; McCaughey, W. J.; Crooks, S. R.; McEvoy, J. D., Effects of short term exposure of unmedicated pigs to sulphadimidine contaminated housing. *Veterinary Record*, **1994**, 134 (17), 450-1.

22. Hamscher, G.; Pawelzick, H. T.; Sczesny, S.; Nau, H.; Hartung, J., Antibiotics in dust originating from a pig-fattening farm: a new source of health hazard for farmers*? Environmental Health Perspectives*, **2003**, 111 (13), 1590-1594.

23. Food and Agriculture Ogranization of the United Nations (FAO). *Residues of some veterinary drugs in animals and foods*, in *Food and Nutrition Paper, Monograph prepared by the 36th meeting of the Joint FAO/WHO Expert Committee on Food Additives,* Vol. 41/3, 1990. https://www.fao.org/docrep/014/T0459E/T0459E.pdf (accessed Apr 25, 2021).

24. Mutschler, E.; Geisslinger, G.; Kroemer, H. K.; Ruth, P.; Schäfer-Korting, M., *Mutschler Arzneimittelwirkungen: Lehrbuch der Pharmakologie und Toxikologie*, Wiss. Verlag-Ges.: Stuttgart, Germany, 2008.

25. United Nations Environment Programme and the World Health Organization, *State of the science of endocrine sisrupting chemicals – 2012*, 2013. https://www.who.int/ceh/publications/endocrine/en/ (accessed on Apr 20, 2021).

26. Studio, G. T.; Aktories, K.; Förstermann, U.; Hofmann, F. B.; Starke, K., *Allgemeine und spezielle Pharmakologie und Toxikologie: Begründet von W. Forth, D. Henschler, W. Rummel*, Elsevier Health Sciences Germany: München, Germany, 2013.

27. Council Directive 96/22/EC of 29 April 1996 concerning the prohibition on the use in stockfarming of certain substances having a hormonal or thyrostatic action and of beta-agonists, and repealing Directives 81/602/EEC, 88/146/EEC and 88/299/EEC 1996 (OJ L 125, 23.5.1996, p.3), as lastly amended by Directive 2008/97/EC of the European Parliament and of the Council (OJ L 318, 28.11.2008, p.9).

28. Committee for Medicinal Products for Veterinary Use, *European public MRL assessment report (EPMAR) altrenogest (equidae and porcine species)*, EMA/CVMP/487477/2011, 2012. https://www.ema.europa.eu/documents/mrl-report/altrenogest-equidae-porcine-species-european-public-maximum-residue-limit-assessment-report-epmar_en.pdf. (accessed on Apr 25, 2021).

29. European Medicines Agency Veterinary Medicines Division, *Outcome of a referral procedure under Article 35 of Directive 2001/82/EC (EMEA/V/A/095), scientific conclusions of the Committee (Annex II)*, EMA/462829/2016, 2016. https://www.ema.europa.eu/en/medicines/veterinary/referrals/altrenogest (accessed on Apr 25, 2021).

30. Bundesanstalt für Landwirtschaft und Ernährung, *Bericht zur Markt- und Versorgungslage Fleisch 2020*, 2020. https://www.ble.de/DE/BZL/Daten-Berichte/Fleisch/fleisch_node.html (accessed on Apr 20, 2021).

31. Statistisches Bundesamt destatis.de GENESIS-Online Datenbank, *Tabelle 41311-0001 Gehaltene Tiere: Deutschland, Jahre, Tierarten*, 2020. https://www-genesis.destatis.de/genesis/online (accessed on Apr 25, 2021).

32. Statistisches Bundesamt destatis.de GENESIS-Online Datenbank, *Tabelle 41311-0003 Betriebe: Deutschland, Jahre, Tierarten*, 2020. https://www-genesis.destatis.de/genesis/online (accessed on Apr 25, 2021).

33. Statistisches Bundesamt destatis.de GENESIS-Online Datenbank, *Tabelle 41311-0004 Betriebe: Bundesländer, Jahre, Tierarten*, 2020. https://www-genesis.destatis.de/genesis/online (accessed on Apr 25, 2021).

34. Statistisches Bundesamt destatis.de GENESIS-Online Datenbank, *Tabelle 41311-0002 Gehaltene Tiere: Bundesländer, Jahre, Tierarten*, 2020. https://www-genesis.destatis.de/genesis/online (accessed on Apr 25, 2021).

35. Kraeling, R. R.; Dziuk, P. J.; Pursel, V. G.; Rampacek, G. B.; Webel, S. K., Synchronization of estrus in swine with allyl trenbolone (RU-2267*). Journal of Animal Science*, **1981**, 52, 831-835.

36. Stark, M., Duldungsverhalten, Ovulationsverlauf und Konzeptionsergebnisse von Jung- und Altsauen nach Ovulationssynchronisation in verschiedenen Behandlungsvarianten. Dissertation, Universität Leipzig, Germany, 1999.

37. Beckjunker, J., Untersuchungen zur Nutzung von Altrenogest (Regumate®) und Gonadotropinen zur Zyklussteuerung von Alt- und Jungsauen mit negativem Trächtigkeitsbefund. Dissertation, Universität Leipzig, Germany, 2007.

38. Brüssow, K. P.; Schneider, F.; Kanitz, W.; Ratky, J.; Kauffold, J.; Wahner, M., Studies on fixed-time ovulation induction in the pig. *Society of Reproduction and Fertility supplement*, **2009**, 66, 187-95.

39. Schnurrbusch, U.; Hühn U., *Fortpflanzungssteuerung beim weiblichen Schwein*, Gustav Fischer Verlag: Stuttgart, Germany, 1994.

40. Brüssow, K.-P.; Wähner, M., Biotechnische Fortpflanzungssteuerung beim weiblichen Schwein. *Züchtungskunde*, **2005**, 77 (2/3), 157-170.

41. Hörning, B., *Zum Einsatz von Hormonen in der intensiven Sauenhaltung – Kurzfassung – Studie im Auftrag des Bund für Umwelt und Naturschutz Deutschland (BUND)*, Berlin, Germany, 2013. https://www.agrarheute.com/sites/default/files/media/597710/597710.pdf (accessed on Apr 25, 2021).

42. Hühn, U., Biotechnische Fortpflanzungssteuerung in der Sauenhaltung. *Veterinär Spiegel*, **2005**, 2, 41-43.

43. Soede, N. M.; Langendijk, P.; Kemp, P., Reproductive cycles in pigs. *Animal Reproduction Science*, **2011**, 124 (3–4), 251-258.

44. Petri, M.S., *Untersuchung zum Verhalten von Schwermetallen, Antibiotika und Hormonen in landwirtschaftlich genutztem Boden*. Dissertation, Justus-Liebig-Universität Gießen, Germany, 2019.

45. Frey, H. H.; Löscher, W., *Lehrbuch der Pharmakologie und Toxikologie für die Veterinärmedizin: 131 Tabellen*, Georg Thieme Verlag: Stuttgart, Germany, 2010.

46. Ganten, D.; Köhrle, J.; Ruckpaul, K., *Molekularmedizinische Grundlagen von para- und autokrinen Regulationsstörungen*, Springer-Verlag: Heidelberg, Germany, 2006.

47. Brinkley, H. J., Endocrine signaling and female reproduction. *Biology of Reproduction*, **1981**, 24 (1), 22-43.

48. Noguchi, M.; Yoshioka, K.; Itoh, S.; Suzuki, C.; Arai, S.; Wada, Y.; Hasegawa, Y.; Kaneko, H., Peripheral concentrations of inhibin A, ovarian steroids, and gonadotropins associated with follicular development throughout the estrous cycle of the sow. *Reproduction*, **2010**, 139 (1), 153-161.

49. Peters, A. R., Endocrine manipulation--toxicological frontiers. *Journal of Reproduction and Fertility / Supplement*, **1992**, 45, 193-201.

50. Machnik, M.; Hegger, I.; Kietzmann, M.; Thevis, M.; Guddat, S.; Schänzer, W., Pharmacokinetics of altrenogest in horses. *Journal of Veterinary Pharmacology and Therapeutics*, **2007**, 30 (1), 86-90.

51. Intervet Deutschland GmbH, *Fachinformation in Form der Zusammenfassung der Merkmale des Tierarzneimittels Regumate 4 mg/ml Lösung zum Eingeben für Schweine (Jungsauen); Altrenogest*. Unterschleißheim, Germany, 2016.

52. van Leeuwen, J. J. J.; Martens, M. R. T. M.; Jourquin, J.; Driancourt, M. A.; Kemp, B.; Soede, N. M., Variation in LH pulsatility during 24 h after a postweaning altrenogest treatment in relation to follicle development in primiparous sows. *Animal Reproduction Science*, **2011**, 126 (1), 101-107.

53. Guthrie, H. D.; Bolt, D. J., Pituitary and ovarian hormone secretion and ovulation in gilts injected with gonadotropins during and after oral administration of progesterone agonist (altrenogest). *Biology of Reproduction*, **1985**, 33 (3), 679-689.

54. Kraeling, R. R.; Webel, S. K., Current strategies for reproductive management of gilts and sows in North America. *Journal of Animal Science and Biotechnology*, **2015,** 6 (1), 3.

55. Soede, N. M.; Bouwman, E. G.; Langendijk, P.; Van Der Laan, I.; Kanora, A.; Kemp, B., Follicle development during luteal phase and altrenogest treatment in pigs. *Reproduction in Domestic Animals*, **2007**, 42 (3), 329-332.

56. Stevenson, J. S.; Davis, D. L., Estrous synchronization and fertility in gilts after 14- or 18-day feeding of altrenogest beginning at estrus or diestrus. *Journal of Animal Science*, **1982**, 55, 119-123.

57. Redmer, D. A.; Day, B. N., Ovarian activity and hormonal patterns in gilts fed allyl trenbolone. *Journal of Animal Science*, **1981**, 53, 1088-1094.

58. Martinat-Botté, F.; Bariteau, F.; Forgerit, Y.; Macar, C.; Poirier, P.; Terqui, M., Synchronization of oestrus in gilts with altrenogest: effects on ovulation rate and foetal survival. *Animal Reproduction Science*, **1995**, 39 (4), 267-274.

59. van Leeuwen, J. J. J.; Martens, M. R. T. M.; Jourquin, J.; Driancourt, M. A.; Wagner, A.; Kemp, B.; Soede, N. M., Follicle size and reproductive hormone profiles during a post-weaning altrenogest treatment in primiparous sows. *Reproduction, Fertility and Development*, **2015,** 27 (2), 304-312.

60. Martinat -Botté, F.; Bariteau, F.; Forgerit, Y.; Macar, C.; Poirier, P.; Terqui, M., Control of reproduction with a progestagen—altrenogest (Regumate) in gilts and at weaning in primiparous sows: effect on fertility and litter size. *Reproduction in Domestic Animals*, **1994,** 29 (4), 362-365.

61. Guthrie, H. D.; Meckley, P. E.; Young, E. P.; Hartsock, T. G., Effect of altrenogest and lutalyse on parturition control, plasma progesterone, unconjugated estrogen and 13,14-dihydro-15-ketoprostaglandin F2α in sows. *Journal of Animal Science*, **1987,** 65, 203-211.

62. van Leeuwen, J. J. J.; Williams, S. I.; Kemp, B.; Soede, N. M., Post-weaning altrenogest treatment in primiparous sows; the effect of duration and dosage on follicular development and consequences for early pregnancy. *Animal Reproduction Science*, **2010**, 119 (3-4), 258-64.

63. Fernández, L.; Diez, C.; Ordóñez J. M.; Carbajo, M., Reproductive performance in primiparous sows after postweaning treatment with a progestagen. *Journal of Swine Health and Production*, **2005**, 13, 28–30.

64. Committee for Veterinary Medicinal Products, *Altrenogest - summary report*, EMA/MRL/831/01/FINAL, 2002. https://www.ema.europa.eu/en/medicines?search_api_views_fulltext=altrenogest (accessed on Apr 25, 2021).

65. Food and Drug Administration, *Freedom of information summary, supplemental original new animal drug application,* NADA 141-222, 2003. https://www.federalregister.gov/documents/2003/10/31/03-27390/new-animal-drugs-altrenogest (accessed on Apr 25, 2021).

66. Xiao, H.; Sun, P.; Sun, F.; Qiu, J.; Wang, J.; Wang, J., Pharmacokinetics of altrenogest in gilts. *Journal of Veterinary Pharmacology and Therapeutics*, **2019**, 42 (6), 660-664.

67. Li, Y.; Yang, H.; Xia, L.; Wang, S.; Bu, S., Comparative pharmacokinetic study of two kinds of altrenogest oral solutions for sows. *Animal Biotechnology*, **2020**, Mar 17,1-7.

68. Li, Y.; Yang, H.; Qiu, S.; Wang, S.; Xia, L.; Bu, S., Investigations on residual elimination of altrenogest oral solution in pigs and the withdrawal time. Animal *Biotechnology*, **2019**, Dec 08, 1-8.

69. Lampinen-Salomonsson, M.; Beckman, E.; Bondesson, U.; Hedeland, M., Detection of altrenogest and its metabolites in post administration horse urine using liquid chromatography tandem mass spectrometry—increased sensitivity by chemical derivatisation of the glucuronic acid conjugate. *Journal of Chromatography B*, **2006**, 833 (2), 245-256.

70. Sumpter, J. P., Feminized responses in fish to environmental estrogens. *Toxicology Letters*, **1995,** 82–83, 737-742.

71. Kolodziej, E. P.; Harter T.; Sedlak, D. L., Dairy wastewater, aquaculture, and spawning fish as sources of steroid hormones in the aquatic environment. *Environmental Science & Technology*, **2004,** 38 (23), 6377-6384.

72. Purdom, C. E.; Hardiman, P. A.; Bye, V. V. J.; Eno, N. C.; Tyler, C. R.; Sumpter, J. P., Estrogenic effects of effluents from sewage treatment works. *Chemistry and Ecology*, **1994,** 8 (4), 275-285.

73. Tyler, C. R.; Jobling, S., Roach, sex, and gender-bending chemicals: the feminization of wild fish in English rivers. *BioScience*, **2008,** 58 (11), 1051-1059.

74. Gross-Sorokin, M. Y.; Roast, S. D.; Brighty, G. C.; Assessment of feminization of male fish in English rivers by the environment agency of England and Wales. *Environmental Health Perspectives*, **2006**, 114 (Suppl 1), 147-151.

75. Jenkins, R.; Angus, R. A.; McNatt, H.; Howell, W. M.; Kemppainen, J. A.; Kirk, M.; Wilson, E. M., Identification of androstenedione in a river containing paper mill effluent. *Environmental Toxicology and Chemistry*, **2001,** 20 (6), 1325-1331.

76. Orlando, E. F.; Kolok, A. S.; Binzcik, G. A.; Gates, J. L.; Horton, M. K.; Lambright, C. S.; Gray, L. E.; Soto, A. M.; Guillette, L. J., Endocrine-disrupting effects of cattle feedlot effluent on an aquatic sentinel species, the fathead minnow. *Environmental Health Perspectives*, **2004,** 112 (3), 353-358.

77. Sone, K.; Hinago, M.; Itamoto, M.; Katsu, Y.; Watanabe, H.; Urushitani, H.; Tooi, O.; Guillette, L. J. Jr.; Iguchi, T., Effects of an androgenic growth promoter 17beta-trenbolone on masculinization of Mosquitofish (Gambusia affinis affinis). *General and Comparative Endocrinology*, **2005,** 143 (2), 151-60.

78. European Medicines Agency Veterinary Medicines Division, *Questions and answers on veterinary medicinal products containing altrenogest to be administered orally to pigs and horses*, EMA/462829/2016, 2016. https://www.ema.europa.eu/en/medicines/veterinary/referrals/altrenogest (accessed on Apr 25, 2021).

79. Zeilinger, J.; Steger-Hartmann, T.; Maser, E.; Goller, S.; Vonk, R.; Länge, R., Effects of synthetic gestagens on fish reproduction. *Environmental Toxicology and Chemistry*, **2009,** 28 (12), 2663-70.

80. Committee for Medicinal Products for Veterianry Use, *Revised guideline on environmental impact assessment for veterinary medicinal products in support of the vich guidelines gl6 and gl 38*, EMEA/CVMP/ERA/418282/2005-Rev.1, 2008. https://www.ema.europa.eu/en/environmental-impact-assessment-veterinary-medicinal-products-support-vich-guidelines-gl6-gl38 (accessed on Apr 25, 2021).

81. Di Nica, V.; Menaballi, L.; Azimonti, G.; Finizio, A., RANKVET: A new ranking method for comparing and prioritizing the environmental risk of veterinary pharmaceuticals. *Ecological Indicators*, **2015,** 52, 270-276.

82. Food and Drug Administration, *FDA animal drug safety communication: FDA highlights potential health risks to people exposed to altrenogest products for horses or pigs*, 2018. https://www.fda.gov/AnimalVeterinary/NewsEvents/CVMUpdates/ucm612395.html (accessed Apr 20, 2021).

83. Newman, A., *Regu-Mate (R) (altrenogest) management practices and potential for exposure in equine care givers*. Master thesis, Tufts University, U.S., 2013.

84. Ferrey, M. L.; Heiskary, S.; Grace, R.; Hamilton, M. C.; Lueck, A., Pharmaceuticals and other anthropogenic tracers in surface water: A randomized survey of 50 Minnesota lakes. *Environmental Toxicology and Chemistry*, **2015**, 34 (11), 2475-2488.

85. Golovko, O.; Sauer, P.; Fedorova, G.; Kroupova, H. K.; Grabic, R., Determination of progestogens in surface and waste water using SPE extraction and LC-APCI/APPI-HRPS. *Science of the Total Environment*, **2018**, 621, 1066-1073.

86. Sauer, P.; Borik, A.; Golovko, O.; Grabic, R.; Stanova, A. V.; Valentova, O.; Stara, A.; Sandova, M.; Kocour Kroupova, H., Do progestins contribute to (anti-)androgenic activities in aquatic environments? *Environmental Pollution*, **2018,** 242 (Pt A), 417-425.

87. Wammer, K. H.; Anderson, K. C.; Erickson, P. R.; Kliegman, S.; Moffatt, M. E.; Berg, S. M.; Heitzman, J. A.; Pflug, N. C.; McNeill, K.; Martinovic-Weigelt, D.; Abagyan, R.; Cwiertny, D. M.; Kolodziej, E. P., Environmental photochemistry of altrenogest: photoisomerization to a bioactive product with increased environmental persistence via reversible photohydration. *Environmental Science & Technology*, **2016**, 50 (14), 7480-7488.

88. Pflug, N. C.; Patterson, E. V.; Martinović-Weigelt, D.; Kolodziej, E. P.; Gloer, J. B.; McNeill, K.; Cwiertny, D. M.; Wammer, K. H., Intramolecular [2 + 2]

photocycloaddition of altrenogest: Confirmation of product structure, theoretical mechanistic insight, and bioactivity assessment. *The Journal of Organic Chemistry*, **2019**, 84 (17), 11366-11371.

89. Yang, X.; Dai, X.; Zhang, Y.; Lin, H.; Wang, J.; He, Z.; Li, Y., Sorption, desorption, and transformation of synthetic progestins in soil and sediment systems. *Geoderma*, **2020,** 362, 114141.

90. Kaufmann, A., High-resolution mass spectrometry for bioanalytical applications: Is this the new gold standard? *Journal of Mass Spectrometry*, **2020,** 55 (9), e4533.

91. Marazuela, M.D., Chapter 18 - Determination of veterinary drug residues in foods by liquid chromatography-mass spectrometry: basic and cutting-edge applications, in *Liquid Chromatography (Second Edition)*, Elsevier: Amsterdam, Netherlands, 2017; pp. 539-570.

92. Abdallah, H.; Arnaudguilhem, C.; Jaber, F.; Lobinski, R., Multiresidue analysis of 22 sulfonamides and their metabolites in animal tissues using quick, easy, cheap, effective, rugged, and safe extraction and high resolution mass spectrometry (hybrid linear ion trap-Orbitrap). *Journal of Chromatography A*, **2014,** 1355, 61-72.

93. Hurtaud-Pessel, D.; Jagadeshwar-Reddy, T.; Verdon, E., Development of a new screening method for the detection of antibiotic residues in muscle tissues using liquid chromatography and high resolution mass spectrometry with a LC-LTQ-Orbitrap instrument. *Food Additives & Contaminants: Part A: Chemistry, Analysis, Control, Exposure & Risk Assessment*, **2011,** 28 (10), 1340-51.

94. Schymanski, E. L.; Jeon, J.; Gulde, R.; Fenner, K.; Ruff, M.; Singer, H. P.; Hollender, J., Identifying small molecules via high resolution mass spectrometry: communicating confidence. *Environmental Science & Technology*, **2014,** 48 (4), 2097-2098.

95. Patel, S.; Ahmed, S., Emerging field of metabolomics: big promise for cancer biomarker identification and drug discovery. *Journal of Pharmaceutical and Biomedical Analysis*, **2015,** 107, 63-74.

96. Castro-Puyana, M.; Pérez-Míguez, R.; Montero, L.; Herrero, M., Application of mass spectrometry-based metabolomics approaches for food safety, quality and traceability. *Trends in Analytical Chemistry*, **2017,** 93, 102-118.

97. Nicholson, J. K.; Lindon, J. C.; Holmes, E., 'Metabonomics': understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica*, **1999,** 29 (11), 1181-1189.

98. Fiehn, O.; Kopka, J.; Dörmann, P.; Altmann, T.; Trethewey, R. N.; Willmitzer, L., Metabolite profiling for plant functional genomics. *Nature Biotechnology*, **2000**, 18 (11), 1157-1161.

99. Hermo, M. P.; Saurina, J.; Barbosa, J.; Barrón, D., High-resolution mass spectrometry applied to the study of metabolome modifications in various chicken tissues after amoxicillin administration. *Food Chemistry*, **2014,** 153, 405-413.

