

**Untersuchungen von Kaffee mit Kernspin-
resonanzspektroskopie unter Anwendung von
zielgerichteten und nicht zielgerichteten Verfahren**

Zur Erlangung des akademischen Grades einer

DOKTORIN DER NATURWISSENSCHAFTEN

(Dr. rer. nat.)

von der KIT-Fakultät für Chemie und Biowissenschaften

des Karlsruher Instituts für Technologie (KIT)

genehmigte

DISSERTATION

von

M. Sc. Vera Gottstein

1. Referent: Prof. Dr. Mirko Bunzel

2. Referent: Prof. Dr. Thomas Kuballa

Tag der mündlichen Prüfung: 09.02.2023

Eidesstattliche Versicherung

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbstständig verfasst habe und keine anderen als die angegebenen Hilfsmittel verwendet habe. Die wörtlich oder sinngemäß aus anderen Werken übernommenen Inhalte habe ich als solche kenntlich gemacht. Darüber hinaus erkläre ich, dass die Arbeit nicht anderweitig als Prüfungsarbeit oder als Dissertation bei einer anderen Fakultät verwendet wird oder wurde.

Karlsruhe, den 09.01.2023

Vera Gottstein

Inhaltsverzeichnis

1	Einleitung.....	1
1.1	Warenkunde Kaffee.....	2
1.1.1	Definition, Geschichte und Welthandel	2
1.1.2	Die Kaffeepflanze	3
1.1.3	Anbau von Kaffee	4
1.1.3.1	Biologisch zertifizierter Anbau.....	6
1.1.4	Ernte von Kaffee	7
1.1.5	Verarbeitung der Rohkaffeebohnen.....	8
1.1.6	Der Röstprozess.....	10
1.1.6.1	Maillard-Reaktion.....	10
1.1.6.2	Veränderung der Kaffeeinhaltsstoffe während des Röstprozesses.....	12
1.1.6.3	Röstverfahren	14
1.1.7	Chemische Zusammensetzung von Röstkaffee.....	15
1.1.7.1	Lipide.....	15
1.1.7.2	Stickstoffhaltige Verbindungen	16
1.1.7.3	Kohlenhydrate	17
1.1.7.4	Melanoidine	18
1.1.7.5	Organische Säuren	19
1.1.7.6	Phenolische Komponenten.....	20
1.1.7.7	Minorkomponenten.....	21
1.1.8	Kaffeedefekte.....	21
1.2	Kernspinresonanz-Spektroskopie.....	22
1.2.1	Physikalische Grundlagen und Messprinzip der NMR	23
1.2.1.1	Chemische Frequenzen.....	25
1.2.1.2	Spin-Spin-Kopplung	26
1.2.2	Grundlagen der quantitativen NMR.....	27
1.2.2.1	Methoden zur Quantifizierung.....	29
2	Zielsetzung	32
3	Material und Methoden.....	33
3.1	Methodenentwicklung für den Metabolomics-Ansatz.....	33
3.2	Untersuchungsmaterialien und Probenvorbereitung.....	34
3.3	Entwicklung einer nicht zielgerichteten Methode für die Untersuchung der lipophilen Metaboliten.....	35
3.3.1	Extraktionsmethode	36

3.3.2	Vorbereitung der Messlösung für die $^1\text{H-NMR}$ -spektroskopische Messung.....	37
3.3.3	Optimierung der Messparameter für die $^1\text{H-NMR}$ -spektroskopische Messung.....	38
3.4	Entwicklung einer nicht zielgerichteten Methode für die Untersuchung der hydrophilen Metaboliten	39
3.4.1	Extraktionsmethode	39
3.4.2	Vorbereitung der Messlösung für die $^1\text{H-NMR}$ -spektroskopische Messung.....	41
3.4.3	Optimierung der Messparameter für die $^1\text{H-NMR}$ -spektroskopische Messung.....	41
3.5	Nicht zielgerichtete Untersuchung mittels Verfahren der uni- und multivariaten Statistik.....	43
3.5.1	Methoden zur Datenvorbehandlung der $^1\text{H-NMR}$ -Spektren.....	43
3.5.2	Kruskal-Wallis Test zur Untersuchung der Signifikanz	46
3.5.3	Unüberwachte (unsupervised) multivariate Datenanalyse	47
3.5.4	Überwachte (supervised) multivariate Datenanalyse.....	49
3.5.5	Validierung der Klassifizierungsmodelle.....	51
3.5.6	Datenfusionierung der $^1\text{H-NMR}$ -spektroskopischen Daten der lipophilen und hydrophilen Metaboliten.....	52
3.6	Identifizierung der Kaffeemetaboliten mittels NMR-Spektroskopie	53
3.7	Zielgerichtete Analyse zur Quantifizierung verschiedener Kaffeemetaboliten	54
4	Ergebnisse und Diskussion.....	57
4.1	Probenmaterial	57
4.2	Entwicklung einer nicht zielgerichteten Methode für die Untersuchung der lipophilen Metaboliten.....	57
4.2.1	Extraktionsmethode	57
4.2.2	Vorbereitung der Messlösung für die $^1\text{H-NMR}$ -spektroskopische Messung.....	61
4.2.3	Optimierung der Messparameter für die $^1\text{H-NMR}$ -spektroskopische Messung.....	63
4.2.4	Überprüfung der Reproduzierbarkeit der Extraktion und der Messung	64
4.3	Entwicklung einer nicht zielgerichteten Methode für die Untersuchung der hydrophilen Metaboliten	65
4.3.1	Entwicklung der Extraktionsmethode	65
4.3.2	Vorbereitung der Messlösung für die $^1\text{H-NMR}$ -spektroskopische Messung.....	68
4.3.3	Optimierung der Messparameter für die $^1\text{H-NMR}$ -spektroskopische Messung.....	69
4.3.4	Überprüfung der Reproduzierbarkeit der Extraktion und der Messung	72
4.4	Nicht zielgerichtete Untersuchung durch Verfahren der uni- und multivariaten Statistik.....	73

4.4.1	Datenvorbehandlung der ¹ H-NMR-Spektren.....	73
4.4.2	Erstellung von Klassifizierungsmodellen basierend auf der Analyse der lipophilen und hydrophilen Kaffeemetaboliten zur Abgrenzung von Kaffee nach Kontinenten.....	74
4.4.2.1	Modellerstellung und interne Validierung.....	74
4.4.2.2	Externe Validierung.....	77
4.4.3	Erstellung von Klassifizierungsmodellen basierend auf der Analyse der lipophilen und hydrophilen Kaffeemetaboliten zur Abgrenzung von Kaffee nach ausgewählten Ländern	78
4.4.3.1	Modellerstellung und interne Validierung.....	78
4.4.3.2	Externe Validierung.....	83
4.4.4	Univariate Analyse der ¹ H-NMR-spektroskopischen Daten mittels Kruskal-Wallis Test.....	84
4.4.5	Identifizierung möglicher lipophiler Metaboliten als Ursache für die Abgrenzung von Kaffee nach ausgewählten Ländern.....	86
4.4.6	Identifizierung möglicher hydrophiler Metaboliten als Ursache für die Abgrenzung von Kaffee nach ausgewählten Ländern.....	90
4.5	Datenfusionierung der ¹ H-NMR-spektroskopischen Daten der lipophilen und hydrophilen Metaboliten für die Abgrenzung von Kaffee nach ausgewählten Ländern	93
4.6	Erstellung eines Klassifizierungsmodells basierend auf der Analyse der lipophilen und hydrophilen Kaffeemetaboliten zur Abgrenzung von Kaffee nach verschiedenen Anbaumethoden	96
4.7	Erstellung eines Klassifizierungsmodells basierend auf der Analyse der lipophilen und hydrophilen Kaffeemetaboliten zur Abgrenzung von Kaffee nach dem Röstgrad und dem Röstverfahren	98
4.8	Identifizierung lipophiler Kaffeemetaboliten mittels NMR-Spektroskopie.....	101
4.9	Identifizierung hydrophiler Kaffeemetaboliten mittels NMR-Spektroskopie	103
4.10	Zielgerichtete Analyse zur Quantifizierung ausgewählter Kaffeemetaboliten.....	106
4.10.1	Entwicklung einer Methode zur Quantifizierung ausgewählter Kaffeemetaboliten durch ¹ H-NMR-Spektroskopie und PULCON	106
4.10.2	Entwicklung einer HPLC-DAD-Referenzmethode zur Quantifizierung ausgewählter Kaffeemetaboliten.....	109
4.10.3	Vergleich der ¹ H-NMR-spektroskopischen und HPLC-DAD-Quantifizierungsmethode.....	110
4.10.4	Anwendung der ¹ H-qNMR-spektroskopischen Methode auf Kaffeeproben...	112
4.10.5	Anwendung der ¹ H-qNMR-spektroskopischen Methode auf Kaffeeproben mit verschiedenen Defekten.....	114
4.11	Zusammenfassende Diskussion	116
5	Zusammenfassung.....	124

Inhaltsverzeichnis

6	Literaturverzeichnis.....	126
7	Chemikalien und Instrumente	146
8	Arbeitsvorschriften.....	155
9	Anhang.....	174
10	Publikationsliste	339

Abkürzungsverzeichnis

16-OMC	16-O-Methylcafestol
1D	Eindimensional
¹H-NMR	Protonen-Kernspinresonanz-Spektroskopie (engl. Proton Nuclear Magnetic Resonance Spectroscopy)
2D	Zweidimensional
3-CQA	3-Caffeoylchinasäure (engl. 3-Caffeoylquinic Acid)
4-CQA	4-Caffeoylchinasäure (engl. 4-Caffeoylquinic Acid)
5-CQA	5-Caffeoylchinasäure (engl. 5-Caffeoylquinic Acid)
AGP	Arabinogalactan-Proteine
ANOVA	Varianzanalyse (engl. Analysis of Variance)
aq	Wässrig (engl. Aqueous)
AQ	Akquisitionszeit
ATP	Adenosintriphosphat
BG	Bestimmungsgrenze
CGA	Chlorogensäure (engl. Chlorogenic Acid)
CGAL	Chlorogensäurelacton (engl. Chlorogenic Acid Lactone)
CoA	CoenzymA
COSY	Engl. Correlation Spectroscopy
CQA	Caffeoylchinasäure (engl. Caffeoylquinic acid)
D1	Relaxationsverzögerung
diCQA	Dicafeoylchinasäure (engl. Dicafeoylquinic acid)
DIN	Deutsches Normungsinstitut
DMS	Dimethylsulfon
DMSO	Dimethylsulfoxid
DS	Anzahl Dummy Scans
EM	Engl. Exponential Function
EtOH	Ethanol
ERETIC	Engl. Eletronic Reference to Acess <i>in vivo</i> Concentrations
EW	Einwaage
FA	Fettsäuren (engl. Fatty Acids)
FID	Freier Induktionsabfall (engl. Free Induction Decay)
FT	Fourier-Transformation
FQA	Feruloylchinasäure (engl. Feruloylquinic Acid)
GABA	γ-Aminobuttersäure
GC	Gaschromatographie
GC-FID	Gaschromatographie gekoppelt mit Flammenionisationsdetektor
GC-C-IRMS	Gaschromatographie gekoppelt mit der Isotopenverhältnis-Massenspektroskopie
GC-MS	Gaschromatographie gekoppelt mit Massenspektrometrie
g.g.A.	Geschützte geographische Angabe
HMBC	Engl. Heteronuclear Multiple Bond Correlation
HMF	5-Hydroxymethylfurfural
HPLC	Hochleistungsflüssigkeitschromatographie (engl. High Performance Liquid Chromatography)
HPLC-DAD	Hochleistungsflüssigkeitschromatographie gekoppelt mit Diodenarray-Detektor

HPLC-MS	Hochleistungsflüssigkeitschromatographie gekoppelt mit Massenspektrometrie
HSQC	Engl. Hetero Nuclear Single Quantum Coherence
HSQC-TOCSY	Engl. Hetero Nuclear Single Quantum Coherence-Total Correlation Spectroscopy
IARC	Engl. International Agency for Research and Cancer
IStd	Interner Standard
KV	Kreuzvalidierung
LB	Engl. Line Broadening Faktor
LD	Lineare Diskriminanzfunktion
LD1	Erste lineare Diskriminanzfunktion
LDA	Lineare Diskriminanzanalyse
MCKV	Monte Carlo Kreuzvalidierung
MeOH	Methanol
MG	Molekulargewicht
MS	Massenspektrometrie
MUFA	Einfach ungesättigte Fettsäuren (engl. Mono Unsaturated Fatty Acids)
NAD	Nicotinamidadenin-Dinucleotid
NG	Nachweisgrenze
NMP	N-Methylpyridinium
NMR	Kernspinresonanz-Spektroskopie (engl. Nuclear Magnetic Resonance Spectroscopy)
NOESY	Engl. Nuclear Overhauser Effect Spectroscopy
NS	Anzahl Scans (Engl. Number of Scans)
NUS	Engl. Non Uniform Sampling
O1	Senderfrequenz Offset
OTA	Ochratoxin A
PC	Hauptkomponenten (eng. Principal Component)
PCA	Hauptkomponentenanalyse (engl. Principal Component Analysis)
PLS	Regression der partiellen kleinsten Quadrate (engl. Partial Least Square)
PLS-DA	Partielle kleinste Quadrate-Diskriminanzanalyse (engl. Partial Least Square Discriminant Analysis)
PUFA	Mehrfach ungesättigte Fettsäuren (engl. Poly Unsaturated Fatty Acids)
PULCON	Engl. Pulse Length Based Concentration Determination
QDA	Quadratische Diskriminanzanalyse
qNMR	Quantitative Kernspinresonanz-Spektroskopie (engl. Quantitative Nuclear Magnetic Resonance Spectroscopy)
QuantRef	Engl. Quantification Reference
RF	Radiofrequenzpuls
RG	Engl. Receiver Gain
SI	Engl. Size of Real Spectrum
SIMCA	Engl. Soft Independent Modeling of Class Analogy
S/N	Signal-zu-Rauschverhältnis (engl. Signal to Noise Ration)
Std	Standard
SW	Spektrale Weite
T1	Longitudinale Relaxationszeit
T2	Transversale Relaxationszeit
TAG	Triacylglyceride
TD	Anzahl der Datenpunkte (engl. Time Domain)

TMS	Tetramethylsilan
TOCSY	Engl. Total Correlation Spectroscopy
TSP	Trimethylsilylpropionat
UF	Ultrafiltration
UFA	Ungesättigte Fettsäuren (Unsaturated Fatty Acids)
VF	Verdünnungsfaktor
Vgl.	Vergleiche
WATERGATE	Engl. Water Supression by Gradient Tailored Excitation
WDW	Fensterfunktion (engl. Window Fuction)
WET	Engl. Water Supression Enhanced Through T_1 Effects

1 Einleitung

Authentizität von Lebensmitteln ist ein zunehmend wichtiges Thema in der Lebensmittelüberwachung. Ein Lebensmittel gilt als authentisch, wenn es mit den Angaben auf dem Etikett übereinstimmt. In diesem Zusammenhang gewinnt der Begriff Lebensmittelbetrug oder Food Fraud an Bedeutung. Unter Lebensmittelbetrug wird das Inverkehrbringen von Lebensmitteln verstanden, deren Originalzustand nicht mit deren Auslobung übereinstimmt. Dies können Auslobungen bezüglich der Tier- oder Pflanzenart, der Produktionsweise, der geographischen Herkunft oder wertgebender Bestandteile sein. Lebensmittelbetrug ist eine Ordnungswidrigkeit bis Straftat und der jährliche Schaden wird in der europäischen Union auf 30 Milliarden Euro geschätzt. Somit ist es von großem Interesse, mittels geeigneter Analyseverfahren die Authentizität der Lebensmittel zu überprüfen.

Lebensmittelbetrug im Bereich Kaffee umfasst fälschliche Auslobungen bezüglich der Art, des geographischen Ursprungs, des biologischen Anbaus oder auch des verwendeten Röstverfahrens. Mögliche Analyseverfahren zur Überprüfung der Authentizität umfassen chromatographische, spektroskopische und spektrometrische Methoden. In einem zielgerichteten Ansatz erfolgt dabei die Identifizierung und Quantifizierung definierter Inhaltsstoffe als Markersubstanzen. Ein Beispiel für die Überprüfung der Authentizität von Kaffee bezüglich der Art mit einem zielgerichteten Ansatz ist die Quantifizierung von 16-O-Methylcafestol mittels Protonen-Kernspinresonanz Spektroskopie ($^1\text{H-NMR}$). Diese Verbindung dient zur Überprüfung des nicht deklarierten Zusatzes von Kaffee der Art *Coffea canephora* Pierre ex A. zu hochwertigerem *Coffea arabica* L. Bei einem nicht zielgerichteten Ansatz liegt der Fokus nicht auf einem bestimmten Inhaltsstoff, sondern es werden so viele verschiedene Metaboliten wie möglich erfasst, ohne diese zu identifizieren oder zu quantifizieren. Mit der Erfassung möglichst vieler Metaboliten, die eine Molekülgröße <1000 kDa aufweisen, wird dieser Ansatz auch als Metabolomics bezeichnet. Eine geeignete Analyseverfahren zur Erfassung vieler verschiedener Metaboliten stellt die $^1\text{H-NMR}$ dar. Die Untersuchung authentischer Lebensmittelproben mittels $^1\text{H-NMR}$ in Kombination mit multivariater Datenanalyse ermöglicht die Erstellung von Klassifizierungsmodellen. Mit diesen kann die Authentizität neuer unbekannter Lebensmittelproben überprüft werden. Bisher wurde die $^1\text{H-NMR}$ in Kombination mit einer multivariaten Datenanalyse erfolgreich zur Überprüfung der Authentizität von biologisch erzeugten Eiern, der Tierart oder auch des geographischen Ursprungs von Wein angewendet. Bis auf die Untersuchung der Authentizität der Kaffeeart in einem zielgerichteten Ansatz fehlen in der Lebensmittelüberwachung zuverlässig etablierte Analysemethoden, um die Überprüfung der Authentizität von Kaffee zu ermöglichen.

Diese Arbeit beschäftigt sich mit der Analyse von gerösteten Kaffeeproben der Art *Coffea arabica* L. mittels $^1\text{H-NMR}$. In Kombination mit einer multivariaten Datenanalyse soll die Möglichkeit der Unterscheidung von Kaffee bezüglich des geographischen Ursprungs, der Anbaumethode und des angewendeten Röstverfahrens überprüft werden.

1.1 Warenkunde Kaffee

1.1.1 Definition, Geschichte und Welthandel

Nach der Anlage zu §1 und §2 der Verordnung über Kaffee, Kaffee- und Zichorien-Extrakte vom 15.11.2001 werden die Samen von Pflanzen der Gattung *Coffea*, welche von Frucht- und Samenschale befreit und ungeröstet sind, als Rohkaffee bezeichnet. Röstkaffee oder Kaffee ist demnach gerösteter Rohkaffee, ungemahlen oder gemahlen, mit einem Wassergehalt von höchstens 50 Gramm pro Kilogramm [1]. Aus dem gemahlene Röstkaffee kann durch Extraktion mit heißem Wasser das Kaffeegetränk hergestellt werden, welches ebenfalls als Kaffee bezeichnet wird [2].

Als Urheimat der Kaffeeart *Coffea arabica* L. (Arabica) wird das abessinische Hochland von Äthiopien und die der Kaffeeart *Coffea canephora* Pierre ex A. (Robusta) der große Kongobogen in Afrika betrachtet [3–5]. Die erste schriftliche Erwähnung von Arabica als Heilmittel gab es anfangs des 11. Jahrhunderts von Avicenna, einem Heilkundigen und Philosophen aus Arabien. Aus einem Manuskript aus dem Jahr 1578 geht hervor, dass Kaffee von Äthiopien nach Arabien gebracht und seit Mitte des 14. Jahrhunderts als Getränk dort konsumiert wurde [4, 6]. In der zweiten Hälfte des 15. Jahrhundert breitete sich die Kaffeebohne im gesamten arabischen Großreich aus. Von da aus gelangte die Kaffeebohne nach Kleinasien, Syrien, Ägypten und das südöstliche Europa. Im Jahr 1592 erschien in Venedig das Schriftstück „De Plantis Aegypti Liber“, welches eine detaillierte Zeichnung und Beschreibung der Kaffeeepflanze und –frucht enthält. In der gleichen Zeit wurden die heilkundlichen Aspekte der Kaffeebohne von Spezialisten wie dem deutschen Botaniker Leonhard Rauwolf beschrieben [6, 7].

Anfang des 17. Jahrhundert kamen die ersten Kaffeebohnen über holländische Handelsreisende nach Europa und die Kaffeeepflanze wurde von den Kolonialmächten in alle Gebiete gebracht, die sich für den Kaffeeanbau eigneten [5, 6]. Das Produkt Kaffee wurde wirtschaftlich bedeutend und mit der beginnenden Industrialisierung zum Heißgetränk für alle gesellschaftlichen Schichten [6]. Auch fanden zu dieser Zeit erste kontrollierte Studien über die Chemie und Biologie von Kaffee statt [7]. Im Jahr 1819 untersuchten Wolfgang Goethe und Friedrich Runge Kaffee mittels Papierchromatographie und entdeckten die Substanz Koffein [8]. Im 20. Jahrhundert wurden weitere Kaffeeinhaltsstoffe wie Chlorogensäure (Chlorogenic Acid, CGA), Kaffeesäure und Trigonellin entdeckt und im 21. Jahrhundert weiter auf ihre antioxidativen Eigenschaften untersucht [7]. Obwohl Robusta bereits 1879 durch den französischen Botaniker Pierre beschrieben wurde, wurde diese Kaffeeart erst im 20. Jahrhundert angebaut und verwendet [6].

Heutzutage ist der Gesamtkonsum von Kaffee in Europa am höchsten, gefolgt von Asien und Nordamerika [9]. In den letzten 20 Jahren stieg die Kaffeeproduktion um knapp 80% an, und Kaffeebohnen werden heutzutage in einer Vielzahl an Ländern weltweit angebaut [10], wobei der Anbau von Kaffee zur Wirtschaft von über 50 Ländern beiträgt [11]. Allerdings sind die zehn größten Anbauländer verantwortlich für ungefähr 80% der Weltproduktion [12]. Im Kalenderjahr 2019/20 waren die Hauptanbauländer Brasilien, Vietnam, Kolumbien und Indonesien [10]. Dabei belief sich die Kaffeeproduktion auf 57% Arabica und 43% Robusta [10]. Die dritte wirtschaftlich relevante Art, *Coffea liberica* W. Bull ex Hiern (Liberica), trägt weniger als 1% zur Weltproduktion bei [7]. Arabica wird vorwiegend in Lateinamerika, Äthiopien und Kenia angebaut wohingegen Robusta hauptsächlich aus Vietnam, Brasilien und Uganda stammt [13]. Der Rohkaffeepreis ist abhängig von der

Qualität und des Herkunftslandes, jedoch für Arabica höher im Vergleich zu Robusta [14]. Dieser Preisunterschied ist darin begründet, dass Arabica allgemein als qualitativ höherwertig angesehen wird [13]. Der Rohkaffee wird dabei in vier Qualitätsgruppen unterteilt. Die Klasse „Colombian Milds“ hat die höchste Qualität und die nass aufbereiteten Arabica-Bohnen stammen aus Kolumbien, Kenia und Tansania. Andere nass aufbereitete Arabica-Bohnen mit geringer Säure („Other Milds“) stammen aus vielen Ländern Zentralamerikas. „Brazilian and Other Arabicas“ sind trocken aufbereitete Bohnen mit dem Ursprung aus Brasilien, Äthiopien und Yemen. Die Klasse mit der niedrigsten Qualitätsstufe beinhaltet Robusta-Bohnen [15].

1.1.2 Die Kaffeepflanze

Kaffee gehört zu der Gattung *Coffea*, welche wiederum zu der Familie der *Rubiaceae* zählt und von der aktuell 129 Arten bekannt sind [16]. Diese werden in die Rubriken *Eucoffea*, *Mascarocoffea*, *Agrocoffea* und *Paracoffea* unterteilt. Die Rubrik *Eucoffea* ist die einzige mit wirtschaftlicher Bedeutung, da zu ihr die Arten Arabica und Robusta zählen [3, 7]. Es wird vermutet, dass Arabica, welcher tetraploid ist ($2n = 44$), aus einer natürlichen Hybridisierung von *Coffea canephora* und *Coffea eugenoides* abstammt [5, 7, 16]. Bis auf Arabica sind alle anderen Arten fast ausschließlich diploid ($2n = 22$) [3, 7]. Von den Arten Arabica und Robusta sind eine Vielzahl botanischer Varietäten bekannt. Zudem gibt es natürliche Hybride zwischen Arabica und Robusta, wie beispielsweise Catimor, Sarchimor, Icatu oder S795 [3].

Die modernen Varietäten von Arabica stammen von den zwei ursprünglichen Varietäten Typica und Bourbon ab, wobei Bourbon eine natürliche Mutante von Typica ist [3, 16, 17]. Varietäten wie Maragogype oder Pulma Hidalgo stammen vom Typica ab, wohingegen Caturra, SL-34 oder Bourbon Pointu von Bourbon abstammen [15, 16]. Aus Kreuzungen zwischen Typica und Bourbon entstand Mundo Novo und aus weiteren Kreuzungen schließlich Catuai und Pacamara [16–18]. Die Varietät S795 gilt als natürliche Kreuzung zwischen Arabica und Liberica [7].

Arabica Kaffeepflanzen können bis zu 6 m hoch werden, während Robusta Kaffeepflanzen eine Höhe von bis zu 10 m erreichen. Bei der diploiden Art Liberica kann der Kaffeebaum auf eine Höhe von bis zu 18 m wachsen. Die Hauptwurzel der Arabica-Pflanze dringt bis zu drei Meter tief in den Boden ein, wohingegen die der Robusta-Pflanze sehr nahe an der Oberfläche liegen [3, 4, 7, 15]. Die Form der Kaffeepflanze ist abhängig von der Varietät. Im Allgemeinen besteht der Kaffeebaum aus einem aufrechten Haupttrieb (Abbildung 1-1(A)) mit primären, sekundären und tertiären Seitenästen. Die Kaffeeflüte besteht aus einer weißen Blumenkrone mit einem Kelch, fünf Staubgefäßen und den Fruchtknoten [15]. Die Blütezeit ist je nach geographischem Gebiet unterschiedlich. Während der Trockenzeit ruhen die Blütenknospen und die Blüte beginnt erst, wenn diese Trockenperiode unterbrochen ist [19]. Beispielsweise blühen die Pflanzen im Südosten Brasiliens, wo ausgeprägte Trocken- und Regenzeiten herrschen, zwei- bis dreimal im Jahr. Dahingegen können sie im äquatorialen, regenreichen Costa Rica bis zu fünfmal im Jahr blühen [3].