100. Morales-Gutiérrez, F. J.; Barbosa, J.; Barrón, D., Metabolic study of enrofloxacin and metabolic profile modifications in broiler chicken tissues after drug administration. *Food Chemistry*, **2015,** 172, 30-39.

101. Dervilly-Pinel, G.; Weigel, S.; Lommen, A.; Chereau, S.; Rambaud, L.; Essers, M.; Antignac, J.-P.; Nielen, M. W. F.; Le Bizec, B., Assessment of two complementary liquid chromatography coupled to high resolution mass spectrometry metabolomics strategies for the screening of anabolic steroid treatment in calves. *Analytica Chimica Acta*, **2011,** 700 (1), 144-154.

102. Regal, P.; Anizan, S.; Antignac, J.-P.; Le Bizec, B.; Cepeda, A.; Fente, C., Metabolomic approach based on liquid chromatography coupled to high resolution mass spectrometry to screen for the illegal use of estradiol and progesterone in cattle. *Analytica Chimica Acta*, **2011,** 700 (1), 16-25.

103. Wu, Y.; Bi, Y.; Bingga, G.; Li, X.; Zhang, S.; Li, J.; Li, H.; Ding, S.; Xia, X., Metabolomic analysis of swine urine treated with β2-agonists by ultra-high performance liquid chromatography-quadrupole time-of-flight mass spectrometry. *Journal of Chromatography A*, **2015,** 1400, 74-81.

104. Arias, M.; Chevallier, O. P.; Graham, S. F.; Gasull-Gimenez, A.; Fodey, T.; Cooper, K. M.; Crooks, S. R. H.; Danaher, M.; Elliott, C. T., Metabolomics reveals novel biomarkers of illegal 5-nitromimidazole treatment in pigs. Further evidence of drug toxicity uncovered. *Food Chemistry*, **2016,** 199, 876-884.

105. Labine, L. M.; Simpson, M. J., The use of nuclear magnetic resonance (NMR) and mass spectrometry (MS)–based metabolomics in environmental exposure assessment. *Current Opinion in Environmental Science & Health*, **2020,** 15, 7-15.

106. Rodríguez-Mozaz, S.; Serra-Compte, A.; Gil-Solsona, R.; Álvarez-Muñoz, D., Chapter 10 - Environmental metabolomics and xenometabolomics for the assessment of exposure to contaminant mixtures, in *Environmental Metabolomics*, Elsevier: Amsterdam, Netherlands, 2020; p. 283-310.

107. David, A.; Lange, A.; Abdul-Sada, A.; Tyler, C. R.; Hill, E. M., Disruption of the prostaglandin metabolome and characterization of the pharmaceutical exposome in fish exposed to wastewater treatment works effluent as revealed by nanoflow-nanospray mass spectrometry-based metabolomics. *Environmental Science & Technology*, **2017**, 51 (1), 616-624.

108. Al-Salhi, R.; Abdul-Sada, A.; Lange, A.; Tyler, C. R.; Hill, E. M., The xenometabolome and novel contaminant markers in fish exposed to a wastewater treatment works effluent. *Environmental Science & Technology*, **2012**, 46 (16), 9080-9088.

109. Xian, F.; Hendrickson, C. L.; Marshall, A. G., High resolution mass spectrometry. *Analytical Chemistry*, **2012**, 84 (2), 708-719.

110. Marshall, A. G.; Hendrickson, C. L., High-resolution mass spectrometers. *Annual Review of Analytical Chemistry*, **2008,** 1, 579-99.

111. Thermo Scientific, Q Exactive Series Operating Manual. 1344742 Revision B, 2015.

112. AB Sciex GmbH, SCIEX X500 QTOF System User guide. RUO-IDV-05-2334-C, 2016.

113. Eliuk, S.; Makarov, A., Evolution of orbitrap mass spectrometry instrumentation. *Annual Review of Analytical Chemistry*, **2015,** 8 (1), 61-80.

114. Makarov, A., Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Analytical Chemistry*, **2000,** 72 (6), 1156-1162.

115. Kingdon, K.H., A method for the neutralization of electron space charge by positive ionization at very low gas pressures. *Physical Review*, **1923,** 21 (4), 408-418.

116. Hu, Q.; Noll, R. J.; Li, H.; Makarov, A.; Hardman, M.; Graham Cooks, R., The Orbitrap: a new mass spectrometer. *Journal of Mass Spectrometry*, **2005,** 40 (4), 430-443.

117. Cameron, A. E.; Eggers, D.F. Jr., An ion ``Velocitron''. *Review of Scientific Instruments*, **1948,** 19 (9), 605-607.

118. Pareige, C.; Lefebvre-Ulrikson, W.; Vurpillot, F.; Sauvage, X., Chapter Five - Time-of-Flight mass spectrometry and composition measurements, in *Atom Probe Tomography*, Academic Press: London, UK, 2016; p. 123-154.

119. Sage, A. B.; Taylor, P., Commercial strategies in non-target and suspected target screening of water samples, in *Assessing Transformation Products of Chemicals by Non-Target and Suspect Screening − Strategies and Workflows Volume 2*, American Chemical Society, 2016; p. 131-141.

120. Pluskal, T.; Castillo, S.; Villar-Briones, A.; Oresic, M., MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics*, **2010**, 11, 395.

121. Gemperline, P., Practical guide to chemometrics, CRC Press: Boca Ration, U.S., 2006.

122. Wold, S.; Esbensen, K.; Geladi, P., Principal component analysis. *Chemometrics and Intelligent Laboratory Systems*, **1987,** 2 (1), 37-52.

123. Pearson, K., LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, **1901,** 2 (11), 559-572.

124. Steliopoulos, P., Validierung PCA-gestützter Analysemethoden zur Authentizitätskontrolle von Lebensmitteln. *Journal für Verbraucherschutz und Lebensmittelsicherheit*, **2013,** 8 (1), 71-77.

125. Dixon, S. J.; Brereton, R. G., Comparison of performance of five common classifiers represented as boundary methods: Euclidean Distance to Centroids, Linear Discriminant Analysis, Quadratic Discriminant Analysis, Learning Vector Quantization and Support Vector Machines, as dependent on data structure. *Chemometrics and Intelligent Laboratory Systems*, **2009,** 95 (1), 1-17.

126. Broadhurst, D. I.; Kell, D. B., Statistical strategies for avoiding false discoveries in metabolomics and related experiments. *Metabolomics*, **2006,** 2 (4), 171-196.

127. Altman, D. G.; Bland, J. M., Diagnostic tests. 1: Sensitivity and specificity. *British Medical Journal*, **1994**, 308 (6943), 1552.

128. Szymanska, E.; Saccenti, E.; Smilde, A. K.; Westerhuis, J. A., Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics*, **2012**, 8 (Suppl 1), 3-16.

129. Barker, M.; Rayens, W., Partial least squares for discrimination. *Journal of Chemometrics*, **2003,** 17 (3), 166-173.

130. Bohm, D. A.; Stachel, C. S.; Gowik, P., Validated determination of eight antibiotic substance groups in cattle and pig muscle by HPLC/MS/MS. *Journal of AOAC International*, **2011,** 94 (2), 407-419.

131. Chambers, M. C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; Hoff, K.; Kessner, D.; Tasman, N.; Shulman, N.; Frewen, B.; Baker, T. A.; Brusniak, M.-Y.; Paulse, C.; Creasy, D.; Flashner, L.; Kani, K.; Moulding, C.; Seymour, S. L.; Nuwaysir, L. M.; Lefebvre, B.; Kuhlmann, F.; Roark, J.; Rainer, P.; Detlev, S.; Hemenway, T.; Huhmer, A.; Langridge, J.; Connolly, B.; Chadick, T.; Holly, K.; Eckels, J.; Deutsch, E. W.; Moritz, R. L.; Katz, J. E.; Agus, D. B.; MacCoss, M.; Tabb, D. L.; Mallick, P., A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology*, **2012,** 30 (10), 918-920.

132. Dunn, W. B.; Broadhurst, D.; Begley, P.; Zelena, E.; Francis-McIntyre, S.; Anderson, N.; Brown, M.; Knowles, J. D.; Halsall, A.; Haselden, J. N.; Nicholls, A. W.; Wilson, I. D.; Kell, D. B.; Goodacre, R. & The Human Serum Metabolome Consortium, Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nature Protocols*, **2011,** 6 (7), 1060-1083.

133. Valderrama, L.; Valderrama, P., Nondestructive identification of blue pen inks for documentoscopy purpose using iPhone and digital image analysis including an approach for interval confidence estimation in PLS-DA models validation. *Chemometrics and Intelligent Laboratory Systems*, **2016**, 156, 188-195.

134. Chong, J.; Wishart, D. S.; Xia, J., Using MetaboAnalyst 4.0 for comprehensive and integrative metabolomics data analysis. *Current Protocols in Bioinformatics*, **2019**, 68 (1), e86.

135. Kind, T.; Fiehn, O., Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC Bioinformatics*, **2007,** 8 (1), 105.

136. Royal Society of Chemistry, *ChemSpider*. 2019. https://www.chemspider.com (accessed on Apr 25, 2021).

137. Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B. A.; Thiessen, P. A.; Yu, B.; Zaslavsky, L.; Zhang, J.; Bolton, E. E., PubChem 2019 update: improved access to chemical data. *Nucleic Acids Research*, **2018,** 47 (D1), D1102-D1109.

138. 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., METLIN: A technology platform for identifying knowns and unknowns. *Analytical chemistry*, **2018,** 90 (5), 3156-3164.

139. Wishart, D. S.; Feunang, Y. D.; Marcu, A.; Guo, A. C.; Liang, K.; Vázquez-Fresno, R.; Sajed, T.; Johnson, D.; Li, C.; Karu, N.; Sayeeda, Z.; Lo, E.; Assempour, N.; Berjanskii, M.; Singhal, S.; Arndt, D.; Liang, Y.; Badran, H.; Grant, J.; Serra-Cayuela, A.; Liu, Y.; Mandal, R.; Neveu, V.; Pon, A.; Knox, C.; Wilson, M.; Manach, C.; Scalbert, A., HMDB 4.0: the human metabolome database for 2018. *Nucleic Acids Research*, **2018**, 46 (D1), D608-d617.

140. Perkins, D. N.; Pappin, D. J. C.; Creasy, D. M.; Cottrell, J. S., Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis*, **1999,** 20 (18), 3551-3567.

141. Geer, L. Y.; Marchler-Bauer, A.; Geer, R. C.; Han, L.; He, J.; He, S.; Liu, C.; Shi, W.; Bryant, S. H., The NCBI BioSystems database. *Nucleic Acids Research*, **2009**, 38 (suppl 1), D492-D496.

142. Uniprot Consortium, UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Research*, **2018**, 47 (D1), D506-D515.

143. Benjamini, Y.; Hochberg, Y., Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, **1995,** 57 (1), 289-300.

144. Li, S.; Park, Y.; Duraisingham, S.; Strobel, F. H.; Khan, N.; Soltow, Q. A.; Jones, D. P.; Pulendran, B., Predicting network activity from high throughput metabolomics. *PLOS Computational Biology*, **2013,** 9 (7), e1003123.

145. Fawcett, T., ROC Graphs: notes and practical considerations for researchers. *Machine Learning*, **2004,** 31, 1-38.

146. Xia, J.; Broadhurst, D. I.; Wilson, M.; Wishart, D. S., Translational biomarker discovery in clinical metabolomics: an introductory tutorial. *Metabolomics*, **2013**, 9 (2), 280-299.

147. Domínguez-Romero, J. C.; García-Reyes, J. F.; Molina-Díaz, A., Comparative evaluation of seven different sample treatment approaches for large-scale multiclass sport drug testing in urine by liquid chromatography–mass spectrometry. *Journal of Chromatography A*, **2014,** 1361, 34-42.

148. Commission Decision 2002/657/EC of 14th august 2002 implementing Council Directive 96/23/EC concerning the performance of analytical methods and the interpretation of results (OJ L 221, 17.8.2002, p.8.), as lastly amended by Commission Decision (OJ L 6, 10.1.2004, p. 38).

149. Deutsches Institut für Normung e. V., *DIN 32645:2008-11 Chemical analysis - Decision limit, detection limit and determination limit under repeatability conditions - Terms, methods, evaluation*, Beuth Verlag: Berlin, Germany, 2008.

150. Food and Drug Administration, U.S. Department of Health and Human Services, Center for Drug Evaluation and Research (CDER), Center of Veterinary Medicine (CVM), *Guidance for industry bioanalytical method validation*, 2018. https://www.fda.gov/downloads/Drugs/Guidances/ucm070107.pdf (accessed on Apr 25, 2021).

151. Julicher, B.; Gowik, P.; Uhlig, S., Assessment of detection methods in trace analysis by means of a statistically based in-house validation concept. *Analyst*, **1998,** 123 (2), 173-179.

152. Steliopoulos, P.; Stickel, E.; Haas, H.; Kranz, S., Method validation approach on the basis of a quadratic regression model. *Analytica Chimica Acta*, **2006**, 572 (1), 121-124.

153. Steliopoulos, P. Expectation of the recovery in case of matrix calibration. In *8th international symposium on hormone and veterinary drug residue analysis*, Ghent, Belgium, 2018.

154. Altschul, S. F.; Madden, T. L.; Schaffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, **1997,** 25 (17), 3389-402.

155. Watson, A. D.; Gunning, Y.; Rigby, N. M.; Philo, M.; Kemsley, E. K., Meat authentication via Multiple Reaction Monitoring mass spectrometry of myoglobin peptides. *Analytical Chemistry*, **2015**, 87 (20), 10315-10322.

156. von Bargen, C.; Dojahn, J.; Waidelich, D.; Humpf, H. U.; Brockmeyer, J., New sensitive high-performance liquid chromatography-tandem mass spectrometry method for the detection of horse and pork in halal beef. *Journal of Agricultural and Food Chemistry*, **2013,** 61 (49), 11986-94.

157. Rawlings, N. D.; Barrett, A. J.; Thomas, P. D.; Huang, X.; Bateman, A.; Finn, R. D., The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Research*, **2017**, 46 (D1), D624-D632.

158. Bellof, G.; Granz, S., *Tierproduktion Nutztiere züchten, halten und ernähren*. Vol. 15. Thieme: Stuttgart, Germany, 2018.

159. Statistisches Bundesamt, *Commercial slaughtering*. 2018. https://www.destatis.de/EN/Themes/Economic-Sectors-Enterprises/Agriculture-Forestry-Fisheries/Animals-Animal-Production/Tables/3-commercial-slaughtering-year.html (accessed on Apr 25, 2021).

160. van den Berg, R. A.; Hoefsloot, H. C. J.; Westerhuis, J. A.; Smilde, A. K.; van der Werf, M. J., Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics*, **2006**, 7 (1), 142.

161. Dong, J.; Cai, X.; Zhao, L.; Xue, X.; Zou, L.; Zhang, X.; Liang, X., Lysophosphatidylcholine profiling of plasma: discrimination of isomers and discovery of lung cancer biomarkers. *Metabolomics*, **2010**, 6 (4), 478-488.

162. Dasilva, G.; Muñoz, S.; Lois, S.; Medina, I., Non-targeted LC-MS/MS assay for screening over 100 lipid mediators from ARA, EPA, and DHA in biological samples based on mass spectral fragmentations. *Molecules*, **2019**, 24 (12), 2276.

163. Ricciotti, E.; FitzGerald, G.A., Prostaglandins and inflammation. *Arteriosclerosis, thrombosis, and vascular biology*, **2011**. 31(5): p. 986-1000.

164. Legler, D. F.; Bruckner, M.; Uetz-von Allmen, E.; Krause, P., Prostaglandin E2 at new glance: novel insights in functional diversity offer therapeutic chances. *International Journal of Biochemistry and Cell Biology*, **2010,** 42 (2), 198-201.

165. Dubois, R. N.; Abramson, S. B.; Crofford, L.; Gupta, R. A.; Simon, L. S.; Van De Putte, L. B.; Lipsky, P. E., Cyclooxygenase in biology and disease. *FASEB journal*, **1998,** 12 (12), 1063-73.

166. Vangaveti, V. N.; Jansen, H.; Kennedy, R. L.; Malabu, U. H., Hydroxyoctadecadienoic acids: oxidised derivatives of linoleic acid and their role in inflammation associated with metabolic syndrome and cancer. *European Journal of Pharmacology*, **2016,** 785, 70-76.

167. Law, S.-H.; Chan, M.-L.; Marathe, G. K.; Parveen, F.; Chen, C.-H.; Ke, L.-Y., An updated review of lysophosphatidylcholine metabolism in human diseases. *International Journal of Molecular Sciences*, **2019**, 20 (5), 1149.

168. Zhang, F.; Jia, Z.; Gao, P.; Kong, H.; Li, X.; Chen, J.; Yang, Q.; Yin, P.; Wang, J.; Lu, X.; Li, F.; Wu, Y.; Xu, G., Metabonomics study of atherosclerosis rats by ultra fast liquid chromatography coupled with ion trap-time of flight mass spectrometry. *Talanta*, **2009,** 79 (3), 836-844.

169. Adachi, J.; Asano, M.; Yoshioka, N.; Nushida, H.; Ueno, Y., Analysis of phosphatidylcholine oxidation products in human plasma using quadrupole time-of-flight mass spectrometry. *Kobe Journal Medical Science*, **2006**, 52 (5), 127-40.

170. Ung, P.; Winkler, D.A., Tripeptide motifs in biology: targets for peptidomimetic design. *Journal of Medicinal Chemistry*, **2011**, 54 (5), 1111-25.

171. Wang, Z.; Ying, Z.; Bosy-Westphal, A.; Zhang, J.; Schautz, B.; Later, W.; Heymsfield, S. B.; Müller, M. J., Specific metabolic rates of major organs and tissues across adulthood: evaluation by mechanistic model of resting energy expenditure. *The American Journal of Clinical Nutrition*, **2010**, 92 (6), 1369-1377.

172. Knights, K. M.; Rowland, A.; Miners, J. O., Renal drug metabolism in humans: the potential for drug-endobiotic interactions involving cytochrome P450 (CYP) and UDP-glucuronosyltransferase (UGT). *British Journal of Clinical Pharmacology*, **2013,** 76 (4), 587-602.

173. Hintikka, L.; Kuuranne, T.; Leinonen, A.; Thevis, M.; Schänzer, W.; Halket, J.; Cowan, D.; Grosse, J.; Hemmersbach, P.; Nielen, M. W. F.; Kostiainen, R., Liquid chromatographic–mass spectrometric analysis of glucuronide-conjugated anabolic steroid metabolites: method validation and interlaboratory comparison. *Journal of Mass Spectrometry*, **2008,** 43 (7), 965-973.

174. Hegger, I., *Untersuchung zur Pharmakokinetik des Arzneistoffes Altrenogest hinsichtlich der Dopingrelevanz beim Pferd*. Dissertation, Tierärztliche Hochschule Hannover, Germany, 2005.

175. Steliopoulos, P.; Stickel, E., Estimation of performance characteristics of a confirmation method for thyreostats in plasma by means of a weighted least-squares approach. *Analytica Chimica Acta*, **2007,** 592 (2), 181-186.

176. Taylor, P.; Scarth, J. P.; Hillyer, L. L., Use of in vitro technologies to study phase II conjugation in equine sports drug surveillance. *Bioanalysis*, **2010,** 2 (12), 1971-1988.

177. Hackett, L. P.; Dusci, L. J.; Ilett, K.; Chiswell, G. M., Optimizing the hydrolysis of codeine and morphine glucuronides in urine. *Therapeutic Drug Monitoring*, **2002,** 24, 652-657.

178. Adedeji, A. O.; Pourmohamad, T.; Chen, Y.; Burkey, J.; Betts, C. J.; Bickerton, S. J.; Sonee, M.; McDuffie, J. E., Investigating the value of urine volume, creatinine, and cystatin C for urinary biomarkers normalization for drug development studies. *International Journal of Toxicology*, **2019**, 38 (1), 12-22.

179. Biswas, S.; Kranz, W. L.; Shapiro, C. A.; Snow, D. D.; Bartelt-Hunt, S. L.; Mamo, M.; Tarkalson, D. D.; Zhang, T. C.; Shelton, D. P.; van Donk, S. J.; Mader, T. L., Effect of rainfall timing and tillage on the transport of steroid hormones in runoff from manure amended row crop fields. *Journal of Hazardous Materials*, **2017,** 324, 436-447.

180. U.S. Environmental Protection Agency (EPA), *Chemistry Dashboard.* 2020. https://comptox.epa.gov/dashboard/DTXSID3048863 (accessed on Apr 25, 2021).

181. Yang, X.; Zhao, H.; Cwiertny, D. M.; Kolodziej, E. P., Sorption and transport of trenbolone and altrenogest photoproducts in soil–water systems. *Environmental Science: Processes & Impacts*, **2019**, 21 (10), 1650-1663.

182. Balmer, M. E.; Goss, K.-U.; Schwarzenbach, R. P., Photolytic transformation of organic pollutants on soil surfaces – an experimental approach. *Environmental Science & Technology*, **2000**, 34 (7), 1240-1245.

183. Schiffman, C.; Petrick, L.; Perttula, K.; Yano, Y.; Carlsson, H.; Whitehead, T.; Metayer, C.; Hayes, J.; Rappaport, S.; Dudoit, S., Filtering procedures for untargeted LC-MS metabolomics data. *BMC bioinformatics*, **2019,** 20 (1), 334-334.

184. Hrboticky, N.; Leiter, L. A.; Anderson, G. H., Menstrual cycle effects on the metabolism of tryptophan loads. *American Journal of Clinical Nutrition*, **1989,** 50 (1), 46-52.