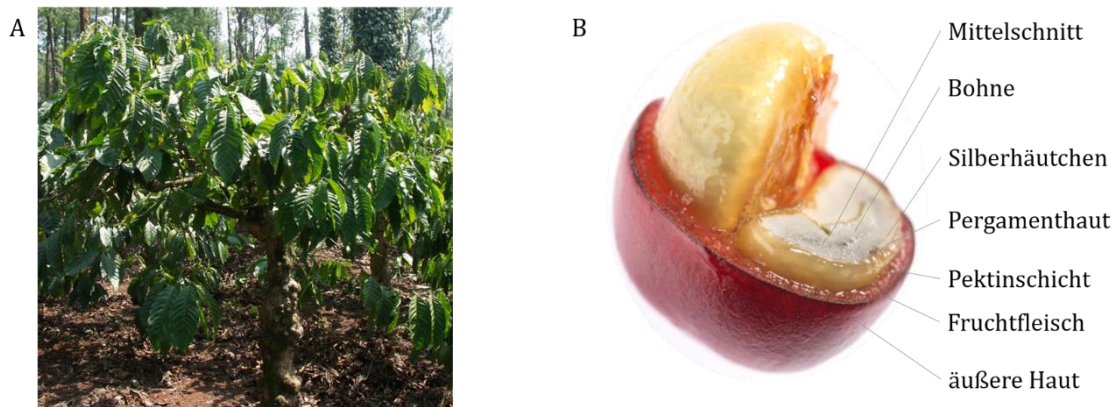


Abbildung 1-1: (A) Foto einer Kaffeepflanze (Bereitstellung des Fotos durch Coffee Consulate, von Luca Siermann); (B) Querschnitt einer reifen Kaffeekirsche mit Beschriftung der einzelnen Bestandteile (modifiziert nach [20])

Die Selbstbestäubung bei Arabica oder Fremdbestäubung bei Robusta findet innerhalb von sechs Stunden nach der Blüte statt [7, 19]. Nach der Bestäubung entwickelt sich eine 10-15 mm lange Frucht, die zwei Samen enthält [7]. Die Frucht wird Kaffeekirsche genannt und besteht aus der Schale, dem Fruchtfleisch, der Pergamenthaut, dem Silberhäutchen und zwei eiförmigen Samen (Abbildung 1-1(B)). Die Schale besteht aus einer einzelligen Schicht, welche mit einer wachsartigen Substanz überzogen ist. Die Pergamenthaut ist eine dünne, papierartige Polysaccharidhülle, wohingegen die Samenschale, auch Silberhäutchen genannt, hauptsächlich aus Cellulose und Hemicellulosen besteht [7]. Die Samen enthalten des Endosperm und nehmen den größten Teil der Frucht ein [7, 16]. Bei Arabica sind die Kaffeekirschen rot oder gelb, dagegen sind bei Robusta mehrere Farbtöne möglich. Die Form der Kaffeekirsche variiert je nach *Coffea* Art. Die Größe der Frucht wird von der gepflanzten Sorte und den Anbaubedingungen beeinflusst [3].

1.1.3 Anbau von Kaffee

Die Bedingungen, unter welchen Kaffee angebaut wird, haben einen großen Einfluss auf die Qualität des Kaffees und die Resistenz gegenüber Krankheiten oder Schädlingen. Kaffee wächst am besten in porösen Böden mit einem pH-Wert von sieben. Arabica wächst bei Temperaturen von 15-25 °C in einer Höhe von 1000-2800 m und bei jährlichen Niederschlagsmengen von 1200-1800 mm. Dahingegen wächst Robusta bei Höhen unter 1000 m und bei jährlichen Niederschlagsmengen von 2000-2500 mm [3, 15].

Der Anbau von Kaffee beginnt mit den Setzlingen, welche durch Samen oder vegetatives Klonen aus orthotropen Zweigen erzeugt werden [3]. Die meisten Kaffeefarmen verfügen über eine Aufzuchtstation, in welcher die Setzlinge aufgezogen werden, bevor sie nach sechs bis zwölf Monaten auf der Farm ausgepflanzt werden [16]. Anschließend braucht die Kaffeepflanze etwa drei Jahre, um reife Früchte zu tragen. Eine Kaffeepflanze kann 10-50 Jahre alt werden, jedoch geht die Fruchteproduktion deutlich früher zurück [3, 21].

Während Länder wie Brasilien und Vietnam intensiven Kaffeeanbau mit großen Plantagen betreiben, wird in vielen anderen Ländern wie Äthiopien, Kolumbien oder Mexiko der Kaffee in kleinen Familienplantagen angebaut [19, 22-25]. Die Pflanzendichte auf einer Farm ist abhängig von der gepflanzten Sorte, der Bodenbeschaffenheit und dem Klima. Länder wie Brasilien, Kolumbien und Mexiko tendieren zu einer höheren Pflanzendichte [3].

Die Anbaustile von Kaffee können in Sonnenanbau, Schattenanbau, Terrassenanbau und Gartenanbau eingeteilt werden. Traditionell wurde Kaffee im Schatten anderer Bäume angebaut, wobei der Schatten ein für Kaffee natürliches Umfeld darstellt (Abbildung 1-2(A)) [4, 15, 23]. Vorteile des Schattenanbaus liegen darin, dass Klima- und Standortbedingungen verbessert und die Sonneneinstrahlung verringert werden, was ein übermäßiges Wachstum verhindert [5, 15, 26]. Dadurch sind die Kaffeepflanzen weniger Stressfaktoren ausgesetzt, wodurch wiederum eine bessere Qualität des Kaffees, insbesondere in der chemischen Zusammensetzung, den organoleptischen Eigenschaften und der Bohnengröße entstehen kann [15, 27]. Ein weiterer Vorteil ist die Erhaltung der Biodiversität [23, 26]. Aufgrund der Erhaltung der Biodiversität besteht die Möglichkeit, diesen Kaffee als „Bird-Friendly“ auszuloben und höhere Preise zu verlangen [23, 28]. Jedoch bringt der Schattenanbau auch Nachteile wie einen geringeren Ertrag mit sich [15]. Zudem stehen die Schattenbäume mit den Kaffeepflanzen bezüglich des Nährstoff- und Wasservorkommens im Boden in Konkurrenz [15]. Bei modernen Plantagen überwiegen meist die Nachteile, sodass bei diesen Plantagen der Sonnenanbau betrieben wird (Abbildung 1-2(B)) [15]. Das ist beispielsweise auf vielen Plantagen in Brasilien und Costa Rica der Fall, wo der Fokus auf einem größtmöglichen Ertrag liegt [3, 15, 16]. Diese Anbauweise kann mit einem Qualitätsverlust des Kaffees verbunden sein [16]. Ein weiterer Nachteil dieses homogenen, weniger biodiversen Systems ist der vermehrte Einsatz von Düngemitteln und Pestiziden aufgrund des höheren Nährstoffbedarfs und der Eindämmung von Krankheits- und Schädlingserregern [15]. Infolge des gestiegenen Verbraucherbewusstseins und des Klimawandels geht der weltweite Trend jedoch wieder zu einem nachhaltigeren Anbaustil zurück [15].

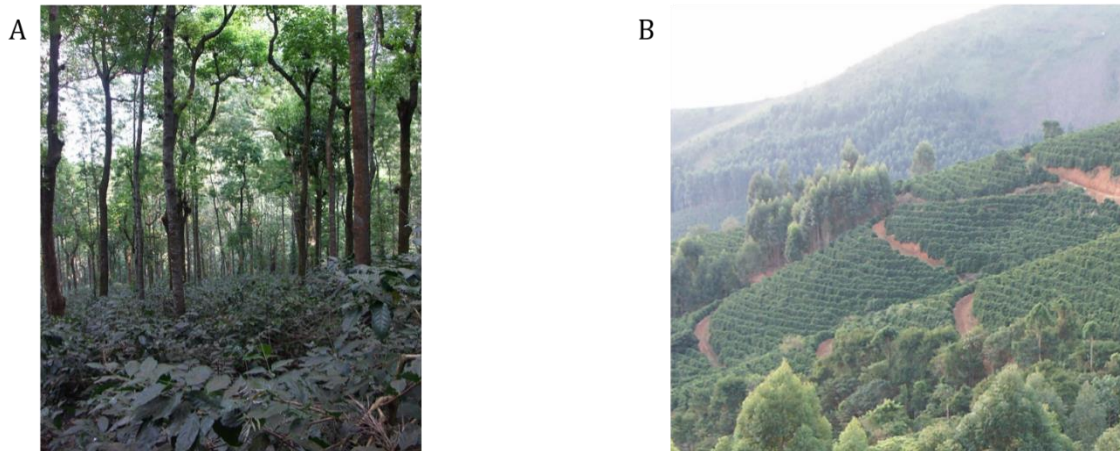


Abbildung 1-2: Veranschaulichung verschiedener Anbaustile. (A) Schattenanbau; (B) Sonnenanbau (Bereitstellung der Fotos durch Coffee Consulate, von Luca Siermann)

In Brasilien werden hauptsächlich die Arabica Varietäten Mundo Novo, Bourbon und Catuai angebaut [24, 29]. In Kolumbien dagegen werden Typica, Caturra und Castillo angepflanzt [16]. Die größte Artenvielfalt findet sich in Äthiopien wieder. Hier wachsen viele ursprüngliche Arabica Varietäten [4, 16]. In Asien wird hauptsächlich Robusta angebaut. In einzelnen Gebieten wachsen jedoch auch Arabica Varietäten wie S795, Bourbon und Typica [16]. Viele Gebiete Afrikas bauen Bourbon an, wobei dieser bevorzugt in Form von Schattenanbau und in kleinen Gärten, weniger in größeren Plantagen angepflanzt wird [16, 25].

Klimatische Veränderungen haben einen großen Einfluss auf den Anbau von Kaffee. Steigende Durchschnittstemperaturen und sinkende Niederschläge in tiefer gelegenen Gebieten verringern dort die Anbaumöglichkeiten von Arabica [30]. Bis 2050 könnte die

optimale Höhe für den Anbau von Arabica in Zentralamerika von 1200 m auf 1600 m ansteigen [30]. Derzeit geeignete Gebiete für den Arabica Anbau sind das brasilianische Minas Gerais, Mittelamerika und das äthiopische Hochland [21]. In Zukunft werden durch den Klimawandel die aktuellen brasilianischen Anbauggebiete weniger geeignet sein [21]. Der brasilianische Kaffeesektor könnte sich an die steigenden Temperaturen anpassen, indem Anbauggebiete in südliche Bundesstaaten verlegt werden [21, 31]. In weiteren Teilen Lateinamerikas und in Afrika werden sich höhere Lagen in Zukunft besser eignen [21]. Robusta ist resistenter gegenüber höheren Temperaturen, sodass es möglich sein könnte, dass der Anbau von Robusta-Kaffee die Produktion von Arabica-Kaffee teilweise ersetzt [21]. Eine weitere Alternative könnte die Entwicklung neuer Arabica Varietäten sein, welche sich besser an die durch den Klimawandel hervorgerufene abiotische Belastungen anpassen können [5].

1.1.3.1 Biologisch zertifizierter Anbau

Viele der synthetisch eingesetzten Agrarchemikalien sind weit verbreitet und werden in der modernen Landwirtschaft routinemäßig eingesetzt [15]. Allerdings ist in den letzten Jahren die Nachfrage für Lebensmittel, welche nicht mit synthetischen Chemikalien behandelt wurden, gestiegen [13, 15]. Auch das wachsende Bewusstsein für die schädlichen Auswirkungen dieser Chemikalien auf die Umwelt haben zu neuen Trends in der Landwirtschaft geführt. Es resultierten neue Vorgehensweisen mit dem Ziel, synthetische Chemikalien nicht mehr anzuwenden und die umweltbedingten Beeinträchtigungen zu minimieren [15]. Die Beweggründe der Kaffeebauern hin zu einer organischen Landwirtschaft sind, dass eine Produktion ohne synthetische Agrarchemikalien sicherer für ihre Familien und die Umwelt ist, die Ausgaben für synthetische Stoffe mindert und zum Schutz des Wassers beiträgt [13]. Die organische Produktion ist ein wachsender Sektor und Verbraucher sind gewillt, einen höheren Preis für diese Produkte zu zahlen [13, 15, 28]. Laut Artikel 5 und 11 der VO(EU) 2018/848 ist der Einsatz von synthetischen Chemikalien und Praktiken wie die Bestrahlung zu Schädlingsbekämpfung und der Einsatz von Gentechnik in der organischen Landwirtschaft verboten [32]. Der Ertrag biologisch erzeugter Ernten kann so bis zu 50% geringer sein im Vergleich zum konventionellen Anbau [15]. Damit eine biologische Farm wirtschaftlich rentabel ist, müssen demnach für biologisch erzeugte Produkte höhere Preise verlangt werden [15, 28].

Die ökologische Kaffeeproduktion umfasst länderübergreifende Zertifizierungsverfahren, die sicherstellen sollen, dass der Kaffee für den Verbraucher tatsächlich von zertifizierten Anbauflächen stammt [28]. Jedoch erzeugen diese Zertifizierungsverfahren und die von der bio-Zertifizierung geforderten Überwachungssysteme erhebliche Kosten für Erzeuger [28]. Wie bereits beschrieben, betreiben einige Farmen nachhaltigen Schattenanbau. In diesem Anbaustil ist die biologische Produktion oft mit inbegriffen, da die Vielfalt von Organismen in biodiversen Ökosystemen den Befall von Schädlingen und Krankheiten bei Kaffeepflanzen vermindern und den Einsatz von synthetischen Chemikalien reduzieren oder überflüssig machen [15]. Dies ist jedoch nicht zwingend mit einer bio-Zertifizierung verbunden [27, 28]. Auch haben viele Landwirte beispielsweise in Afrika keinen Zugang zu den Mitteln und der Infrastruktur, um teure Pestizide und Fungizide zu verwenden, und gestalten ihren Anbau daher natürlich [19]. Ein weiterer Nachteil des Zertifizierungsverfahrens ist, dass einem bio-zertifizierten Erzeuger das Zertifikat entzogen wird, wenn ein Kaffeebauer in einer benachbarten Parzelle synthetische Chemikalien einsetzt [28].

1.1.4 Ernte von Kaffee

In vielen Ländern ist die Ernte ein länger andauernder Vorgang, da die Kaffeepflanzen mehrmals im Jahr blühen und sich die Kaffeekirschen an einer Kaffeepflanze in unterschiedlichen Reifestadien befinden können (Abbildung 1-3) [3, 7]. Beginn der Ernte ist oft der Zeitpunkt, bei dem der Großteil der Kaffeekirschen reif ist [7]. Die Haupternte findet in Brasilien von Mitte Juni bis Ende September statt. In Kolumbien gibt es zwei Ernteperioden, wobei die Haupternte von Oktober bis Dezember andauert und die Nebenernte von April bis Juni. In Äthiopien wird hauptsächlich im Zeitraum November bis Februar geerntet [15].



Abbildung 1-3: Foto einer Kaffeepflanze mit Kaffeekirschen verschiedener Reifestadien (Bereitstellung des Fotos durch Coffee Consulate von Dr. Steffen Schwarz)

Die Ernte der Kaffeekirschen ist unter anderem von der Verarbeitungsmethode abhängig. Bei der nassen oder halbtrockenen Aufarbeitung erfolgt die Ernte mit dem Ziel, so viel reife Kaffeekirschen wie möglich zu erhalten. Bei der trockenen Aufarbeitung hingegen werden alle Kaffeekirschen gleichzeitig geerntet mit dem Kompromiss, dass sich darunter auch unreife Früchte befinden [15].

Die Ernte kann manuell oder maschinell erfolgen. Beispielsweise in Minas Gerais in Brasilien wird zunehmend auf mittelgroßen bis großen Farmen eine maschinelle Ernte eingesetzt, um Betriebskosten zu senken [3, 5, 15]. Kaffeebauern in diesen Gebieten fahren mit großen Maschinen durch die Reihen der Kaffeepflanzen. Die Bäume werden dabei so lange geschüttelt, bis sich die Früchte lösen, oder die Äste werden mit einem Gerät abgestreift, das einem flexiblen Kamm ähnelt [7, 15, 16]. Durch diese Art der Ernte werden alle Früchte mit unterschiedlichem Reifestadium gleichzeitig gepflückt. Nach der maschinellen Ernte müssen die Früchte sortiert werden, um die reifen von den überreifen und unreifen Kaffeekirschen zu trennen. Bei diesem Vorgang werden auch zusätzlich abgefallene Blätter oder Äste heraussortiert [16].

Ein großer Teil des Kaffees wird immer noch von Hand geerntet, da Maschinen in bergigen Anbaugebieten nur schwer eingesetzt werden können. Eine zeitsparende Methode der Handlese wird als „Stripping“ bezeichnet und besteht darin, mit einer Bewegung alle Kirschen gemeinsam von einem Zweig zu lösen. Allerdings führt auch diese Erntemethode zu einer Mischung aus überreifen, reifen und unreifen Kaffeekirschen, welche anschließend sortiert werden müssen [16]. Diese Art der Ernte wird bevorzugt von Kaffeebauern in Brasilien und Äthiopien durchgeführt [7, 15]. In Brasilien ist zunehmend das mechanische Stripping verbreitet, bei welchem die Ernte der Kaffeekirschen mit einer in der Hand gehaltenen Erntemaschine erfolgt [5]. Der qualitativ hochwertigste Kaffee entsteht durch die Handlese. Hier ernten Pflücker jede Kaffeekirsche einzeln bei optimaler Reife [15, 16]. Ein wichtiger Aspekt beim Pflücken der Kaffeekirschen ist es, ungewünschte Kontamination, wie Staub oder Schmutz zu vermeiden [7]. Daher werden die Kaffeekirschen bevorzugt in Säcken geerntet, anstatt diese auf den Boden fallen zu lassen [15].

1.1.5 Verarbeitung der Rohkaffeebohnen

Nach der Ernte werden die Kaffeekirschen verarbeitet, wobei die Rohkaffeebohnen von der restlichen Frucht getrennt werden [7]. Nach der Verarbeitung werden die Rohkaffeebohnen nach Größe und Anzahl von Defekten sortiert [7, 15].

Die Verarbeitung der Rohkaffeebohnen hat einen großen Einfluss auf die spätere Qualität des Kaffees [24]. Es gibt verschiedene Methoden der Verarbeitung, alle beinhalten jedoch die Schritte Säubern, Trennen, Lagern, Verarbeiten und Klassifizieren [24]. Die Rohkaffeeverarbeitung kann mit Hilfe der trockenen, nassen oder der halbtrockenen Methode erfolgen (Abbildung 1-4). Für Arabica wird die nasse Methode bevorzugt angewandt [33]. Bei der nassen Methode werden nur reife Kaffeekirschen verarbeitet, wohingegen die trockene Methode für unreife, reife und überreife Früchte verwendet werden kann [34].

Die trockene Methode ist die älteste Methode der Rohkaffeeverarbeitung und findet oft Verwendung, wenn die Kaffeekirschen mit der maschinellen Ernte oder der Stripping Methode geerntet wurden [7, 15]. Nach der Ernte werden die Kaffeekirschen mit der Hand grob nach ihrem Reifegrad sortiert und auf dem Boden oder auf speziellen Trocknungstischen ausgebreitet, um in der Sonne zu trocknen. Die Kaffeekirschen müssen während der Trocknung regelmäßig gewendet werden, um Schimmelbildung zu vermeiden. Nachdem der Wassergehalt auf 10-12% gesunken ist, wird die Schale und das Fruchtfleisch der Kaffeekirsche mechanisch entfernt [7, 15, 16]. Zurück bleiben die grünen Kaffeebohnen mit dem Silberhäutchen, welche anschließend sortiert werden [7, 15, 16]. Vorteil der trockenen Methode ist der geringe Bedarf an Wasser. Obwohl diese Methode standardgemäß für Robusta eingesetzt wird, wird sie auch in Äthiopien und in vielen brasilianischen Plantagen für die Verarbeitung von Arabica angewendet [7, 15, 16, 33]. Durch die trockene Aufarbeitungsmethode erhalten Kaffees ein fruchtiges Aroma [16].

Bei der nassen Aufbereitung werden die reifen Kaffeekirschen zunächst im Floating-Becken sortiert. Die reifen Kaffeekirschen sinken auf den Boden und die unreifen Früchte schwimmen an der Wasseroberfläche. Anschließend wird das Fruchtfleisch der reifen Kirschen im sogenannten Entpulper entfernt. Nach diesem Schritt behalten die Bohnen jedoch noch die Pergamenthaut mit Resten des Fruchtfleisches. Nach dem Entpulper gelangen die Kaffeebohnen in einen Tank mit frischem Wasser, in welchem sich das restliche Fruchtfleisch von der Bohne durch Fermentation ablöst [7, 15, 16]. Die Fermentation findet natürlich oder durch Zusatz von Mikroorganismen und Enzymen statt und dauert 12-36 h [7, 15]. Sie wird von Faktoren wie der Umgebungstemperatur beeinflusst [16]. Mit steigender Temperatur steigt auch die Geschwindigkeit des Fermentationsprozesses. Nach der Fermentation werden die Kaffeebohnen gewaschen und getrocknet. Dazu werden die Bohnen auf dem Boden oder auf Trocknungstischen ausgebreitet und in der Sonne bis zu einem Wassergehalt von 11-12% getrocknet. Die trockenen Bohnen werden mit Hilfe einer Mühle von ihrer Pergamenthaut und den Silberhäutchen befreit und sortiert. Durch die nasse Aufbereitung erhalten die Bohnen eine höhere Qualität, ein großer Nachteil ist jedoch der hohe Wasserbedarf [7, 16]. Die nasse Aufarbeitung findet auf Kaffeepiantagen wie zum Beispiel in Kolumbien, Mexiko, El Salvador, Guatemala, Kenia und Tansania Anwendung, bei welchen die Kaffeekirschen handverlesen geerntet werden [2, 7, 15].

Die halbtrockene Methode ist üblich in Indonesien und Brasilien und wird auch als Pulped Natural bezeichnet. Nach der Ernte werden die Kaffeekirschen im Floating-Becken sortiert, die Kaffeebohnen im Entpulper vom Fruchtfleisch getrennt und kurz getrocknet,

wobei der Fermentierungsprozess entfällt [7, 15, 16]. Anstatt die Bohnen bis zu einem Wassergehalt von 10-12% zu trocknen, ist die Trocknung bei dieser Methode bereits bei einem Wassergehalt von 30-35% beendet. Anschließend wird die noch an den Bohnen haftende Pergamenthaut entfernt und die Rohkaffeebohnen nochmals getrocknet [7, 15].

Ein spezieller, hochpreisiger Kaffee aus Indonesien ist der Luwak Kaffee, bei dem die reifen Kaffeekirschen von der Zibetkatze gefressen werden. Der Verdauungstrakt der Katze entfernt das Fruchtfleisch und die Fermentation findet durch die im Tier vorhandenen Mikroorganismen statt. Die mit dem Kot ausgeschiedenen Rohkaffeebohnen werden eingeweicht, gewaschen und anschließend in der Sonne getrocknet [5].

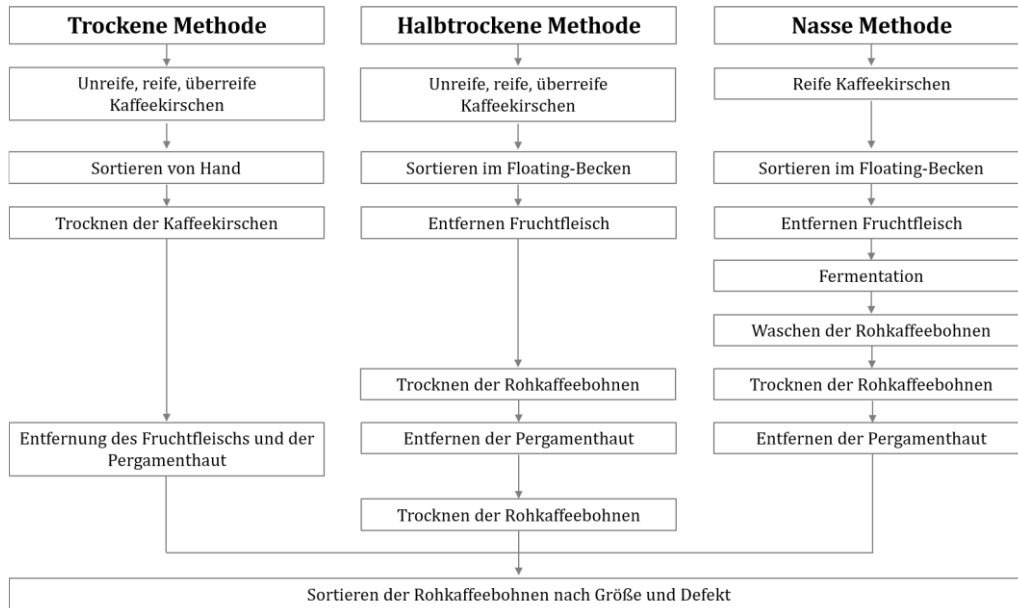


Abbildung 1-4: Flussdiagramm mit den verschiedenen Schritten der trockenen, halbtrockenen und nassen Methode der Rohkaffeeverarbeitung (modifiziert nach [7, 15, 16, 24, 33])

Durch die verschiedenen Methoden der Rohkaffeeverarbeitung wird die Bildung unterschiedlicher Metaboliten angeregt. So ist der Anteil der freien Aminosäuren wie Glutaminsäure in nass aufgearbeiteten Rohkaffeebohnen höher als in trocken aufgearbeiteten Bohnen [34]. Trocken aufbereitete Bohnen enthalten dagegen einen höheren Anteil der nicht-proteinogenen Aminosäure γ -Aminobuttersäure (GABA) [34]. GABA wird als Stressmetabolit durch enzymatische Decarboxylierung von Glutaminsäure während der Trocknung gebildet [7, 34]. Auch ist der Gehalt von Trigonellin in trocken aufbereiteten Kaffeebohnen höher. Durch die Fermentation während der nassen Aufbereitung sinkt der Gehalt an Glucose und Fructose, wohingegen Saccharose nicht beeinflusst wird und der Gehalt der Polysaccharide ansteigt [35, 36]. Die Menge von Glucose und Fructose bei der halbtrockenen Aufbereitung liegt zwischen der der trockenen und nassen Aufbereitung [35].

Die Zusammensetzung und Konzentration der CGA wird maßgeblich vom Reifestadium der Kaffeekirsche beeinflusst, wodurch sich auch ein Unterschied dieser in den verschiedenen Aufarbeitungsmethoden widerspiegelt [7, 35]. Während der nassen Verarbeitung steigt jedoch das Verhältnis zwischen Caffeoylechinolchinasäure-Isomere (Caffeoylechinolchinasäure, CQA) und Dicafeoylchinasäuren (diCQA) [36]. Neben dem Gehalt an Säuren ist auch der Fettgehalt bei nass aufbereiteten Kaffeebohnen höher, wobei die Fettsäurezusammensetzung nicht beeinflusst wird [36, 37]. Durch die nasse Aufarbeitung kann sich die Konzentration von CGA

aufgrund der Lagerung der Kaffeebohnen in Wasser verringern [38]. Weiter erfolgt während des natürlichen Fermentationsprozesses der nassen Aufarbeitung die Bildung von organischen Carbonsäuren wie beispielsweise Essigsäure und Milchsäure [39]. Die Proteine werden durch die trockene oder nasse Aufarbeitung nicht beeinflusst, jedoch hat der Verdauungsprozess der Zibetkatze einen Einfluss auf die Proteinzusammensetzung beim Luwak-Kaffee [35].

1.1.6 Der Röstprozess

Durch den komplexen Röstprozess bei Temperaturen von 200-250 °C löst sich das an den Bohnen anhaftende Silberhäutchen ab, und der Rohkaffee entwickelt sich zu einem aromareichen Röstkaffee, aus welchem das Kaffeegetränk nach dem Mahlen der Kaffeebohnen hergestellt wird. In Röstkaffee wurden über 300 verschiedene Verbindungen identifiziert, darunter auch Pyrrole, Oxazole, Furane, Imidazole und Pyrazine [37]. Der Röstprozess lässt sich in verschiedene Abschnitte wie das Trocknen, die Färbung, den ersten Crack, die Entwicklung, den zweiten Crack und die Abkühlung einteilen [5, 16]. Während der Trocknung nehmen die Kaffeebohnen die Hitze des Rösters auf und Wasser verdunstet. Anschließend beginnen in der Färbung die ersten Bräunungsreaktionen, in welchen Kohlenstoffdioxid freigesetzt wird. Durch die Entwicklung des Gases und das Verdampfen von Wasser steigt der innere Druck in der Bohne. Dieser Druck entlädt sich mit dem ersten Crack und das Volumen der Kaffeebohne nimmt zu. Ab diesem Punkt entwickeln sich die für Röstkaffee typischen Aromastoffe. Auch in dieser Phase werden wieder Gase freigesetzt, die den Innendruck der Kaffeebohnen weiter ansteigen lassen. Mit dem zweiten Crack werden diese freigesetzt. An diesem Punkt liegt eine dunkle Kaffeeröstung vor und Kaffeeöl tritt aus der Oberfläche der Bohnen aus [2, 5, 16].

Während der Röstung ändert sich die Farbe der Kaffeebohnen von einem grünlichen Grau über gelb, orange, braun, dunkelbraun bis hin zu fast schwarz und zeigt damit den Röstgrad an [5]. Die Volumenzunahme aufgrund der Gasentwicklung beträgt um bis zu den Faktor zwei, die Trocknung und die chemischen Reaktionen während des Röstprozesses führen zu einer Veränderung in der Struktur des Bohnengewebes [5]. Die sehr kompakte und dichte Struktur der Rohkaffeebohnen wird durch das Rösten zerstört und führt zur Bildung von Zellen mit gasgefülltem Hohlraum, wodurch die Porosität der Bohnen zunimmt [5]. Der Wassergehalt der Kaffeebohnen nimmt von ursprünglich 10-12% auf ungefähr 2,5% ab. Auch das Gewicht der Kaffeebohne sinkt während des Röstens um 12-20% [5].

Neben der Temperatur ist der Wassergehalt und die Wasseraktivität bei den verschiedenen Röstabschnitten wichtig für die physikalischen und chemischen Veränderungen in der Kaffeebohne [5]. Die wichtigsten chemischen Reaktionen, die während des Röstens von Kaffee ablaufen, sind die Maillard-Reaktion, Pyrolyse und Karamellisierung [5]. Zucker und Proteine reagieren im Rahmen der Maillard-Reaktion zu gewünschten Komponenten, wie zu braun gefärbten Pigmenten und aromaaktiven Verbindungen oder Bitterstoffen, aber auch zu Verbindungen mit potentiell mutagenen Eigenschaften [2, 37]. Neben der Maillard-Reaktion führt auch die Karamellisierung zu braun gefärbten Produkten und verschiedenen heterocyclischen Verbindungen wie 5-Hydroxymethylfurfural (HMF) [2, 37].

1.1.6.1 Maillard-Reaktion

Die Maillard-Reaktion, benannt nach dem französischen Wissenschaftler Luis Camille Maillard, fasst zahlreiche Reaktionen zwischen reduzierenden Zuckern und Aminosäuren,

Peptiden und Proteinen, die eine freie Aminogruppe besitzen, zusammen [7, 40]. Im Rahmen dieser nicht enzymatischen Bräunungsreaktion entstehen zahlreiche Aromastoffe sowie braun gefärbte Melanoidine (Abbildung 1-5) [41–43]. Die Entstehung dieser Produkte ist stark abhängig von den Reaktionsbedingungen wie pH-Wert der Matrix, der Temperatur und dem Wassergehalt [42].

In der Anfangsphase der Maillard-Reaktion führt die nucleophile Addition von einer Aminokomponente an die Carbonylfunktion reduzierender Zucker zur Bildung von N-substituierten Glycosylaminen und weiter zu Iminen (Schiff'sche Basen) [2]. Über ein 1,2-Enaminol entstehen 1-Amino-1-desoxy-2-ketosen (Amadori-Verbindungen) oder 2-Amino-2-desoxy-1-alosen (Heyns-Verbindungen) [2]. Bei der Reaktion von Aldosen mit einer Aminokomponente entsteht die Amadori-Verbindung, bei der Reaktion von Ketosen die Heyns-Verbindung [2, 43]. In der Zwischenphase der Maillard-Reaktion finden zahlreiche Dehydratisierungen und Fragmentierungen von Zuckern und Abbau von Aminosäuren sowie der Strecker-Abbau statt [43]. Die 1,2-Enolisierung oder 2,3-Enolisierung der Amadori-Verbindung, gefolgt von der Abspaltung der Aminokomponente oder einem Wassermolekül führt zur Bildung von α -Dicarbonylverbindungen (Desoxyosone) [43, 44]. Folgeprodukte dieser Desoxyosone und ihrer Endiol-Form sind unter anderem aliphatische Säuren, HMF, Furfural, Furan-, Pyrrol- und Pyrin-Derivate [2, 45]. Des Weiteren können Desoxyosone mit Aminosäuren im Rahmen des Strecker Abbaus unter Bildung von Aldehyden, Kohlenstoffdioxid und α -Aminoketonen reagieren [42, 43]. Weitere Produkte des Strecker Abbaus sind Pyrazine, Dehydropyrazine und Tetrahydropyrazine [46]. Die Endphase der Maillard-Reaktion führt über Aldolkondensationen, Polymerisierungen, Zyklisierungen, Quervernetzungen und Isomerisierungen zur Bildung von braunen, heterocyclischen, stickstoffhaltigen Polymeren, welche als Melanoidine bezeichnet werden [42–44, 47]. Die Struktur der Kaffeemelanoide ist sehr komplex und vielfältig [47]. Zudem sind verschiedene Verbindungen wie Polysaccharide, Proteine, CGA und Arabinogalactan-Proteine (AGPs) an der Bildung von Kaffeemelanoide beteiligt [47–50]. Untergeordnet könnten auch Galactomannane an der Bildung dieser Melanoide beteiligt sein [48]. Die Bildung insbesondere hochmolekularer Melanoide nimmt dabei mit dem Röstgrad zu [51].

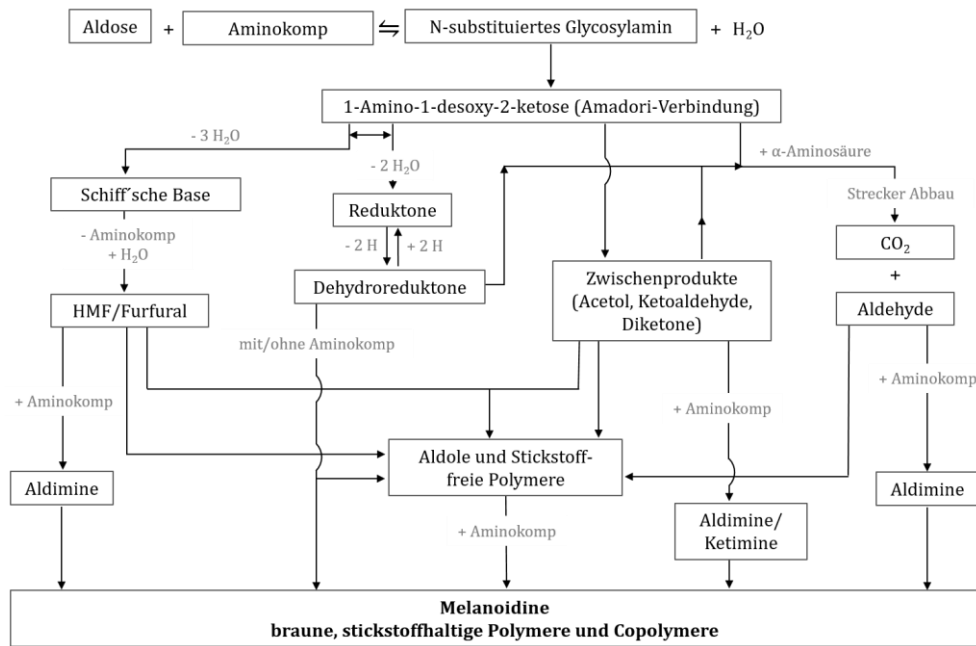


Abbildung 1-5: Übersicht bekannter Reaktionen ausgehend von der Amadori-Verbindung im Rahmen der Maillard-Reaktion nach dem Hodge-Schema. Aminokomp: Aminokomponente (modifiziert nach [43])

1.1.6.2 Veränderung der Kaffeeinhaltsstoffe während des Röstprozesses

Rohkaffeebohnen bestehen hauptsächlich aus Kohlenhydraten (55-61%), Lipiden (6-11%), Proteinen (6-10%), CGA (7-10%), Mineralstoffen (4%), Koffein (1-2%), Trigonellin (1-3%) und freien Aminosäuren (<1%) [47, 52]. Davon werden Verbindungen wie Mono- und Disaccharide, Proteine, freie Aminosäuren, Trigonellin und CGA als Vorläuferstrukturen für die während des Röstens entstehenden Aromen betrachtet [5].

Mannose, Galactose, Arabinose und Glucose sind die Hauptzucker in Rohkaffeebohnen und liegen zu einem großen Teil als Galactomannan, hochverzweigtes Typ II Arabinogalactan und Cellulose vor [37, 52]. Der größte Teil der Arabinogalactane liegt kovalent gebunden mit Proteinen als AGP vor [47]. Pectine, Xylane und Xyloglucane nehmen nur einen kleinen Teil dieser Polysaccharide ein [52, 53]. Neben den Polysacchariden kommt das Disaccharid Saccharose mit einer Konzentration von 8% in Rohkaffeebohnen vor. Saccharose gilt als die wichtigste Vorläuferstruktur für die Bildung von Aroma, Geschmack und Farbe [5]. Saccharose wird zu Beginn der Röstung zu Glucose und Fructose gespalten, welche neben weiteren Monosacchariden eine wichtige Rolle in der Maillard-Reaktion und Karamellisierung darstellen. Aus den Arabinogalactan-Seitenketten wird Arabinose freigesetzt, welche in die Bildung von Melanoidinen involviert ist [47, 52, 54, 55]. Neben der bevorzugten Abspaltung der Arabinose-Seitenketten erfolgt auch teilweise die Hydrolyse des Galactan-Rückgrats der Arabinogalactane [52, 56]. Die daraus freigesetzte Galactose wird während der Röstung anschließend abgebaut [57]. Die Protein-Polysaccharidbindungsstelle der AGPs bleibt während des Röstprozesses bestehen, obwohl ein starker Abbau des Polysaccharidanteils stattfindet [57, 58]. Auch die Galactomannane werden teilweise abgebaut und Galactose wird aus den Seitenketten abgespalten, wohingegen das Rückgrat der Mannane stabiler ist [52, 53]. Der Abbau von Xyloglucanen, welche ein Teil des Cellulose-Xyloglucan-Netzwerks sind, führt zu einem Verlust der Festigkeit der Zellwand [52]. Auch die in Rohkaffee enthaltenen Pectine werden während der Röstung stark abgebaut [53]. Polysaccharide tragen nicht nur zur Aromabildung bei, sondern tragen bei

dem gebrühten Kaffee auch zu verschiedenen Aspekten des Mundgefühls bei, wie beispielsweise die Viskosität oder die Schaumstabilität der Crema [57, 59]. Während des Röstens werden bis zu 40% der Polysaccharide abgebaut, wobei Cellulose unverändert bleibt und Arabinogalactan stärker abgebaut wird als Galactomannan [5, 60]. Der Abbau der Polysaccharide wird zusätzlich bei geringer Röstung von der Varietät beeinflusst [53].

Die in Röstkaffee enthaltenen Furane sind unter anderem auf den Abbau von Saccharose und die Pyrolyse von Arabinogalactanen zurückzuführen [55]. Reaktionsprodukte der Kohlenhydrate im Rahmen der Maillard-Reaktion sind neben HMF auch Acrylamid und Furfurylalkohol [7]. Furfurylalkohol und Acrylamid sind von der internationalen Agentur für Krebsforschung (International Agency for Research and Cancer, IARC) als möglich und wahrscheinlich krebserregend für den Menschen eingestuft [61, 62]. HMF ist von der IARC noch nicht klassifiziert worden, jedoch deuten Tierversuche eine krebserregende Wirkung von HMF an [63]. Zu Beginn der Röstung steigt die Konzentration von Acrylamid zunächst, nimmt jedoch mit steigendem Röstgrad wieder ab [64–66]. Im Gegensatz dazu sind die höchsten Gehalte von Furfurylalkohol und HMF in stark gerösteten Kaffeebohnen vertreten [66].

In Rohkaffeebohnen ist ein hoher Anteil von CGA enthalten, welche Ester aus Hydroxyzimtsäure-Isomeren mit Chinasäure sind [67]. Dabei stellen Kaffeesäure, Ferulasäure oder *p*-Cumarsäure die Hydroxyzimtsäure-Isomere dar [7]. Der Röstprozess fördert die Isomerisierung, die Hydrolyse in einen phenolischen und einen nicht-phenolischen Anteil sowie den weiteren Abbau der CGA [47]. Den Hauptteil der CGA in Kaffee nehmen CQA ein und von den verschiedenen Isomeren wird zunächst 5-CQA abgebaut, gefolgt von 4- und 3-CQA [67]. In Rohkaffeebohnen ist hauptsächlich 5-CQA enthalten, wohingegen in gerösteten Kaffeebohnen das Verhältnis von 5-CQA, 4-CQA und 3-CQA ähnlich ist [68]. Während der Röstung werden durch Dehydrierung der Chinasäure zunächst Lactone (Chlorogenic Acid Lactone, CGAL) gebildet, welche unter anderem für den bitteren Geschmack in gebrühtem Kaffee verantwortlich sind [5, 67, 69]. Die Wasserabspaltung aus dem Chinasäurering erfordert eine *syn*-1,3-diaxiale Konfiguration der Hydroxyl- und Carboxylgruppen. Dadurch können nur CGA-Isomere, die in Position fünf keine Hydroxyzimtsäuresubstituenten besitzen, CGAL bilden. In Röstkaffee sind daher die Lactone von 3-CQA und 4-CQA (3-Caffeoyl- und 4-Caffeoylchinasäure- γ -Lacton) vorhanden [38, 67]. Bei fortgeschrittenen Röstphasen werden sowohl CGA als auch CGAL teilweise weiter zu flüchtigen und nichtflüchtigen phenolischen Verbindungen abgebaut [5, 67]. Zu Abbauprodukten der CGA und CGAL gehören Phenylindane, welche zu dem bitteren Geschmack von Kaffee beitragen [69]. Weitere phenolische Abbauprodukte können 2-Methoxyphenol, 4-Ethyl-2-methoxyphenol und 4-Vinyl-2-methoxyphenol sein, welche dem Kaffee das typische rauchige, holzige Röstaroma verleihen [5]. Aus dem phenolischen Teil reagiert Kaffeesäure über das Intermediat 4-Vinylcatechol zu Chatechol, 4-Methylchatechol und 4-Ethylcatechol und weiter zu bitter schmeckenden Oligomeren von 4-Vinylcatechol [69, 70]. Weitere Kondensationsreaktionen der Abbauprodukte von Kaffeesäure führen zu ligninähnlichen Verbindungen [70]. Zusätzlich kann Catechol auch aus durch den Abbau der Chinasäure-Einheit der CGA und im Rahmen der Maillard-Reaktion aus Kohlenhydraten und Aminosäuren entstehen [71]. Weitere Abbauprodukte der Chinasäure sind Pyrogallol und Hydroxyhydrochinon [71]. Ferulasäure wird über 4-Vinylguaiacol zu Guaiacol, 4-Methylguaiacol und 4-Ethylguaiacol umgewandelt [70].

Aliphatische organische Säuren wie Zitronen-, Äpfel-, Essig- oder Ameisensäure sind wichtig für den sauren Geschmack von Kaffee [5]. Zitronen- und Äpfelsäure, die neben anderen organischen Säuren wie Essig- oder Milchsäure bereits in Rohkaffeebohnen vorhanden sind, werden während des Röstens abgebaut. Dahingegen werden Essig-, Ameisen-, Glycol-, Bernstein- und Milchsäure durch den thermischen Abbau von Kohlenhydraten wie Saccharose zunächst gebildet [5, 7, 37–39, 45, 45, 72, 73]. In fortgeschrittenen Phasen des Röstprozesses sinkt die Konzentration von Essig- und Ameisensäure aufgrund ihrer Flüchtigkeit, sodass der saure Geschmack mit steigendem Röstgrad sinkt [5, 45, 68, 72]. Dagegen werden Milch- und Glycolsäure auch bei Rösttemperaturen von >240°C weiter durch den Abbau von Saccharose gebildet [37].

Freie Aminosäuren werden bei Rösttemperaturen von >180 °C vollständig abgebaut [74]. Das Aminosäureprofil hat einen großen Einfluss auf das Röstaroma, da die Aminosäuren Schlüsselkomponenten in der Maillard-Reaktion und im Strecker-Abbau darstellen [5, 75]. Hauptsächlich sind in Rohkaffeebohnen L-Glutaminsäure, L-Asparaginsäure, L-Alanin und GABA vertreten [7, 74]. Diketopiperazine, Kondensationsprodukte aus zwei freien Aminosäuren, tragen zum bitteren Geschmack von Röstkaffee bei [5, 37]. Weitere wichtige stickstoffhaltige Verbindungen sind Koffein und Trigonellin. Während die Konzentration von Koffein mit steigender Röstung nur leicht sinkt, wird Trigonellin hauptsächlich zu N-Methylpyridinium (NMP) und Nicotinsäure, weiter auch zu Nicotinamid, Pyrazinen, Pyridinen und Pyrrolen abgebaut [35, 64, 72, 75–79]. Dabei steigt vor allem die Konzentration des thermischen Abbauproduktes NMP mit steigendem Röstgrad der Kaffeebohnen [78]. Die Abbauprodukte NMP, Nicotinsäure und Nicotinamid sind weniger relevant für die Bildung als Aromastoffe, haben jedoch vermutlich positive Effekte auf die menschliche Gesundheit [19, 35].

In der Lipidfraktion können während der Röstung Oxidationsreaktionen insbesondere ungesättigter Fettsäuren (Unsaturated Fatty Acids, UFA) auftreten, welche zur Bildung von Aldehyden, Ketonen oder Furanen führen [2, 7, 79]. Diese Verbindungen können mit anderen Kaffeebestandteilen wie Amininen zu heterozyklischen Aromaverbindungen weiterreagieren [2, 7]. Bei den in Kaffeebohnen vorhandenen Diterpenen treten Eliminierungsreaktionen von Wasser auf und es entstehen Dehydroditerpene, deren Konzentration mit steigendem Röstgrad zunimmt [7, 80]. Die in Rohkaffeebohnen enthaltenen Tocopherole werden während des Röstens teilweise abgebaut [7]. Die in geringen Mengen in Rohkaffeebohnen enthaltenen Phospholipide Phosphatidylcholin und Phosphatidylinositol werden während des Röstprozesses zu einem großen Teil abgebaut [81]. Der Fettgehalt steigt mit zunehmendem Röstgrad aufgrund der Gewichtsveränderung der Kaffeebohne relativ an, allerdings wird die Lipidzusammensetzung während des Röstprozesses nur geringfügig verändert [2, 65].

Nach der Röstung findet ein langsamer Desorptionsprozess statt und es tritt weiter hauptsächlich Kohlenstoffdioxid aus den Kaffeebohnen aus [82]. Daher ist es üblich, die gerösteten Kaffeebohnen in einer Verpackung mit einem Ventil aufzubewahren, durch welches das Gas entweichen kann [5].

1.1.6.3 Röstverfahren

Zu Beginn der Röstung findet eine endotherme Phase statt, in welcher Energie eingebracht werden muss, um Wasser zu verdampfen und die chemischen Reaktionen einzuleiten. Dabei kann die Wärmeübertragung des Rösters auf die Bohnen in Form von

Konvektion, Leitung und Strahlung erfolgen [5, 83]. Auf dieser Grundlage wurden zahlreiche verschiedene Röstverfahren entwickelt [83]. In der Kaffeeröstung wird der traditionelle Drehtrommelröster für diskontinuierliche aber auch kontinuierliche Röstprozesse eingesetzt [83]. In diesem Verfahren werden die Bohnen in einer horizontalen rotierenden Trommel oder in einem vertikalen festen Trommelröster mit heißen Gasen vermischt [83, 84]. Die Wärmeübertragung findet in allen drei Formen statt. Wärmeleitung tritt zwischen Röster und Kaffee oder zwischen den Bohnen untereinander auf. Bei der Konduktion wird die Wärme von den heißen Wänden des Rösters auf die Bohnen übertragen, wohingegen bei der Konvektion die Wärme von heißen Gasen transportiert wird [5]. In einem Trommelröster wird ein erheblicher Teil der Wärme durch Konduktion übertragen [5]. Eine weitere Röstmethode ist die Wirbelschicht-Röstung, bei welcher hauptsächlich Konvektion stattfindet [5, 84, 85]. Daneben wurden jedoch alternative Verfahren entwickelt, in welchen die Wärme mittels Infrarot, Mikrowellen oder Heißdampf übertragen wird [5].

Die Effizienz des Röstprozesses und die Qualität von geröstetem Kaffee wird von Faktoren wie der Röstzeit, der Gaszusammensetzung und Rösttemperatur beeinflusst [83]. Die typische Röstzeit in einer Trommelröstung beträgt 8-20 Minuten, wobei diese bei der Wirbelschicht-Röstung auf bis zu zwei Minuten verkürzt werden kann [5]. Bei diesen Kurzzeit-Röstverfahren mit einer Dauer von 1-4 Minuten wird die Aufheizphase durch eine bessere Wärmeübertragung stark abgekürzt und die dadurch bedingte, puffartige Verdampfung des Wassers führt zu einer größeren Volumenzunahme der Bohnen [2, 37].

Der angestrebte Röstgrad ist bedingt durch den Verbraucher sehr unterschiedlich. Verbraucher in den USA und Mitteleuropa bevorzugen hell gerösteten Kaffee, wohingegen Verbraucher in Frankreich, Italien und den Balkanländern dunkel gerösteten Kaffee (Espresso) favorisieren [2].

1.1.7 Chemische Zusammensetzung von Röstkaffee

1.1.7.1 Lipide

Im Allgemeinen besitzen Arabica Varietäten mit durchschnittlich 15% einen höheren Fettgehalt im Vergleich zu Robusta mit 10%. [7, 86]. Der Hauptteil der Lipide befindet sich in Form von Kaffeeöl innerhalb der Zellen und nur ein kleiner Teil haftet als Kaffee-Wachs an der Bohnenoberfläche [86]. Die Lipidfraktion besteht zu 75,0 - 78,8% aus Triacylglyceriden (TAG), gefolgt von bis zu 20% Diterpenfettsäureestern, Mono- und Diacylglyceriden, freien Fettsäuren (Fatty Acids, FA) und freien Diterpenen. Sterole und Sterolester (β -Sitosterol, Stigmasterin, Campesterol, Δ^5 -Avenasterol, Δ^7 -Avenasterol), γ -Tocopherol und Phosphoglycerolipide sind mit geringeren Anteilen vertreten [2, 7, 11, 37, 80, 86, 87]. In Arabica sind hauptsächlich die Diterpene Kahweol und Cafestol (Abbildung 1-6) und in geringen Mengen die durch die Röstung entstandenen Abbauprodukte Dehydrokahweol und Dehydrocafestol (Abbildung 1-6) enthalten [7, 80]. In Robusta befinden sich geringere Mengen Kahweol und zusätzlich das thermisch stabile Diterpen 16-O-Methylcafestol (16-OMC), das in der Praxis als Markersubstanz zur Unterscheidung von Arabica und Robusta herangezogen wird [37, 80, 88-92, 92, 93]. Allerdings ist 16-OMC in deutlich geringerer Konzentration auch in Arabica detektierbar [93, 94].

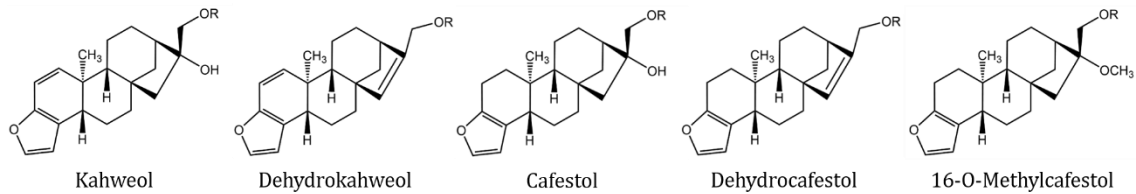


Abbildung 1-6: Strukturformeln der in Kaffee enthaltenen Diterpene Kahweol, Dehydrokahweol, Cafestol, Dehydrocafestol und 16-O-Methylcafestol. R: Fettsäure

Sowohl die freie Form von Kahweol und Cafestol, als auch deren Palmitatester werden im menschlichen Darm schnell absorbiert und verstärken die Aktivität des Cholesterin-Ester-Transferproteins. Dadurch steigt der Cholesterinspiegel im Serum von Blut an [80, 86, 95]. Andererseits erhöht Cafestol die Aktivität der Gluathion-S-Transferase, das die Bindung einer Vielzahl von elektrophilen Substanzen katalysiert und somit ein wichtiges entgiftendes Enzymsystem darstellt [86].