185. Carretti, N.; Florio, P.; Bertolin, A.; Costa, C. V. L.; Allegri, G.; Zilli, G., Serum fluctuations of total and free tryptophan levels during the menstrual cycle are related to gonadotrophins and reflect brain serotonin utilization. *Human Reproduction*, **2005,** 20 (6), 1548-1553.

186. Johnson, R. S.; Martin, S. A.; Biemann, K.; Stults, J. T.; Watson, J. T., Novel fragmentation process of peptides by collision-induced decomposition in a tandem mass spectrometer: differentiation of leucine and isoleucine. *Analytical Chemistry*, **1987,** 59 (21), 2621-2625.

187. Roepstorff, P.; Fohlman, J., Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomedical Mass Spectrometry*, **1984**, 11 (11), 601-601.

188. Meusinger, R., *NMR-Spektren richtig ausgewertet: 100 Übungen für Studium und Beruf*, Springer-Verlag: Heidelberg, Germany, 2009.

189. Cooper, W.J.; Zika, R.G., Photochemical formation of hydrogen peroxide in surface and ground waters exposed to sunlight. *Science*, **1983,** 220 (4598), 711-712.

190. Goncharuk, V.; Vakulenko, V.; Shvadchina, Y.; Sova, A.; Sitnichenko, T.; Kalinichenko, I., Formation and decomposition of hydrogen peroxide during UV-radiation, ozonization, and O-3/UV treatment of river water. *Journal of Water Chemistry and Technology*, **2008**, 30, 335-343.

191. Mumma, R. O.; Hoiberg, C. P.; Weber, W. W., Preparation of sulfate esters. the synthesis of steroid sulfates by a dicyclohexylcarbodiimide-mediated sulfation. *Steroids*, **1969,** 14 (1), 67-74.

192. Liesenfeld, S.; Steliopoulos, P.; Wenig, S.; Gottstein, V.; Hamscher, G., Comprehensive LC-HRMS metabolomics analyses for the estimation of environmental inputs of altrenogest in pig breeding. *Chemosphere*, **2022**, 287, in-press (https://doi.org/10.1016/j.chemosphere.2021.132353).

6 Appendix A: Tables

Table S1. Overview of cited residue studies for ALT in gilts in the course of the European and American drug administration.

Study (dosing)	Withdrawal (days)	Analyte	Liver (ppb)	Kidney (ppb)	Muscle (ppb)	Fat (ppb)
A (20 mg/d 18 days) [64]	0 (6 hours)	Total residue concentration of radio-labeled ALT	476	210	<2	<2
	5		105	23	<2	<2
	10		n.a.	n.a.	<2	<2
	15		54	<15	<2	<2
	30		<30	<15	<2	<2
	60		<30	<15	<2	<2
	179		<30	<15	<2	<2
B (20 mg/d 18 days) [65]	15	Total residue concentration of radio-labeled ALT	60.3	15.7	5.2	2.5
	30		26.1	5.7	1.8	1.4
	60		15.6	1.9	<1	<1
	179		10.1	<1	<1	<1
C (20 mg/d 18 days) [28]	0 (4.5hours)	ALT	196	11.6	6.7	58.7
	7		0.74	0.26	n.a.	n.a.
	15		0.25	n.a.	n.a.	n.a.
D (20 mg/d 18 days) [28]	0 (4.5 hours)	Total residue concentration of radio-labeled ALT	1444	372	30	91
	7		122	75	7	4
	15		62	12	4	2
E (20 mg/d 18 days) [28]	1	ALT	85.37	9.16	4.7	55.27
	7		< 1.25	< 1.25	< 1.25	< 1.25
	14		< 1.25	< 1.25	< 1.25	< 1.25
	21		< 1.25	< 1.25	< 1.25	< 1.25
F (20 mg/d 18 days) [28]	7	ALT	1.045–2.519	n.a.	n.a.	n.a.
	14		0.639–1.471	n.a.	n.a.	n.a.
	21		< 0.229	n.a.	n.a.	n.a.
G (20 mg/d 7 days) [65]	1	Total residue concentration of radio-labeled ALT	432	122	11	33
	2		172	96	20	14
	3		190	77	10	6
	5		105	39	8	6
	7		46	24	4	<LOD
	10		48	22	3	<LOD
	15		36	12	4	<LOD
H (20 mg/d 18 days) [68]	0 (6 hours)	ALT	241	41	41	137
	1		73	20	6	34
	5		3	<LOD	<LOD	<LOD
	10		<LOD	<LOD	<LOD	<LOD
	20		<LOD	<LOD	<LOD	<LOD

LOD: limit of detection; n.a.: not available.

Table S2. Analyte spectrum detected in the positive *porcine* samples used for PCA, multiple findings of analytes per sample listed.

	Muscle samples			Kidney samples		
Substance group	> MRL	> MRL, < $CC\alpha$	< MRL[a]	> MRL	> MRL, < $CC\alpha$	< MRL[a]
Aminoglycosides						
Dihydrostreptomycin				2		
Spectinomycin						1
β-lactam antibiotics						
Ampicillin	2			1		
Benzylpenicillin	1		2	2		
Macrolides						
Tilmicosin	1			1		
Tulathromycin	1					1
Quinolones						
Marbofloxacin	1			1		
Sulfonamides/ Diaminopyrimidines						
Sulfamethazine	3			1		2
Trimethoprim	2		1			2
Tetracyclines						
Chlortetracycline	2		1	1		2
Doxycycline	5		2	3		3
Oxytetracycline	5	1		6		
Tetracycline			1			

MRL: maximum residue limit, $CC\alpha$: decision limit at and above which it can be concluded with an error probability of α ($\alpha = 0.05$) that MRL is exceeded, [a]samples with findings below MRL in muscle samples were considered as positive samples (drug-treated, infected animals), since MRL was exceeded in kidney samples of the respective animals or vice versa.

Table S3. Analyte spectrum detected in the positive *bovine* samples used for PCA, multiple findings of analytes per sample listed.

	Muscle samples			Kidney samples		
Substance group	> MRL	> MRL, < $CC\alpha$	< MRL[a]	> MRL	> MRL, < $CC\alpha$	< MRL[a]
Aminoglycosides						
Dihydrostreptomycin						1
Kanamycin						1
β-lactam antibiotics						
Amoxicillin	1		1	2		
Benzylpenicillin			8	8		
Corticosteroids						
Dexamethasone	2		2	4		
Macrolides						
Tulathromycin	1			1		
Tetracyclines						
Oxytetracycline	1				1	
Tetracycline	2					1
Quinolones						
Sum of enrofloxacin / ciprofloxacin			1	1		
Marbofloxacin			1	1		

MRL: maximum residue limit, $CC\alpha$: decision limit at and above which it can be concluded with an error probability of α ($\alpha = 0.05$) that MRL is exceeded, [a]samples with findings below MRL in muscle samples were considered as positive samples (drug-treated animals), since MRL was exceeded in kidney samples of the respective animals or vice versa.

Table S4. Spiking level of QC samples for muscle and kidney samples.

Analyte	Muscle samples (µg/kg)	Kidney samples (µg/kg)
Cephalosporins		
Cefalexin	200	1000
Cefapirin	50	100
Ceftiofur	1000	6000
Cefquinome	50	200
Desacetylcefapirin	50	100
Desfuroylceftiofur	1000	6000
Coccidiostats		
Halofuginone	10	30
Corticosteroids		
Betamethasone	0.75	0.75
methylprednisolone	10	10
Prednisolone	4	4
Diaminopyrimidines		
Trimethoprim	100	100
Lincosamides		
Lincomycin	100	1500
Pirlimycin	100	400
Macrolides		
3-O-Acetyltylosin[a]	50	50
Azithromycin	20	20
Erythromycin	200	200
Gamithromycin	100	100
Josamycin	20	20
Oleandomycin	20	20
Spiramycin	200	600
Tildipirosin	1200	10 000
Tilmicosin	50	1000
Tulathromycin	600	6000
Tylosin	100	100
Tylvalosin	50	50
Pleuromutilins		
8-α-Hydroxymutilin[b]	100	100
Tiamulin	100	100
Valnemulin	50	100
Quinolones		
Ciprofloxacin	100	200
Danofloxacin	200	400
Difloxacin	400	800
Enrofloxacin	100	200
Flumequine	200	1500
Marbofloxacin	150	150
Nalidixic acid	40	40
Oxolinic acid	100	150

Analyte	Muscle samples (µg/kg)	Kidney samples (µg/kg)
Sarafloxacin	20	20
Sulfonamides		
Dapsone	20	20
Sulfachlorpyrazine	100	100
Sulfachlorpyridazine	100	100
Sulfadiazine	100	100
Sulfadimethoxine	100	100
Sulfadoxine	100	100
Sulfamerazine	100	100
Sulfamethazine	100	100
Sulfamethoxazole	100	100
Sulfamethoxypyridazine	100	100
Sulfanilamide	100	100
Sulfaquinoxaline	100	100
Sulfathiazole	100	100
Tetracyclines		
Chlortetracycline	100	600
Doxycycline	100	600
Oxytetracycline	100	600
Tetracycline	100	600
β-lactam antibiotics		
Ampicillin	50	50
Benzylpenicillin	50	50
Cloxacillin	300	300
Dicloxacillin	300	300
Nafcillin	300	300
Oxacillin	300	300
Phenoxymethylpenicillin	25	25
internal standards		
Amoxicillin-d4	100	100
Benzylpenicillin-d7	100	100
Demeclocycline	100	100
Enrofloxacin-d5	100	100
Oxolinic acid-d5	100	100
Roxithromycin	100	100
Sarafloxacin-d8	100	100
Sulfadimethoxine-d6	100	100
Sulfadoxine-d3	100	100

[a]: marker residue for tylvalosin, [b]: marker residue for tiamulin.

Table S5. Parameters used for data processing in MZmine 2 for muscle samples.

Parameters	Q-Orbitrap Focus	Q-TOF X500R
Crop filtering		
Time (min)	0.5–12.5	0.5–12.5
MS Level	1	1
Spectrum type	Centroid	Centroid
m/*z* range	75-1055	75-1055
Mass detection		
MS Level	1	1
Polarity	Positive	Positive
Noise level	1.0×10^5	1.0×10^3
Chromatogram builder		
Minimum time span (min)	0.05–1.5	0.05–1.5
Minimum height	5.0×10^5	5.0×10^3
m/*z* tolerance (Da or ppm)	0.002 or 5	0.01 or 25
Smoothing	15	15
Chromatogram deconvolution (Noise amplitude algorithm)		
Minimum peak height	5.0×10^5	5.0×10^3
Peak duration range (min)	0.05–1.5	0.05–1.5
Amplitude of noise	1.0×10^5	1.0×10^3
Isotope filtering		
m/*z* tolerance (Da or ppm)	0.002 or 5	0.01 or 25
Retention time tolerance (min)	0.01	0.01
Monotonic shape	Yes	Yes
Maximum charge	3	3
Representative isotope	Lowest m/z	Lowest m/z
Alignment (join aligner)		
m/*z* tolerance (Da or ppm)	0.005 or 10	0.02 or 50
Weight for *m*/*z* (%)	10	10
Retention time tolerance (min)	0.2	0.2
Weight for RT (%)	10	10
Require same charge state	Yes	Yes
Compare isotope pattern	Yes	Yes
Isotope *m*/*z* tolerance (Da or ppm)	0.005 or 10	0.02 or 50
Minimum absolute intensity	5.0×10^3	5.0×10^2
Minimum score (%)	65	65
Gap filling (Peak finder algorithm)		
Intensity tolerance (%)	20	20
m/*z* tolerance (Da or ppm)	0.005 or 10	0.02 or 50
Retention time tolerance (min)	0.05	0.05
Retention time correction	Yes	Yes
Duplicate filtering		
m/*z* tolerance (Da or ppm)	0.005 or 10	0.02 or 50
Retention time tolerance (min)	0.2	0.2
Row filtering		
Minimum peaks in a row	5	5
Post-processing		
Background correction	$> 1.0 \times 10^7$	$> 1.0 \times 10^6$

Table S6. Parameters used for data processing in MZmine 2 for kidney samples.

Parameters	Q-Orbitrap Focus	Q-TOF X500R
Crop filtering		
Time (min)	0.5–12.5	0.5–12.5
MS Level	1	1
Spectrum type	Centroid	Centroid
m/*z* range	75–1055	75–1055
Mass detection		
MS Level	1	1
Polarity	Positive	Positive
Noise level	1.0×10^5	1.0×10^3
Chromatogram builder		
Minimum time span (min)	0.05–1.5	0.05–1.5
Minimum height	3.0×10^5	5.0×10^3
m/*z* tolerance (Da or ppm)	0.002 or 5	0.01 or 25
Smoothing	15	15
Chromatogram deconvolution (Noise amplitude algorithm)		
Minimum peak height	3.0×10^5	5.0×10^3
Peak duration range (min)	0.05–1.5	0.05–1.5
Amplitude of noise	1.0×10^5	1.0×10^3
Isotope filtering		
m/*z* tolerance (Da or ppm)	0.002 or 5	0.01 or 25
Retention time tolerance (min)	0.01	0.01
Monotonic shape	Yes	Yes
Maximum charge	3	3
Representative isotope	Lowest m/z	Lowest m/z
Alignment (join aligner)		
m/*z* tolerance (Da or ppm)	0.005 or 10	0.02 or 50
Weight for *m*/*z* (%)	10	10
Retention time tolerance (min)	0.2	0.2
Weight for RT (%)	10	10
Require same charge state	Yes	Yes
Compare isotope pattern	Yes	Yes
Isotope *m*/*z* tolerance (Da or ppm)	0.005 or 10	0.02 or 50
Minimum absolute intensity	5.0×10^3	5.0×10^2
Minimum score (%)	65	65
Gap filling (Peak finder algorithm)		
Intensity tolerance (%)	20	20
m/*z* tolerance (Da or ppm)	0.005 or 10	0.02 or 50
Retention time tolerance (min)	0.05	0.05
Retention time correction	Yes	Yes
Duplicate filtering		
m/*z* tolerance (Da or ppm)	0.005 or 10	0.02 or 50
Retention time tolerance (min)	0.2	0.2
Row filtering		
Minimum peaks in a row	5	5
Post-processing		
Background correction	$> 1.0 \times 10^7$	$> 1.0 \times 10^6$

Table S7. Comparison of Q-Orbitrap and Q-TOF data regarding significant analytes (*m/z* at retention time in min) responsible for the difference between positive and control *porcine* muscle samples, 20 out of the top 25 Q-Orbitrap hits were found in Q-TOF data; positive fold changes indicate higher levels in positive samples compared to control samples.

Q-Orbitrap data			Q-TOF data		
m/z (@min)	*p*-adj.	FC	*m/z* (@min)	*p*-adj.	FC
393.2239 (@8.55)	5.8×10^{-15}	76.7	393.226 (@8.35)	6.0×10^{-17}	23.3
320.1599 (@2.09)	2.1×10^{-14}	3.7	320.161 (@1.99)	1.1×10^{-14}	3.0
375.2133 (@8.93)	2.7×10^{-14}	62.3	375.211(@8.90)	4.6×10^{-9}	6.1
570.3394 (@6.75)	4.0×10^{-14}	37.5	570.342(@6.52)	3.4×10^{-17}	21.6
348.2737 (@8.81)	4.0×10^{-14}	12.0	348.275 (@8.64)	1.5×10^{-15}	7.5
404.2637 (@7.10)	5.2×10^{-14}	68.1	n.d.		
386.2529 (@8.23)	1.4×10^{-13}	76.3	n.d.		
480.1642 (@1.19)	1.4×10^{-13}	56.3	480.166 (@1.13)	1.1×10^{-6}	36.1
335.2209 (@8.54)	2.4×10^{-13}	97.7	335.222 (@8.34)	3.4×10^{-17}	20.4
552.3286 (@7.64)	2.4×10^{-13}	66.6	552.310 (@7.11)	2.5×10^{-15}	108.4
375.2136 (@11.14)	4.9×10^{-13}	117.2	375.216 (@10.99)	8.2×10^{-17}	101.3
391.2083 (@8.18)	9.9×10^{-13}	17.4	391.213 (@8.08)	8.6×10^{-15}	28.9
372.2738 (@9.64)	2.2×10^{-12}	41.6	n.d.		
534.3184 (@10.20)	2.2×10^{-12}	109.9	534.321 (@9.98)	2.0×10^{-15}	31.4
506.1800 (@3.94)	2.2×10^{-12}	36.5	506.182 (@3.71)	1.6×10^{-12}	12.9
388.2686 (@8.19)	2.5×10^{-12}	163.6	n.d.		
173.117 (@8.87)	2.7×10^{-12}	22.1	173.118 (@8.65)	2.9×10^{-12}	10.8
170.081 (@1.57)	2.9×10^{-12}	14.4	n.d.		
204.1050 (@4.95)	2.9×10^{-12}	13.1	204.105 (@4.81)	2.8×10^{-13}	12.3
536.3337 (@9.99)	5.0×10^{-12}	58.2	536.336 (@9.65)	4.9×10^{-10}	47.7
301.2157 (@9.63)	5.0×10^{-12}	61.8	301.216 (@9.45)	3.4×10^{-17}	60.0
552.3288 (@8.87)	1.4×10^{-11}	404.8	552.329 (@8.57)	3.5×10^{-16}	434.7
394.2275 (@8.55)	2.0×10^{-11}	6.8	394.226 (@8.21)	2.2×10^{-9}	2.5
1004.4665 (@6.13)	2.2×10^{-11}	0.3	1004.471 (@5.95)	3.1×10^{-9}	0.2
335.2209 (@9.69)	3.9×10^{-11}	21.6	335.222 (@9.54)	3.4×10^{-17}	4.0

FC: fold change, *p*-adj.: multiple test corrected *p*-value based on the Benjamini Hochberg false discovery rate, n.d.: not detected.

Table S8. Mean peak areas of spiked analytes at MRL level with relative standard deviation (RSD) in *bovine* muscle QC samples before and after normalization by sum of all peaks (n = 4).

Analyte	Q-Orbitrap			Q-TOF		
	Mean (a.u. x 10^7)	RSD before norm. (%)	RSD after norm. (%)	Mean (a.u. x 10^4)	RSD before norm. (%)	RSD after norm. (%)
Cephalosporins						
Cefalexin	1.56	29.58	9.45	15.48	10.10	4.93
Cefapirin	0.35	46.23	17.01	9.51	8.97	6.35
Ceftiofur	9.84	33.66	15.18	15.27	23.18	8.58
Cefquinome	1.47	33.69	15.83	2.10	23.39	10.80
Desacetylcefapirin	0.65	33.30	15.97	7.96	13.15	3.60
Desfuroylceftiofur	0.34	47.72	31.01	17.47	33.74	19.36
Coccidiostats						
Halofuginone	n.d.			n.d.		
Corticosteroids						
Betamethasone	0.32	14.42	20.08	n.d.		
Methylprednisolone	n.d.			n.d.		
Prednisolone	n.d.			n.d.		
Diaminopyrimidines						
Trimethoprim	12.22	33.17	18.94	34.95	22.79	7.99
Lincosamides						
Lincomycin	8.07	39.48	10.29	72.35	17.44	3.30
Pirlimycin	5.62	25.47	17.48	59.68	19.23	4.94
Macrolides						
3-O-Acetyltylosin[a]	1.63	27.54	15.61	24.31	30.90	16.15
Azithromycin	n.d.			12.06	29.98	15.67
Erythromycin	0.05	13.40	26.24	12.90	38.95	28.09
Gamithromycin	2.91	54.10	24.50	41.92	36.57	21.29
Josamycin	1.93	21.30	19.42	34.60	27.08	12.45
Oleandomycin	0.91	26.96	31.74	11.69	28.12	14.72
Spiramycin	3.96	18.26	19.74	40.03	31.48	16.34
Tildipirosin	28.74	32.93	9.92	274.03	14.30	6.48
Tilmicosin	2.54	25.38	30.64	17.86	40.84	25.51
Tulathromycin	7.12	21.96	15.95	94.44	19.33	5.91
Tylosin	6.31	30.78	11.86	89.98	30.11	15.12
Tylvalosin	2.21	24.05	21.58	38.81	26.42	12.31
Pleuromutilins						
8-α-Hydroxymutilin[b]	n.d.			n.d.		
Tiamulin	18.23	27.23	11.86	18.89	23.13	8.59
Valnemulin	5.56	29.33	7.98	n.d.		
Quinolones						
Ciprofloxacin	3.35	34.12	20.59	27.06	18.55	6.06
Danofloxacin	15.69	42.58	19.31	59.27	16.17	7.83
Difloxacin	33.15	49.66	20.02	156.06	17.32	2.50
Enrofloxacin	7.50	25.71	10.90	34.75	23.69	9.57
Flumequine	n.d.			43.42	22.52	7.97

Analyte	Q-Orbitrap			Q-TOF		
	Mean (a.u. x 10^7)	RSD before norm. (%)	RSD after norm. (%)	Mean (a.u. x 10^4)	RSD before norm. (%)	RSD after norm. (%)
Marbofloxacin	8.24	49.99	21.22	51.86	13.49	2.64
Nalidixic Acid	9.59	27.51	4.81	9.87	35.65	22.43
Oxolinic Acid	10.77	51.14	23.65	20.20	30.12	15.36
Sarafloxacin	0.32	55.06	33.10	4.77	30.75	18.17
Sulfonamides						
Dapsone	0.11	56.12	29.48	n.d.		
Sulfachlorpyrazine	n.d.			4.28	26.42	12.00
Sulfachlorpyridazine	1.17	36.27	9.53	4.58	21.42	7.79
Sulfadiazine	1.03	44.12	14.31	5.38	16.73	3.55
Sulfadimethoxine	7.62	45.46	21.74	18.19	24.81	12.12
Sulfadoxine	4.85	31.24	8.07	16.53	27.95	14.61
Sulfamerazine	2.30	22.96	12.31	13.31	19.55	7.11
Sulfamethazine	5.89	43.52	27.76	18.28	27.34	14.73
Sulfamethoxazole	1.23	12.00	44.38	8.11	25.58	10.97
Sulfamethoxypyridazine	3.14	46.00	21.73	11.51	22.64	9.46
Sulfanilamide	n.d.			n.d.		
Sulfaquinoxaline	2.71	40.79	17.21	8.51	22.27	8.78
Sulfathiazole	0.70	50.03	21.28	5.48	10.22	5.56
Tetracyclines						
Chlortetracycline	0.84	36.45	20.77	13.90	21.95	7.82
Doxycycline	1.72	18.23	18.81	20.98	30.58	16.39
Oxytetracycline	0.99	46.83	17.05	22.83	14.71	1.27
Tetracycline	1.99	43.16	43.36	22.60	21.14	7.24
β-Lactam antibiotics						
Ampicillin	0.82	40.44	15.47	12.70	10.31	6.00
Benzylpenicillin	n.d.			n.d.		
Cloxacillin	0.13	19.20	43.58	30.89	21.02	6.45
Dicloxacillin	n.d.			17.08	25.02	11.18
Nafcillin	3.57	34.41	16.59	30.34	33.12	18.63
Oxacillin	0.15	40.17	21.99	37.27	20.48	6.16
Phenoxymethylpenicillin	n.d.			n.d.		
Internal standards						
Amoxicillin-d4	n.d.			n.d.		
Benzylpenicillin-d7	5.62	35.09	29.24	n.d.		
Demeclocycline	0.90	58.65	28.66	12.03	21.80	7.05
Enrofloxacin-d5	7.10	28.23	20.10	29.79	18.69	7.52
Oxolinic Acid-d5	13.29	35.26	8.99	25.41	27.19	12.86
Roxithromycin	6.59	25.94	16.08	11.56	10.42	16.45
Sarafloxacin-d8	1.05	25.14	17.37	26.80	21.77	6.79
Sulfadimethoxine-d6	5.62	38.04	15.26	19.47	25.64	12.29
Sulfadoxine-d3	5.71	30.00	5.60	20.12	29.46	15.31

n.d.: not detected; [a]: marker residue for tylvalosin, [b]: marker residue for tiamulin

Table S9. Mean peak areas of spiked analytes at MRL level with measurement uncertainty (MU) and relative standard deviation (RSD) in *porcine* kidney QC samples (*n* = 8).

Analyte	Q-Orbitrap			Q-TOF		
	Mean (a.u. x 10^7)	MU (a.u. x 10^7)	RSD (%)	Mean (a.u. x 10^5)	MU (a.u. x 10^5)	RSD (%)
Cephalosporins						
Cefalexin	17.56	1.57	10.97	2.97	0.32	13.31
Cefapirin	1.68	0.09	6.34	0.92	0.08	11.09
Ceftiofur	219.35	17.34	9.70	65.05	7.03	13.25
Cefquinome	0.66	0.09	16.82	0.33	0.02	8.24
Desacetylcefapirin	1.50	0.25	20.29	0.64	0.08	14.91
Desfuroylceftiofur	5.58	1.15	25.35	6.74	0.85	15.49
Coccidiostats						
Halofuginone	0.30	0.03	12.57	n.d.		
Corticosteroids						
Betamethasone	n.d.			n.d.		
Methylprednisolone	n.d.			n.d.		
Prednisolone	n.d.			1.61	0.23	17.61
Diaminopyrimidines						
Trimethoprim	16.34	1.15	8.61	1.94	0.37	23.59
Lincosamides						
Lincomycin	81.18	26.43	39.94	39.25	3.07	9.59
Pirlimycin	22.11	2.10	11.65	n.d.		
Macrolides						
3-O-Acetyltylosin[a]	1.37	0.1	8.70	2.02	0.39	23.39
Azithromycin	1.70	0.18	13.28	0.68	0.20	35.35
Erythromycin	1.50	0.79	64.50	0.63	0.22	42.59
Gamithromycin	2.55	0.28	15.12	n.d.		
Josamycin	1.90	0.13	8.62	3.25	0.56	21.27
Oleandomycin	n.d.			0.30	0.14	55.59
Spiramycin	5.93	0.47	9.72	5.70	1.25	26.87
Tildipirosin	105.12	6.23	7.27	87.01	8.23	11.60
Tilmicosin	64.84	2.03	3.84	42.38	10.54	30.50
Tulathromycin	46.07	2.47	6.59	0.83	0.20	29.87
Tylosin	n.d.			8.75	1.72	24.07
Tylvalosin	1.58	0.12	9.10	3.27	0.49	18.47
Pleuromutilins						
8-α-Hydroxymutilin[b]	n.d.			n.d.		
Tiamulin	25.92	1.21	5.73	17.38	3.13	22.07
Valnemulin	6.27	0.40	7.78	8.03	1.46	22.27
Quinolones						
Ciprofloxacin	10.81	0.59	6.67	n.d.		
Danofloxacin	41.42	2.99	8.87	9.46	1.31	16.99
Difloxacin	71.33	8.20	14.09	19.47	3.19	20.11
Enrofloxacin	21.47	0.97	5.56	5.58	0.86	18.99
Flumequine	303.39	13.94	5.64	29.26	4.48	18.80
Marbofloxacin	11.32	2.19	23.71	3.42	0.46	16.49

Analyte	Q-Orbitrap			Q-TOF		
	Mean (a.u. x 10^7)	MU (a.u. x 10^7)	RSD (%)	Mean (a.u. x 10^5)	MU (a.u. x 10^5)	RSD (%)
Nalidixic Acid	18.57	0.88	5.81	1.50	0.26	20.99
Oxolinic Acid	64.88	7.04	13.31	2.67	0.61	27.97
Sarafloxacin	0.79	0.16	24.10	0.29	0.04	15.55
Sulfonamides						
Dapsone	n.d.			n.d.		
Sulfachlorpyrazine	n.d.			0.44	0.10	27.68
Sulfachlorpyridazine	1.14	0.34	36.07	n.d.		
Sulfadiazine	1.92	0.23	14.64	0.39	0.06	19.21
Sulfadimethoxine	13.04	1.09	10.25	2.07	0.45	26.82
Sulfadoxine	10.54	0.68	7.97	1.41	0.29	25.18
Sulfamerazine	4.36	0.45	12.61	0.51	0.09	21.82
Sulfamethazine	8.29	0.65	9.60	1.21	0.19	19.45
Sulfamethoxazole	2.28	0.53	28.50	0.61	0.15	29.84
Sulfamethoxypyridazine	3.66	0.50	16.79	0.58	0.14	30.20
Sulfanilamide	n.d.			0.11	0.02	23.24
Sulfaquinoxaline	7.29	1.03	17.32	1.05	0.20	23.93
Sulfathiazole	1.38	0.29	25.31	n.d.		
Tetracyclines						
Chlortetracycline	10.57	2.48	28.79	5.10	0.99	23.78
Doxycycline	22.01	1.19	6.61	13.82	2.13	18.89
Oxytetracycline	14.38	2.22	18.95	7.46	0.86	14.20
Tetracycline	13.59	1.04	9.40	0.17	0.05	37.09
β-Lactam antibiotics						
Ampicillin	1.23	0.16	16.18	0.73	0.08	12.88
Benzylpenicillin	n.d.			0.08	0.04	59.37
Cloxacillin	n.d.			3.17	0.44	16.87
Dicloxacillin	n.d.			2.23	0.23	12.78
Nafcillin	9.37	1.06	13.84	2.56	0.67	31.98
Oxacillin	n.d.			3.89	0.59	18.60
Phenoxymethylpenicillin	n.d.			0.06	0.02	45.24
Internal standards						
Amoxicillin-d4	0.32	0.04	15.89	2.23	0.44	24.00
Benzylpenicillin-d7	0.29	0.11	48.29	1.71	0.34	24.74
Demeclocycline	1.41	0.12	10.71	2.31	0.49	26.04
Enrofloxacin-d5	8.70	0.37	5.22	0.15	0.06	46.38
Oxolinic Acid-d5	42.02	3.75	10.96	2.54	0.43	20.55
Roxithromycin	9.27	0.52	6.87	0.24	0.05	23.59
Sarafloxacin-d8	4.89	0.24	5.93	3.42	0.68	24.44
Sulfadimethoxine-d6	1.35	0.19	17.12	0.95	0.11	14.49
Sulfadoxine-d3	13.41	0.65	5.96	10.70	1.64	18.75

n.d.: not detected; [a]: marker residue for tylvalosin, [b]: marker residue for tiamulin

Table S10. Mean peak areas of spiked analytes at MRL level in *bovine* kidney QC samples. For Q-Orbitrap data with measurement uncertainty (MU) and relative standard deviation (RSD) are given and for Q-TOF data RSD before and after normalization by sum of all peaks are given (*n* = 4).

Analyte	Q-Orbitrap			Q-TOF		
	Mean (a.u. x 10^7)	MU (a.u. x 10^7)	RSD (%)	Mean (a.u. x 10^5)	RSD before norm. (%)	RSD after norm. (%)
Cephalosporins						
Cefalexin	27.42	4.07	8.07	33.28	31.70	6.42
Cefapirin	1.66	0.08	2.65	6.14	25.62	12.57
Ceftiofur	8.83	0.58	3.55	32.86	35.88	5.11
Cefquinome	0.89	0.13	7.70	2.34	8.76	34.32
Desacetylcefapirin	2.02	0.47	12.54	6.32	21.79	16.88
Desfuroylceftiofur	24.50	6.18	13.73	99.18	45.20	6.47
Coccidiostats						
Halofuginone	0.32	0.09	14.99	0.82	41.36	11.50
Corticosteroids						
Betamethasone	n.d.			n.d.		
Methylprednisolone	n.d.			0.72	46.39	7.95
Prednisolone	0.19	0.10	28.04	5.65	17.62	18.87
Diaminopyrimidines						
Trimethoprim	17.32	1.56	4.90	13.45	22.52	15.35
Lincosamides						
Lincomycin	106.44	7.04	3.60	n.d.		
Pirlimycin	28.23	0.24	0.47	77.36	36.04	7.33
Macrolides						
3-O-Acetyltylosin[a]	1.48	0.11	4.08	0.54	61.51	27.33
Azithromycin	2.00	0.63	17.07	2.17	40.76	6.42
Erythromycin	0.34	0.09	14.95	4.13	44.90	7.72
Gamithromycin	2.56	0.17	3.60	9.03	54.93	15.62
Josamycin	2.33	0.06	1.40	14.90	35.43	4.15
Oleandomycin	0.61	0.18	16.19	2.11	57.60	17.95
Spiramycin	3.53	0.65	10.03	14.58	60.87	21.62
Tildipirosin	1.53	0.35	7.42	536.99	31.17	6.96
Tilmicosin	61.95	1.35	1.19	144.47	62.05	22.12
Tulathromycin	56.00	5.25	5.10	2.54	42.59	7.67
Tylosin	7.44	0.25	1.83	31.72	41.46	3.41
Tylvalosin	2.01	0.23	6.32	17.19	37.57	11.79
Pleuromutilins						
8-α-Hydroxymutilin[b]	n.d.			n.d.		
Tiamulin	28.63	1.98	3.76	82.89	34.18	4.34
Valnemulin	6.16	0.62	5.47	34.18	35.18	4.22
Quinolones						
Ciprofloxacin	11.55	1.82	8.57	14.57	36.09	6.07
Danofloxacin	38.22	6.58	9.36	53.52	27.55	10.55
Difloxacin	90.58	4.67	2.80	87.27	31.44	8.46
Enrofloxacin	17.53	2.64	8.20	23.84	29.44	8.12

Analyte	Q-Orbitrap			Q-TOF		
	Mean (a.u. x 10^7)	MU (a.u. x 10^7)	RSD (%)	Mean (a.u. x 10^5)	RSD before norm. (%)	RSD after norm. (%)
Flumequine	293.87	3.07	0.57	146.56	25.67	10.71
Marbofloxacin	11.72	1.51	7.00	20.44	26.02	11.88
Nalidixic Acid	20.56	1.47	3.90	7.45	27.36	11.30
Oxolinic Acid	36.80	0.99	1.46	12.94	26.76	13.40
Sarafloxacin	1.02	0.31	16.80	2.02	29.00	13.47
Sulfonamides						
Dapsone	0.38	0.07	9.79	0.16	32.00	36.37
Sulfachlorpyrazine	0.13	0.04	18.24	2.03	15.94	19.94
Sulfachlorpyridazine	1.01	0.53	28.70	0.86	36.79	12.98
Sulfadiazine	2.63	0.22	4.48	3.10	31.04	21.61
Sulfadimethoxine	12.16	0.83	3.73	8.72	33.54	3.95
Sulfadoxine	10.23	1.49	7.95	5.44	26.27	12.02
Sulfamerazine	5.57	0.12	1.21	3.88	32.68	11.22
Sulfamethazine	8.67	1.19	7.49	4.76	38.19	7.20
Sulfamethoxazole	2.48	0.23	5.03	2.69	20.92	20.35
Sulfamethoxypyridazine	5.17	0.12	1.28	2.72	26.47	15.36
Sulfanilamide	n.d.			0.70	51.47	22.32
Sulfaquinoxaline	4.00	0.35	4.74	3.25	34.73	15.33
Sulfathiazole	1.97	0.22	6.07	1.70	22.91	18.63
Tetracyclines						
Chlortetracycline	12.44	0.54	2.38	25.05	38.73	4.67
Doxycycline	23.28	6.18	14.46	50.42	39.05	11.11
Oxytetracycline	13.93	0.90	3.50	47.01	31.82	5.99
Tetracycline	15.25	1.84	6.58	30.50	37.27	1.16
β-Lactam antibiotics						
Ampicillin	1.68	0.07	2.25	5.29	23.71	12.97
Benzylpenicillin	n.d.			n.d.		
Cloxacillin	n.d.			19.31	27.99	10.47
Dicloxacillin	n.d.			11.74	24.72	16.15
Nafcillin	3.68	1.05	15.54	n.d.		
Oxacillin	n.d.			23.16	26.16	10.15
Phenoxymethylpenicillin	n.d.			0.66	26.02	10.73
Internal standards						
Amoxicillin-d4	0.18	0.03	9.08	n.d.		
Benzylpenicillin-d7	n.d.			1.65	39.14	11.39
Demeclocycline	1.34	0.54	21.83	3.20	33.05	16.91
Enrofloxacin-d5	7.31	1.03	7.66	12.25	27.95	9.84
Oxolinic Acid-d5	29.79	1.75	3.19	11.52	29.72	11.70
Roxithromycin	9.68	0.51	2.88	56.08	40.21	3.53
Sarafloxacin-d8	4.45	0.28	3.48	7.39	28.70	15.03
Sulfadimethoxine-d6	13.85	1.50	5.90	9.44	32.98	4.41
Sulfadoxine-d3	12.56	1.15	4.96	6.07	34.52	10.64

n.d.: not detected; [a]: marker residue for tylvalosin, [b]: marker residue for tiamulin

7 Appendix B: Figures

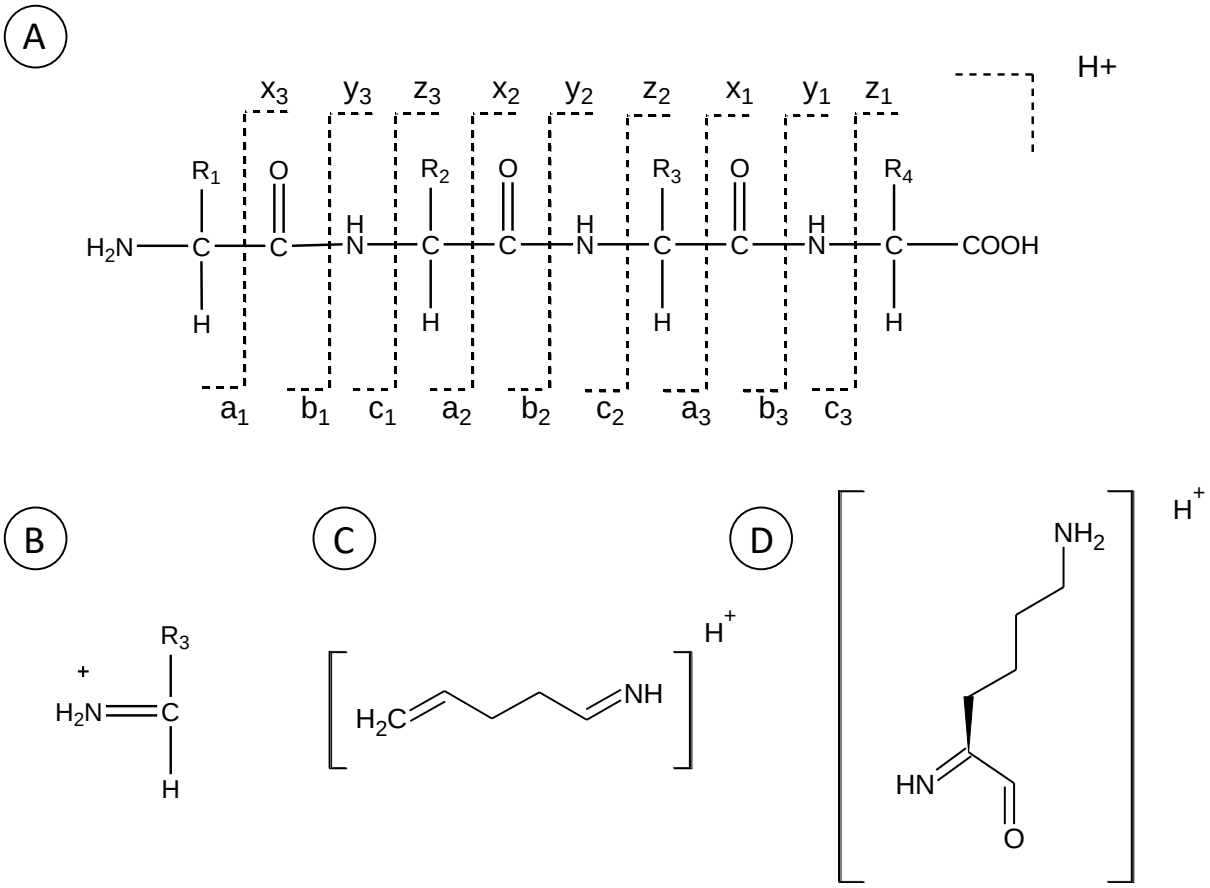


Figure S1. A: nomenclature for fragment ions according to [186, 187]. If the charge is retained on the N-terminus one distinguishes a-,b- or c-fragments, whereas x-, y- and z-fragments are observed if the charge is retained on the C-terminus. The subscript indicates the number of peptides of the fragment [140]. B: immonium ions can be formed by a combination of a type a and type y cleavage. They are labeled with the one letter code for the amino acid [140]. C: example for an immonium related ion of lysine with *m/z* 84. D: example for an immonium related ion of lysine with *m/z* 129.

8 Appendix C: Supplementary data

8.1 Identification of transformation products and metabolites of ALT

8.1.1 Photo-isomerization products of ALT

LC-HRMS data acquisition

In order to identify light-induced elimination products of ALT, a full scan with a resolving power of 70,000 (FWHM at *m*/*z* of 200) in the range of *m*/*z* 50–750 was carried out. Data-dependent fragmentation was performed with a resolving power of 17,500 (FWHM at *m*/*z* of 200) using a collision energy of 35 eV. Mass tolerance was set to 5 ppm. All other LC-MS/MS conditions were identical to the quantitative method (see Section 3.2.2.2). Xcalibur software was used for qualitative analysis (Thermo Scientific, Idstein Germany). In order to identify photo-products, full MS data was processed using MZmine 2 [120]. Raw files were converted to mzXML format, imported into MZmine 2 and chromatograms were crop filtered (1-13 min). Masses were detected with centroid algorithm at a noise level of 1.0×10^7. Chromatogram building was carried out with minimum time span of 0.1 min, minimum height of 2.0×10^7 and mass accuracy 0.005 Da. Chromatograms were smoothed (filter width 15). For peak chromatogram deconvolution Wavelets (XCMS) algorithm was used. Signal to noise threshold was set to 20, wavelet scales were set to the range of 0.1–10, 0.05–2.00 time span was chosen for peak duration. Peaks were aligned with the RANSAC aligner with 0.01 *m*/*z* mass accuracy and 0.2 min time tolerance (10% relative retention time tolerance), 100,000 iterations, number of points was set to 20 and a threshold of 0.1 min was chosen. Gap filling was performed with 2% tolerance at 0.01 *m*/*z* mass accuracy and 0.1 min (retention time) tolerance.

NMR data acquisition

^{1}H NMR data acquisition was performed by Svenja Wenig (CVUA Karlsruhe). Data evaluation was performed by myself. Text section regarding NMR data acquisition was provided by Svenja Wenig.

^{1}H NMR spectra were acquired using a Bruker Avance III HD Ascend 400 spectrometer (BrukerBiospin, Rheinstetten, Germany) equipped with a 5-mm broadband observe (BBO) and broadband inverse (BBI) probe with Z-gradient coils, using a Bruker Automatic Sample Changer (Sample Xpress). Chemical shifts were reported in ppm in reference to the solvent

signal of CD_3OD at δ^1H 3.35 and $\delta^{13}C$ 49.3. Assignments for ALT and the resulting photoproduct were carried out using ^1H, ^1H JRES, ^{13}C, ^{13}C DEPT90 and ^{13}C DEPT135, ^1H-^1H COSY, ^1H-^{13}C HSQC, ^1H-^{13}C HMBC, and ^1H-^1H TOCSY spectra.