Die Fettsäureverteilung ähnelt dem eines Pflanzenöls [7]. Die Hauptfettsäuren sind Linolsäure (C18:2 $\Delta^{9c,12c}$) und Palmitinsäure (C16:0) mit Anteilen von 30,0 – 50,4% und 16,8 – 38,6%. In den TAG ist Palmitinsäure vorwiegend an Position 1 und 3 und Linolsäure an Position zwei und drei an Glycerin gebunden [7, 87]. Stearinsäure (C18:0) und Ölsäure (C18:1 Δ^{9c}), sind mit 4,5 – 13,1% und 7,6 – 18,9% in ähnlichen Verhältnissen vertreten. Die Minorfettsäuren Myristinsäure (C14:0), Palmitoleinsäure (C16:1 Δ^{9c}), α -Linolensäure (C18:3 $\Delta^{9c,12c,15c}$), Arachinsäure (C20:0) und Behensäure (C22:0) sind mit Anteilen von 0,1 – 2,3%, 0,2 – 4,0%, 0,3 – 6,0%, 0,7 – 6,7% und 0,5 – 0,6% vorhanden [37, 87]. Diterpen- und Sterolester enthalten bevorzugt gesättigte FA, und die ungesättigten FA liegen vermehrt als freie FA oder gebunden an TAG vor [37, 87].

Das Kaffee-Wachs besteht hauptsächlich aus Carbonsäure-5-Hydroxytryptaminen, welche eine antioxidative Wirkung haben und teilweise durch die Röstung abgebaut werden [37, 80]. Dabei sind Arachinsäure, Behensäure und Lignocerinensäure (C24:0) mit der primären Aminogruppe der 5-Hydroxytryptamine verknüpft [37].

1.1.7.2 Stickstoffhaltige Verbindungen

Die in Rohkaffee enthaltenen Proteine werden während des Röstprozesses ab- und teilweise in Melanoidine eingebaut [7]. In Röstkaffee konnten Proteine mit einem Molekulargewicht <14 kDa und >200 kDa nachgewiesen werden [7, 47, 96]. Prolin, Glutaminsäure und Leucin sind in Proteinen zu größeren Teilen vertreten, gefolgt von Alanin, Asparaginsäure, Phenylalanin und Isoleucin [74].

Die wichtigste stickstoffhaltige Verbindung in Kaffee ist das Purinalkaloid Koffein (1,3,7-Trimethyl-3,7,dihydro-1H-purin-2,6-dion, Abbildung 1-7) [7]. Im Vergleich zu Robusta (2,0%) besitzen Arabica Varietäten einen geringeren Koffeingehalt von 1,0% [97–99]. Die Biosynthese von Koffein verläuft im Rahmen des Purin-Stoffwechsel hauptsächlich über Xanthosin [97]. Durch eine Methylierung und Abspaltung von Ribose entsteht 7-Methylxanthin, aus welcher durch Methylierung Theobromin gebildet wird. Das Produkt einer weiteren Methylierung ist Koffein [37, 97]. Das Purinalkaloid liegt assoziiert mit CQA in einem 1:1 π -Komplex vor und dient in jungen Blättern, Früchten und Blütenknospen als chemische Abwehr vor Insektenlarven [99, 100]. In den Samenschalen wird Koffein freigesetzt, um die Keimung anderer Samen zu hemmen [38, 97]. Beim Menschen wird Koffein schnell durch den Gastrointestinaltrakt absorbiert und von der Leber zu Verbindungen wie

Paraxanthin, Theophyllin und Theobromin metabolisiert [101]. Koffein durchquert ungehindert die Blut-Hirn Schranke und stimuliert das Nervensystem [97, 101]. Negative kurzfristig auftretende Wirkungen von Koffein sind Beschleunigung des Herzschlags, Störungen des Magen-Darm-Trakts, Angstzustände, erhöhter Blutdruck und Schlaflosigkeit [102]. Zudem ist die Aufnahme von Koffein negativ mit dem Kalzium-Level im menschlichen Körper assoziiert [95].

Trigonellin (1-Methylpyridin-1-ium-3-carboxylat, Abbildung 1-7) ist ein weiterer in Kaffee vorkommender, stickstoffhaltiger Sekundärmetabolit [100]. In Rohkaffeebohnen kommt das Pyridinalkaloid mit einem Gehalt von 1-3% bezogen auf das Trockengewicht vor und wird wie in Kapitel 1.1.6.2 beschrieben teilweise während des Röstprozesses abgebaut [7, 99]. In Röstkaffee ist diese Verbindung noch zu 0,1-1% vertreten [38]. Die direkte Vorläuferstruktur von Trigonellin ist Nicotinsäure (Abbildung 1-7), die im Pyridinnucleotid-Zyklus aus Nicotinamidadenin-Dinucleotid (NAD) und Nicotinsäure-Mononucleotid entsteht [97, 99, 103]. Mittels der Trigonellin-Synthase wird Nicotinsäure methyliert und Trigonellin gebildet [99]. Trigonellin wird als Reaktion auf biotische und abiotische Stressbedingungen in der Kaffeepflanze angereichert und dient als Reserveverbindung für die NAD-Biosynthese [100, 103]. Studien zeigten, dass Trigonellin hypoglykämische, neuroprotektive, antiinvasive, krebshemmende und antibakterielle Wirkung besitzt [104]. Zudem hat Trigonellin eine stimulierende Wirkung auf den zellulären Energiestoffwechsel, indem es die Nutzung von Glucose als Energiesubstrat für die Synthese von Adenosintriphosphat (ATP) induziert [105]. Die gleiche Wirkung hat auch das durch den Röstprozess aus Trigonellin gebildete NMP (Abbildung 1-7), wobei sich der Wirkungsmechanismus von dem von Trigonellin unterscheidet [105]. Daneben zeichnet sich NMP durch eine gewichtsreduzierende Wirkung, eine chemopräventative Wirkung, eine antioxidative Kapazität und eine Hemmung der Magensäuresekretion aus [101, 105].

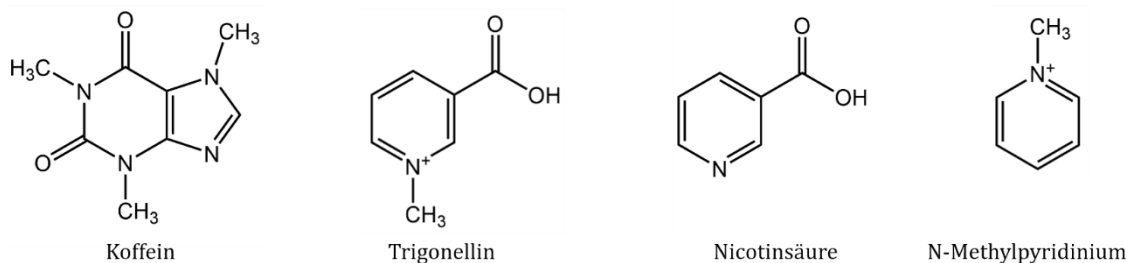


Abbildung 1-7: Strukturformeln der in Kaffee vorkommenden stickstoffhaltigen Verbindungen Koffein, Trigonellin, Nicotinsäure und N-Methylpyridinium

1.1.7.3 Kohlenhydrate

Die niedermolekularen Kohlenhydrate werden während der Röstung fast vollständig abgebaut und sind nur noch in geringen Mengen in Röstkaffeebohnen enthalten [58]. Die in Rohkaffeebohnen enthaltenen Arabinogalactane und Galactomannane werden während des Röstprozesses bis zu 60% und 36% abgebaut [53]. Der Abbau der Polysaccharide erfolgt durch die Hydrolyse des Polysaccharidgerüsts, was zu einer Verringerung des Molekulargewichtes führt. Weiter wird durch die Abspaltung von Zuckern aus den Seitenketten eine Verringerung der Verzweigung und dadurch eine bessere Löslichkeit der Polymere bewirkt [56, 58]. Die Hälfte des Polysaccharidanteils in Röstkaffee nehmen Galactomannane ein, die andere Hälfte besteht zu ähnlichen Anteilen aus Arabinogalactan und Cellulose [7]. Arabinogalactane besitzen in Kaffee eine große Heterogenität, was insbesondere den Verzweigungsgrad und die Monosaccharid-zusammensetzung der Seitenketten betrifft

[57]. Typ II Arabinogalactane bestehen zum größten Teil aus β -1,3-Galactorhamnosylresten, von welchen einige an O6 durch Arabinose- und Galactosereste substituiert sind [52, 57]. Die negativ geladenen Glucuronosylreste treten als nicht reduzierende Endeinheiten an einer 1,6-verknüpften Galactosylseitenkette auf [58]. Die Arabinogalactane sind in Kaffee hauptsächlich in Form von AGPs kovalent an Proteine gebunden [58]. Durch den Röstprozess werden bevorzugt die Arabinosereste an den Arabinogalactan Seitenketten abgebaut, was zu einem veränderten Arabinose/Galactose-Verhältnis und zu einer geringeren Verzweigung abhängig vom Röstgrad führt [96]. Jedoch bleibt auch nach dem Abbau der Arabinogalactan-Einheit die Protein-Polysaccharidbindung der AGP bestehen [7, 53]. Kaffeemannane bestehen zum größten Teil aus linearen Ketten von β -1,4-Mannosylresten mit einzelnen Galactoseeinheiten, welche mit einem Mannosylrest an Position C6 verknüpft sind [52, 58]. Zusätzlich können die Galactomannane acetyliert sein [52, 106]. Die Mannane liegen dabei als Mischung aus substituierten und unsubstituierten Polymeren vor [58]. Galactomannane sind während des Röstens stabiler als Arabinogalactane, jedoch wird auch bei diesen Polymeren der Polymerisierungs- und Verzweigungsgrad verringert [52, 53, 96, 106].

Die von Menschen aufgenommenen Kaffeepolysaccharide zählen zu den löslichen Ballaststoffen [7, 107]. Der Anteil dieser löslichen Ballaststoffe sowie deren Zusammensetzung wird unter anderem von der Varietät und von der Methode der Rohkaffeeaufarbeitung beeinflusst [107]. Die Polysaccharide werden von Darmbakterien fermentiert, was zur Bildung von kurzkettigen FA führt, welche positive Auswirkungen auf den Körper haben [7]. Dabei werden Mannane intensiver abgebaut als Arabinogalactane [7]. Zudem beeinflussen Kaffeepolysaccharide die Zusammensetzung der Mikrobiota des Darms [7]. AGPs könnten eine antitussive Aktivität besitzen [7].

1.1.7.4 Melanoidine

Röstkaffee besteht, abhängig vom Röstgrad, zu 25-30% aus Melanoidinen [7, 51]. Die Struktur der Melanoidine ist noch nicht vollständig aufgeklärt, da sie aufgrund ihrer Komplexität und Vielseitigkeit nicht direkt analysiert werden können [47]. So besitzen einige Melanoidine einen anionischen Charakter, wohingegen andere hydrophob sind [47]. Es sind verschiedene Melanoidin-Fractionen vorhanden, die sich unter anderem im Hinblick auf den Gesamtkohlenhydratgehalt und strukturelle Merkmale der Kohlenhydrate, wie ein unterschiedlicher Verzweigungsgrad der Arabinogalactane und die Beibehaltung der AGP, unterscheiden [70]. Es wird angenommen, dass Kohlenhydrate, Aminosäuren, Proteine und phenolische Verbindungen Komponenten der Melanoidine sind [47]. Arabinogalactane in Form von AGPs und untergeordnet Galactomannane sind kovalent in den Melanoidinen eingebunden, wobei die AGP-Fraktion einen großen Teil der hochmolekularen Melanoidine darstellt [47, 50, 53, 106]. Neben den AGP sind auch weitere Aminosäuren und Proteinfragmente in den Melanoidinen eingebunden, deren Aminosäurezusammensetzung zu großen Teilen aus Alanin, Asparagin/Asparaginsäure, Glutamin/Glutaminsäure und Glycin besteht [47, 51, 70]. Zusätzlich sind auch Maillard-Produkte von Lysin in den Melanoidinen enthalten [47]. Als phenolische Komponente sind CGA, Hydroxycimtsäureisomere und Chinasäure sowie deren Abbauprodukte kovalent und auch nicht kovalent in die Melanoidine eingebunden [41, 47, 49]. Studien deuten an, dass eine mögliche kovalente Bindung über einen Ester mit den Hydroxycimtsäuren oder der Chinasäure erfolgen könnte [51, 70]. Mögliche Bindungsstellen in den Melanoidinen könnten Proteine der AGPs oder Arabinose-Reste der Arabinogalactane sein können [49, 51, 70]. Nicht kovalente Bindungen könnten in Form von ionischen Wechselwirkungen auftreten oder Wasserstoffbrücken sein [70].

Die über das Kaffeegetränk aufgenommenen Melanoidine sind weitgehend resistent gegenüber dem menschlichen Magen und Dünndarm [47]. In *vitro* haben diese Polymere eine antioxidative Kapazität, was jedoch auch auf nicht kovalent gebundene, niedermolekulare Verbindungen wie CGA zurückgeführt werden kann [47]. Der antioxidative Effekt beruht auf der radikalfangenden Eigenschaft, sowie der Fähigkeit, mit Metallen einen Chelat-Komplex einzugehen [47]. Die Aufnahme von Kaffee-Melanoidinen kann zudem eine hemmende Wirkung gegenüber der Aktivität von Matrixmetallproteasen haben und so vermutlich zum Schutz vor Dickdarmkrebs beitragen [47]. Zusätzlich haben Melanoidine eine antikariogene Aktivität indem sie die Biofilmbildung von *Streptococcus mutans* hemmen [47]. Negative Auswirkungen der Melanoidine könnte die im Magenlumen stattfindende Reaktion dieser mit Nitrit und Thiocyanat aus dem Speichel sein. Das dabei entstehende Stickstoffmonoxid kann im Magen zur Hemmung des mikrobiellen Wachstums, der Schleimhautbildung und der Magenmobilität beitragen [47].

1.1.7.5 Organische Säuren

Die in Rohkaffee hauptsächlich vorkommenden organischen Säuren sind Zitronensäure, Chinasäure und Äpfelsäure [37, 108]. Neben diesen sind noch zahlreiche weitere Minorsäuren wie Bernsteinsäure, Milchsäure und Oxalsäure vertreten [7, 109]. Chinasäure wird aus 3-Dehydrochinasäure synthetisiert, welches ein Zwischenprodukt des Shikimisäure-Stoffwechsels ist [109]. Äpfelsäure und Zitronensäure sind Intermediate des Tricarbonsäurezyklus und Milchsäure wird aus Brenztraubensäure, dem Endprodukt der Glycolyse, gebildet [109]. Oxalsäure entsteht in Kaffee wahrscheinlich über die Oxidation von Glyoxylsäure [109]. Chinasäure kommt nur zu 1% als freie Säure vor und während des Röstprozesses steigt der Gehalt bedingt durch den Abbau von CGA leicht an [37, 75]. Wie in Kapitel 1.1.6.2 beschrieben, werden im Röstprozess zusätzlich Ameisen-, Essig-, Glycol und Milchsäure gebildet, wobei die Konzentration von Ameisen- und Essigsäure bei fortgeschrittenem Röstgrad und anschließender Lagerung des Röstkaffees wieder sinkt [37, 38, 45]. Der Abbau von Zitronen- und Äpfelsäure während des Röstens führt bei Zitronensäure zu Itaconsäure, Citraconsäure und Masaconsäure und bei Äpfelsäure zu Fumar- und Maleinsäure [2, 37, 38]. Tabelle 1-1 gibt einen Überblick über die in Röstkaffee vorkommenden organischen Säuren mit dem entsprechenden Konzentrationsbereich [2, 37]. Neben diesen kommen weitere Minorsäuren wie die Abbauprodukte von Zitronen- und Äpfelsäure nur in geringen Mengen vor [2, 37].

Tabelle 1-1: Gehalte von organischen Säuren in kommerziellen, gerösteten Kaffeeproben, analysiert mittels HPLC-UV. Angegeben ist der Konzentrationsbereich von 17 Kaffeeproben [37].

Organische Säure	Konzentrationsbereich [g/kg]
Chinasäure	8,9-15
Zitronensäure	4,3-7,0
Essigsäure	3,6-5,5
Ameisensäure	1,8-2,5
Bernsteinsäure	1,9-8,0
Glycolsäure	1,7-4,9
Äpfelsäure	1,0-3,9
Milchsäure	0,0-1,8

Die organischen Säuren tragen zum sauren Geschmack des Kaffeegetränks bei [37]. Einen Beitrag zum Säuregrad liefern vor allem Zitronen-, Essig-, Ameisen-, China- und Äpfelsäure [37].

1.1.7.6 Phenolische Komponenten

Die phenolischen Komponenten in Kaffee sind hauptsächlich CGA, welche in Rohkaffeebohnen einen Anteil von bis zu 10% besitzen [2]. Die Esterbindung der CGA kann an Position 3,4 und/oder 5 der Chinasäure stattfinden (Abbildung 1-8(A)) [7]. In Rohkaffeebohnen ist das Hauptisomer 5-CQA mit einem relativen Anteil von >50%. 4-CQA und 3-CQA nehmen einen Anteil von jeweils 10% ein (Abbildung 1-8(B)). Zudem existieren die drei diCQA 3,4-diCQA, 3,5-diCQA und 4,5-diCQA, die zusammen einen relativen Anteil von 15-20% einnehmen. Daneben sind die drei Feruloylchinasäuren (Feruloylquinic Acid, FQA) 3-FQA, 4-FQA und 5-FQA vertreten [7]. Die Biosynthese dieser Sekundärmetaboliten beginnt mit der Umwandlung von Phenylalanin in *p*-Cumaroyl-CoA, wobei Zimtsäure und *p*-Cumaronsäure als Zwischenstufen fungieren [110]. Die Umwandlung von *p*-Cumaroyl-CoenzymA (CoA) zu Caffeoyl-CoA erfolgt durch eine Hydroxylierung und die weitere Bildung von Feruloyl-CoA aus Caffeoyl-CoA durch eine Methylierung. Ausgehend von diesen Intermediaten erfolgt die Konjugation mit Chinasäure zu CGA, während CoA abgespalten wird [110, 111]. Die Chinasäure wird dabei aus 3-Dehydrochinasäure gebildet, welche wiederum aus Phosphoenolpyruvat mit Dihydroxyacetonphosphat als Zwischenprodukt entsteht [111].

CGA besitzen in Pflanzen mehrere Funktionen. Es wird angenommen, dass 5-CQA als Zimtsäurevorrat für die Ligninbiosynthese akkumuliert werden kann [110]. Zusätzlich werden CGA mit Schutz vor UV-Licht, Beitrag zur mikrobiellen Resistenz und einer Resistenz gegen tierische Schädlinge in Verbindung gebracht [110].

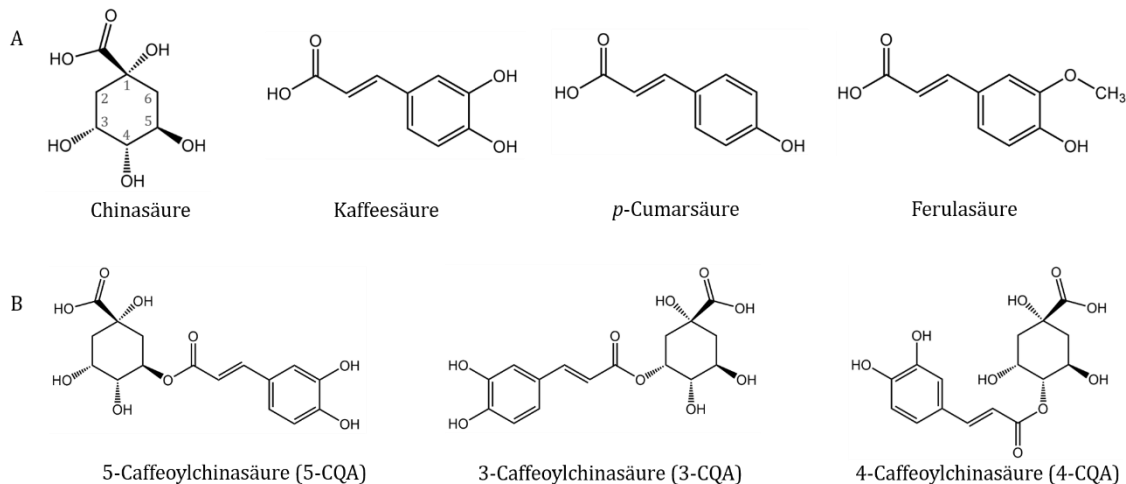


Abbildung 1-8: (A) Strukturformeln wichtiger Monomere der CGA; (B) Strukturformel der in Kaffee hauptsächlich vorkommenden Chlorogensäuren. CGA: Caffeoylchinasäure

Während des Röstens werden CGA abhängig vom Röstgrad bis zu 90% abgebaut, resultierend in einem breiten Produktspektrum, was von einfachen Phenolen zu strukturell komplexen Kondensationsprodukten reicht [2, 54, 70, 110]. In Röstkaffee sind CGA mit einem Anteil von ungefähr 3% vorhanden und bestehen zu 85% aus CQAs, 7% aus diCQAs und 5% aus FQAs [7]. Die beim Röstprozess entstehenden Hauptbitterstoffe CQL kommen mit einem Anteil von 0,2-0,4% bezogen auf die Trockenmasse in gerösteten Kaffeebohnen vor [7].

Die potentiellen gesundheitlichen Vorteile des Kaffeekonsums wie eine geringere Inzidenz verschiedener chronischer und degenerativer Krankheiten werden teilweise auf die CGA zurückgeführt [110]. So können CGA und deren thermischen Abbauprodukte Radikale

abfangen und haben daher eine antioxidative Wirkung [38, 112]. Des Weiteren besitzen CGA sowie deren Metaboliten eine entzündungshemmende und neuroprotektive Wirkung [110].

1.1.7.7 Minorkomponenten

Der flüchtige Anteil von Röstkaffee ist sehr komplex zusammengesetzt und beinhaltet über 850 Verbindungen, von denen jedoch nicht alle zum Aroma beitragen [2]. Zu diesen Verbindungen zählen Kohlenwasserstoffe, Alkohole, Aldehyde, Ketone, Ester, Pyrazine, Pyrrole, Pyridine, schwefelhaltige Verbindungen, Furane, Phenole und Oxazole [37]. Die Zusammensetzung dieser Verbindungen ist vom geographischen Ursprung, von der Rohkaffeeverarbeitung und insbesondere vom Röstgrad abhängig [37, 75]. Die wichtigsten Aromastoffe in Kaffee sind 2-Furfurylthiol, 4-Vinylguaiacol, Alkylpyrazine, Furanone, Acetaldehyd, Propanal und Methylpropanal [2, 37, 75]. Diese Verbindungen entstehen während des Röstens durch den Abbau von Lipiden, Kohlenhydraten, Proteinen, Aminosäuren und CGA sowie im Rahmen der Maillard-Reaktion (Kapitel 1.1.6.1 und 1.1.6.2) [7].

Die Mineralstoffe bestehen zu mehr als 99% aus den Makroelementen Kalium, Calcium und Magnesium [2, 113]. Von diesen Makroelementen ist Kalium mit durchschnittlich 15,6 mg/kg dominant [113]. Zu den Mikroelementen in Röstkaffee zählen Eisen, Mangan, Kupfer, Zink und Nickel [113]. Kobalt, Chrom, Blei und Cadmium sind nur in Spuren vorhanden [113].

Neben den gesundheitsfördernden Kaffeeinhaltsstoffen existieren auch toxikologisch relevante Verbindungen wie Mykotoxine [7]. Mykotoxine sind niedermolekulare Gifte, welche als Sekundärmetaboliten von einigen Fadenpilzen produziert werden [7]. Kaffee ist vor allem mit dem Mykotoxin Ochratoxin A (OTA) kontaminiert, das sich schädigend auf die Nieren auswirkt [7]. Andere Mykotoxine wie Aflatoxine kommen selten vor [7]. Der Gehalt von OTA in Kaffee ist gering und überschreitet nur selten den Grenzwert von 5 µg/kg [7].

1.1.8 Kaffeedefekte

Kaffeedefekte sind Schäden oder Mängel in oder an den Rohkaffeebohnen und können in verschiedene Klassen eingeteilt werden [15]. Feldgeschädigte Bohnen besitzen Defekte, die während des Anbaus entstehen, beispielsweise durch eine Schädlingsplage wie den Kaffeekirschenbohrer oder Stress der Pflanze aufgrund klimatischer Bedingungen, Wasser- oder Nährstoffmangel [15]. Die von der Verarbeitung geschädigten Bohnen beinhalten Defekte, die durch unzureichende Verarbeitungsschritte bei der Rohkaffeeverarbeitung wie dem Entfernen des Fruchtfleisches, Waschen oder Trocknen entstehen können [15]. Auch können Schäden bedingt durch eine falsche Lagerung der Rohkaffeebohnen oder durch die Mitlagerung von Schädlingen vorhanden sein [15]. Durch unzureichende Reinigungsverfahren nach dem Entfernen der Pergamenthaut können neben den Rohkaffeebohnen noch weitere getrocknete Teile der Kaffeefrucht oder der Kaffeepflanze vorliegen, welches auch als Defekt bezeichnet wird [15].

Der Kaffeekirschenbohrer kommt ursprünglich aus Afrika und ist ein Schädling für Kaffee, der vor allem in geringeren Höhen angebaut werden. Das Weibchen gräbt einen Tunnel in das Innere der unreifen Kaffeekirsche, legt dort ihre Eier ab und die Larven fressen das Äußere des Bohnengewebes [15]. Je nachdem, wie stark die Bohnen angefressen werden, wird zwischen „Eaten Beans“ und „Well Eaten Beans“ unterschieden. „Dark Brown Beans“ werden durch den Angriff der Antesiawanze verursacht, die den Saft aus den unreifen

Früchten saugt. Dieser Defekt wird auch durch überreife Früchte verursacht [15]. Sogenannte „Quaker“ sind nicht vollständig ausgebildete Bohnen, was auf Wachstumsprobleme oder auf einen Nährstoffmangel der Pflanze zurückzuführen ist [15]. Bedingt durch Wachstumsstörungen entstehen fehlentwickelte Bohnen, die abnormal geformt sind („Malformed“) [114].