^1H NMR spectra were acquired at 400.13 MHz using the standard pulse program noesygppr1d (1D NMR spectra) with presaturation of the water signal. ^{13}C NMR experiments as well as the ^{13}C distortionless enhancement by polarization transfer (DEPT) experiments were performed at 100.6 MHz. The DEPT90 and DEPT135 spectra were acquired using the pulse programs **dept90** and **dept135** with the following parameters: number of scans 2048, 64k data points, spectral width 20161.291 Hz, transmitter frequency offset (observe nucleus ^{13}C) 9055.15 Hz, transmitter frequency offset (second nucleus ^1H) 2000.65 Hz. The CH coupling constant (CNST2) used was 145 Hz. Two-dimensional ^1H-^1H COSY (correlation spectroscopy) spectra were acquired using the pulse program **cosygpmfphpp** with the following parameters: number of scans 64, size of fid 2048 and 256 in f2 and f1, respectively, sweep width 8802.818 Hz, and transmitter offset frequency 1800.59 Hz. Two-dimensional inverse-detected ^1H-^{13}C heteronuclear correlation (HSQC, heteronuclear single quantum correlation) NMR spectra were acquired using the pulse program **hsqcedetgpsisp2.3** with the following parameters: number of scans 64, number of dummy scans 32, size of fid 2048 and 256 in f2 and f1 dimensions, respectively, sweep width 4401.409 Hz and 20161.291 Hz in ^1H and ^{13}C dimensions, respectively, and transmitter offset frequency 1800.59 Hz and 9055.15 Hz in ^1H and ^{13}C dimensions, respectively. The HMBC (heteronuclear multiple bond correlation) NMR spectra were acquired using the pulse program **hmbcgplpndqf** with the following parameters: number of scans 64, number of dummy scans 16, size of fid 2048 and 128 in ^1H and ^{13}C dimensions, respectively, sweep width 4401.409 Hz and 20161.291 Hz in ^1H and ^{13}C dimensions, respectively, and transmitter offset frequency 1800.59 Hz and 9055.15 Hz in ^1H and ^{13}C dimensions, respectively. The ^1H-^1H TOCSY (total correlation spectroscopy) NMR spectra were acquired using the pulse program **mlevphpp** with the following parameters: number of scans 32, number of dummy scans 16, size of fid 2048 and 128 in f2 and f1, respectively, sweep width 8802.818 Hz and 20161.291 Hz and transmitter offset frequency 1800.59 Hz.

Synthesis of the primary photo-isomerization product of ALT

An aqueous solution of ALT (0.25 ng/mL) was prepared and stored at room temperature with light exposure to facilitate the photo-isomerization process (n = 20). Reaction progress was checked via LC-HRMS. After complete photo-isomerization the aqueous solutions were extracted with 20 mL MTBE, centrifuged (10 min, ~2,500 × g) and the supernatants were transferred to new reaction tubes. The process was repeated and the extracts combined. The extracts were dried *in vacuo* (120 min, 55 °C, 1 mbar) and reconstituted in 1 mL methanol-d4.

Identification

In order to identify the isomerization product observed under light exposure in the stability tests, an aqueous solution of ALT was stored at room temperature with light exposure. After 14 days ALT was almost completely isomerized to its primary photo-isomer. The peak area of the photo-isomer was 91% of the sum of peak areas in the extracted ion chromatogram of *m/z* 311.2006 (see Figure S2).

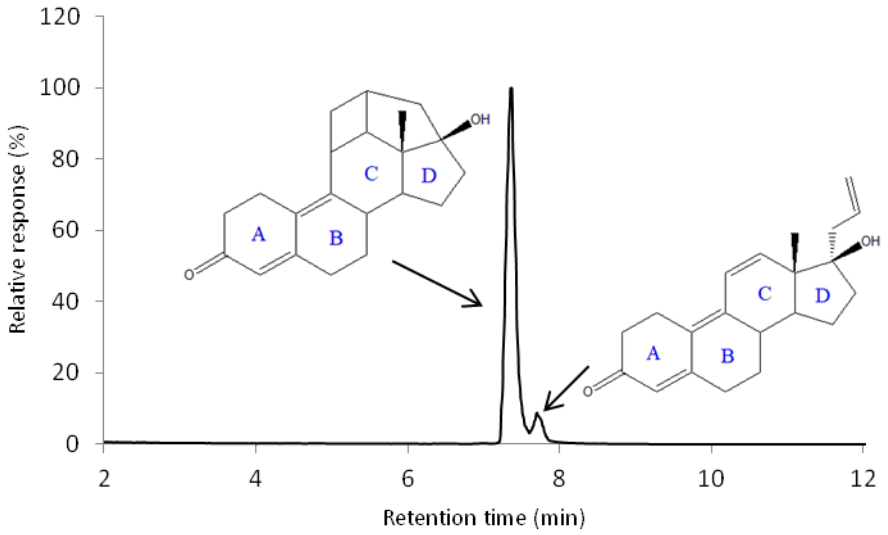


Figure S2. Extracted ion chromatogram of ALT (@ 7.69 min) and its primary photo-product (@ 7.32 min) after storage of ALT solution over 14 days at room temperature with light exposure (*m/z* 311.2006 $[M+H]^+$).

In order to identify the primary photo-isomer extensive NMR analysis was performed.

The two doublets (6.59 ppm, 6.49 ppm; J = 10.08 Hz) of the endocyclic olefin group in ring C as well as the singlet (5.79 ppm) of the olefin in ring A and the signals of the vinyl group (6.01 ppm (m), 5.14 ppm (d; J = 10.28 Hz), 5.09 ppm (d; J = 17.16 Hz)) were the most characteristic ^{1}H NMR resonances of ALT. The corresponding carbons showed signals with chemical shifts of 118.6 to 144.0 ppm in the ^{13}C NMR spectrum (see Table S11).

However, the ^{1}H NMR spectrum of the photo-isomer only showed a singlet at 5.69 ppm attributable to the olefin group in ring A and new signals in the alkane region (3.60 ppm, 2.77 ppm, 2.89 ppm, 3.08 ppm and 1.60 ppm, (see Figure S3). Furthermore, DEPT135 and DEPT90 data only showed one signal of a CH carbon in the alkene region at 122.15 ppm (see Figure S4 and Table S12). This indicated the loss of the endocyclic olefin and vinyl group, resulting from an intramolecular photochemical [2+2] cycloaddition to form a new cyclobutane and cyclopentane ring (see Figure S5). Hereafter, the primary photo-product is referred to as ALT cycloaddition product (ALT-CAP) [87]. The phase-sensitive ^{1}H-^{13}C HSQC (heteronuclear single quantum correlation) was used to assign the relationship between the signals of CH_2/CH/CH_3 groups and the signals of their respective protons (see Figure S6). Furthermore, two-dimensional data obtained from ^{1}H-^{13}C HMBC, ^{1}H-^{1}H COSY and ^{1}H-^{1}H TOCSY helped to verify the structure of the molecule. ^{1}H-^{13}C HMBC provides information on correlation between protons and carbons separated by two or three bonds [188]. ^{1}H-^{1}H COSY indicates the coupling of two protons within a small number of chemical bonds, whereas ^{1}H-^{1}H TOCSY reveals correlations in larger interconnected spin systems [188]. The network of spin couplings in the new cyclobutane ring is shown in the ^{1}H-^{1}H TOCSY spectrum (see Figure S7).

All described observations are consistent with the environmental photochemistry of ALT already described in the literature [87].

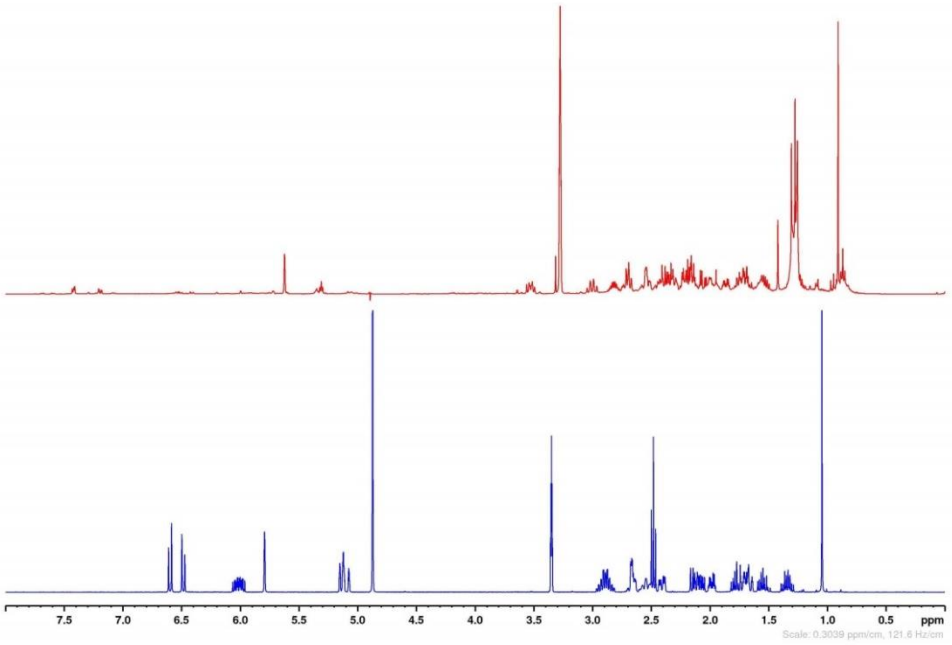


Figure S3. ^{1}H-NMR spectrum of ALT (blue) and mixture of ALT photo-products (red) in methanol-d4.

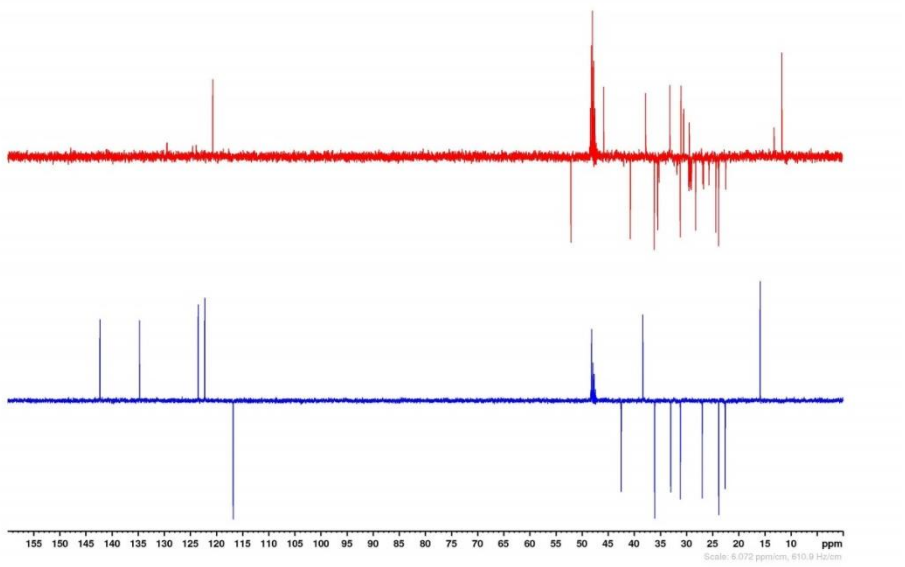


Figure S4. DEPT135 spectrum of ALT (blue) and mixture of ALT photo-products (red) in methanol-d4.

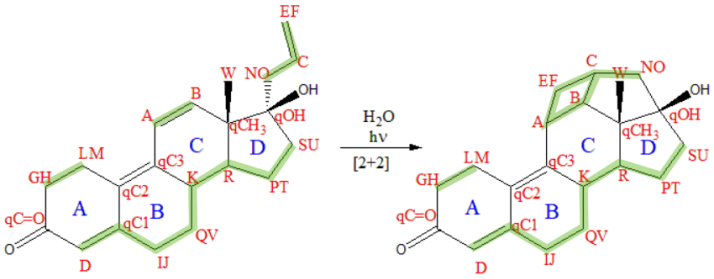


Figure S5. Proposed pathway of the photo-isomerization of ALT via [2+2] cycloaddition to form the primary photo-product ALT-CAP (modified from [87]); green shaded bonds belong to a network of spin couplings according to ^{1}H-^{1}H TOCSY data.

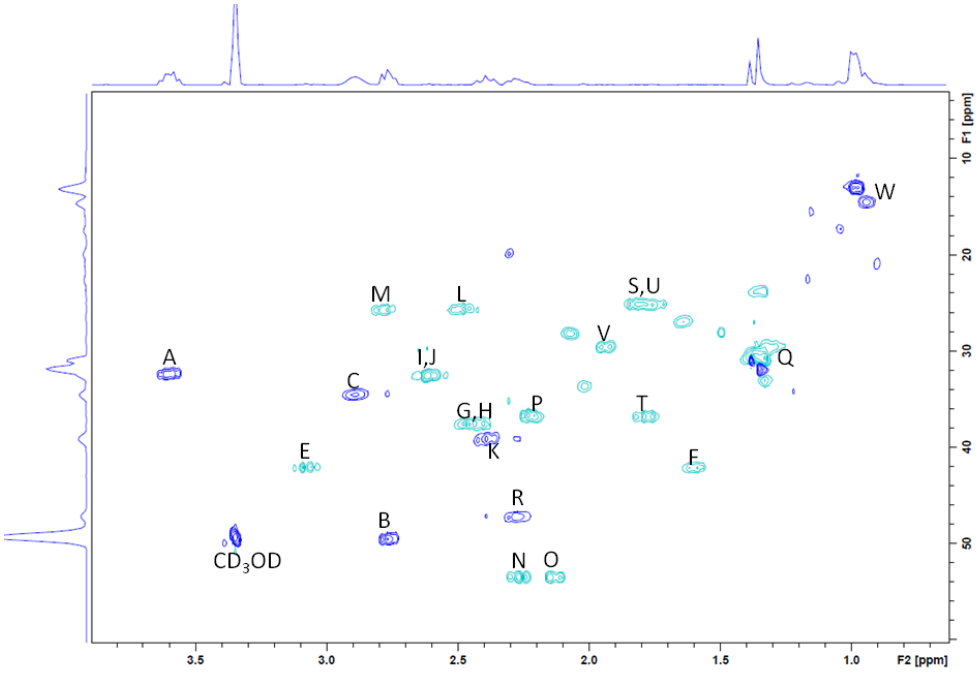


Figure S6. ^{1}H-^{13}C HSQC spectrum of altrenogest primary photo-product ALT-CAP in methanol-d4 with structural assignments according to Figure S5; section from 0.5 ppm to 4.0 ppm.

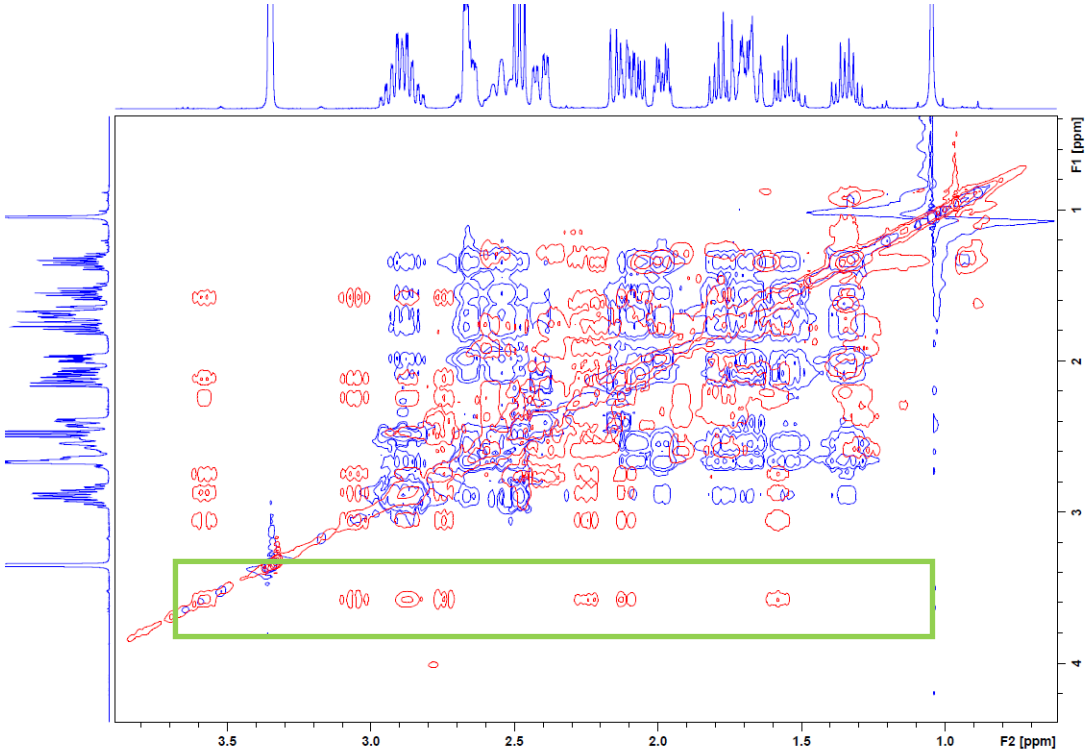


Figure S7. ^{1}H-^{1}H TOCSY spectrum of ALT (blue) and its primary photo-product (ALT-CAP, red) in methanol-d4; section from 0.5 ppm to 4.0 ppm. The spin coupling network of the new cyclobutane ring in the primary photo-product can be confirmed by the spin couplings of the protons at 3.60 ppm and 3.08 ppm with 2.89, 2.77, 2.27, 2.13, 1.60 ppm.

Table S11. NMR data for ALT in methanol-d4 with structural assignments.

	$\delta^{1}H^{a,d}$	$\delta^{13}C^{b,d}$	DEPT135/DEPT90	COSY/(TOCSY)	HSQC	HMBC (2/4-bond (H-C)	HMBC (3-bond (H-C)
A	6.59 (d, *J* = 10.08 Hz)	125.1	CH	B	B	qC3	qCH3, C(K),qC2
B	6.49 (d, *J* = 10.08 Hz)	144.0	CH	A	A	qCH3	qC3
C	6.01 (m)	136.4	CH	E,F,N,O			
D	5.79 (s)	123.9	CH	I,J,(Q,V,K,R,P,T,S,U)		C(L,M)	C(I,J),qC2
E	5.14 (d, J_{cis} = 10.28)[c]	118.6	CH2	C,F,N,O			C(N,O)
F	5.09 (d, J_{trans} = 17.16)[c]	118.6	CH2	C,E,N,O			C(N,O)
G	2.89 (m)	25.4	CH2	H,L,M	H,L,M	C(L,M),qC1	qC2
H	2.89 (m)	25.4	CH2	G,L,M	G,L,M	C(L,M),qC1	qC2
I	2.67 (m)	32.9	CH2	D,J,Q,V,(K,R,P,T,S,U)		C(Q,V),qC1	K,C(D)
J	2.67 (m)	32.9	CH2	D,I,Q,V,(K,R,P,T,S,U)		C(Q,V),qC1	K,C(D)
K	2.56 (m)	40.1	CH	Q,V,R, (D,I,J,P,T,S,U)			
L	2.48 (t, *J* = 7.32 Hz)[c]	37.8	CH2	G,H,M	G,H,M	C(G,H),qC2	
M	2.48 (t, *J* = 7.32 Hz)[c]	37.8	CH2	G,H,L	G,H,L	C(G,H),qC2	
N	2.41 (m)	44.2	CH2	C,O,E,F		qOH, C(C)	C(E,F)
O	2.13 (m)	44.2	CH2	C,N,E,F		C[C],qOH	C(E,F)
P	2.05 (m)	34.7	CH2	U,T, (D,I,J,Q,V,K,R,S)		C(N,O),C(S,U)	qOH ,qCH3
Q	1.99 (m)	28.67	CH2	I,J,V,K,R,(D,P,T,S,U)			qC1,qC3
R	1.78 (m)	49.9	CH	K,P,T,S,U(D,I,J,P,T,Q,V)		C(K),qCH3	qC3,C(W)
S	1.72 (m)	24.31	CH2	U,T,P, (D,I,J,Q,V,K,R)		qOH	qCH3
T	1.68 (m)	34.7	CH2	P,U,(D,I,J,Q,V,K,R,S)		C(R),C(S,U)	qOH
U	1.55 (m)	24.31	CH2	P,T, (D,I,J,Q,V,K,R,S)		C(P,T),C(K)	C(R)
V	1.34 (m)	28.67	CH2	Q,I,J,K(R,D,P,T,S,U)		C(I,J),C(K)	
W	1.05 (s)	17.6	CH3			qCH3	qOH,C(B)
qC=O							
qC1		160.4					
qC2		128.2					
qC3		144.9					
qCH3		50.9				B,CH3	A
qOH		82.9					

[a]400.13 MHz, [b]100.6 MHz, [c]assigned using ^{1}H-^{1}H JRES data, [d]assigned using ^{1}H-^{13}C HSQC data.

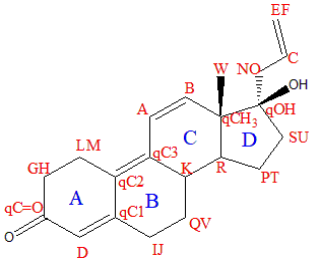
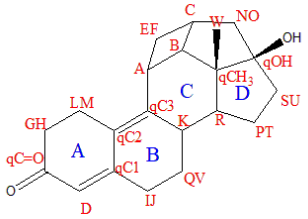


Table S12. NMR data for the photo-isomer of ALT ([2+2] cycloaddition product; ALT-CAP) in methanol-d4 with structural assignments.

	$\delta^{1}H^{a,c}$	$\delta^{13}C^{b,c}$	DEPT135/DEPT90	COSY/(TOCSY)	HSQC	HMBC (2/4-bond (H-C)	HMBC (3-bond (H-C)
A	3.60 (m)	32.46	CH	E,F,B, (C,N,O)	E	qC3,C(B)	qC2,qCH3,C(K)
B	2.77 (m)	49.6	CH	A,C, (E,F,N,O)	C	C(W),qCH3,C(C)	C[R],C(E,F)
C	2.89 (m)	34.54	CH	B,E,F,N,O,(A)	B	C(B)	qCH3,C(A)
D	5.69 (s)	122.15	CH	J,I, (Q,V,K,R,P,T,S,U)			qC2, C(G,H), C(I,J)
E	3.08 (m)	42.13	CH2	A,C,F, (B,N,O)	A		C(B),qC3,C(O,N)
F	1.6 (m)	42.13	CH2	A,C,E, (B,N,O)		C(A), C(C)	qC3,C(N,O)
G	2.48 (m)	37.61	CH2	H,L,M		C(L,M)	qC2
H	2.4 (m)	37.61	CH2	G,L,M		C(L,M)	qC2
I	2.61 (m)	32.63	CH2	D,J,Q,V, (K,R,P,T,S,U)		C(D),C(K)	
J	2.61 (m)	32.63	CH2	D,I,Q,V, (K,R,P,T,S,U)		C(D),C(K)	
K	2.39 (m)	39.17	CH	Q,R,V, (D,I,J,P,T,S,U)	R		qC2
L	2.48 (m)	25.78	CH2	G,H,M		C(G,H)	
M	2.78 (m)	25.78	CH2	G,H,L		qC2,C(G,H)	qC1
N	2.27 (m)	53.54	CH2	C,O		qOH,C(C)	qCH3,C(E,F),C(S,U)
O	2.13 (m)	53.54	CH2	C,N		qOH, C(C)	C(B)
P	2.22 (m)	36.86	CH2	R,T,S,U, (D,I,J,Q,V,K)			qOH,qCH3
Q	1.94 (m)	29.62	CH2	I,J,Q, (D,K,R,P,T,S,U)			qC1,qC3
R	2.28 (m)	47.28	CH	K,P,T, (D,I,J,Q,V,S,U)	K	qCH3	qOH,C(S,U)
S	1.78 (m)	25.22	CH2	P,T,U, (D,I,J,Q,V,K,R)		qOH,C(P,T)	C[R]
T	1.78 (m)	36.86	CH2	P,R,S,U, (D,I,J,Q,V,K)			qCH3,qOH
U	1.78 (m)	25.22	CH2	P,T,S, (D,I,J,Q,V,K,R)		qOH	qCH3
V	1.29 (m)	29.62	CH2	I,J,V,K, (D,R,P,T,S,U)		qC1	C(K),C(I,J)
W	0.96 (s)	13.29	CH3			qCH3	qOH,C(B)
qC=O							
qC1		161.1				I,J	L,M,Q,V
qC2		126.7				D,A,L,M,K	
qC3		156.1				A,K	L,M,V,E
qCH3		54.6				B,W	A,C,N,U,S,T
qOH		95.6				N,O,S,U,W	



[a]400.13 MHz, [b]100.6 MHz, [c]assigned using ^{1}H-^{13}C HSQC data.

Detection of secondary photo-products of ALT

The most abundant NMR signals can be explained by the formation of the primary photo-product ALT-CAP. However, other ^{1}H NMR and ^{13}C NMR signals of lower intensity were also detected. New signals with chemical shifts in the range of conjugated double bonds (7–8 ppm) and in the range of alcohols (4.5–5.5 ppm) were observed.

Therefore, full MS data of the photo-product mixture of ALT was checked for further characteristic *m*/*z* signals of possible photo-products. Data processing using MZmine 2 revealed, that the five following *m*/*z* ratios were the most abundant: 327.1951 (@ 6.19 min), 304.2607 (@ 9.93 min), 282.2788 (@ 9.93 min), 226.9513 (@ 1.75 min), 202.1800 (@ 1.76 min). However, in comparison to blank and standard solutions only the *m*/*z* ratio 327.1951 was uniquely detected in the photo-isomerization product mixture.

A secondary photo-product with the precursor mass of *m*/*z* 327.1963 ($[M+H]^+$) was reported before [87]. It is described as a result from a photo-hydration process of ALT-CAP, which reverts back to ALT-CAP in the dark [87]. However, the addition of water to ALT-CAP (*m*/*z* 311.2006 ($[M+H]^+$), as described by the authors, leads to a precursor mass of *m*/*z* 329.2111 ($[M+H]^+$). Therefore, the observed *m*/*z* ratio does not concur with the proposed molecule structure, but with oxygenated species of ALT-CAP.

In the extracted ion chromatogram of the *m*/*z* 327.1951 at least three peaks were observed (see Figure S8, top). The product ion spectrum of the peak at 6.0 min is shown Figure S8 (bottom). Some fragments can be explained by the basic chemical structure of ALT-CAP, but the exact position of the hydroxyl group cannot be determined. Thus, it is proposed that the oxidation of the primary photo-product (ALT-CAP) results in various secondary photo-products of minor concentrations. The earlier retention time relative to ALT and its primary photo-product is in line with the presence of more polar functional groups resulting from oxidation, i.e., hydroxylated products.

The reaction process could be explained by a free radical chain reaction with hydroperoxide intermediates. The presence of hydrogen peroxide in aqueous solutions under ultraviolet radiation is well described [189, 190]. The formation of hydrogen peroxide involves $OH^{\bullet}$, $HO_2^{\bullet}$ free radicals and the superoxide anion $O_2^{\bullet -}$, which can be formed via photochemical reduction of oxygen by natural humic materials [189, 190]. These free radicals are capable of oxidizing a wide spectrum of organic compounds [190].

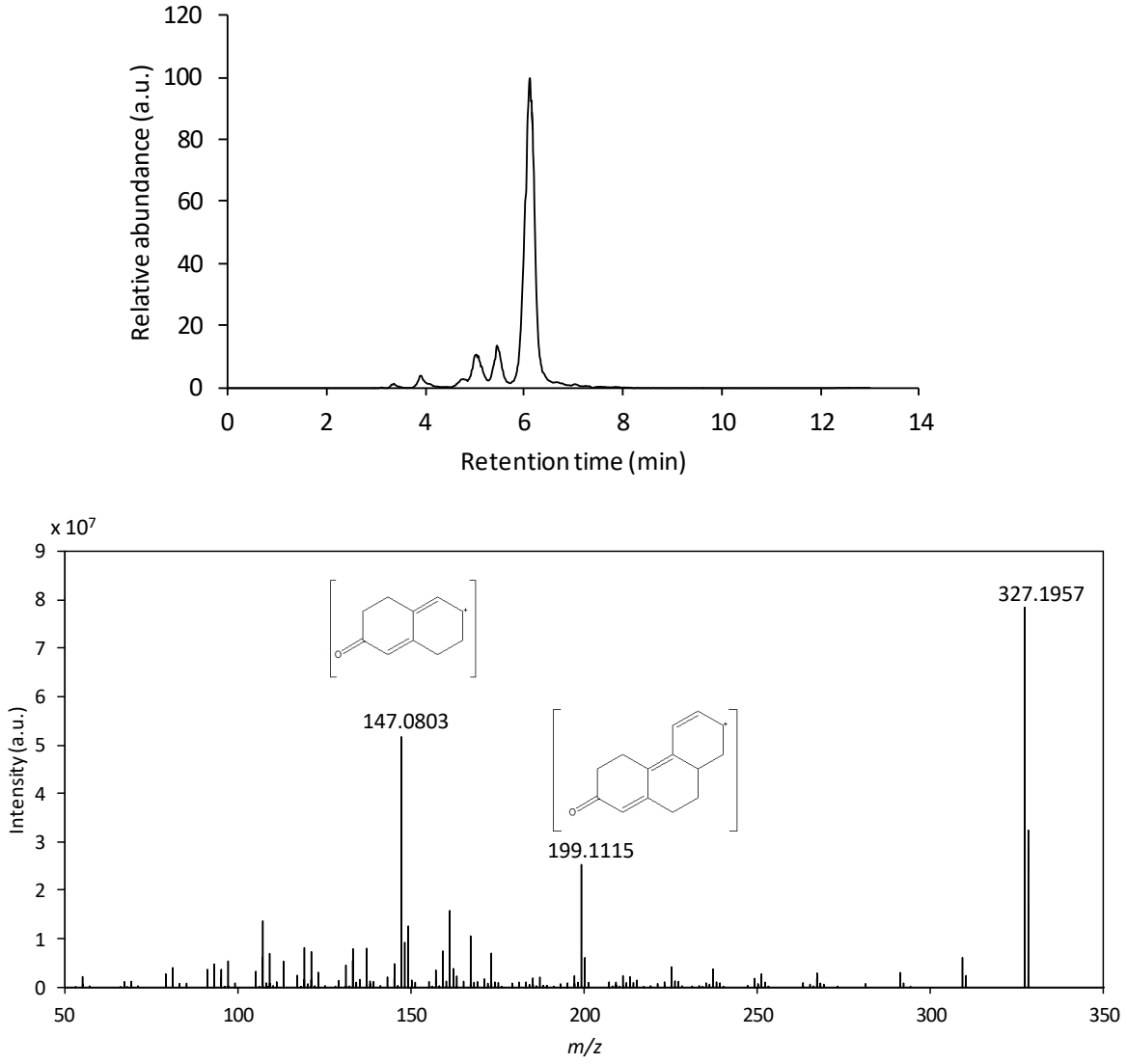


Figure S8. Top: extracted ion chromatogram of the secondary photo-isomerization products of ALT (*m/z* 327.1951 $[M+H]^+$); bottom: product ion spectrum of secondary photo-product of ALT-CAP (@ 6.0 min; *m/z* 327.1951 $[M+H]^+$) with proposal of fragmentation products.

In conclusion, the peak observed at 7.3 min was identified as primary photo-product ALT-CAP. Besides the primary photo-product several minor secondary hydroxylated products of ALT-CAP were observed.

8.1.2 ALT glucuronides

LC-HRMS data acquisition

In order to identify the ALT glucuronides, a full scan with a resolving power of 70,000 (FWHM at *m*/*z* of 200) in the range of *m*/*z* 50–750 was carried out. Data-dependent fragmentation was performed with a resolving power of 17,500 (FWHM at *m*/*z* of 200) using a collision energy of 35 eV. Mass tolerance was set to 5 ppm. All other LC-HRMS conditions were identical to the quantitative method (see Section 3.2.2.2). Xcalibur software was used for qualitative analysis (Thermo Scientific, Idstein Germany).

Synthesis of ALT glucuronides

Synthesis was performed according to [176]. In brief, ALT (500 µg, 1.6 µmol), uridine 5′-diphosphoglucuronic acid (UDA) (1 mg, 1.7 µmol) and S9 rat liver fractions (500 µg protein) were incubated in 50 mM Tris–HCl buffer (pH 7.4) with 10 mM $MgCl_2$ at 37 °C overnight (18 h) in a final reaction volume of 500 µL. 1.0 mL of methanol (-20°C) was added to stop the reaction and the sample was centrifuged (21,500 x g, 5 min, 5°C) for protein precipitation. The supernatant was evaporated to dryness at a temperature of 50°C under nitrogen gas flow, reconstituted in 0.5 mL methanol/water (60:40 (v/v)) and filtered (0.45 µm filter) prior to LC-HRMS analysis for reaction monitoring.

Synthesis of ALT glucuronides was confirmed by LC-HRMS analysis. The extracted ion chromatogram with the accurate mass of *m*/*z* 487.2327 showed two peaks at 5.3 min and 6.2 min corresponding to the glucuronides of the photo-isomer ALT-CAP and of ALT, respectively (see Figure S9). The earlier retention time of the more polar glucuronides compared to the parent molecule (7.3 and 7.6 min) is consistent with a reversed phase retention mechanism. Product ion spectra obtained with a collision energy of 35 eV showed characteristic fragments of ALT and the characteristic neutral loss of 176.0309 Da for the glucuronic acid moiety was observed. Yield of the reaction could not be assessed because the product could not be purified in sufficient yield for NMR analysis. ALT is poorly soluble in water, but the *in vitro* synthesis with the help of the enzymatic activity of S9 rat liver fractions can only be performed in aqueous buffer. However, the mass spectrometric data is very helpful for the identification of glucuronic acid conjugation products of ALT.

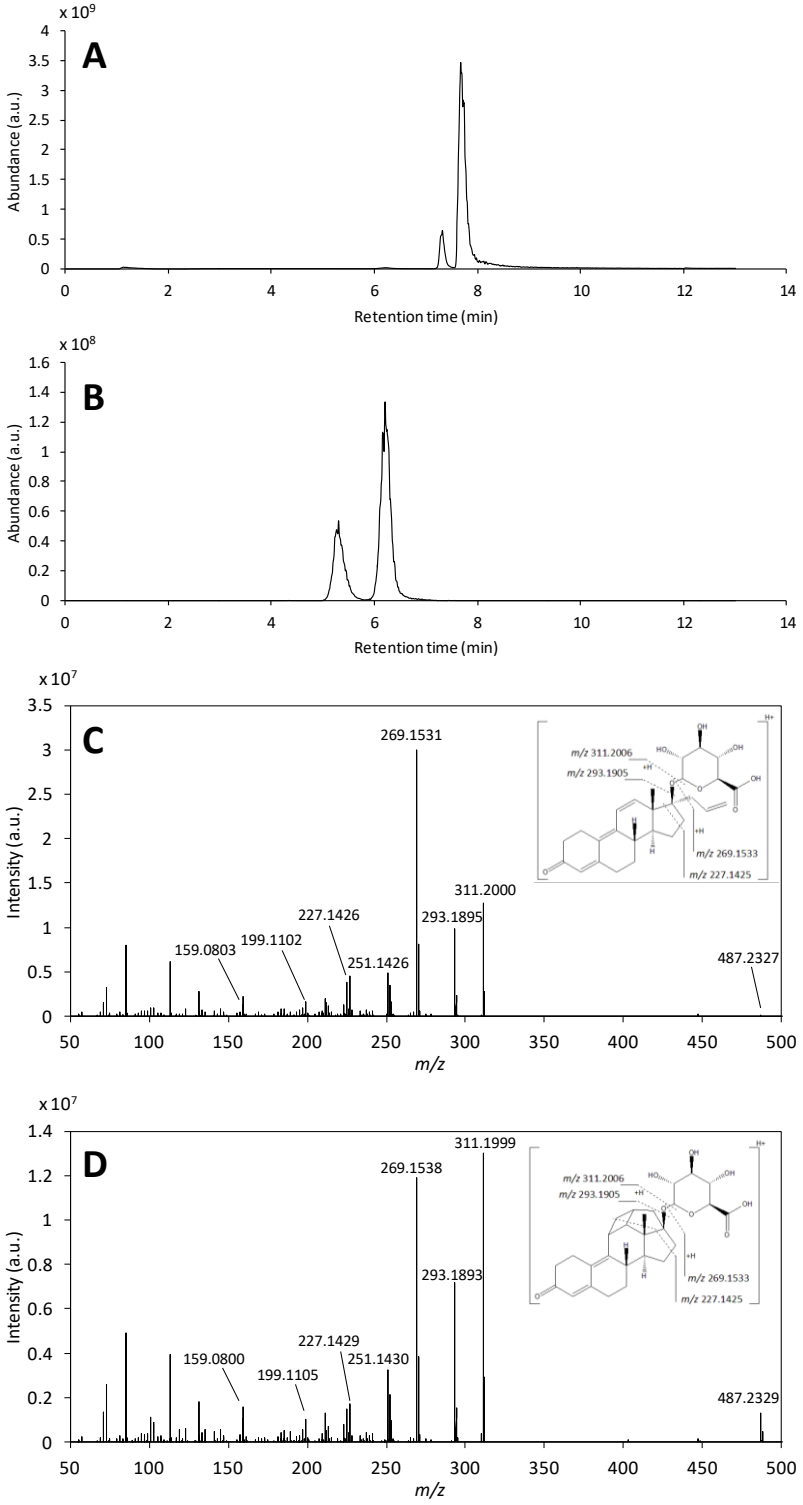


Figure S9. A: extracted ion chromatogram (*m/z* 311.2006) of ALT isomers (@ 7.31 min (ALT-CAP) and 7.68 min (ALT)); B: extracted ion chromatogram (*m/z* 487.2327) of ALT-CAP glucuronide (@ 5.29 min) and ALT glucuronide (@ 6.21 min); C/D: product ion spectrum of ALT glucuronide (precursor ion *m/z* 487.2327; $[M+H]^+$ @ 6.21 min (C) and 5.29 min (D)) following incubation of S9 rat liver fractions with ALT.

8.1.3 ALT sulfate

LC-HRMS data acquisition for the identification of ALT sulfate

In order to identify ALT sulfate, a full scan with a resolving power of 70,000 (FWHM at *m/z* of 200) in the range of *m/z* 50–750 was carried out in negative ion mode. Data-dependent fragmentation was performed with a resolving power of 17,500 (FWHM at *m/z* of 200) using a collision energy of 35 eV. Mass tolerance was set to 5 ppm. All other LC-HRMS conditions were identical to the quantitative method (see Section 3.2.2.2). Xcalibur software was used for qualitative analysis (Thermo Scientific, Idstein Germany).

Synthesis of ALT sulfate

The sulfate conjugate of ALT was synthesized by dicyclohexylcarbodiimide-mediated sulfation under mild conditions [191].

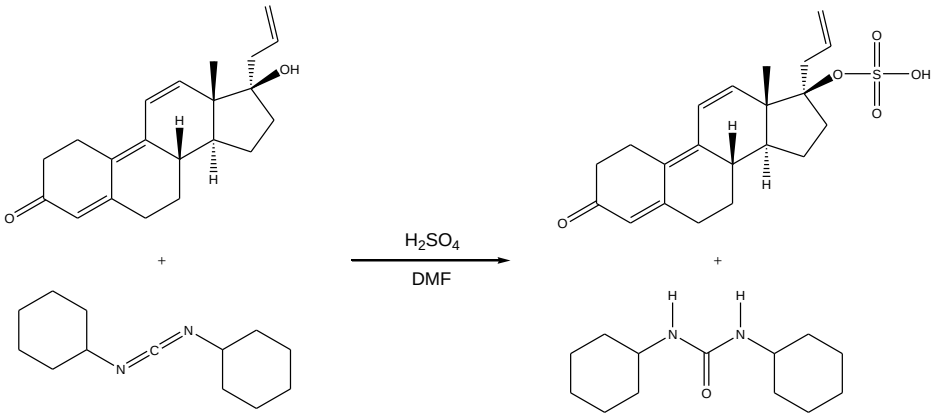


Figure S10. Reaction mechanism of sulfation of ALT; modified according to [191].

In brief, ALT (31 mg, 0.1 mmol), dissolved in 1.2 mL dimethylformamide (DMF), was added to N,N'-dicyclohexylcarbodiimide (103 mg, 0.5 mmol), dissolved in 1.6 mL DMF in an ice bath. After the addition of H_2SO_4 (1.8 µL, 0.15 mmol) the reaction mixture was diluted in 1.2 mL DMF, kept at 0°C and occasionally shaken for 15 min. After centrifugation (10 min, 2,500 x g) to separate the precipitated dicyclohexylurea from the reaction mixture, an aliquot of the supernatant was diluted with methanol (50:50 (v/v)) for further clean-up. The solution was transferred to a SPE cartridge (OASIS WAX®, 6 mL, 200 mg, purchased from Waters, Eschborn, Germany), preconditioned with 4 mL of methanol and 4 mL of methanol/DMF (50:50 (v/v)). After a washing step with 4 mL of methanol and 4 mL of a mixture of 2% HCOOH in methanol and 1 min drying *in vacuo*, the analyte was eluted with 4 mL of 5% NH_4OH in methanol. The eluates were collected, combined and evaporated to dryness at a

temperature of 50°C under nitrogen gas flow. The synthesis product was reconstituted in 0.5 mL methanol/water (60:40 (v/v)) with 1% HCOOH and filtered (0.45 µm filter). After dilution by factor 1000 the samples were transferred to a HPLC vial prior to LC-HRMS analysis for reaction monitoring.

Synthesis of ALT sulfate was confirmed by LC-HRMS analysis. The extracted ion chromatogram of the accurate mass of *m*/*z* 389.1433 showed a peak at 3.83 min, which was assigned to ALT sulfate. A slight fronting of the peak was observed. Furthermore, a small peak at 2.86 min was observed. Hence, a weak isomerization cannot be excluded. However, a pronounced isomerization of ALT did not occur since the synthesis was performed in organic solvents and a pronounced photo-isomerization is mainly observed in aqueous solutions.
The earlier retention time of the more polar sulfate compared to the parent drug (7.67 min) is consistent with a reversed phase retention mechanism. Product ion spectra obtained with a collision energy of 35 eV showed characteristic fragments of ALT sulfate [69]. The fragment with the *m*/*z* 96.9590 was observed as main fragment, which corresponds to the mass-to-charge ratio of HSO_4^-. Further characteristic ions were detected at minor intensities.
Unfortunately, the stability of ALT sulfate was very low in organic and aqueous solutions leading to an elimination of ALT sulfate in 2–4 hours. Thus, it was not possible to synthesize a purified product in good yield. However, the mass spectrometric data is very valuable for the identification of sulfate conjugation products of ALT.