„Black Beans“ sind schwarze Rohkaffeebohnen, die durch einen Pilzbefall der Kirsche, aber auch durch falsche Bewässerung und Erntetechniken entstehen können [15, 114]. „Sour Beans“ oder „Brown Beans“ sind das Ergebnis einer Fehlfermentation bei der Rohkaffeeverarbeitung [15, 114]. Wurden die Rohkaffeebohnen zu schnell oder zu heiß getrocknet, bilden sich „Dark Green Beans“ [114]. Die „Faded Beans“ sind verblasste Bohnen, welche bei der nassen Aufarbeitung zu lange im Fermentationsbecken verblieben sind [114]. „Bleached Beans“ sind ebenfalls ausgebleicht, der Grund hierfür ist jedoch Regenfall während der Trocknung [114]. Bohnen, die während der nassen Aufarbeitung im Entpulper angeschnitten oder gequetscht wurden, werden als „Pulper Cuts“ bezeichnet. „Brocken Beans“ entstehen meistens während der Entfernung der Pergamenthaut [15].

Durch unzureichende Reinigungsverfahren können Reste der Kaffeekirsche wie die Schale, die Pergamenthaut oder Fragmente des getrockneten äußeren Perikarps, aber auch Reste der Kaffeepflanze wie kleine Zweige zu den Rohkaffeebohnen gelangen [15, 114]. Diese Defekte werden „Husks“, „Shells“, „Parchment“, „Pods“ und „Twigs“ genannt [114].

Die Defekte haben unterschiedlich starke Auswirkungen auf das sensorische Profil und werden daher in der ISO 10470:2004 Green Coffee Defect Reference Chart verschieden gewichtet [115]. Jedoch ist die Klassifizierung der Kaffeedefekte weltweit nicht einheitlich geregelt und folgt unterschiedlichen Richtlinien. Die International Coffee Organization empfiehlt eine Höchstzahl von 86 Defekten/300 g Arabica Kaffee. Die Anzahl an Defekten richtet sich dabei nach dem Einfluss der Mängel auf den Masseverlust und die sensorische Beurteilung [116]. Die Methode der Speciality Coffee Association of America umfasst verschiedene Kaffeequalitäten, welche sich nach Anzahl der Fehler in 350 g Rohkaffee richten. Die brasilianische Klassifizierung für Rohkaffee verwendet eine Klassifizierung mit vielen Einstufungen, die sich danach richten, wie viele Defekte in 300 g Rohkaffee vorhanden sind [117]. Nach der New York Methode werden die Kaffeedefekte in Bezug auf den Defekt der „Black Beans“ gewichtet, wobei sieben Klassifizierungsstufen entstehen, welche als NY2 bis NY8 bezeichnet werden. Die schlechteste Klassifizierungsstufe ist NY8, bei welcher 450 „Black Beans“ in 300 g Rohkaffee enthalten sein dürfen. Die Einteilung der anderen Defekte wird im Hinblick auf die Gleichwertigkeit der „Black Beans“ bewertet. So entsprechen beispielsweise zwei „Sour Beans“ einem Äquivalent der „Black Beans“ [118].

1.2 Kernspinresonanz-Spektroskopie

Die Kernspinresonanz-Spektroskopie (Nuclear Magnetic Resonance Spectroscopy, NMR) ist eine Analysentechnik, welche in den letzten Jahren vermehrt bei der Analyse und Authentizitätskontrolle von Lebensmitteln eingesetzt wurde. Die NMR wurde im Jahr 1945 erstmals von zwei unabhängig arbeitenden Arbeitsgruppen der Harvard- und Stanford-Universität angewandt. Die erste mutmaßliche Anwendung der NMR in der Lebensmittelwissenschaft fand im Jahr 1957 statt, bei der mittels NMR mit niedriger Auflösung die Feuchtigkeit in Lebensmitteln gemessen wurde [119]. Die mittlerweile angewandte hochauflösende eindimensionale (1D) NMR kann in einem einzigen Experiment strukturelle und quantitative Informationen verschiedenster

Lebensmittelinhaltsstoffe liefern. Diese Experimente benötigen keine Trennung der verschiedenen Lebensmittelbestandteile oder eine aufwändige Probenvorbereitung [119]. In Fällen, in denen die Lebensmittelproben zu komplex sind und es zu Signalüberlappungen in der 1D-NMR kommt, können homonukleare und heteronukleare zweidimensionale (2D)-NMR-Experimente angewandt werden [119]. Die Kombination der NMR mit multivariaten statistischen Verfahren erlaubt eine effiziente Auswertung der oft komplexen Datensätze [119]. Diese multivariaten statistischen Verfahren werden in Kapitel 3.5 genauer beschrieben. Eine häufig verwendete Form der NMR ist die eindimensionale Protonen NMR (^1H -NMR). In der 2D-NMR werden häufig die Kerne ^1H und ^{13}C gemessen [119, 120].

Für die NMR-Messung kommen heutzutage supraleitende Kryomagnete zum Einsatz, welche durch flüssiges Helium gekühlt sind und höhere Magnetfeldstärken und damit auch verbesserte Empfindlichkeiten im Vergleich zu Permanentmagneten erreichen [120]. Die Durchführung der NMR-Experimente finden mit dem sogenannten Impulsverfahren statt. Dazu werden mit einem Hochfrequenzimpuls alle Kerne einer Sorte gleichzeitig angeregt [120].

1.2.1 Physikalische Grundlagen und Messprinzip der NMR

Atomkerne haben eine quantenmagnetische Eigenschaft, die als Spin bezeichnet und durch welche ein magnetisches Moment μ verursacht wird. Diese beiden vektoriellen Größen sind proportional zueinander und ihre Beziehung wird durch die Proportionalitätskonstante γ beschrieben. Diese Konstante wird auch als gyromagnetisches Verhältnis bezeichnet und ist für jedes Isotop charakteristisch. Kerne mit einer Kernspinquantenzahl I von Null haben kein magnetisches Moment und können folglich auch nicht mit der NMR detektiert werden. Dies trifft auf die beiden Atomkerne ^{12}C und ^{16}O zu, welche beide NMR inaktiv sind. Für die NMR aktiven Kerne ^1H sowie ^{13}C ist $I = \frac{1}{2}$. Die ^1H Kerne haben ein hohes gyromagnetisches Verhältnis und eine natürlich vorkommende Häufigkeit von 99,99%, wodurch diese in der NMR empfindlich detektiert werden. Die geringe natürliche Häufigkeit von ^{13}C Isotopen von 1,1% zusammen mit einem geringeren gyromagnetischen Verhältnis setzt die Empfindlichkeit dieser stark herab.

Werden NMR aktive Kerne in ein statisches Magnetfeld B_0 gebracht, richten sich die Spins aus. Dabei können $2I+1$ mögliche Spinzustände angenommen werden. Für die beiden Kerne ^1H und ^{13}C sind daher je zwei Zustände möglich, die als α - und β -Zustand bezeichnet werden. Die Spins richten sich dabei parallel (α -Zustand) oder antiparallel (β -Zustand) zum Magnetfeld B_0 aus (Abbildung 1-9(A)). Die beiden Zustände haben unterschiedliche Energieniveaus, deren Energieunterschied laut Formel 1 definiert ist [119, 120]:

$$\Delta E = \gamma B_0 \frac{h}{2\pi} \quad (1)$$

Der Energieunterschied der beiden Zustände ist somit von Magnetfeldstärke als auch von dem gyromagnetischen Verhältnis abhängig. Ein kleiner Überschuss der Spins befindet sich in dem energieärmeren α -Zustand, woraus sich eine makroskopische Magnetisierung M_0 in Feldrichtung ergibt (Abbildung 1-9(C)).

Durch den Einfluss von B_0 präzedieren die Kerndipole um die Richtungsachse (z-Achse) des Magnetfeldes mit einer bestimmten Frequenz (Abbildung 1-9(B)). Diese Larmorfrequenz ν_L ist proportional zur magnetischen Flussdichte und dem gyromagnetischen Verhältnis (Formel 2) [119, 120]:

$$\nu_L = \left| \frac{\gamma}{2\pi} \right| B_0 \quad (2)$$

Wird nun zusätzlich Energie in Form einer elektromagnetischen Welle (Radiofrequenzpuls, RF) eingestrahlt, deren Frequenz der Larmorfrequenz entspricht, ist die Resonanzbedingung erfüllt und es sind Übergänge zwischen den Zuständen möglich.

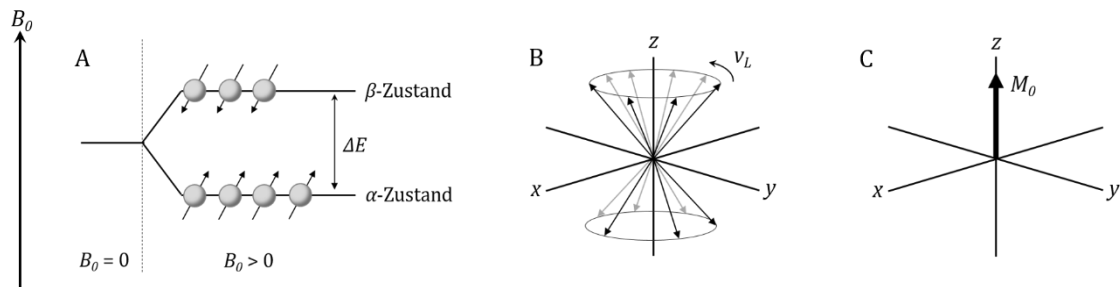


Abbildung 1-9: (A) Besetzung der Energiezustände für Kerne mit einer Kernspinqantenzahl von $\frac{1}{2}$ mit Aufspaltung in den α - und β -Zustand bei angelegtem statischen Magnetfeld B_0 . Im Gleichgewichtszustand befindet sich ein geringer Überschuss im α -Zustand. (B) Verteilung der Spinzustände für Kerne mit einer Kernspinqantenzahl von $\frac{1}{2}$ auf dem Doppelpräzessionskegel. Die Kerndipole präzedieren mit der Larmorfrequenz ν_L um die Achse von B_0 , welches in $+z$ -Richtung zeigt. (C) aus der Verteilung der Spinzustände resultierende makroskopische Magnetisierung M_0 (modifiziert nach [119–121])

In einem sich um die z -Achse rotierenden Koordinatensystem mit den Achsen x' und y' wird durch die Einstrahlung des RF die makroskopische Magnetisierung um einen bestimmten Winkel θ zur z -Achse ausgelenkt. Ein RF von 90°_x lenkt die Magnetisierung M_0 in die $x'y'$ -Ebene aus, woraus die Quermagnetisierung M_y resultiert. Der Magnetisierungsvektor präzediert mit der Larmorfrequenz um die z -Achse und erzeugt eine oszillierende Spannung in der Detektorspule. Die elektrischen Signale der Spule werden als Freier Induktionsabfall (Free Induction Decay (FID)) bezeichnet. Nach Einstrahlen des RF finden Relaxationsprozesse und somit der Übergang in den thermischen Grundzustand statt. Die mit dem FID aufgenommene Zeitdomäne kann durch eine mathematische Operation, der Fourier-Transformation (FT), in eine Frequenzdomäne umgewandelt werden, und ein Spektrum mit den typischen NMR-Signalen wird erzeugt [119–121].

Bei den beschriebenen Relaxationsprozessen wird überschüssige Energie wieder abgegeben und die Kernspins kehren den Grundzustand zurück. Gemäß der Bloch'schen Gleichung können Relaxationsprozesse mit zwei verschiedenen Relaxationszeiten beschrieben werden, die longitudinale und transversale Relaxationszeit. Die longitudinale Relaxation beschreibt die Rückkehr der makroskopischen Magnetisierung in Richtung z -Achse (Abbildung 1-10). Die Energie der Spins wird dabei in Form von Wärme an die Umgebung abgegeben und der Vorgang wird daher auch als Spin-Gitter-Relaxation bezeichnet [121]. Diese Energie ist jedoch so gering, dass die dadurch bedingten Temperaturänderungen in der Probe nicht nachweisbar sind. Die Zeit, die dieser Prozess benötigt, wird als T_1 -Relaxationszeit bezeichnet.

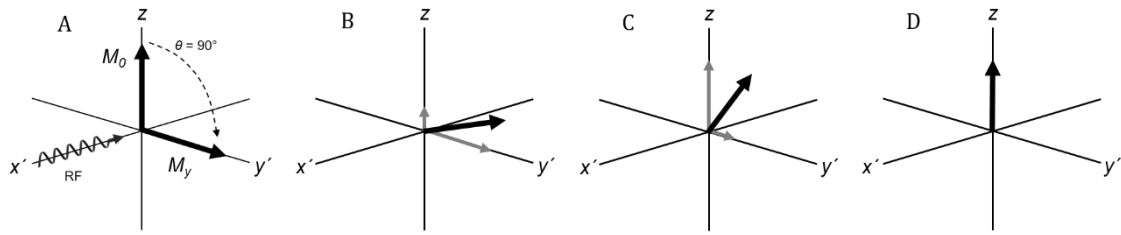


Abbildung 1-10: Verlauf der longitudinalen Relaxation in einem sich um die z -Achse rotierenden Koordinatensystem. (A) Auslenkung der makroskopischen Magnetisierung (M_0) durch den Radiofrequenzpuls (RF) um den Winkel $\theta = 90^\circ$. (B) Es resultiert die Quermagnetisierung M_y und (C) die Wiederherstellung der makroskopischen Magnetisierung M_0 in Richtung z -Achse. (D) Widerhergestellter Gleichgewichtszustand (modifiziert nach [119–121])

Die transversale Relaxation beschreibt den Verlust der Phasenkohärenz zwischen den Kernspins (Abbildung 1-11). Im Grundzustand präzedieren die Kernspins wieder mit beliebiger Phasenverschiebung zueinander um die z -Achse. Die transversale Relaxation findet durch einen gegenseitigen Energieaustausch mit nicht angeregten Kernen statt, woraus sich der Name Spin-Spin-Relaxation ergibt. Die Zeit für die transversale Relaxation wird T_2 -Relaxationszeit genannt [119, 121].

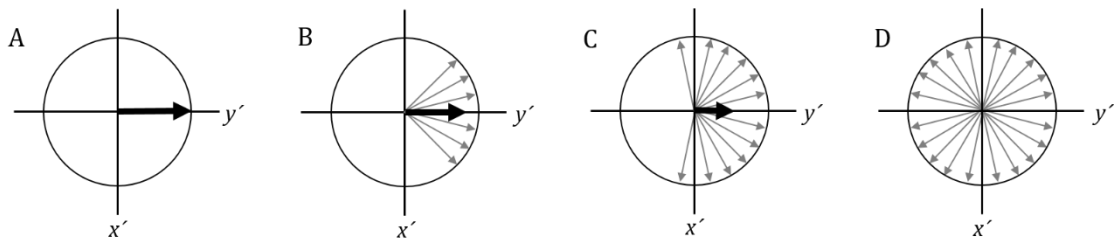


Abbildung 1-11: Verlauf der transversalen Relaxation in einem sich um die z -Achse rotierenden Koordinatensystem. (A) Die Quermagnetisierung nach einem Radiofrequenzpuls von $\theta = 90^\circ$. (B) und (C), das Auffächern der einzelnen Kerndipole in der Transversalebene; (D) der wiederhergestellte Grundzustand. Die Kerndipole präzedieren mit beliebiger Phasenverschiebung. (Modifiziert nach [120, 121])

Da T_2 kürzer ist als T_1 , wird letzteres als Maß zur Rückkehr in den Grundzustand verwendet. Die T_1 von Protonen liegt bei mittelgroßen, organischen Molekülen im Bereich von 0,3-5,0 s, wohingegen diese bei ^{13}C im Bereich von einigen Sekunden bis hin zu mehreren zehn Sekunden liegt [121, 122]. Um sicherzugehen, dass die Spins nach einem RF von 90° wieder vollkommen relaxiert sind, sollte eine Repetitionszeit von $5 \cdot T_1$ eingehalten werden [121]. Die Bestimmung der T_1 -Relaxationszeit kann mit dem Inversion Recovery Experiment durchgeführt werden, worauf in Kapitel 3.3.2 genauer eingegangen wird.

1.2.1.1 Chemische Frequenzen

Die Resonanzfrequenzen in der NMR werden von der Umgebung des beobachteten Kerns beeinflusst. Der Atomkern ist von der Elektronendichte der Bindungselektronen umgeben, welche von der Bindung und ihrer relativen Position innerhalb des Moleküls beeinflusst wird. Wird das Atom dem angelegten Magnetfeld B_0 ausgesetzt, löst dies in der Elektronenwolke eine Drehbewegung aus. Diese Bewegung erzeugt ein lokales Magnetfeld, das dem äußeren Magnetfeld B_0 entgegengesetzt ist, womit das am Kernort wirkende Magnetfeld B_{eff} sich von dem angelegten Feld B_0 leicht unterscheidet (Formel 3) [119–121]:

$$B_{eff} = B_0 - \sigma B_0 = (1 - \sigma)B_0 \quad (3)$$

Die Abschirmungskonstante σ ist abhängig von der Dichte und Verteilung der Elektronenwolke und nimmt für schwere Kerne einen höheren Wert an. Die Elektronendichte wird von der chemischen Umgebung innerhalb eines Moleküls beeinflusst, und die Veränderung der lokalen Feldstärke führt zu einer geringfügig veränderten Resonanzfrequenz (Formel 4):

$$\nu_1 = \left| \frac{\gamma}{2\pi} \right| (1 - \sigma) B_0 \quad (4)$$

Chemisch nicht äquivalente Kerne besitzen demnach eine unterschiedliche Resonanzfrequenz und erzeugen verschiedene Resonanzsignale.

Da die Resonanzfrequenz ν_1 und magnetische Flussdichte B_0 miteinander verknüpft sind, gibt es keinen absoluten Maßstab. Aufgrund dessen wird ein relativer Maßstab verwendet, welcher die Differenz zwischen einem Resonanzsignal einer Substanz und das einer Referenz-Verbindung wiedergibt. Zur Erzeugung des Referenzsignals wird meist Tetramethylsilan (TMS) oder Trimethylsilylpropionat (TSP) in die Messlösung hinzugegeben. Da jedoch auch die Frequenzdifferenz von B_0 abhängig ist, wurde die dimensionslose Größe δ als chemische Verschiebung definiert (Formel 5). Die chemische Verschiebung wird in ppm ausgedrückt und bezeichnet die Position eines bestimmten Resonanzsignals im NMR-Spektrum in Bezug auf das Signal der Referenzsubstanz, welche den Beginn der δ -Skala definiert. Dabei wird das Signal von TMS oder TSP gleich 0 ppm gesetzt [119, 120]. Es gilt:

$$\delta_{\text{Substanz}} [\text{ppm}] = \frac{\nu_{\text{Substanz}} - \nu_{\text{Referenz}}}{\nu_{\text{Referenz}}} 10^6 \quad (5)$$

Die chemische Verschiebung wird stark von benachbarten Substituenten des zu betrachtenden Kerns beeinflusst. Des Weiteren hat unter anderem die räumliche Anisotropie einen Einfluss auch die chemische Verschiebung [119, 120].

1.2.1.2 Spin-Spin-Kopplung

Die magnetischen Momente benachbarter Kerne wechselwirken miteinander, sodass es zur Aufspaltung von Signalen kommt, was als Feinstruktur bezeichnet wird. Diese Wechselwirkung kann direkt über den Raum oder indirekt über Bindungselektronen stattfinden.

Durch die Wechselwirkung zweier Kerne entsteht ein Zweispinsystem und es kommt zur Aufspaltung in zwei Dubletts, wobei die Abstände jeder Resonanzlinie eines Dubletts für beide Kerne gleich und durch die skalare Kopplungskonstante $J_{1,2}$ gekennzeichnet sind. Durch die Wechselwirkung mit mehreren Kernen kann es zu Triplets, Quartetts oder Multiplettstrukturen kommen. In Spektren erster Ordnung ist die Zahl der Signale eines Multipletts M abhängig von der Anzahl der Nachbarn (p) und kann sich bei einer Kernspinzahl I von $\frac{1}{2}$ nach Formel 6 berechnen lassen [120]. Das Intensitätsmuster dieser Multiplizität ergibt sich nach dem Pascalschen Dreieck (Formel 6) [119]. Es gilt:

$$M = p + 1 \quad (6)$$

Bei Wechselwirkungen von chemisch nicht äquivalenten Kernen ergibt sich eine Multiplizität von $(n+1)(n'+1)$, mit n und n' für die Anzahl der nicht äquivalenten Kerne ist [119]. Für einen Wert von 1 für jeweils n und n' würde sich daraus ein Dublett vom Dublett ergeben [119].

1.2.2 Grundlagen der quantitativen NMR

Die quantitative NMR (qNMR) ermöglicht ohne aufwändige Probenvorbereitung die simultane Bestimmung mehrerer Zielanalyten aus verschiedenen Substanzklassen [119]. Daher und aufgrund weiterer Vorteile wurde die NMR in den letzten Jahren zur quantitativen Analyse von Lebensmitteln, Naturprodukten und Pharmazeutika eingesetzt [119].

Die qNMR basiert auf der direkten Proportionalität zwischen der Fläche eines Signals I_x und der Anzahl der Kerne N_x , welche das Signal erzeugen (Formel 7):

$$I_x = K_S \cdot N_x \quad (7)$$

Die Spektrometerkonstante K_S wird von Faktoren wie dem Spektrometer, dem Probenkopf, der Pulsanregung, der Temperatur und der Probe selbst beeinflusst [123, 124].

Für genaue und präzise qNMR-Messungen sollten einige experimentelle Akquisitions- und Prozessierungsparameter optimiert werden [119, 123]. Zunächst sollte bei der Wahl der NMR-Röhrchen darauf geachtet werden, dass diese einen Durchmesser mit möglichst geringer Varianz besitzen. Bevor die Messung gestartet wird, sollte die Magnetfeldhomogenität optimiert werden (Shimming) und die Abstimmung und Anpassung der Spektrometerfrequenz (Tuning und Matching) erfolgen.

Bei quantitativen Analysen werden vorzugsweise Einzelpuls-NMR-Experimente verwendet, die aus einer Relaxationsverzögerung (Relaxation Delay, D) und einem 90° Puls, gefolgt von der Signalerfassung bestehen. Ist der dynamische Bereich zwischen Lösungsmittel- und Analyt-Signal zu groß, werden Lösungsmittelunterdrückungstechniken angewandt.

Für die Anregung der Kerne kann ein Pulswinkel von 90° oder kleiner verwendet werden. Die Pulsleistung sollte ausreichend hoch und die Pulslänge so gering sein, um alle Kerne innerhalb der spektralen Breite gleichmäßig anzuregen. Die Wirksamkeit des Pulses kann abhängig von den chemisch-physikalischen Eigenschaften der Probe variieren. Daher sollte die Pulslänge, die ungefähr $10 \mu\text{s}$ beträgt, bei jeder Probe kalibriert werden. Durch einen 90° Puls wird die maximale Intensität der Signale erreicht im Vergleich zu geringeren Pulswinkeln [119, 123]. Allerdings benötigen die Kerne bei einem 90° Puls eine längere Zeit, um in den Gleichgewichtszustand zurückzukehren. Für einen 90° Puls wird eine Repetitionszeit von $5 \cdot T_1$ vorgeschlagen, um 99,3% der Gleichgewichts-Magnetisierung zu erreichen. Bei einem Puls von beispielsweise 30° sollte dahingegen eine Repetitionszeit von $3 \cdot T_1$ eingehalten werden. Die Repetitionszeit beinhaltet die Relaxationsverzögerung und die Akquisitionszeit und richtet sich nach dem Analyten mit der der längsten T_1 . Zur Quantifizierung von Metaboliten, die bei kürzeren Repetitionszeiten gemessen wurde, können die Relaxationseffekte durch einen T_1 -Korrekturfaktor kompensiert werden. Die Zeit zur Aufnahme des FID (Akquisitionszeit) sollte ausreichend lang gewählt werden, um ein Abschneiden des FID und die daraus resultierenden Intensitätsverzerrungen zu vermeiden [119, 123, 125]. Mit zunehmender Akquisitionszeit werden zudem mehr Datenpunkte zum

Definieren des Spektrums verwendet. Die Länge der Akquisitionszeit wird auf Grundlage der spektralen Breite und der erforderlichen digitalen Auflösung gewählt [122]. Messparameter wie die Empfängerverstärkung (Receiver Gain), die digitale Auflösung und die Temperatur sollten vor der Messung optimiert und auch über mehrere Messungen hinweg konstant gehalten werden [119, 123].

Für präzise quantitative Ergebnisse in der ^1H -NMR sollte ein Signal-zu-Rauschverhältnis (Signal to Noise ration, S/N) von 250:1 vorliegen. Um bei gleichbleibender Magnetfeldstärke ein besseres S/N zu erreichen, kann die Anzahl der Scans erhöht werden [123].

Neben den Aufnahmeparametern sollten auch die Prozessierungsparameter optimal gewählt werden. Beim Windowing wird eine Filterfunktion auf den FID angewandt, um das S/N oder die Auflösung der Signale zu verbessern. Durch die Anwendung einer exponentiellen Multiplikation wird der vordere Teil des FID verstärkt gewichtet. Die NMR-Signalintensität nimmt während der Akquisitionszeit ab, wohingegen die Rauschintensität konstant bleibt. Somit enthält der vordere Teil des FID hauptsächlich die Informationen und eine Gewichtung dieses Teils kann die Empfindlichkeit erhöhen. Die Linienverbreiterung der Signale wird durch die Abklingrate der Exponentialfunktion bestimmt, wobei für diesen Parameter Werte von $L_b = 0,3-1,0$ Hz empfohlen werden.

Anschließend kann bei der Prozessierung ein Zero Filling angewandt werden, in welcher die Anzahl der Datenpunkte des FID künstlich erhöht und so die digitale Auflösung des Spektrums verbessert wird. In Routineexperimenten wird die Anzahl der Datenpunkte verdoppelt. Grund hierfür ist die Tatsache, dass die Information im FID nach der FT in zwei Teile (Real- und Imaginärteil) aufgeteilt werden. Wird nach der FT nur der Realteil betrachtet, kommt es zum Informationsverlust. Durch das Zero Filling um die Verdopplung der Anzahl der Datenpunkte wird diese verlorene Information wiedergewonnen. Eine weitere Erhöhung der Anzahl der Datenpunkte wirkt sich nicht weiter auf eine Verbesserung der digitalen Auflösung aus. Eine Verdopplung der Datenpunkte vor der Aufnahme des FID wäre mit einer längeren Messzeit und vermehrten Aufnahme des Rauschens verbunden, was in einem schlechteren S/N resultieren würde. Mit der Anwendung des Zero Fillings kann die digitale Auflösung verbessert werden, ohne die Messdauer zu erhöhen und das S/N zu verschlechtern [121].