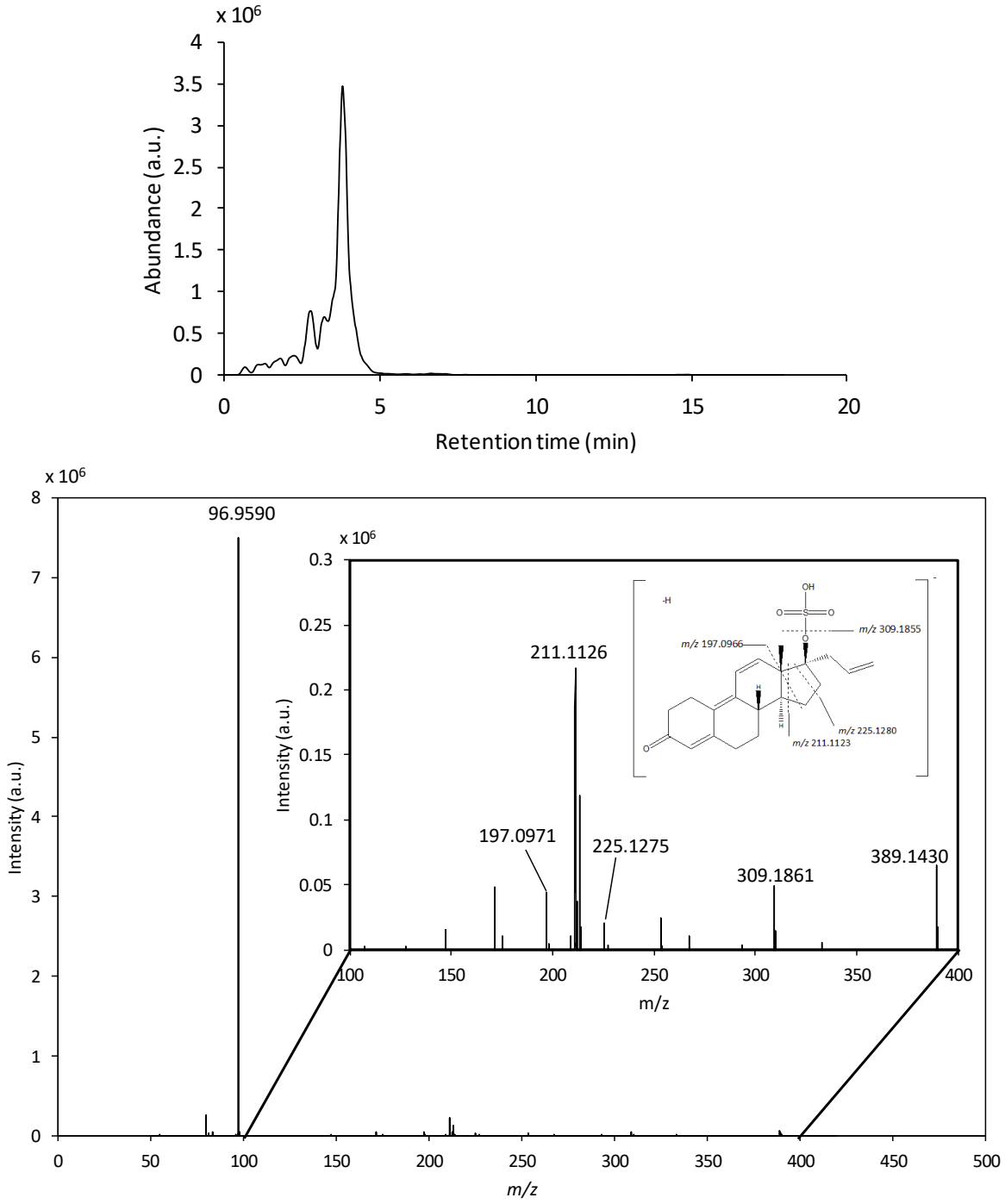


Figure S11. Top: extracted ion chromatogram of ALT sulfate (*m/z* 389.1433); bottom: product ion spectrum of peak eluting at 3.83 min (precursor ion *m/z* 389.1433; [M-H]-) with characteristic fragments of ALT sulfate.

8.2 Method validation for quantitative ^{1}H NMR determination of creatinine in urine

The ^{1}H NMR spectrum of creatinine shows two characteristic resonances, i.e., a singlet with a chemical shift of δ = 3.05 ppm corresponding to the methyl group and a singlet at δ = 4.06 ppm (CH_2 group) relative to referencing substance TSP. The ratio of the integrals is 3:2 corresponds to the number of protons responsible for the signals.

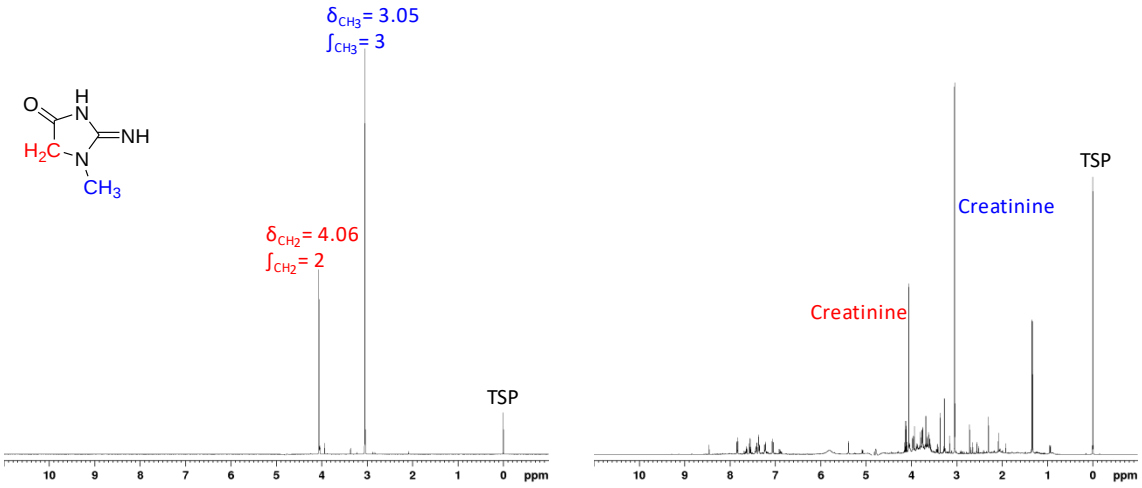


Figure S12. Left: ^{1}H NMR spectrum of creatinine standard at a concentration of 5 mg/mL. Right: representative urine ^{1}H NMR spectrum.

Decision limit, detection limit and determination limit under repeatability conditions according to DIN 32645 were calculated by regressing measured values on actual values. A urine sample was spiked with creatinine in the concentration range from 0.05–1 mg/mL and creatinine was quantified via standard addition. The creatinine concentration of the urine sample was added to the spiking concentration and the sum was taken as actual value. The external calibration curve ranged from 0.05–10 mg/mL creatinine to cover the working range. Recovery was calculated on a function of concentration and was about 90% in the working range. Quantification of creatinine in urine samples was performed via external calibration using signal at δ = 3.05 ppm.

Linearity of the method was confirmed via external standards in the range of 0.05 to 10 mg/mL creatinine (see Figure S13). The decision limit and the determination limit were 0.03 mg/mL and 0.1 mg/mL, respectively. The measurement uncertainty of results corrected for recovery was below 2% within the working range.

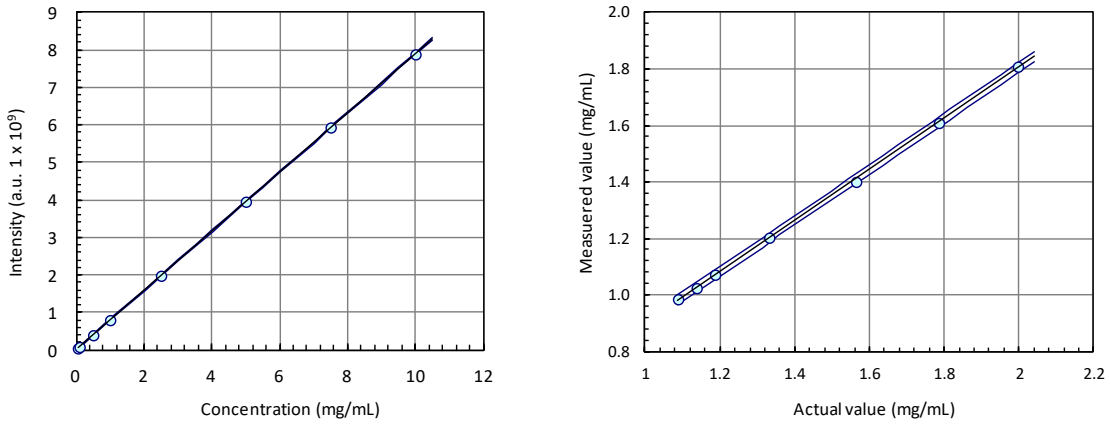


Figure S13. Left: external calibration curve covering the working range from 0.05 mg/mL to 10 mg/mL. Right: OLS regression to determine decision limit, detection limit and determination limit according to DIN 32645, blue line: prediction band (PI: p = 0.95).

Table S13. Performance characteristics according to DIN 32645.

	Creatinine (mg/mL)		
DIN 32645	**Decision limit**	**Detection limit**	**Determination limit**
	0.03	0.06	0.1

9 Appendix D: Instrumentation, materials and chemicals

9.1 Instrumentation

9.1.1 Liquid chromatography high resolution mass spectrometry

UHPLC-Q-Orbitrap:

Dionex Ultimate 3000 RS UHPLC (Thermo Scientific, Idstein, Germany) consisting RS binary pump, WPS-3000 RS autosampler and TCC-3000 RS column compartment coupled with Q-Orbitrap Q Exactive Focus (Thermo Scientific, Idstein, Germany) equipped with a heated electrospray (HESI-II) source.

Column (project 1): Phenomenex Luna Omega C18 polar column (1.6 µm, 100 Å, 100 × 2.1 mm; Phenomenex, Aschaffenburg, Germany).
Column (project 2): Phenomenex Prodigy C18 column (5 µ, ODS (3), 100 Å, 150 mm x 3 mm; Phenomenex, Aschaffenburg Germany).

UHPLC-Q-TOF:

Agilent 1290 Infinity UHPLC system, consisting of a G4220A binary pump, G4226A autosampler and G1316C thermostated column compartment (Agilent Technologies, Waldbronn, Germany) coupled with X500R QTOF (Sciex, Darmstadt, Germany) mass spectrometer equipped with a Turbo V source operating in positive electrospray ionization (ESI) mode.

Column: Phenomenex Luna Omega C18 polar column (1.6 µm, 100 Å, 100 × 2.1 mm; Phenomenex, Aschaffenburg, Germany).

9.1.2 Liquid chromatography (semi–preparative)

Agilent 1100 Infinity HPLC system, consisting of a G1379A degasser, G1312A binary pump, G1313A autosampler, G1353A diode array detector with an G1364A fraction collector (Agilent Technologies, Waldbronn, Germany).

Column: Phenomenex Luna Omega C18 polar column (1.6 µm, 100 Å, 100 × 2.1 mm; Phenomenex, Aschaffenburg, Germany).

9.1.3 Nuclear magnetic resonance (NMR) spectroscopy

Bruker Avance III HD Ascend 400 spectrometer (BrukerBiospin, Rheinstetten, Germany) equipped with a 5-mm broadband observe (BBO) and broadband inverse (BBI) probe with Z-gradient coils, using a Bruker Automatic Sample Changer (Sample Xpress). ^{1}H NMR spectra were acquired at 400.13 MHz.

9.1.4 General laboratory equipment

Analytical balance:	Satorius Cubis MSA225P-1CE-DI (Satorius, Göttingen, Germany)
Centrifuges:	Mikro 200R (Hettich, Tuttlingen, Germany) Rotina 380 (Hettich, Tuttlingen, Germany)
Folded filters:	MN 615 ¼ (Macherey-Nagel, Düren, Germany)
Funnel:	Various sizes (VWR, Darmstadt, Germany)
HPLC-Vials:	WE1905 (Ziemer, Langerwehe, Germany)
Laboratory scale:	Satorius BP 4100S (Satorius, Göttingen, Germany)
pH-meter:	InoLab pH 7110 (Xylem Analytics, Weilheim, Germany)
Piston pipette for different volumes:	Various models (Eppendorf, Hamburg, Germany)
Sample concentrator:	DRI Block Heater Digital DB-3D (Techne, Staffordshire, UK)
Shaking machine:	Universalschüttler SM30B (Edmund Bühler, Bodelshausen, Germany)
Shaking water bath:	Julabo SW 22 (Julabo, Seelbach, Germany)
SPE vacuum manifold system:	J.T. Baker SPE 24G (VWR, Darmstadt, Germany)
Test tubes:	1,5 mL, 15 mL and 50 mL (Sarstedt, Nümbrecht, Germany)
Test tube shaker:	Hei-Mix Multi-Reax (Heidolph, Schwabach, Germany)
Thermoshaker:	Thermomixer comfort (Eppendorf, Hamburg, Germany)
Ultrasonic bath:	Sonorex RK162 (Bandelin, Berlin, Germany)
Vacuum centrifuge:	RVC 2-33IR (Christ, Osterrode am Harz, Germany)

9.2 Chemicals

All chemicals and solvents used were of highest analytical grade and purchased from Merck (Darmstadt, Germany). Water with a resistance of >18.0 MΩ/cm was prepared in-house by an arium pro VF Ultrapure Water System (Satorius, Göttingen, Germany). All analytical standards of veterinary drugs were purchased from Sigma-Aldrich (Steinheim, Germany), LGC (Wesel, Germany), and Biozol (Eching, Germany) or were provided by the German National Reference Laboratory for veterinary drug residue analysis (Berlin, Germany). Lysophosphatidylcholine 17:0 was supplied by Sigma-Aldrich (Steinheim, Germany) and tripeptide prolylphenylalanylglycine (PFG) was custom-synthesized with a purity >98% (trifluoroacetate salt) by BioCat GmbH (Heidelberg, Germany).
S9 liver fractions from male rat (Sprague-Dawley) were purchased from Sigma Aldrich (Steinheim, Germany).

Stock solutions were prepared as follows:

- cephalosporines, β-lactam antibiotics and tetracyclines at a concentration of 200 µg/mL in methanol,
- corticosteroids and halofuginone at a concentration of 100 µg/mL in methanol,
- quinolones at a concentration of 100 µg/mL in acetonitrile:methanol (1:1 (v/v)),
- sulfonamides/trimethoprim at a concentration of 100 µg/mL in acetonitrile and
- lincosamides, macrolides, pleuromutilins and internal standards at a concentration of 1000 µg/mL in methanol.

Mix standards were prepared by dilution with water/methanol (90:10 (v/v)).

Stock solutions of altrenogest and altrenogest-d5 were prepared at a concentration of 100 µg/mL in methanol. All standard solutions were stored at 4°C.

9.3 Buffers and solutions

Acetonitril (0.1% formic acid):	1 L acetonitrile + 1 mL formic acid
Ammonium formate buffer (50 mM, pH 2.6):	1.57 g ammonium formate dissolved in 500 mL water, pH adjusted to pH 2.6 with formic acid
EDTA-McIlvaine buffer:	31.5 g citric acid + 35.6 g Na_2HPO_4 x H_2O + 93 g titriplex III dissolved in water, pH was adjusted to pH 4 with phosphoric acid
Dipotassium phosphate buffer (1.5 M, pH 7.4) in D_2O containing 5 mM TSP:	40.827 g K_2HPO_4, 39 mg NaN_3 and 200 mg trimethylsilylpropanoic acid-2,2,3,3-d4 (TSP) dissolved in 200 mL D_2O, pH adjusted to 7.4 with KOH
Magnesium chloride buffer (150 mM):	3.05 g $MgCl_2 \cdot 6H_2O$ dissolved in 100 mL water
Methanol (0.1% formic acid):	1 L methanol + 1 mL formic acid
Sodium acetate buffer (50 mM, pH 4.8):	4.1 g sodium acetate dissolved in 1 L of water, pH adjusted to pH 4.8 with acetic acid
Tris-HCl buffer (50mM, pH 7.4):	5 mL Tris-HCl buffer (1M, pH 7.4) + 95 mL water
Water (0.1% formic acid):	1 L water + 1 mL formic acid

10 Appendix E: MATLAB codes

10.1 Exemplary code for PCA with subsequent QDA , Monte Carlo cross-validation and PLS-DA

```
clear
clc

P = readmatrix('data.txt');
yQDA = readmatrix('y.txt');

T = P(:,2:71);

min = min(min(T));
min = min/2;

n = length(P);

%Replace 0 by half of minimum

for j = 1:70
    for i = 1:n
        if T(i, j)> 0
            T(i,j)= T(i,j);
            else T(i,j)= min;
        end
    end
end

% Normalization by the sum of all peaks
SumPeaks = sum(T);

NormT2 = T./SumPeaks;

% Log-transform data
logNormT = log(NormT2);

% Transpose data
X = logNormT';

% Standardize data
m = mean(X);
s = std(X);

for j = 1:70
    for i = 1:n
       Xstd (j , i) = (X(j , i) - m(1 , i))/s(1 , i);
    end
end
```

```
%PCA
[coeff,score,latent] = pca(Xstd);
varXstd = var(Xstd);
sumvarXstd = sum(varXstd);
explained = latent/sumvarXstd*100;

%scores of control samples
x1 = [score([1,2,4:18,24,31,33:36,38:51,58,65,69,70], 1)];
x2 = [score([1,2,4:18,24,31,33:36,38:51,58,65,69,70], 2)];
x3 = [score([1,2,4:18,24,31,33:36,38:51,58,65,69,70], 3)];

%scores of positive samples
y1 = [score([3,19:23,25:30,32,37,52:57,59:64,66:68], 1)];
y2 = [score([3,19:23,25:30,32,37,52:57,59:64,66:68], 2)];
y3 = [score([3,19:23,25:30,32,37,52:57,59:64,66:68], 3)];

%scores of spiked QC samples
a1 = [score([1,9,24,33,35,43,58,69], 1)];
a2 = [score([1,9,24,33,35,43,58,69], 2)];
a3 = [score([1,9,24,33,35,43,58,69], 3)];

%scores of control QC samples
b1 = [score([2,15,31,34,49,65,70], 1)];
b2 = [score([2,15,31,34,49,65,70], 2)];
b3 = [score([2,15,31,34,49,65,70], 3)];

%scores of positive QC samples
c1 = [score([3,32,37,68], 1)];
c2 = [score([3,32,37,68], 2)];
c3 = [score([3,32,37,68], 3)];

% scores plot

figure(1)
scatter3(x1, x2, x3, 'bo')
hold on
scatter3(y1, y2,y3,'mo')
hold on
scatter3(a1, a2,a3,'b+')
hold on
scatter3(b1, b2,b3, 'bo', 'filled')
hold on
scatter3(c1, c2,c3,'mo', 'filled')
grid on
xlabel('PC1 (12.8%)');
ylabel('PC2 (11.3%)');
zlabel('PC3 (4.5%)')
title ('PCA');

%% QDA
% Classification: negative (0) and positive (1)

y = yQDA(:,1);

Negative = yQDA([1,2,4:18,24,31,33:36,38:51,58,65,69,70],1);
```

```
Positive = yQDA([3,19:23,25:30,32,37,52:57,59:64,66:68],1);

for i = 1:length(Negative(:,1))
    assign_Negative (i,1) = 0;
end

for i = 1:length(Positive(:,1))
    assign_Positive (i,1) = 1;
end

scores_Negative = score([1,2,4:18,24,31,33:36,38:51,58,65,69,70],
[1:3]);

scores_Positive = score([3,19:23,25:30,32,37,52:57,59:64,66:68],
[1:3]);

obj_NegPos = fitcdiscr([scores_Negative ; scores_Positive], [as-
sign_Negative ; assign_Positive], 'DiscrimType', 'quadratic');

% 10 fold CV

cvmodel = crossval(obj_NegPos , 'Kfold' , 10);

NMC = kfoldLoss(cvmodel, 'mode', 'individual')


%% Monte Carlo cross-validation

for a = 1:5000

z = 10; %test set
N = 5000; %repetitions

r = randperm(70, z);

NegativeX = X([1,2,4:18,24,31,33:36,38:51,58,65,69,70],:);
PositiveX = X([3,19:23,25:30,32,37,52:57,59:64,66:68],:);

X_sort = [NegativeX;PositiveX];
X_training = X_sort;
X_training(r,:)=[];

NegativeY = yQDA([1,2,4:18,24,31,33:36,38:51,58,65,69,70],1);
PositiveY = yQDA([3,19:23,25:30,32,37,52:57,59:64,66:68],1);

yQDA_sort = [NegativeY;PositiveY];
y_training = yQDA_sort;
y_training(r,:)=[];

%Standardization of training set

mCV = mean(X_training);
sCV = std(X_training);

for j = 1:60
    for i = 1:n
```

```
       XStd_training (j , i) = (X_training(j , i) - mCV(1 ,
i))/sCV(1 , i);
    end
end

[coeffCV,scoreCV,latentCV] = pca(XStd_training);


for j = 1:z
    for i = 1:n
       XStd_Test (j , i) = (X_sort(r(j) , i) - mCV(1 , i))/sCV(1 ,
i);
    end
end

Scores_Test = XStd_Test*coeffCV;
Scores_Test = Scores_Test(:,[1:3]);

idx0 = y_training == 0;
out0 = sum(idx0(:));

assign_Negative2 = [];

for b = 1:out0
   assign_Negative2(b,1) = 0;
end

idx1 = y_training == 1;
out1 = sum(idx1(:));

assign_Positive2 = [];

for c = 1:out1
   assign_Positive2(c,1) = 1;
end

for d = 1:out0
   scoresTraining_Negative = scoreCV([1:d], [1:3]);
end

for e = out0+1
    scoresTraining_Positive = scoreCV([e:end], [1:3]);
end

obj_NegPosCV = fitcdiscr([scoresTraining_Negative ; scoresTrain-
ing_Positive], [assign_Negative2 ; assign_Positive2], 'DiscrimType',
'quadratic');

Classification = predict(obj_NegPosCV, Scores_Test);

for k = 1:z
    ClassificationCV(a,k)= Classification(k,:);
end

for k = 1:z
```

```
    if (ClassificationCV(a,k) == 0) && (yQDA_sort(r(k),1) == 0)
        tnPLS(a,k)=1;
        tpPLS(a,k)=0;
        fnPLS(a,k)=0;
        fpPLS(a,k)=0;
    elseif (ClassificationCV(a,k) == 1) && (yQDA_sort(r(k),1) == 1)
        tnPLS(a,k)=0;
        tpPLS(a,k)=1;
        fnPLS(a,k)=0;
        fpPLS(a,k)=0;
    elseif (ClassificationCV(a,k) == 1) && (yQDA_sort(r(k),1) == 0)
        tnPLS(a,k)=0;
        tpPLS(a,k)=0;
        fnPLS(a,k)=0;
        fpPLS(a,k)=1;
    elseif (ClassificationCV(a,k) == 0) && (yQDA_sort(r(k),1) == 1)
        tnPLS(a,k)=0;
        tpPLS(a,k)=0;
        fnPLS(a,k)=1;
        fpPLS(a,k)=0;
    end

end

end

% Visualization of cross-validation

tnPLS = sum(sum(tnPLS));
fpPLS = sum(sum(fpPLS));
tpPLS = sum(sum(tpPLS));
fnPLS = sum(sum(fnPLS));

rate(1,1) = tnPLS/(tnPLS + fpPLS); %Specificity
rate(2,1) = fpPLS/(tnPLS + fpPLS); %1-Specificity
rate(3,1) = tpPLS/(tpPLS + fnPLS); %Sensitivity
rate(4,1) = fnPLS/(tpPLS + fnPLS); %1-Sensitivity


figure(2)
bar(rate, 'FaceColor',[0 .5 .5],'EdgeColor',[0 .9
.9],'LineWidth',1.5)
labels = {'true negative', 'false positive', 'true positive', 'false
negative'};
xt = get(gca, 'XTick');
text(xt, rate, labels, 'HorizontalAlignment','center', 'VerticalA-
lignment','bottom')

%% PLS-DA

[XL,YL,XS,YS,beta,PCTVAR,MSE,stats] = plsregress(Xstd,y,2);

% Scores plot
x1PLS = XS([1,2,4:18,24,31,33:36,38:51,58,65,69,70],1);
x2PLS = XS([1,2,4:18,24,31,33:36,38:51,58,65,69,70],2);
```

```
y1PLS = XS([3,19:23,25:30,32,37,52:57,59:64,66:68],1);
y2PLS = XS([3,19:23,25:30,32,37,52:57,59:64,66:68],2);

%spiked QC samples
a1PLS = XS([1,9,24,33,35,43,58,69], 1);
a2PLS = XS([1,9,24,33,35,43,58,69], 2);

%control QC samples
b1PLS = XS([2,15,31,34,36,49,65,70], 1);
b2PLS = XS([2,15,31,34,36,49,65,70], 2);

%positive QC samples
c1PLS = XS([3,32,37,68], 1);
c2PLS = XS([3,32,37,68], 2);


figure(3); % PLS-DA scores plot
scatter(x1PLS,x2PLS,'bo');
hold on
scatter(y1PLS,y2PLS,'mo');
hold on
scatter(a1PLS,a2PLS,'bo', 'filled');
hold on
scatter(b1PLS,b2PLS,'b+');
hold on
scatter(c1PLS,c2PLS,'mo', 'filled');
hold off
xlabel('Component 1');
ylabel('Component 2');
title('PLS-DA');
lgd = legend({'control','positive'});


%% Monte Carlo cross-validation

for a = 1:5000

z = 10; %test set
N = 5000; %repetitions


r = randperm(70, z);


X_training = X;
X_training(r,:)=[];

y_training = y;
y_training(r,:)=[];

%Standardization of training set

m = mean(X_training);
s = std(X_training);

for j = 1:60
```

```
    for i = 1:n
       XStd_training (j , i) = (X_training(j , i) - m(1 , i))/s(1 ,
i);
    end
end


[P,q,Scores,mu,BETA] = plsregress(XStd_training, y_training, 2);



for j = 1:z
    for i = 1:n
       XStd_Test (j , i) = (X(r(j) , i) - m(1 , i))/s(1 , i);
    end
end

for j=1:z
    L(j,1) = 1 ;
end

XSi = [L,XStd_Test];

for k = 1:z
YM(a,k) = XSi(k,:)*BETA;
end

% Discriminant analysis

for k = 1:z
if YM(a,k) > 0.5
   YMT(a,k) = 1;
else YMT(a,k) = 0;
end
end


for k = 1:z

    if (YMT(a,k) == 0) && (y(r(k),1) == 0)
        tnPLS(a,k)=1;
        tpPLS(a,k)=0;
        fnPLS(a,k)=0;
        fpPLS(a,k)=0;
    elseif (YMT(a,k) == 1) && (y(r(k),1) == 1)
        tnPLS(a,k)=0;
        tpPLS(a,k)=1;
        fnPLS(a,k)=0;
        fpPLS(a,k)=0;
    elseif (YMT(a,k) == 1) && (y(r(k),1) == 0)
        tnPLS(a,k)=0;
        tpPLS(a,k)=0;
        fnPLS(a,k)=0;
        fpPLS(a,k)=1;
    elseif (YMT(a,k) == 0) && (y(r(k),1) == 1)
        tnPLS(a,k)=0;
```

```
        tpPLS(a,k)=0;
        fnPLS(a,k)=1;
        fpPLS(a,k)=0;
    end

end

% Test for over-fitting
for k = 1:z
    for j = 1:z
    SRes(a,k) = (y(r(j)) - YM(a,k)).^2;
    end
end

end

SSRes = sum(SRes);
SSRes = sum(SSRes);

MSE_CV = SSRes/(z*N)

if MSE(2,3) < MSE_CV
    OF = {'yes'}
   else OF = {'no'}
end

% Visualization of cross-validation

tnPLS = sum(sum(tnPLS));
fpPLS = sum(sum(fpPLS));
tpPLS = sum(sum(tpPLS));
fnPLS = sum(sum(fnPLS));

rate(1,1) = tnPLS/(tnPLS + fpPLS); %Spezificity
rate(2,1) = fpPLS/(tnPLS + fpPLS); %1-Spezificity
rate(3,1) = tpPLS/(tpPLS + fnPLS); %Sensitivity
rate(4,1) = fnPLS/(tpPLS + fnPLS); %1-Sensitiviy

figure(4)
bar(rate, 'FaceColor',[0 .5 .5],'EdgeColor',[0 .9
.9],'LineWidth',1.5)
labels = {'true negative', 'false positive', 'true positive', 'false
negative'};
xt = get(gca, 'XTick');
text(xt, rate, labels, 'HorizontalAlignment','center', 'VerticalA-
lignment','bottom')
```

10.2 Exemplary code for batch normalization based on QC samples

```
clear
clc

P = readmatrix('data.txt');
yQDA = readmatrix('y.txt');
```

```
u = readmatrix('Factor.txt'); %Mean values of all variables from QC
samples from day 2 of measurement divided by mean values of all var-
iables from QC samples from day 1

n = length(P);

T = P(:,2:71);

min = min(min(T));
min = min/2;

n = length(P);

%Replace 0 by half of minimum

for j = 1:70
    for i = 1:n
        if T(i, j)> 0
            T(i,j)= T(i,j);
            else T(i,j)= min;
        end
    end
end



%% Normalization based on QC samples


for j = 1:70
    for i = 1:n
       SkalT_QC (i , j) = (T(i , j)/ u(i , 1));
    end
end

SkalT2 =[T(:,[1:34]), SkalT_QC(:,[35:70])];

SumPeaks2 = sum(SkalT2);

for j = 1:70
    for i = 1:n
       NormSkalT2 (i , j) = (SkalT2(i , j)/SumPeaks2(1 , j));
    end
end


% Log-transform data
logNormSkalT2 = log(NormSkalT2);

% Transpose data
X2 = logNormSkalT2';

% Standardize data
m2 = mean(X2);
s2 = std(X2);
```

```
for j = 1:70
    for i = 1:n
       Xstd2 (j , i) = (X2(j , i) - m2(1 , i))/s2(1 , i);
    end
end


% PCA
[coeff,score,latent] = pca(Xstd2);
varXstd = var(Xstd2);
sumvarXstd = sum(varXstd);
explained = latent/sumvarXstd*100;


% negative samples
x1 = [score([1,2,4:18,24,31,33:36,38:51,59,66,69,70], 1)];
x2 = [score([1,2,4:18,24,31,33:36,38:51,59,66,69,70], 2)];


% positive sammples
y1 = [score([3,19:23,25:30,32,37,52:58,60:65,67:68], 1)];
y2 = [score([3,19:23,25:30,32,37,52:58,60:65,67:68], 2)];


% spiked QC samples
a1 = [score([1,9,24,33,35,43,59,69], 1)];
a2 = [score([1,9,24,33,35,43,59,69], 2)];


% negative QC samples
b1 = [score([2,15,31,34,36,49,66,70], 1)];
b2 = [score([2,15,31,34,36,49,66,70], 2)];


% positive QC samples
c1 = [score([3,32,37,53,68], 1)];
c2 = [score([3,32,37,53,68], 2)];


% Scores plot

figure(1)
scatter(x1, x2, 'bo')
hold on
scatter(y1, y2,'mo')
hold on
scatter(a1, a2,'bo', 'filled')
hold on
scatter(b1, b2,'b+')
hold on
scatter(c1, c2,'mo', 'filled')
grid on
xlabel('PC1 (12.2%)');
ylabel('PC2 (6.1%)');
title ('PCA');
```

10.3 Exemplary code for normalization based on correction curve fitted to the QC samples

```
clear
clc

% Quadratic regression

P2 = readmatrix('QC_samples.txt');
n = length(P2);

T = P2(:,[2:end]);

min = min(min(T));
min = min/2;

% Replace 0 by min/2
for i = 1:n
    for j = 1:13
        if P2(i, j)> 0
            P2(i, j)= P2(i, j);
            else P2(i, j)= min;
        end
    end
end

y2 = P2(:, [2:end]);
y2 = y2';

X2 = readmatrix('Regression_x.txt'); % run number

QC_spiked = y2([1:4],:);
X_QC_spiked = X2([1:4],:);

QC_negative = y2([5:8],:);
X_negative = X2([5:8],:);

QC_positive = y2([9:12],:);
X_positive = X2([9:12],:);

X_QC_spiked = X_QC_spiked(:,2);

X_negative = X_negative(:,2);

X_positive = X_positive(:,2);


for i=1:n

p_QCspiked (:,i) = polyfit(X_QC_spiked,QC_spiked(:,i),2);

end


for i=1:n
```

```
p_QCnegative(:,i) = polyfit(X_negative,QC_negative(:,i),2);

end


for i=1:n

p_QCpositive(:,i) = polyfit(X_positive,QC_positive(:,i),2);

end


% Matrix of coefficient

QM_Coeff1 = [p_QCspiked(1,:);p_QCnegative(1,:); p_QCpositive(1,:)];

mean_Coeff1 = mean(QM_Coeff1);

QM_Coeff2 = [p_QCspiked(2,:);p_QCnegative(2,:); p_QCpositive(2,:)];

mean_Coeff2 = mean(QM_Coeff2);

QM_Coeff3 = [p_QCspiked(3,:);p_QCnegative(3,:); p_QCpositive(3,:)];

mean_Coeff3 = mean(QM_Coeff3);

p_QCtotal = [mean_Coeff1;mean_Coeff2;mean_Coeff3];


Xkor = [1:44];

for i=1:n
    for j = 1:44
    ykor(i,j) = polyval(p_QCtotal(:,i),Xkor(:,j));
    end
end

min = min(min(ykor));
min = min/2;

% replace 0 by min/2

for j = 1:44
    for i = 1:n
        if ykor(i, j)> 0
            ykor(i,j)= ykor(i,j);
            else ykor(i,j)= min;
        end
    end
end


%% Normalization based on coefficient

R = readmatrix('data.txt'); % in Excel alle 0 replaced by min/2
```

```
yQDA = readmatrix('y.txt');

n = length(R);

R = R(:,2:45);

for i = 1:n
    for j = 1:44
        Rkor(i,j) = (R(i,j)/ykor(i,j));
    end
end

Rkor1 = Rkor;
Rkor1(:,23) = []; % Delete since samples was contaminated

% Log-transform data
logNormRkor = log(Rkor1);

%Transpose data
X = logNormRkor';

%Standardize data
m = mean(X);
s = std(X);

for j = 1:43
    for i = 1:n
       Xstd (j , i) = (X(j , i) - m(1 , i))/s(1 , i);
    end
end


%PCA
[coeff,score,latent] = pca(Xstd);
varXstd = var(Xstd);
sumvarXstd = sum(varXstd);
explained = latent/sumvarXstd*100;

% negative samples
x1 = [score([1,2,4:16,18,20:22,24,31,42,43], 1)];
x2 = [score([1,2,4:16,18,20:22,24,31,42,43], 2)];


% positive samples
y1 = [score([3,19,23,25:30,32:41], 1)];
y2 = [score([3,19,23,25:30,32:41], 2)];


% spiked QC samples
a1 = [score([1,9,24,42], 1)];
a2 = [score([1,9,24,42], 2)];


% negative QC samples
b1 = [score([2,15,31,43], 1)];
b2 = [score([2,15,31,43], 2)];
```

```
% positive QC samples
c1 = [score([3,19,38,41], 1)];
c2 = [score([3,19,38,41], 2)];


% Scores plot

figure(1)
scatter(x1, x2, 'bo')
hold on
scatter(y1, y2,'mo')
hold on
scatter(a1, a2,'b+')
hold on
scatter(b1, b2, 'bo', 'filled')
hold on
scatter(c1, c2,'mo', 'filled')
grid on
xlabel('PC1 (21.8%)');
ylabel('PC2 (17.8%)');
title ('PCA');
```

List of abbreviations

Abbreviations	Explanation
ADI	Acceptable daily intake
AGC	Automatic gain control
ALT	Altrenogest
ALT-CAP	Altrenogest cycloaddition product
AUROC	Area under the receiver operating characteristic curve
BiGG	Biochemical, Genetic and Genomic knowledge base
BLASTp	protein Basic Local Alignment Search Tool
CCα	Decision limit according to Commission Decision 2002/657/EC
CCβ	Detection capability according to Commission Decision 2002/657/EC
CD	Commission decision
CE	Collision energy
CL	Decision limit according to DIN 32645
COSY	Correlation spectroscopy
COX	cyclooxygenase
CVMP	Committee for Medicinal Products for Veterinary Use
DC	Direct current
DEPT	Distortionless enhancement by polarization transfer
DoE	Design of experiments
DP	Declustering potential
EMA	European Medicines Agency
FDA	Food and Drug Administration
FET	Fisher's exact test
FFT	Fast Fourier Transformation
FSH	Follicle-stimulating hormone
FWHM	Full width at half maximum
GnRH	Gonadotropin-releasing hormone
HCD	Higher energy collisional dissociation
H-ESI	Heated electro spray ionization
HMBC	Heteronuclear multiple bond correlation
HMDB	Humane Metabolome Database
HPLC	High performance liquid chromatography
HSQC	Heteronuclear single quantum correlation
IDA	Information dependent acquisition
ID ratio	Identification ratios
KEGG	Kyoto Encyclopedia of Genes and Genomes
LC-HRMS	Liquid chromatography coupled with high resolution mass spectrometry
LDA	Linear discriminant analysis
LH	Luteinizing hormone
LOD	Detection limit according to DIN 32645
LOQ	Determination limit according to DIN 32645
Lyso-PC	Lysophosphatidylcholine
MRL	Maximum residue limit

Abbreviations	Explanation
MSEP	Mean square error of prediction
MU	Measurement uncertainty
NMC	Number of missclassifications
NMR	Nuclear magnetic resonance
NOAEL	No observed adverse effect level
NRCP	National residue control plan
NSAIDs	Non-steroidal anti-inflammatory drugs
OLS	Ordinary least-squares regression
Q	Quadrupole
QC	Quality control
QDA	Quadratic discriminant analysis
QqQ	Triple quadrupole mass spectrometer
PC	Principal component
PCA	Principal components analysis
PEC	Predicted environmental concentration
PFG	Prolylphenylalanylglycine
PLS-DA	Partial least squares-discriminant analysis
PNEC	Predicted no effect concentration
PRESS	Predicted residual sum of squares
PRM	Parallel reaction monitoring
RF	Radiofrequency
ROC	Receiver Operating Characteristic
RSD	Relative standard deviation
SD_r	Standard deviation of repeatability
SD_{wiR}	Standard deviation of within-laboratory reproducibility
SIM	Selected ion monitoring
SPE	Solid phase extraction
TOCSY	Total correlation spectroscopy
TOF	Time-of-flight
U_{Cr}	Urinary creatinine
UHPLC	Ultra high performance liquid chromatography
UNEP	United Nations Environment Programme
VICH	Veterinary International Conference on Harmonization
WHO	World Health Organization
WWTP	Waste water treatment plant

List of tables

List of figures

List of tables in Appendixes

List of figures in Appendixes

List of publications

Articles in peer-reviewed journals

Liesenfeld, S.; Steliopoulos, P.; Hamscher, G., Comprehensive metabolomics analysis of non-targeted LC-HRMS data provides valuable insights regarding the origin of veterinary drug residues. *Journal of Agricultural and Food Chemistry*, **2020**, 68 (44), 12493-12502.

Liesenfeld, S.; Steliopoulos, P.; Wenig, S.; Gottstein, V.; Hamscher, G., Comprehensive LC-HRMS metabolomics analyses for the estimation of environmental inputs of altrenogest in pig breeding. *Chemosphere*, **2022**, 287, in-press (https://doi.org/10.1016/j.chemosphere.2021.132353).

Conference abstracts

Liesenfeld, S.; Steliopoulos, P.; Hamscher, G., *Entwicklung und Validierung einer LC-MS/MS-Methode zur quantitativen Bestimmung von Altrenogest in Schweineurin*. Poster presentation at Arbeitstagung der LChG RV Südwest, Gießen, Germany, Mar 6–Mar 7, 2018.

Liesenfeld, S.; Steliopoulos, P.; Hamscher, G., *In-house validation study of a confirmatory method for altrenogest in swine urine and plasma samples using LC-MS/MS and in-vitro synthesis of altrenogest glucuronide*. Poster presentation at 8th International symposium on Hormone and Veterinary Drug Residue Analysis, Ghent, Belgium, May 22–May 25, 2018.

Liesenfeld, S.; Steliopoulos, P.; Hamscher, G., *Entwicklung und Validierung einer LC-MS/MS-Methode zur quantitativen Bestimmung von Altrenogest in Gewebeproben vom Schwein*. Oral presentation at Arbeitstagung der LChG RV Südwest, Stuttgart-Hohenheim, Germany, Mar 5–Mar 6, 2019.

Liesenfeld, S.; Steliopoulos, P.; Hamscher, G., *Entwicklung einer Auswerteprozedur zur systematischen Analyse von Orbitrap-Daten aus Metabolismus-Studien*. Poster presentation at 48. Deutscher Lebensmittelchemikertag, Dresden, Germany, Sept 16–Sept18, 2019.

Acknowledgments

I would like to express my sincere gratitude to Prof. Dr. Gerd Hamscher for giving me the opportunity to do research in this interesting scientific field and to prepare this thesis in collaboration between his working group and the Chemisches und Veterinäruntersuchungsamt (CVUA) Karlsruhe. Especially, I would like to thank him for his dedicated supervision and his continous guidance.

I would like to thank Prof. Dr. Mirko Bunzel for agreeing to act as second referee and Prof. Dr. Bernhard Spengler as well as Prof. Dr. Melanie Hamann for agreeing to act as members of the examination committee.

Furthermore, I would like to thank Dr. Stephan Walch and Dr. Gerhard Marx for giving me the opportunity to engage in this research project and perform all pratical analyses in the laboratories of the CVUA Karlsruhe.

My special thanks go to Dr. Panagiotis Steliopoulos whose guidance, support and encouragement has been invaluable to me throughout this time. I am grateful for all you have taught me, starting with new insights in chemometrics, a structured approach in all scientific questions and last but not least a precise presentation of my thoughts (not only on scientific topics). Thank you, Pana, for being a great mentor!

I would like to thank Dr. Tanja Zacharias for giving me the opportunity to collect urine samples from ALT-treated gilts at the Landesanstalt für Schweinezucht Boxberg and all team members for their help and all the effort they put into the successful sample collection.

I would like to thank Svenja Wenig and Vera Gottstein for their help and support with all NMR-related questions and for performing the NMR analyses.

I am extending my thanks to my colleagues in the Devision 3 at CVUA Karlsruhe: Dr. Gudrun Kopf for her valuable input in analytical questions and in all topics related to the NRCP management, Dr. Marika Pritzl for being a great sparring partner during the urine sample collection on a freezing day in December, Anke Rullmann for being a great friend and always listening to me when I needed someone to talk to as well as Uta Kraus and Franziska Scharmann for their encouraging words at all times.

Further, I would like to thank the whole lab team of the Department of Veterinary Drug Residue Analysis at CVUA Karlsruhe for their helpful tips concerning all practical labwork, their help in sample management, their willingness to share their expertise with me and their encouraging words at all times. Thank you to Ingrid Arnold-Götz, Karina Dammert, Ursula Ehrstein, Susanne Eichhorn, Tanja Grünbacher, Heike Haas, Tina Herbster, Eleonora Horst, Rita Oberüber, Susanne Reichert, Corinna Schmidt, Michael Stoverock and Silke Tulatz.

I would also like to thank Constanze Sproll, Iris Hundeck-Dudenhöffer and Ingrid Kübel for their continuously provided encouragement and their help with the LC-TOF-analysis.

Finally, I would like to thank my friends and family for their loving support and especially my husband David! Thank you for being extremely supportive of me at every step of the way, for your patience and understanding during the hard times of this thesis and for providing happy distractions when my mind needed some rest. I am truly thankful for having you in my life!