Bei der Phasenkorrektur werden Fehler korrigiert, welche aufgrund des Rechenalgorithmus der diskreten FT entstehen. Die Phasenverschiebung der Signale kann einerseits aufgrund von Off-Resonanz-Effekten entstehen. Dieser Effekt beschreibt eine ineffiziente Pulsleistung, um alle Kernspins in der Probe gleichzeitig anzuregen. Bei ^1H -Kernen ist dieser Phasenfehler klein und kann leicht durch eine Phasenkorrektur 0. Ordnung behoben werden. Andererseits ist nach dem RF eine kurze zeitliche Verzögerung notwendig, bevor ein unverzerrter FID aufgenommen werden kann. Diese Zeitverzögerung ist notwendig, um die Spektrometerelektronik zu schützen und führt zu Phasenfehlern, welche mit einer Korrektur 1. Ordnung behoben werden können [119, 121]. Nach der FT enthält das NMR-Spektrum einen Real- und Imaginärteil, in welchen auch beide Fehler auftreten können. Durch die Phasenkorrektur wird der absorptive Teil im realen Spektrum maximiert, wobei der imaginäre Teil den dispersiven Charakter besitzt [119, 121]. Die Phasenkorrektur kann entweder manuell oder automatisch durchgeführt werden. Die manuelle Phasenkorrektur wird für eine genaue Signalintegration empfohlen. Die automatische Phasenkorrektur ist für die Korrektur mehrerer Spektren in metabolischen Studien von entscheidender Bedeutung [119].

Eine gerade Basislinie ermöglicht eine genaue Integration der Signale und eine Erkennung kleiner Signale. Ein Basislinienfehler könnte insbesondere bei metabolomischen Analysen fatal sein, da die NMR-Spektren viele kleine, aber womöglich statistisch relevante Signale beinhalten. Eine verzerrte Basislinie kann das Resultat von beschädigten Datenpunkten sein, die aus einer hohen Signalverstärkung oder einer unvollständigen Erholung der Empfängerelektronik nach dem RF entstehen. Ein verkürzter FID kann auch zu Verzerrungen in der Basislinie führen. Es gibt verschiedene Verfahren zur Basislinienkorrektur basierend auf einer Polynom-, Sinus- oder Exponentialfunktion. Eine häufig durchgeführte Basislinienkorrektur besteht darin, die Grundlinie des Spektrums an ein Polynom anzupassen und dann dieses vom Spektrum zu subtrahieren [119, 121].

Die Integration der Signale ist einer der wichtigsten Schritte in der qNMR. Bei der Integration verschiedener Signale in mehreren Proben sollte diese immer auf die gleiche Weise und mit gleichen Integrationsgrenzen durchgeführt werden. Der Integrationsbereich hat einen großen Einfluss auf die quantitative Genauigkeit. Um 99% der Gesamtfläche abzudecken, sollte der Integralbereich in beide Richtungen mindestens das 20-fache der Signalbreite betragen. Dies kann jedoch aufgrund benachbarter Signale schwierig sein und daher muss ein Kompromiss eingegangen werden [123, 125]. Bei stark überlappenden Signalen ist es möglich, die Integration über Kurvenanpassungsmethoden wie beispielsweise Line-Shape-Fitting Verfahren durchzuführen [123, 125, 126]. Eine Signaltrennung kann unter anderem auch durch die Optimierung des Lösungsmittels, der Probenkonzentration, pH-Wert der Lösung, Ionenkonzentration oder der Messtemperatur erreicht werden [125].

1.2.2.1 Methoden zur Quantifizierung

Die qNMR kann als primäre Verhältnismethode eingestuft werden, von welcher es drei verschiedene Ansätze gibt: die Verhältnismessung, die Gehaltsmessung und die Reinheitsbestimmung. Die Verhältnismessung wird als relative Methode angesehen, währenddessen die beiden anderen Ansätze als absolute Methoden betrachtet werden [124].

Die Verhältnismessung von zwei Verbindungen X und Y wird unter Verwendung des Verhältnisses der Signalintegrale I_x/I_y des molaren Verhältnisses n_x/n_y und unter Berücksichtigung der Anzahl der beteiligten Kerne (N_x und N_y) durchgeführt (Formel 8) [125]. Da die Spektrometerkonstante für beide Verbindungen gleich ist, wurde diese in Formel 8 aufgehoben:

$$\frac{n_x}{n_y} = \frac{I_x}{I_y} \cdot \frac{N_y}{N_x} \quad (8)$$

Bei der absoluten Methode kann der Gehalt c des Analyten X unter Verwendung eines Standards (Std), der Integrale I_x und I_{Std} , der Anzahl der beteiligten Kerne N , der Molmasse der Verbindungen MG , des Gewichts m in der Probe und des bekannten Gehaltes des Standards c_{Std} bestimmt werden (Formel 9) [125].

$$c_x = \frac{I_x}{I_{Std}} \cdot \frac{N_{Std}}{N_x} \cdot \frac{MG_x}{MG_{Std}} \cdot \frac{m_{Std}}{m_x} \cdot c_{Std} \quad (9)$$

Bei dieser Methode muss die gravimetrische Zugabe eines internen Standards (IStd) erfolgen. Die Signalintensität von IStd und Analyt sollte eine vergleichbare Höhe aufweisen. Des Weiteren sollte der IStd in sehr reiner Form verfügbar sein, stabil, chemisch inert, nicht flüchtig und nicht hygroskopisch sein. Interferenzen zwischen Analyt und IStd sollten vermieden werden. Daher ist es vorteilhaft, wenn der IStd ein einfaches Signalspektrum aufweist. Die longitudinale Relaxationszeit T_1 sollte maximal der des Analyten entsprechen, da sonst die Messzeit erhöht werden müsste [123, 125]. Der Einsatz eines externen Standards ist auch möglich. Hierfür kann ein spezielles NMR-Röhrchen mit einer zusätzlichen koaxialen Kapillare verwendet werden, welche mit der Standardlösung gefüllt ist. Dabei sollten Standard und Analyt im gleichen Lösungsmittel gelöst sein und das Volumen des Röhrchens und der Kapillare sollte einheitlich sein und muss kalibriert werden [125]. Die effektive Konzentration des Standards der koaxialen Kapillare sollte kalibriert werden, indem diese in das NMR-Röhrchen gegeben wird, welches eine Lösung eines Primärstandards enthält [123]. Durch dieses Verfahren kann allerdings das Shimming der Probe erschwert sein und die Sensitivität wird aufgrund des geringeren Einheitsvolumen der Analytlösung herabgesetzt [123].

Daneben existiert die Möglichkeit, zwei NMR-Röhrchen zu verwenden, von denen eins mit dem Analyten und eins mit der Standardlösung befüllt ist. Ein weiteres mögliches Verfahren ist die Verwendung eines elektronischen Referenzsignals (Electronic Reference to Access *in vivo* Concentrations, ERETIC) oder die Pulse Length Based Concentration Determination (PULCON) [125].

Die ERETIC Methode dient der Bestimmung absoluter Konzentrationen und nutzt hierfür statt eines Std ein künstlich erzeugtes Referenzsignal. Das hochfrequente Referenzsignal wird während der Erfassungszeit über eine freie Spule eingespeist. Die vollständige Kontrolle über die Intensität und Frequenz des erzeugten Signals ermöglicht die Platzierung in einen freien Spektralbereich, womit eine Überlappung mit Analytsignalen vermieden werden kann. Allerdings muss dieses erzeugte Referenzsignal separat gegen einen Std kalibriert werden und erfordert eine Neuordnung der Spektrometeranschlüsse [125]. Nach der Kalibrierung des Referenzsignals kann die Konzentration des Analyten berechnet werden (Formel 10) [123]:

$$c_x = K \cdot [ERETIC] \cdot \frac{I_x}{I_{ERETIC}} \quad (10)$$

In Formel 10 steht c_x für die Konzentration des Analyten, K beinhaltet die Molmasse des Analyten und die Anzahl der Protonen des zu betrachtenden Analytsignals, $ERETIC$ ist die kalibrierte Konzentration des Referenzsignals und I_x bzw. I_{ERETIC} stellen die Integrale der entsprechenden Signale dar [123].

Ein weiteres Verfahren, welche keine Neuordnung der Spektrometeranschlüsse erfordert, und daher einfach zu implementieren ist, ist die PULCON Methode. Die Konzentrationsbestimmung erfolgt bei dieser Methode durch Korrelation der absoluten Signalintegrale von Analyt und einem separat gemessenen Std. Durch die Messung eines separaten Std entfällt die Suche nach einem geeigneten Std, welcher nicht mit Analytsignalen überlagert. Auch werden mögliche Wechselwirkungen zwischen Std und Analyten umgangen. Mögliche Signalbeeinflussungen, die bei unterschiedlichen physikalischen Eigenschaften der Lösungen auftreten können, werden durch das Prinzip der Reziprozität

gelöst, womit die Korrelation der Signalstärken in zwei unterschiedlichen NMR-Spektren möglich ist [126]. Das Prinzip der Reziprozität besagt, dass die Länge des 90° oder 360° Pulses für eine Probe in einer gegebenen Spule umgekehrt proportional zu der erreichbaren Empfindlichkeit ist [126]. Somit kann unter Berücksichtigung der Pulslänge für Analyt- und Standardlösung die Konzentration c_x des Analyten nach Formel 11 berechnet werden:

$$c_x = f_T \cdot c_{Std} \cdot \frac{I_x \cdot T_x \cdot \theta_{360}^x \cdot n_{Std}}{I_{Std} \cdot T_{Std} \cdot \theta_{360}^{Std} \cdot n_x} \quad (11)$$

In Formel 11 berücksichtigt f_T Schwankungen der Signalintensitäten bei der Verwendung unterschiedlicher Experimente oder Parametereinstellungen für Analyt- und Standardlösung und kann experimentell bestimmt werden. I steht für das Integral, T für die Messtemperatur, n für die Anzahl der Scans und θ_{360} für die Länge des 360° Puls [126]. Werden für die Messung von Std und Analyt die gleiche Messtemperatur und Anzahl der Scans verwendet, so heben sich diese Parameter gegenseitig auf.

Bei der Aufnahme der NMR-Spektren sollte ein Tuning und Matching und eine 90° Puls-längenkalibrierung für jede Probe durchgeführt werden. Zusätzlich sollte für Messungen der gleiche Receiver Gain verwendet werden, da sonst eine unterschiedliche Skalierung berücksichtigt werden muss.

Vorteil dieser Methode ist, dass die Bestimmung der Analyten von der Ionenkonzentration unabhängig ist [126]. Zudem erfordert diese Methode keine spezielle Software oder anspruchsvolle Experimente und ist gleichzeitig sehr robust [126]. Die PULCON-Methode wurde bereits in verschiedenen Matrices wie alkoholfreien Erfrischungsgetränken, Honig und Säften im Rahmen einer Multikomponentenanalyse angewandt [127]. In Kaffee wurde die Auswirkung der Röstung auf Kaffeelipide unter anderem mittels PULCON untersucht [128].

2 Zielsetzung

Mit einer sich veränderten Verbrauchererwartung geht einher, dass der Bedarf an biologisch angebauten Produkten mit bekannter geographischer Herkunft steigt. Für Kaffee fehlen in der Lebensmittelüberwachung bisher zuverlässige, etablierte Analysemethoden, um die Überprüfung der Authentizität bezüglich des geographischen Ursprungs und der Anbaumethoden zu ermöglichen. In der Literatur beschriebene Ansätze zur Bestimmung des geographischen Ursprungs von Kaffee verbinden häufig gaschromatographische Analysemethoden gekoppelt mit Massenspektrometrie, Isotopenverhältnis-Massenspektrometrie, schwingungs-spektroskopische Methoden oder Elementanalytik mit Verfahren der multivariaten Datenanalyse. Vorteil der Protonen-Kernspinresonanz-Spektroskopie ($^1\text{H-NMR}$), welche in der Vergangenheit ebenfalls zur Überprüfung der Authentizität von Lebensmitteln eingesetzt wurde, ist eine einfache Probenvorbereitung bei reproduzierbarer Messung. Zudem stellt die $^1\text{H-NMR}$ strukturelle und quantitative Informationen verschiedenster Kaffeeinhaltsstoffe zur Verfügung. Die Kombination von $^1\text{H-NMR}$ mit Verfahren der multivariaten Datenanalyse ermöglicht grundsätzlich die Untersuchung von Authentizitätsparametern wie die Anbaumethode oder die geographische Herkunft. Anhand eines Datensatzes, der durch die Messung authentischer Proben resultiert, können potentielle Unterschiede erkannt und als Ergebnis ein Klassifizierungsmodell erstellt werden.

Das Ziel dieser Arbeit ist die Entwicklung einer $^1\text{H-NMR}$ -basierten, nicht zielgerichteten Methode zur Überprüfung der Authentizität von Kaffee bezüglich des geographischen Ursprungs und der Anbaumethode. Die Betrachtung des geographischen Ursprungs soll zunächst auf Ebene der Kontinente und anschließend, wenn möglich, auf Ebene der Länder stattfinden. Dazu sollen die $^1\text{H-NMR}$ -Spektren verschiedener authentischer Kaffeeproben mittels Verfahren der multivariaten Datenanalyse ausgewertet werden, um mögliche Unterschiede zwischen bestimmten Herkunftsgebieten oder zwischen biologischem und konventionellem Anbau erkennen zu können. Das Resultat ist die Erstellung eines Klassifizierungsmodells, welches auf unbekannte Kaffeeproben angewendet werden kann. Zur umfassenden Analyse der Inhaltsstoffe sollen Methoden zur Extraktion und Bestimmung der lipophilen und hydrophilen Kaffeeinhaltsstoffe entwickelt werden. Bei der Methodenentwicklung liegt der Fokus zunächst auf einer hohen Reproduzierbarkeit der Probenvorbereitung. Daneben sollen möglichst viele Komponenten erfasst werden. Bei der Untersuchung der Kaffeeproben ist es wichtig, dass diese authentisch sind und dass von jedem Herkunftsgebiet oder von jeder Anbaumethode eine ausreichende Anzahl vorhanden ist. Authentisch bedeutet in diesem Zusammenhang, dass die Angaben der Kaffeeproben über Herkunft und Anbaumethode bekannt sind und mit dem Originalzustand der jeweiligen Kaffeeprobe übereinstimmen.

Darüber hinaus soll im Rahmen dieser Arbeit eine zielgerichtete Methode zur Quantifizierung verschiedener hydrophiler Kaffeeinhaltsstoffe mittels $^1\text{H-NMR}$ -Spektroskopie entwickelt werden. Für die Quantifizierung von Koffein, Trigonellin, N-Methylpyridinium, Hydroxymethylfurfural und verschiedenen organische Säuren sollen die NMR-Spektren der für den nicht zielgerichteten Ansatz entwickelten Methode zur Analyse der hydrophilen Inhaltsstoffe herangezogen werden. Zur Überprüfung der Eignung der entwickelten Quantifizierungsmethode soll zusätzlich eine Methodvalidierung durchgeführt werden. Um die Eignung der entwickelten Methode weiter zu überprüfen, soll ein Vergleich mit einer zu entwickelten Referenzmethode basierend auf Hochleistungsflüssigkeitschromatographie und Diodenarray-Detektion (HPLC-DAD) durchgeführt werden.

3 Material und Methoden

3.1 Methodenentwicklung für den Metabolomics-Ansatz

Der Begriff Metabolomics beschreibt die Erfassung möglichst vieler niedermolekularer Stoffwechselverbindungen in einem biologischen System. Das Metabolom umfasst die Gesamtheit dieser niedermolekularen Substanzen, bestehend aus primären und sekundären Stoffwechselverbindungen. Es wird durch Faktoren wie Umwelt, Alter, Art, Krankheiten oder Exposition gegenüber synthetischen Stoffen beeinflusst und bietet somit die Möglichkeit zur Untersuchung unterschiedlicher Fragestellungen wie beispielsweise die Bestimmung des geographischen Ursprungs oder den Anbaubedingungen [129–132].

Allgemein können Metabolomics-Analysen in zielgerichtet oder nicht zielgerichtet unterteilt werden. Bei zielgerichteten Analysen erfolgt die Identifizierung und Quantifizierung bestimmter Metaboliten. Diese Art der Analyse ist wichtig, um das Verhalten einer bestimmten Verbindung oder einer Klasse von Verbindungen in einer Probe unter bestimmten Bedingungen zu beurteilen. Dagegen werden bei der nicht zielgerichteten Analyse so viele verschiedene Metaboliten wie möglich erfasst, ohne diese zu identifizieren und zu quantifizieren. Dadurch kann ein bestimmtes Muster erkannt werden, was auch als chemischer Fingerabdruck bezeichnet wird. Diese Muster können anschließend dazu verwendet werden, um beispielsweise zwei Klassen von Proben zu unterscheiden [131, 132]. Die Unterscheidung der Klassen wird durch die Verwendung von Verfahren der statistischen multivariaten Datenanalyse erreicht [133]. Weiter können mit einer großen Anzahl von Proben eine Datenbank und Klassifizierungsmodelle erstellt werden, die es neben der Unterscheidung von Klassen ermöglicht, unbekannte Proben diesen Klassen zuzuordnen. Darüber hinaus können mit dem nicht zielgerichteten Ansatz neue Metaboliten identifiziert werden, die ursächlich für die Unterschiede der Klassen sind [132].

Der Metabolomics-Ansatz besteht aus verschiedenen Teilen, beginnend mit der Probenvorbereitung einschließlich der Metabolitenextraktion, gefolgt von der analytischen Erfassung der Metaboliten und der Datenauswertung mit Hilfe geeigneter statistischer multivariater Verfahren (Abbildung 3-1) [131, 132]. Wichtig ist, dass alle Schritte des Metabolomics-Ansatzes standardisiert sind. Nur so ist ein direkter Vergleich der zu unterschiedlichen Zeiten generierten Daten gewährleistet [134]. Unter Berücksichtigung biologischer Variationen muss eine ausreichend große Anzahl authentischer Proben untersucht werden, um gesicherte Informationen zu erhalten [133]. Dabei sollten so viele Metadaten der Probe wie möglich vorliegen.

Das Ziel der Extraktion für einen nicht zielgerichteten Ansatz ist es, so viele Metaboliten wie möglich zu erfassen. Aufgrund strukturell unterschiedlicher Metaboliten wird die Verwendung mehrerer Extraktionen mit verschiedenen Extraktionsmitteln und separater Metaboliten-Analyse empfohlen [131, 134]. In dieser Arbeit wurden daher zwei getrennte Extraktionsmethoden zur Untersuchung der lipophilen und hydrophilen Metaboliten entwickelt.

Für die analytische Erfassung eines Metaboloms kann die $^1\text{H-NMR}$ herangezogen werden [132, 135]. Die $^1\text{H-NMR}$ bietet eine simultane Analyse verschiedener Sekundärmetaboliten neben primären Stoffwechselverbindungen in einer kurzen Zeit und ist daher für den zielgerichteten und nicht zielgerichteten Ansatz sehr gut geeignet [133].

Nach der Erfassung des Metaboloms mit der $^1\text{H-NMR}$ gibt es in Abhängigkeit der Art der Metabolomics-Analyse unterschiedliche Vorgehensweisen. Bei der zielgerichteten Analyse werden die NMR-Spektren quantitativ ausgewertet und bestimmte Metaboliten mit entsprechenden Quantifizierungsmethoden (vgl. Kapitel 1.2.2.1) erfasst. In der vorliegenden Arbeit wurde die PULCON-Methode zur Quantifizierung von Trigonellin, Koffein, NMP, HMF, Essigsäure, Ameisensäure, Milchsäure und den CQA im hydrophilen Extrakt angewendet.

Um die NMR-Spektren der Proben im Rahmen einer nicht zielgerichteten Analyse miteinander vergleichen zu können, sind Datenvorbehandlungen nötig [136]. Für die Analyse von einem großen Datensatz, welcher bei der Erfassung des Metaboloms von mehreren Hundert Proben entsteht, sind Verfahren der statistischen uni- und multivariaten Datenanalyse erforderlich. Diese Verfahren beinhalten Methoden wie die Hauptkomponentenanalyse (Principal Component Analysis, PCA) und die Partielle Kleinste Quadrate-Diskriminanzanalyse (Partial Least Square Discriminant Analysis, PLS-DA) [133].

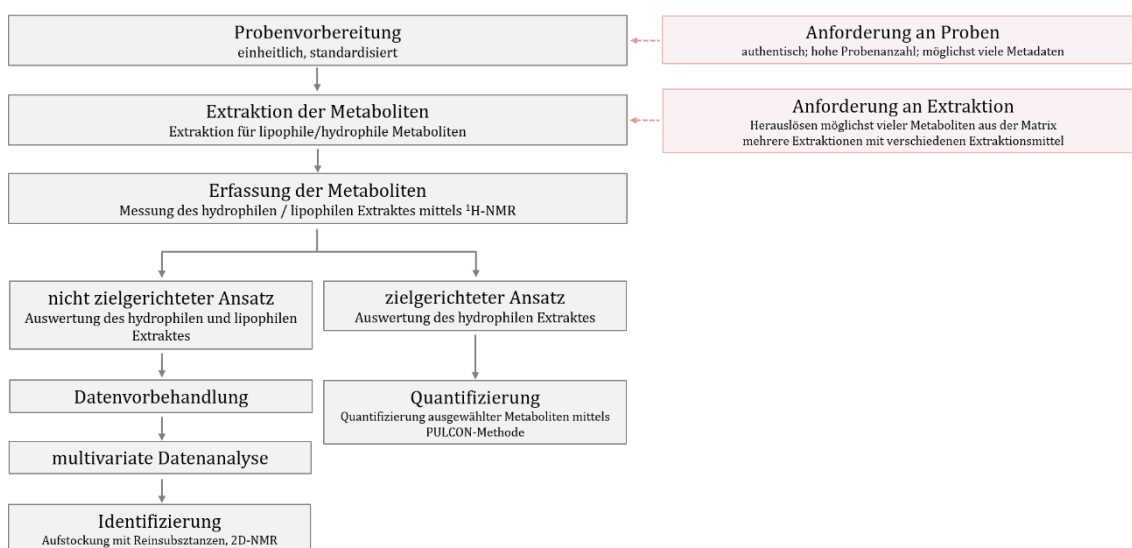


Abbildung 3-1: Schema der verschiedenen Schritte des im Rahmen dieser Arbeit angewendeten Metabolomics-Ansatzes

Der nicht zielgerichtete Metabolomics Ansatz wurde bisher bereits für die Untersuchung der Authentizität von geröstetem Kaffee angewandt. Dabei wurde die Abgrenzung von geröstetem Kaffee aus verschiedenen Kontinenten, Ländern oder Regionen untersucht. Zudem wurde eine Abgrenzung von Kaffee nach biologischem und konventionellem Anbau getestet. Für diese Untersuchungen wurden Methoden wie die $^1\text{H-NMR}$, Gaschromatographie gekoppelt mit Massenspektrometrie (GC-MS), Gaschromatographie gekoppelt mit der Isotopenverhältnis-Massenspektroskopie (GC-C-IRMS), Hochleistungsflüssigkeitschromatographie gekoppelt mit Massenspektrometrie (HPLC-MS) und die Elementanalyse verknüpft mit multivariaten Datenanalysen verwendet [137–150].

3.2 Untersuchungsmaterialien und Probenvorbereitung

In der vorliegenden Arbeit wurden ausschließlich Arabica-Kaffeeproben untersucht. Die verwendeten Arabica-Kaffeeproben wurden von verschiedenen Institutionen zur Verfügung gestellt. Ein Teil der Kaffeeproben wurde von dem Chemischen und Veterinäruntersuchungsamt Karlsruhe im Zeitraum von 2019 bis 2022 erhoben. Ein weiterer Teil der Kaffeeproben war von der Firma Coffee Consulate mit Sitz in Mannheim. Die

Kaffeeproben von der Firma Coffee Consulate kamen überwiegend aus den Ländern Brasilien, El Salvador, Indien und Mexiko. Ein weiterer Teil des Probenmaterials wurde von verschiedenen Röstereien zur Verfügung gestellt. Daneben wurde ein kleiner Teil aus lokalen Röstereien und dem Einzelhandel erworben. Insgesamt wurden für den nicht zielgerichteten Ansatz 603 Arabica-Kaffeeproben untersucht, welche mittels Trommelröster geröstet wurden und wovon 101 biologisch angebaut wurden. Davon kamen 89 Kaffeeproben aus Afrika, 66 Kaffeeproben aus Asien, 176 Kaffeeproben aus Nordamerika und 190 Kaffeeproben aus Südamerika. Von den 190 Kaffeeproben aus Südamerika stammten 116 aus Brasilien und 35 aus Kolumbien. 51 Proben aus Äthiopien trugen zu der Probenanzahl aus Afrika bei. Weitere 82 Proben wurden untersucht, welche biologisch angebaut wurden, jedoch keine Herkunftsangabe besaßen. Zusätzlich wurden neben den 603 trommelgeröstete Proben 85 Rohkaffeebohnen mit dem Heißlufttröster der Firma Ikawa geröstet. Bei dem Rösten der Rohkaffeeproben mittels Heißlufttröster wurde darauf geachtet, dass die Kaffeebohnen am Ende des Röstprozesses einen mittleren Röstgrad aufwiesen. Für den zielgerichteten Ansatz wurden weitere 51 Kaffeeproben untersucht, welche verschiedene Defekte aufwiesen.

Zu den Proben wurden so viele Metadaten wie möglich erfasst, darunter der genaue geographische Ursprung, die Anbaumethode und das verwendete Röstverfahren (Tabelle 9-1). Die Lagerung der Proben erfolgte in einer luftdichten Verpackung mit einem speziellen Ventil, durch welches das nach der Röstung entstehenden Gas weiter austreten konnte. Die Proben wurden dunkel und kühl bei Temperaturen von 5-7 °C gelagert. Es wurden gemahlene Kaffeebohnen als auch ganze Bohnen untersucht, wobei die Kaffeebohnen vor der Extraktion zu einer Korngröße von 0,3 mm gemahlen wurden.

Bedingt durch die Untersuchung bereits gerösteter Kaffeeproben lagen diese zum Zeitpunkt der Erfassung des Metaboloms mit unterschiedlichen Röstgraden vor. Diese Variation wurde berücksichtigt, indem der Farbwert jeder Kaffeeprobe bestimmt wurde. Diese Messung wurde mit dem Farbmesser „Tonino“ der Firma Artisan durchgeführt und beruht auf dem Prinzip der photometrischen Reflektion [151]. Ein Farbwert von 60-80 wurden dabei als dunkle Röstung und Werte von 100-130 dagegen als helle Röstung eingestuft.

Zudem können variierende Wasserkonzentrationen in den Proben zu einer ungenauen zielgerichteten und nicht zielgerichteten Analyse führen [134]. Da in dieser Arbeit handelsüblich gerösteter Kaffee untersucht wurde, wurde der Feuchtegehalt jeder Probe vor der Messung thermogravimetrisch bestimmt.

Die Proben mit den entsprechenden Metadaten sind im Anhang in Tabelle 9-1 und Tabelle 9-2 aufgelistet. Die Schritte der Probenvorbereitung sind in Arbeitsvorschrift 8.1 zusammengefasst.

3.3 Entwicklung einer nicht zielgerichteten Methode für die Untersuchung der lipophilen Metaboliten

Die Entwicklung der nicht zielgerichteten Methode für die Untersuchung der lipophilen Metaboliten bestand darin, zunächst eine optimale Extraktionsmethode sowie im Anschluss optimale Akquisitions- und Prozessierungsparameter für die ¹H-NMR-Messung herauszuarbeiten.

3.3.1 Extraktionsmethode

Gerösteter Kaffee besteht aus einer komplexen Mischung aus Primär- und Sekundärmetaboliten (Kapitel 1.1.7). Pflanzen enthalten eine Vielzahl von Verbindungen mit unterschiedlichen Polaritäten, wobei zwischen drei verschiedenen Polaritäten unterschieden werden kann. Dies sind einerseits die unpolaren Komponenten, welche sich in der Zellwand befinden und Lipide wie TAG, Terpene und freie FA, sowie deren Derivate [135, 152]. Die mittel polaren Verbindungen sind hauptsächlich Sekundärmetaboliten, wie beispielsweise Koffein oder Trigonellin und die wasserlöslichen polaren Komponenten sind primäre Stoffwechselprodukte wie Aminosäuren und Zucker [135]. Um möglichst viele dieser Komponenten zu extrahieren, ist die Wahl des Extraktionsmittels entscheidend. Unter anderem ist die Polarität des Lösungsmittels ein entscheidender Faktor. Die Polarität eines unpolaren organischen Lösungsmittels kann beispielsweise mit Methanol erhöht werden, was dazu führt, dass auch mittel polare Verbindungen extrahiert werden [135]. Für die Untersuchung von Pflanzenmaterial wird die Extraktion mit einem Gemisch aus Chloroform, Methanol und Wasser empfohlen, da dadurch sowohl unpolare, als auch mittel polare Verbindungen extrahiert werden können [131]. Hierbei kann das Mischungsverhältnis der Lösungsmittel variiert werden, um die optimale Polarität des Extraktionsmittels für die entsprechende Matrix zu erreichen. Für die Extraktion von tierischen Lipiden finden hauptsächlich die Methode nach Folch et al. und die nach Blight und Dyer Anwendung, welche unterschiedliche Verhältnisse von Chloroform, Methanol und Wasser einsetzen [153, 154]. Als umweltfreundlichere Alternative zur Methode von Folch et al. wurde in der Methode nach Lin et al. zur Extraktion von tierischen Lipiden eine Mischung aus Ethylacetat und Ethanol in verschiedenen Mischungsverhältnissen angewendet [155].

Die Extraktion nach Folch et al. erfolgt zunächst mit einer Mischung aus Chloroform und Methanol (2:1, v/v), wobei für ein Teil Probe 20 Teile der Extraktionslösung verwendet wird. Durch die nachfolgende Zugabe von Wasser oder einer Salzlösung wird eine Phasentrennung erreicht. Die untere Phase, welche die Lipide enthält, wird anschließend mehrmals mit einem Gemisch aus Chloroform, Methanol und Wasser (3:38:41, v/v/v) gewaschen. Die Zugabe einer Salzlösung hat den Vorteil, dass dadurch die Dissoziation azider Lipide verringert wird und diese dadurch bei der Phasentrennung vermehrt in der unteren Phase vorliegen [154]. Bei der Methode nach Blight und Dyer muss ein Mischungsverhältnis von Chloroform, Methanol und Wasser von 1:2:0,8 (v/v/v) vorliegen, sodass ein einphasiges Extraktionsmittel entsteht. Nach der Extraktion der Probe erfolgt die Zugabe von Chloroform und Wasser, sodass ein Verhältnis von 2:2:1,8 (Chloroform, Methanol, Wasser, v/v/v) erhalten wird. Dies führt zu einer Trennung in eine Chloroform-Phase, welche die Lipide enthält, und in eine Methanol-Wasser-Phase mit polaren Metaboliten [153]. Ein Vorteil der Blight und Dyer Methode gegenüber der Extraktion nach Folch et al. ist ein geringerer Lösungsmittelverbrauch, da für ein Teil Probe nur drei Teile Extraktionsmittel benötigt werden. Neben diesen Standardmethoden existieren modifizierte Methoden, bei welchen das Verhältnis von Probe zu Extraktionsmittel herabgesetzt wurde [156]. Die Methode von Folch et al. sowie die von Blight und Dyer und deren modifizierte Varianten wurden auch zur Extraktion von pflanzlichen Lipiden verwendet [157–160].

Neben der Polarität des Lösungsmittels sollte auch die Toxizität und dessen Einfluss auf die Umwelt berücksichtigt werden [135]. Daher wurde in der Methode nach Lin et al. zur Extraktion von tierischen Lipiden eine Mischung aus Ethylacetat und Ethanol in verschiedenen Mischungsverhältnissen angewendet [155]. Es konnte gezeigt werden, dass bei einer Verwendung von Ethylacetat und Ethanol im Verhältnis 1:1 oder 2:1 (v/v) vergleichbare

Ergebnisse zur Methode nach Folch et al. bezüglich der Extraktionserträge und der Lipidzusammensetzung erzielt wurden [155].

Um eine gesteigerte Ausbeute der Metaboliten zu erzielen, kann bei einem ausgewählten Lösungsmittel die Extraktionszeit oder -temperatur erhöht werden. Jedoch steigt bei der Extraktion mit erhöhter Temperatur auch das Risiko der Artefaktbildung [134]. Bei der Extraktion von Pflanzenmetaboliten im Rahmen eines Metabolomics-Ansatzes wird daher eine Extraktion bei Raumtemperatur empfohlen [134].

Im Rahmen der vorliegenden Arbeit wurden verschiedene Methoden zur Extraktion von Kaffeelipiden verglichen. Dazu wurden die Methoden nach Blight und Dyer, nach Folch et al. und nach Lin et al. miteinander verglichen. Zusätzlich wurde die Methode nach Ackermann et al. getestet, in welcher die Methode nach Folch et al. bezüglich des Probe-zu-Lösungsmittelverhältnisses optimiert wurde. Daneben wurde die Extraktion mit Chloroform, Chloroform und Methanol in einem Mischungsverhältnis von 1:1 (v/v) und die Extraktion in dem toxikologisch weniger bedenklichen Aceton überprüft. Dabei war das Ziel, so viele Metaboliten wie möglich mit einer hohen Reproduzierbarkeit aus der Kaffeematrix herauszulösen. Die verschiedenen Extraktionsmethoden sind in Arbeitsvorschrift 8.2.1 beschrieben. Als optimale Methode für die Untersuchung der lipophilen Metaboliten in der Matrix Kaffee hat sich die Extraktion mit einer Mischung aus Chloroform und Methanol in einem Verhältnis von 1:1 (v/v) herausgestellt. Nach der Wahl des Lösungsmittels wurden weitere Parameter wie die Extraktionszeit und die Verwendung von Glasperlen zur besseren Durchmischung während der Extraktion untersucht (Arbeitsvorschrift 8.2.2)

Vor der Extraktion erfolgte die Einwaage auf ein exaktes Gewicht der Kaffeeproben von 575,0 mg, wobei der unterschiedliche Feuchtegehalt der Kaffeeproben in der Datenvorbehandlung der $^1\text{H-NMR}$ -Spektren berücksichtigt wurde. Die Kaffeeproben wurden mit einer Mischung aus Chloroform und Methanol in einem Verhältnis von 1:1 (v/v) für eine Stunde bei Raumtemperatur mit Zusatz von Glasperlen extrahiert (Arbeitsvorschrift 8.2.3).

3.3.2 Vorbereitung der Messlösung für die $^1\text{H-NMR}$ -spektroskopische Messung

Bei der $^1\text{H-NMR}$ hat die Wahl des Messlösungsmittels einen großen Einfluss auf die chemische Verschiebung und die Auflösung der Metabolitensignale. Die Überlappung von Signalen kann durch ein geeignetes Lösungsmittel oder Lösungsmittelgemisch für die Messung minimiert werden [125].

Bevorzugt sollte der Extrakt direkt analysiert werden, um ein Verdampfen und erneutes Lösen der Metaboliten zu vermeiden. Bei der Verwendung der NMR setzt dies jedoch eine Extraktion mit deuteriertem Lösungsmittel oder die Anwendung von Pulssequenzen zur Lösungsmittelunterdrückung voraus. Hier ist die Auswahl von Lösungsmitteln eingeschränkt, da nicht alle Lösungsmittel in deuterierter Form oder nur zu deutlich höheren Preisen erhältlich sind. Zusätzlich sollten die Auswirkungen auf die Auflösung der NMR-Signale berücksichtigt werden [134]. Das häufig zur Extraktion von Lipiden eingesetzte deuterierte Chloroform (CDCl_3) extrahiert nur unpolare Lipidbestandteile, wodurch im Vergleich zu einem Lösungsmittel höherer Polarität wie die Mischung mit Methanol weniger Informationen im NMR-Spektrum vorhanden sind. Daher ist die Anwendung von reinem CDCl_3 im Rahmen einer nicht zielgerichteten Analyse nicht geeignet [134]. Ein weiterer Nachteil bei der Verwendung von CHCl_3 oder CDCl_3 ist, dass es bei diesen Lösungsmitteln zum Abbau und schlussendlich zur Bildung von Phosgen und Salzsäure kommen kann. Die

Abbauprodukte können unter anderem mit stickstoffhaltigen Verbindungen wie Alkaloiden reagieren [135, 161]. Arana et al. führten eine Extraktion der Kaffeemetaboliten mit Methanol durch und gaben einen kleinen Anteil an deuteriertem Methanol (MeOH- d_4) in die Messlösung [138]. Durch diese Vorgehensweise musste das Extraktionsmittel nicht verdampft und die Metaboliten nicht erneut gelöst werden. Allerdings müssen während der ^1H -NMR-Messung Techniken zur Lösungsmittelunterdrückung angewandt werden [138].

Skiera et al. verwendeten bei der Untersuchung von freien FA in pflanzlichen Ölen ein Lösungsmittelgemisch, bestehend aus CDCl_3 und deuteriertem Dimethylsulfoxid (DMSO- d_6) in einem Verhältnis von 5:1 (v/v). Dabei verlangsamt der Zusatz von DMSO- d_6 den Protonentransfer von OOH-Gruppen oxidierter Speiseöle, was zu schmalen Signalen führt. Wechselwirkungen der OOH-Gruppen mit dem DMSO- d_6 führen zu einer Tieffeldverschiebung [162].

Wie in Kapitel 3.3.1 beschrieben, wurde in der vorliegenden Arbeit eine Mischung aus Methanol und Chloroform im Verhältnis 1:1 (v/v) zur Extraktion der Kaffeemetaboliten verwendet. Eine Extraktion mit CDCl_3 und MeOH- d_4 in diesem Mischungsverhältnis würde bei der Extraktion einer großen Probenanzahl zu stark erhöhten Kosten führen. Daher wurde in der vorliegenden Arbeit ein Aliquot des Extraktes abgenommen, die Lösungsmittel verdampft, der Rückstand ausgewogen und die Metaboliten in einem deuteriertem Lösungsmittel wieder gelöst. Dabei wurden verschiedene deuterierte Lösungsmittelsysteme beim erneuten Lösen der Metaboliten untersucht. Die Lösungsmittelsysteme bestehend aus unterschiedlichen Anteilen an CDCl_3 und MeOH- d_4 sowie das Lösungsmittelgemisch CDCl_3 und DMSO- d_6 5:1 (v/v) wurden hinsichtlich der chemischen Verschiebung und Auflösung der Metabolitensignale bewertet (Arbeitsvorschrift 8.2.4). Die Untersuchungen zeigten die besten Ergebnisse bei der Verwendung des Lösungsmittelgemisches bestehend aus CDCl_3 und MeOH- d_4 im Verhältnis 1:1 (v/v).

Durch die Verwendung von CDCl_3 besteht die Möglichkeit, dass sich Abbauprodukte wie Phosgen oder Salzsäure in dem Lösungsmittelgemisch befinden [161]. Um zu überprüfen, ob ein möglicher Abbau von CDCl_3 eine Auswirkung auf die Erfassung der Kaffeemetaboliten hat, wurden die Abbauprodukte zunächst aus dem CDCl_3 entfernt. Dazu wurde das CDCl_3 mit einer Natriumcarbonat-Lösung ausgeschüttelt und anschließend auf getrocknetem Natriumcarbonat gelagert. Das dadurch entstandene behandelte CDCl_3 sowie das unbehandelte CDCl_3 wurden jeweils mit MeOH- d_4 vermischt und beide Lösungsmittelsysteme miteinander verglichen (Arbeitsvorschrift 8.2.4). Es wurde festgestellt, dass der Abbau von CDCl_3 einen Einfluss auf die Erfassung der Kaffeemetaboliten hat. Aufgrund dessen wurden die Abbauprodukte vor der Verwendung des CDCl_3 im weiteren Verlauf aus diesem entfernt und das daraus resultierende CDCl_3 als „entschärftes CDCl_3 “ bezeichnet (Arbeitsvorschrift 8.2.4).

3.3.3 Optimierung der Messparameter für die ^1H -NMR-spektroskopische Messung

Die Aufnahme präziser und reproduzierbarer Spektren ist bei der qNMR sowie für nicht zielgerichtete Ansätze Voraussetzung [119]. Wie in Kapitel 1.2.2 beschrieben kann ein Pulswinkel von 90° oder kleiner für die Anregung der Kerne verwendet werden. Vorteil eines kleineren Pulswinkels ist eine geringere Repetitionszeit, wodurch eine geringere Messzeit resultiert. Allerdings ist durch eine geringe Auslenkung der Quermagnetisierung auch eine

geringere Intensität der Signale zu erwarten. Zur Optimierung des Pulses wurde ein Pulswinkel von 90° und ein Pulswinkel von 30° getestet (Arbeitsvorschrift 8.2.4).

Zur Bestimmung der nötigen Repetitionszeit wurden in einem Inversion Recovery Experiment die T_1 -Zeiten der in den Kaffeeproben in höheren Konzentrationen vorkommenden Metaboliten ermittelt. Das Inversion Recovery Experiment ist eine zwei-Puls-Sequenz, in der die makroskopische Magnetisierung durch einen 180° Puls invertiert wird. Die makroskopische Magnetisierung wird somit in Richtung der $-z$ -Achse ausgelenkt und kehrt über die x - y -Ebene zurück in das Gleichgewicht entlang der z -Achse. Da die Magnetisierung entlang der z -Achse nicht beobachtet werden kann, wird die Relaxation untersucht, indem die makroskopische Magnetisierung nach einer geeigneten Zeitspanne τ nach dem 180° Puls mit einem 90° Puls wieder in die x - y -Ebene gebracht wird. Bei τ gleich Null richtet sich die makroskopische Magnetisierung in Richtung der $-y$ -Achse aus und erzeugt ein negatives Signal. Diese Impulsfolge wird mit steigenden Werten von τ wiederholt. Ist τ ausreichend lang, tritt zwischen den beiden Pulsen eine vollständige Relaxation auf, sodass sich die makroskopische Magnetisierung nach dem 90° Puls auf der $+y$ -Achse befindet und ein maximal positives Signal aufgezeichnet wird. Durch die halblogarithmische Auftragung der Signalintensität oder Signalfläche gegen die Zeitspanne zwischen den Impulsen τ kann die T_1 -Zeit ermittelt werden [121].

Weiter wurden verschiedene Messtemperaturen im Bereich von 280-300 K untersucht, die Empfängerverstärkung auf die Probe angepasst und die Stabilität der Messlösung überprüft (Arbeitsvorschrift 8.2.4).

Für die Erfassung der lipophilen Metaboliten in dem nicht zielgerichteten Ansatz wurde ein Pulswinkel von 90° und eine Messtemperatur von 290 K gewählt. Um ein Abschneiden des FID zu verhindern, wurde eine Akquisitionszeit von 8 s gewählt. Durch die Erkenntnisse der T_1 -Zeit-Bestimmung wurde eine Relaxationsverzögerung von 15 s verwendet. Die Messparameter für die Erfassung der lipophilen Metaboliten ist in Arbeitsvorschrift 8.2.5 zu finden. Die Überprüfung der Stabilität der Messlösung ergab, dass diese 48 h stabil ist. Da ein NMR-Spektrum sehr reproduzierbar und eine Einfachbestimmung daher ausreichend ist [133], wurden die Kaffeeproben jeweils einfach vermessen.

3.4 Entwicklung einer nicht zielgerichteten Methode für die Untersuchung der hydrophilen Metaboliten

3.4.1 Extraktionsmethode

Neben der Extraktion mit unpolaren organischen Lösungsmitteln wird bei der nicht zielgerichteten Untersuchung von Kaffee in der Literatur auch eine Extraktion, bestehend aus einer Mischung aus Methanol und Wasser angewandt [131]. Durch diese Extraktion können hydrophile Kaffeemetaboliten wie organische Säuren, Aminosäuren, Kohlenhydrate und Alkaloide gewonnen werden. Für die Erfassung der Metaboliten mittels $^1\text{H-NMR}$ wird häufig eine Extraktion in deuteriertem Wasser (D_2O) angewandt [131, 144, 145]. Da bei der Extraktion einer großen Probenmenge mit D_2O ein erheblicher Kostenfaktor berücksichtigt werden sollte, kann die Extraktion in H_2O durchgeführt und anschließend ein kleines Volumen D_2O zur Messlösung hinzugefügt werden [139, 146]. Diese Vorgehensweise erfordert allerdings die Anwendung von Lösungsmittelunterdrückungstechniken bei der Erfassung der Metaboliten mittels $^1\text{H-NMR}$.

Bei der Methodenentwicklung für die Extraktion der hydrophilen Metaboliten wurde der Fokus auf das Verhältnis von Probe zu Extraktionsmittel und auf die Extraktionszeit gelegt. Da bei einer höheren Extraktionstemperatur die Gefahr der Artefaktbildung besteht [134], wurde eine Extraktion bei Raumtemperatur angewandt.

Um die $^1\text{H-NMR}$ -Spektren nach der Erfassung der Metaboliten multivariat auswerten zu können, sollten variierende chemische Verschiebungen eines Metabolitensignals, sogenannte Signal-Shifts, vermieden werden. Diese Signal-Shifts können bedingt durch einen unterschiedlichen pH-Wert oder auch aufgrund unterschiedlicher Konzentrationen einzelner Metaboliten entstehen. Hervorzuheben ist hier der Metabolit Koffein, welcher in wässriger Lösung unter Bildung von Dimeren und höheren Aggregaten selbst assoziiert [163]. Zudem liegt Koffein zusammen mit CGA in einem 1:1 π -Komplex vor [99, 100]. Die Selbstassoziation sowie die Bildung des Komplexes mit CGA wird von der Konzentration beeinflusst und kann zu Shifts der Koffein- und CGA-Signale führen [163, 164]. Das Ausmaß dieser Shifts wurde untersucht, indem verschiedene Mengenverhältnisse von Probe zu Extraktionsmittel überprüft wurden (Arbeitsvorschrift 8.3.1). Um ein quantitatives und reproduzierbares Herauslösen der Metaboliten sicherzustellen, wurden Extraktionen bei unterschiedlichen Extraktionszeiten durchgeführt und zudem geprüft, ob die Verwendung von Glasperlen während der Extraktion eine besseres Herauslösen der Metaboliten ermöglicht (Arbeitsvorschrift 8.3.1).

Da der Abbau der Polysaccharide während der Röstung von der Varietät beeinflusst wird (Kapitel 1.1.6.2), besteht die Möglichkeit, dass diese bei der Extraktion nicht reproduzierbar aus den verschiedenen Kaffeeproben herausgelöst werden. Zusätzlich kann das Vorhandensein von Polysacchariden bei der $^1\text{H-NMR}$ -Messung zu breiteren Signalen führen [133]. Daher wurde nach einer Vorgehensweise gesucht, um die Polysaccharide nach der Extraktion wieder aus dem Extrakt zu entfernen. Für die Entfernung der Polysaccharide aus einer Kaffeematrix wurde bisher keine analytische Methode beschrieben. Bekedam et al. verwendeten für die Isolierung von Melanoidinen eine Diafiltration mit einer Membran, die eine Porengröße von 3 kDa besaß [48]. Somit wäre die Ultrafiltration (UF) mit einer Membran, welche eine Porengröße von 3 kDa aufweist, eine Möglichkeit für die Entfernung der Polysaccharide, welche in Melanoidine eingebunden vorliegen. Die UF ist ein einfaches Probenvorbereitungsverfahren, wobei die Filtration durch Anwendung von Druck oder Zentrifugation erfolgt [165]. Für Metabolomics-Ansätze werden hierfür Membranen mit einer Porengröße von 3-30 kDa verwendet. Bei der Verwendung einer 3 kDa Membran können niedermolekulare Verbindungen mit einem Molekulargewicht < 3 kDa von Makromolekülen wie beispielsweise Polysacchariden abgetrennt werden [165]. Um die Entfernung aller Kaffeepolysaccharide sicherzustellen, wurden in dieser Arbeit Membranen mit verschiedenen Porengrößen von 1, 2, 3 und 5 kDa getestet (Arbeitsvorschrift 8.3.2). Vor der Verwendung der Membranen sollten diese durch Zentrifugation mit Wasser vorgereinigt werden, da in vielen Filtern Glycerin enthalten ist. Ohne ein Vorspülen dieser Filter mit Wasser sind unerwünschte Glycerinsignale im $^1\text{H-NMR}$ -Spektrum vorhanden [166, 167].

Für die Extraktion der hydrophilen Metaboliten wurde H_2O als Lösungsmittel ausgewählt. Die Extraktion erfolgte durch Einwaage von 200,0 mg der Probe, Zusatz von 8,0 mL H_2O und Schütteln für 20 min bei Raumtemperatur. Für die Überprüfung der Eignung der UF wurden insgesamt 75 Kaffeeproben verwendet. Dazu wurde der Extrakt der Proben zunächst membranfiltriert. Ein Teil dieses Extraktes wurde für die Messlösung vorbereitet. Ein weiterer Teil des membranfiltrierten Extraktes wurde mittels Zentrifugenfilter mit einer Porengröße von 1 und 2 kDa zentrifugiert und anschließend für die Messlösung vorbereitet

(Arbeitsvorschrift 8.3.3). Die resultierenden ^1H -NMR-Spektren wurden anschließend multivariat ausgewertet. Bei allen darüber hinaus untersuchten Kaffeeproben wurde der Extrakt membranfiltriert und nicht weiter einer UF unterzogen.

3.4.2 Vorbereitung der Messlösung für die ^1H -NMR-spektroskopische Messung

Bei der Verwendung von wässrigen Lösungsmitteln ist der pH-Wert ein kritischer Faktor, da sich bei niedrigen oder hohen pH-Werten Artefakte bilden können [134]. Für die nicht zielgerichtete Analyse werden die ^1H -NMR-Spektren nach der Erfassung der Metaboliten in der multivariaten Datenanalyse miteinander verglichen. Ein variierender pH-Wert der Proben kann hierbei zu Shifts einzelner NMR-Signale im Spektrum führen. Dies erschwert die Vergleichbarkeit der ^1H -NMR-Spektren in der nachfolgenden multivariaten Datenanalyse. Insbesondere bei Carbonsäuren oder Aminosäuren kann eine pH-Wert abhängige Protonierung oder Deprotonierung zu Shifts der entsprechenden ^1H -NMR-Signale führen [125, 168]. Diese Shifts können durch die Verwendung von Puffern minimiert werden, wobei anorganische Puffer entsprechend dem physiologischen pH-Wert der Probe oder einem pH-Wert von 6,0 empfohlen werden [133, 134, 167]. Die Verwendung von protonierten, organischen Puffern sollte vermieden werden, da die dadurch zusätzlich erzeugten Signale im NMR-Spektrum Metabolitensignale überlagern können [167]. Zusätzlich kann die Ionenstärke der Probe auch zu Veränderungen der chemischen Verschiebung der ^1H -NMR-Signale führen. Durch eine hohe Salzkonzentration kann es zu Schwierigkeiten bei der Abstimmung und Anpassung der Spektrometerfrequenz kommen [123].

Bei bisherigen nicht zielgerichteten Untersuchungen von Kaffee, bei welchen die Extraktion in H_2O oder D_2O durchgeführt wurde, wurde auf eine Verwendung von Puffern verzichtet [150, 169, 170]. Um den auftretenden Signal-Shifts, insbesondere von Signalen der in Kaffee vorkommenden organischen Säuren, entgegenzuwirken und vergleichbare ^1H -NMR-Spektren aller Proben zu erhalten, wurde im Rahmen der vorliegenden Arbeit ein Puffer für die Messlösung verwendet. Es wurden Natriumphosphatpuffer mit verschiedenen pH-Werten und Ionenstärken getestet (Arbeitsvorschrift 8.3.4). Durch die Anwendung eines Natriumphosphatpuffers mit einer Konzentration von 1 mol/L und einem pH-Wert von 6,0 konnten hinsichtlich der chemischen Verschiebung der organischen Säuren vergleichbare ^1H -NMR-Spektren erzielt werden. Die Vorbereitung der Messlösung für die Erfassung der hydrophilen Metaboliten mittels ^1H -NMR ist in Arbeitsvorschrift 8.3.5 festgehalten.

3.4.3 Optimierung der Messparameter für die ^1H -NMR-spektroskopische Messung

Bei der Extraktion der Metaboliten mit H_2O entsteht ein starker Konzentrationsunterschied zwischen dem Wasser und den extrahierten Metaboliten. Da die Signalintensität proportional zur Konzentration ist, wird bei Messung des Extraktes mittels ^1H -NMR ein im Vergleich zu den Metabolitsignalen sehr großes Wassersignal erzeugt. Mit einem intensiven Wassersignal wird eine schnelle Sättigung des Empfängers erreicht, sodass die Empfängerverstärkung herabgesetzt werden muss, was sich besonders nachteilig auf Minorkomponenten auswirkt. Daraus resultiert eine unzureichende Aufnahme der Metabolitensignale. Wird die Empfängerverstärkung nicht herabgesetzt, entstehen bei einer Sättigung des Empfängers störende Spektrallinien und nichtlineare Basislinienoszillationen. Zudem besteht die Möglichkeit, dass Signale von Metaboliten von dem breiten

Wassersignal überlagert werden und es weiter zu Grundlinienverzerrungen kommt [121, 171, 172]. Die Intensität der Lösungsmittelresonanz sollte daher verringert werden, bevor die Resonanzen der Metaboliten den Empfänger erreichen, sodass die Empfängerverstärkung nicht herabgesetzt werden muss [121].

Um das unerwünschte Wassersignal zu minimieren, können in der ^1H -NMR verschiedene Techniken der Lösungsmittelunterdrückung angewendet werden. Die Techniken der Lösungsmittelunterdrückung können in drei große Klassen eingeteilt werden. Methoden auf Basis der Vorsättigung, Methoden auf Basis der Magnetisierungszerstörung und Verfahren, die eine Lösungsmittelsättigung vermeiden [173].

Bei der Vorsättigung findet die Lösungsmittelunterdrückung vor der Anregung und Erfassung der Metabolitenresonanzen statt. Dies ist die einfachste Technik, sie kann leicht zu bestehenden Experimenten hinzugefügt werden und lässt Resonanzen abseits der Vorsättigungsfrequenz unbeeinflusst [121]. Die Vorsättigung kann auf unterschiedliche Weise erfolgen. Zum einen kann eine einfache Vorsättigung der Lösungsmittelresonanz erreicht werden, indem ein verlängerter RF von 1-3 s verwendet wird, sodass eine schmale effektive Bandbreite entsteht [172, 174]. Weiter kann nach der Vorsättigung ein sogenannter Composite Pulse folgen, um mögliche Impulsfehler oder eine ineffiziente Impulsstärke des Sättigungspulses zu korrigieren [119, 175].

Auch kann die Vorsättigung in die Messimpulssequenz eingebaut werden, wie es beim 1D-NOESY (Nuclear Overhauser Effect Spectroscopy) Experiment der Fall ist [172, 173]. Das 1D-NOESY-Experiment ist in Bezug auf Metabolomics-Ansätze die am meisten angewandte Technik der Lösungsmittelunterdrückung [171]. Bei der Pulssequenz folgen nach dem niederfrequenten Vorsättigungspuls zwei unmittelbar aufeinanderfolgende 90° Pulse. In der nachfolgenden Mischzeit kommt es zu einer unterschiedlichen Relaxation, wobei die Spins des Lösungsmittels schneller relaxieren. Durch einen darauffolgenden dritten 90° Puls werden die Metabolitenspins in die $x'y'$ -Ebene ausgelenkt und können detektiert werden, während die Lösungsmittel-Spins aufgrund der schnelleren Relaxation nicht erfasst werden [176]. Des Weiteren werden bei dem NOESY-Experiment nur Signale aufgezeichnet, die aus einem gewünschten Bereich der Probe im NMR-Röhrchen stammen. Dieser Teil der Probe befindet sich im zentralen Bereich der Empfängerspule, in welchem die Kerne die höchste Homogenität des Magnetfeldes erfahren. In diesem Bereich ist die Effektivität des Pulses am größten, was zu schmalen Signalen und einer einfacheren Unterdrückung führt. Die zusätzliche Verwendung eines Phasenzklus in der Pulssequenz führt weiter dazu, dass während der Aufnahme des FID nur erwünschte Signale detektiert werden. Zusätzlich werden durch die Verwendung des Phasenzklus Artefaktsignale herausgelöscht und erscheinen nicht im NMR-Spektrum [176]. Die Wirksamkeit der Vorsättigung ist hier von der Leistung, der Länge und der Position des Pulses abhängig. Auch wird die Effektivität der Unterdrückung bei einem schlechten Shimming herabgesetzt [174]. Eine höhere Leistung des Sättigungspulses ergibt eine bessere Unterdrückung, jedoch werden durch ein größeres effektives Unterdrückungsband die Intensitäten benachbarter Resonanzen reduziert [174].

Das WET (Water Suppression Enhanced Through T_1 Effects) ist ein weiteres Verfahren zur Lösungsmittelunterdrückung. Es nutzt vier selektive hochfrequente Pulse, denen jeweils ein Gradientenpuls folgt. Die Zeit zwischen den vier Pulsen ist identisch. Der Vorteil dieser Pulssequenz ist die Fähigkeit, mehrere Lösungsmittelresonanzen gleichzeitig zu unterdrücken [173]. Das WET-Experiment wird aufgrund der schnellen Art der

Unterdrückung vorwiegend bei gekoppelten Techniken wie der Flüssigkeitschromatographie in Verbindung mit der NMR eingesetzt [171].

Das WATERGATE (Water Suppression by Gradient Tailored Excitation) Experiment ist eine Gradienten-basierte Technik, bei welcher die Resonanzen der Metaboliten durch einen selektiven Puls refokussiert werden, während die Wasserresonanz durch zwei Gradientenpulse zerstört wird [173]. Bei dem Excitation-Sculpting-Experiment wird die WATERGATE-Pulssequenz zwei Mal hintereinander wiederholt, wobei unterschiedliche Gradientenstärken verwendet werden [171, 173]. Diese Methoden der Unterdrückung sind weniger empfindlich gegenüber Fehlern bei Impulsamplituden, jedoch ist der Anregungsbereich um die Lösungsmittelresonanz sehr breit. Dadurch werden Resonanzen nahe der Wasserresonanz signifikant unterdrückt [171, 173].

In der vorliegenden Arbeit wurden die verschiedenen Techniken der Lösungsmittelunterdrückung für die Erfassung der hydrophilen Metaboliten angewendet und verglichen (Arbeitsvorschrift 8.3.4). Bei dem NOESY-Experiment wurde zusätzlich die Leistung des Sättigungspulses, die Relaxationsverzögerung und die Messtemperatur überprüft (Arbeitsvorschrift 8.3.4). Für die Erfassung der wässrigen Metaboliten wurde schließlich das NOESY-Experiment zur Unterdrückung des Wassersignals angewendet. Für die Leistung des Sättigungspulses wurden 25 Hz gewählt. Daneben wurden weitere Messparameter wie die Messtemperatur und die Relaxationsverzögerung optimiert (Arbeitsvorschrift 8.3.4). Die Überprüfung der Stabilität der Messlösung (Arbeitsvorschrift 8.3.4) ergab, dass diese für zehn Tage stabil ist. Für die Erfassung der hydrophilen Metaboliten wurden die Kaffeeproben aufgrund der hohen Reproduzierbarkeit [133] jeweils mit den in Arbeitsvorschrift 8.3.5 beschriebenen Parametern einfach vermessen.

3.5 Nicht zielgerichtete Untersuchung mittels Verfahren der uni- und multivariaten Statistik

3.5.1 Methoden zur Datenvorbehandlung der ^1H -NMR-Spektren

Experimentelle, messtechnische und auswertetechnische Variationen führen in den anschließenden Verfahren der multivariaten Datenanalyse zu störenden Variabilitäten und sollten daher vermieden werden [136]. Um für die multivariate Datenanalyse in dieser Hinsicht vergleichbare ^1H -NMR-Spektren zu erhalten, ist eine adäquate Datenvorbehandlung essentiell.

Neben der Phasen- und Basislinienkorrektur, welche in Kapitel 1.2.2 bereits erklärt wurden, zählen das Alignment, das Skalieren, das Normalisieren, die Bucketierung, die Transformation, eine spektrale Selektion und die Entfernung unerwünschter Lösungsmittelsignale zu den Methoden der Datenvorbehandlung [133, 177].

Für einen Metabolomics-Ansatz sollte zunächst eine automatische Phasenkorrektur erfolgen und anschließend alle ^1H -NMR-Spektren auf Phasenverzerrung untersucht werden. Bei noch bestehenden Phasenverzerrungen sollte die Phase manuell nachgebessert werden [178]. Auch eine verzerrte Basislinie kann die multivariaten Ergebnisse beeinträchtigen und ist daher zu korrigieren. Für einen hohen Probendurchsatz wird eine automatische Basislinienkorrektur empfohlen [179, 180].

Signal-Shifts in den ^1H -NMR-Spektren verschiedener Proben können bedingt durch Veränderungen des pH-Wertes, der Messtemperatur, der Salzkonzentration in der Messlösung

oder durch ein inhomogenes Magnetfeld auftreten [178, 179, 181]. Dies kann zu Variationen in den Spektren der gleichen Produktklasse und so zu fehlerhaften Interpretationen in der multivariaten Datenauswertung führen. Treten die Signal-Shifts systematisch auf, können diese Variationen korrigiert werden, in dem die Spektren gegenüber einem Referenzsignal mit bekannter chemischer Verschiebung ausgerichtet werden. Als Referenzsignal kann das Signal von TSP oder TMS bei $\delta_{\text{H}} = 0,00$ ppm verwendet werden [178, 179, 182]. Bei TSP sollte berücksichtigt werden, dass diese Verbindung beispielsweise mit Proteinen interagieren kann, was sich auf die Signalform und -lage auswirkt [179]. Wird gesamte Spektrum auf ein Referenzsignal ausgerichtet, kann diese Methode zeitintensiver werden, wenn die Größe der Spektren zunimmt. Eine Alternative ist die Unterteilung der ^1H -NMR-Spektren in einzelne Bereiche, auf die anschließend die Ausrichtung angewandt wird [178, 183]. Die Shifts können durch ein Alignment allerdings nur bis zu einem gewissen Grad korrigiert werden. Daher sollten diese durch eine geeignete Probenvorbereitung minimiert werden [180]. In dieser Arbeit wurde ein Alignment auf das gesamte Spektrum angewandt. Als Referenzsignal diente TMS bei den ^1H -NMR-Spektren der lipophilen Metaboliten und TSP bei den ^1H -NMR-Spektren der hydrophilen Metaboliten mit jeweils einer chemischen Verschiebung von $\delta_{\text{H}} = 0,00$ ppm (Arbeitsvorschrift 8.4.1).

Ein ^1H -NMR-Spektrum besitzt in Abhängigkeit der Aufnahmeparameter mindestens 22000 Datenpunkte. Um diese Vielzahl an Datenpunkten zu reduzieren, wird eine Bucketierung der Spektren durchgeführt [177, 184, 185]. Dabei werden die ^1H -NMR-Spektren in kleine Bereiche, sogenannte Buckets, mit einer Größe von 0,01 – 0,05 ppm unterteilt. Die Intensität jedes Buckets wird durch Berechnung der Fläche unter der Kurve bestimmt. Diese Intensitäten werden anschließend als Variablen in der nachfolgenden multivariaten Datenanalyse verwendet. Infolgedessen wird die Menge der Variablen von mindestens 22000 Datenpunkten auf 500 – 1000 Buckets reduziert. Auch können durch die Bucketierung kleinere Signal-Shifts herausgemittelt werden [185]. Es gibt verschiedene Methoden, ein NMR-Spektrum in Buckets einzuteilen. Darunter eine Bucketierung mit gleichem Abstand, eine Gauß'sche Bucketierung und eine dynamisch adaptive Bucketierung. Die Bucketierung mit gleichem Abstand ist die am häufigsten angewandte Technik [166, 178]. Der Nachteil gleich großer Buckets besteht darin, dass sich ein Signal in zwei benachbarte Buckets aufteilen kann. Daher sollten die Buckets so positioniert werden und ausreichend breit sein, um ein oder mehrere Signale aufzunehmen [178, 185]. Nach der Bucketierung werden in der spektralen Selektion die Buckets, in denen kein Metabolit vorkommt oder in denen keine brauchbare Information enthalten ist, entfernt. Dies betrifft Bereiche außerhalb des Fensters von $\delta_{\text{H}} = 0,20$ – 10 ppm sowie Buckets mit Lösungsmittelsignalen. Diese Bereiche sind für die multivariate Datenanalyse nicht von Interesse und könnten darüber hinaus das Ergebnis beeinflussen [166]. Für diese Arbeit wurde eine Bucketierung mit gleichem Abstand angewendet. Für die ^1H -NMR-Spektren der lipophilen Metaboliten wurden 2000 Buckets im Bereich von $\delta_{\text{H}} = 0,50$ – 9,80 ppm verwendet, was einer Breite der Buckets von 0,005 ppm entspricht. Anschließend wurden die Buckets, in welchen Lösungsmittelsignale erschienen, entfernt. Dabei handelte es sich um die Lösungsmittelsignale von Chloroform, Wasser, Methanol und das im Chloroform zur Stabilisierung enthaltene Ethanol. Für die hydrophilen Metaboliten wurde das ^1H -NMR-Spektrum in 1000 Buckets im Bereich von $\delta_{\text{H}} = 0,50$ – 9,80 ppm mit einer Breite der Buckets von 0,009 ppm eingeteilt. Anschließend wurden die Buckets mit dem Wassersignal entfernt (Arbeitsvorschrift 8.4.1). Die Intensitäten der restlichen Buckets wurden als Variablen für die univariate und multivariate Datenanalyse verwendet.

Der nächste Schritt in der Datenvorbehandlung ist die Normalisierung, in welcher eine Korrektur der Konzentrationsunterschiede erfolgt [178]. Die Variationen in der Konzentration können durch Unterschiede in der Probenvorbereitung oder aufgrund von biologischen Unterschieden entstehen [177]. Bei der Probenvorbereitung kann dies eine unterschiedliche Probeneinwaage sein und biologische Variationen können bedingt durch einen unterschiedlichen Wassergehalt auftreten [177, 178, 186]. Dies führt zu Schwankungen in den Signalunterschieden, welche durch Normalisierungsmethoden beseitigt werden können [186]. Es können verschiedene Normalisierungsmethoden, darunter die Flächennormalisierung, die probabilistische Quotientennormalisierung, die Bereichsnormalisierung und die Normalisierung auf einen Referenzmetaboliten angewandt werden. Bei der Flächennormalisierung wird jeder Datenpunkt durch eine gleiche Gesamtfläche, ein Integral oder einen Mittelwert des Spektrums geteilt. Die probabilistische Quotientennormalisierung kann angewendet werden, wenn analytische oder bereichsspezifische Unterschiede im NMR-Spektrum vorliegen. Diese Methode verwendet ein repräsentatives Referenzspektrum und bildet den Quotienten zwischen den Datenpunkten und dem Referenzspektrum. Die Normalisierung dieser Datenpunkte erfolgt mit dem Median der Quotienten [179, 186]. Bei der Bereichsnormalisierung werden die Datenpunkte auf einen neuen Bereich, beispielsweise von $\delta_{\text{H}} = -1,00 - 1,00$ ppm normalisiert [186]. Die Normalisierung auf einen Referenzmetaboliten erfolgt, indem die Intensität aller Signale relativ zur Signalfäche des Referenzmetaboliten ausgedrückt wird. Um einen Metaboliten als Referenz verwenden zu können, muss dieser über alle Proben hinweg in gleicher Konzentration vorliegen. Beispielsweise wird bei Untersuchungen von Urin Kreatinin als Referenzmetabolit verwendet, da dessen Ausscheidungsrate als konstant angesehen wird [178, 186]. Zusätzlich kann jeder Probe ein interner Standard mit bekannter Konzentration zugesetzt und dieser als Referenz verwendet werden. Durch dieses Verfahren können allerdings nur Variationen bedingt durch die Probenvorbereitung und Erfassung der Metaboliten beseitigt werden [179, 185]. Bei der Verwendung von TSP als Referenz sollte zunächst überprüft werden, ob Interaktionen mit Proteinen oder Lipoproteinen vorliegen, da dadurch Variationen in der Signalintensität entstehen können [178]. In der vorliegenden Arbeit wurde sowohl für die Analyse der lipophilen, als auch für die Analyse der hydrophilen Metaboliten eine Normalisierung auf einen internen Standard als Referenz durchgeführt. Bei der Untersuchung der lipophilen Metaboliten wurde Dimethylsulfon (DMS) in bekannter Konzentration jeder Probe zugesetzt (Arbeitsvorschrift 8.2.5) und bei der Analyse der hydrophilen Metaboliten wurde als Referenz TSP verwendet (Arbeitsvorschrift 8.3.5). Zuvor wurde die Interaktion von TSP mit Proteinen oder Lipoproteinen in der Kaffeematrix untersucht und ausgeschlossen. Zusätzlich wurde überprüft, ob biologische Variationen durch eine Normalisierung auf eine einheitliche Fetteinwaage bei der Analyse der lipophilen Metaboliten oder eine Normalisierung bezüglich eines einheitlichen Wassergehaltes entfernt werden können (Arbeitsvorschrift 8.4.1).

Der Konzentrationsbereich der in der Probe vorhandenen Metaboliten kann sich über mehrere Größenordnungen erstrecken. Bei der Untersuchung von Pflanzenmaterial weisen die Konzentrationen von Sekundärmetaboliten in vielen Fällen einen geringeren Wert im Vergleich zu den Primärmetaboliten auf. Darüber hinaus ist die Variation mit der Konzentration verknüpft, sodass Metaboliten mit höherer Konzentration eine stärkere Variation aufweisen und so auch einen stärkeren Einfluss auf die multivariate Datenanalyse haben [166, 181, 185, 187]. Ziel der Skalierung ist es, den Einfluss dominanter Variablen, die für die Analyse möglicherweise nicht erforderlich sind, zu reduzieren [119, 166, 181]. Dies wird

erreicht, indem jede Variable durch einen Skalierungsfaktor dividiert wird, der für jede Variable unterschiedlich ist [188]. Eine richtige Skalierung verbessert den Informationsgehalt der Daten, während eine ungeeignete Skalierung dazu führen kann, dass unerwünschte Teile, welche für die multivariate Datenanalyse nicht relevant sind, hervorgehoben werden. Dies könnte zu einer fehlerhaften Interpretation in der Datenanalyse führen [180]. Es können verschiedene Skalierungsmethoden angewandt werden, darunter die Auto-Skalierung, die Pareto-Skalierung oder die Vast-Skalierung [166, 182, 188]. Die Autoskalierung wird im Metabolomics-Ansatz häufig angewendet und verwendet als Skalierungsfaktor die Standardabweichung. Dadurch tragen Metaboliten mit geringer als auch mit hoher Konzentration gleichermaßen zur Datenanalyse bei [179, 182, 188]. Die Pareto-Skalierung ähnelt der Autoskalierung. Statt der Standardabweichung wird jedoch die Wurzel dieser als Skalierungsfaktor verwendet. Die Vast-Skalierung ist eine Erweiterung der Auto-Skalierung, wobei hier als Skalierungsfaktor die Standardabweichung und der Variationskoeffizient verwendet werden [188, 189]. Ein Nachteil all dieser Methoden ist jedoch, dass dadurch der Einfluss von Variablen, die keine relevanten Informationen enthalten, erhöht werden kann [166]. Aus diesem Grund wurde in der vorliegenden Arbeit auf eine Skalierung verzichtet.

Neben Skalierungsmethoden gibt es mehrere Transformationsansätze wie die log-Transformation. Diese Methode beschreibt eine nichtlineare Umwandlung der Daten, wodurch große Werte im Datensatz stärker reduziert werden, als kleine Werte und besitzt dadurch einen pseudo-Skalierungseffekt [166, 179, 189]. Ein Nachteil dieser Methode ist allerdings, dass der Wert Null nicht verarbeitet werden kann. Auch nimmt die log-Transformation Werte an, die sich minus unendlich nähern, wenn sich der zu transformierende Wert sich Null nähert [188, 189]. In dieser Arbeit wurde eine log-Transformation durchgeführt, wobei aufgrund der beschriebenen Problematik erst Variablen mit einem Wert >1 transformiert wurden (Arbeitsvorschrift 8.4.1).

3.5.2 Kruskal-Wallis Test zur Untersuchung der Signifikanz

Die Anwendung univariater Methoden wie des Kruskal-Wallis Tests im Metabolomics-Ansatz dient zur Überprüfung der Signifikanz der Unterschiede [190]. Im Gegensatz zu multivariaten Methoden untersuchen univariate Methoden jeweils nur eine Variable [189]. Abhängig von der Annahme der Daten bezüglich der Normalverteilung, der klasseninternen Varianz und der Abhängigkeit der Klassen sind verschiedene univariate Tests möglich. Die parametrischen Tests wie der t -Test oder die Varianzanalyse (Analysis of Variance, ANOVA) beruhen auf der Annahme der Normalverteilung oder der Gauß'schen Verteilung der Daten [189]. Eine weitere Annahme der parametrischen Tests ist eine gleiche klasseninterne Varianz [191]. Im Gegensatz dazu gehen die nicht-parametrischen Test von keiner bestimmten Verteilungspopulation aus. Die nicht-parametrischen Tests sind weniger leistungsfähig als parametrische Tests und können somit leichter einen statistisch signifikanten Unterschied übersehen. Allerdings sind Daten eines Metabolomics-Ansatzes häufig nicht normalverteilt, haben eine ungleiche klasseninterne Varianz und einen unterschiedlichen Umfang von Stichproben. Bei diesen Daten gelten nicht-parametrische Tests als leistungsfähiger [191]. In diesem Fall wird der Kruskal-Wallis Test durchgeführt, welcher dem Mann-Whitney U Test ähnelt [192]. Der Unterschied besteht darin, dass der Kruskal-Wallis Test für mehr als zwei Klassen angewendet werden kann [192, 193]. Beim Kruskal-Wallis Test wird zuerst die Nullhypothese aufgestellt, dass zwischen den Klassen kein Unterschied vorhanden ist. Es wird ein Signifikanzniveau α definiert, ab wann die Nullhypothese abgelehnt wird. Dieser

Wert kann beispielsweise auf 5% ($\alpha = 0,05$) festgelegt werden und dient als Maß der Wahrscheinlichkeit für ein falsch-positives Ergebnis. Im Hypothesentest wird die Wahrscheinlichkeit berechnet, mit welcher die Nullhypothese abgelehnt wird. Diese Wahrscheinlichkeit wird als p -Wert ausgedrückt. Liegt der p -Wert unterhalb von α , wird die Nullhypothese folglich zurückgewiesen und es liegt ein signifikanter Unterschied vor [191]. Dieser Hypothesentest wird für jede Variable (Bucket) durchgeführt [191]. Wird der Kruskal-Wallis Test gleichzeitig für alle Buckets eines NMR-Spektrums angewendet, steigt die Wahrscheinlichkeit, dass falsch-positive Ergebnisse erzielt werden [191]. Das bedeutet, bei der Verwendung von einem Signifikanzniveau von $\alpha = 0,05$ steigt bei der Berechnung mehrerer p -Werte in einem Datensatz die Wahrscheinlichkeit, dass mehrere p -Werte $<0,05$ erhalten werden. Die Wahrscheinlichkeit, dass die Nullhypothese fälschlicherweise abgelehnt wird, ist folgendermaßen definiert (Formel 12) [191]:

$$1 - (1 - \alpha)^k \quad (12)$$

Dabei ist k die Anzahl der durchgeführten Hypothesentests und α das Signifikanzniveau. Um den Fehler, welcher durch das multiple Testen entsteht, zu beseitigen, kann das Signifikanzniveau basierend auf der Anzahl der Test angepasst werden. Die Standardmethode zur Anpassung des Signifikanzniveaus ist die Bonferroni-Korrektur, welche das Signifikanzniveau durch die Anzahl der durchgeführten Tests dividiert [189]. Die Anzahl der durchgeführten Tests entspricht in dieser Arbeit der Anzahl der verwendeten Buckets. Die zur Bonferroni-Korrektur ähnliche Šidák-Korrektur führt eine genauere Berechnung des korrigierten Signifikanzniveaus α_{korr} durch (Formel 13) [194, 195]:

$$\alpha_{korr} = 1 - (1 - \alpha)^{1/k} \quad (13)$$

Die Anzahl der durchgeführten Hypothesentests k entspricht hier der Anzahl der verwendeten Buckets und α dem Signifikanzniveau. Zur Überprüfung der Nullhypothese werden die p -Werte mit dem korrigierten Signifikanzniveau α_{korr} verglichen. In dieser Arbeit wurde zur Überprüfung der Signifikanz der Unterschiede ein Kruskal-Wallis Test mit einem Signifikanzniveau von $\alpha = 0,05$ durchgeführt, welches mittels Šidák korrigiert wurde (Arbeitsvorschrift 8.4.2)

3.5.3 Unüberwachte (unsupervised) multivariate Datenanalyse

Nach der Vorbehandlung der Daten besteht der nächste Schritt der multivariaten Datenanalyse darin, die Gesamtstruktur der Daten auf mögliche Muster oder Gruppierungen zu untersuchen [166, 181, 196, 197]. Diese Phase der multivariaten Datenanalyse basiert auf sogenannten unüberwachten Methoden, welche eine unvoreingenommene Sicht auf die Daten ermöglichen. Zu diesen Methoden zählen die PCA oder auch die hierarchische Clusteranalyse [136, 166]. Bei der hierarchischen Clusteranalyse werden Proben nach ihrer Ähnlichkeit gruppiert, wobei die Daten gebündelt als Dendrogramm dargestellt werden. Nachteil dieser Methode ist, dass diese keine Information über den Grund für eine bestimmte Klassenbildung liefert [166, 181]. Die PCA ist die am häufigsten verwendete unüberwachte multivariate Datenanalyse [166]. Die PCA führt zur Projektion der ursprünglichen Daten auf einen niedriger dimensional Raum, während so viele Informationen wie

möglich in den Daten erfasst werden. Dazu werden durch die PCA aus den ursprünglichen Variablen mittels Linearkombination neue, latente Variablen berechnet, die als Hauptkomponenten (Principal Components, PCs) bezeichnet werden [185, 198, 199]. Die PCs werden so berechnet, dass sie nacheinander in absteigender Bedeutung die Varianz der Originaldaten erklären. Sind die ersten PCs für den größten Teil der Varianz verantwortlich, können die restlichen PCs außer Acht gelassen werden [184]. Dadurch kann eine Dimensionsreduktion von mehreren hundert bis tausend Ursprungsvariablen in wenige PCs erreicht werden [166]. Diese sind paarweise orthogonal zueinander und in absteigender Varianz geordnet, sodass jede neue PC einen Teil der Datenvarianz beschreibt, welcher nicht von den vorherigen PCs beschrieben wurde [181, 185, 198, 200, 201].

Im Detail ist die PCA die Lösung eines Eigenwertproblems. Der Ausgangspunkt dieses Verfahrens ist die mittenzentrierte Datenmatrix X , in welcher die Zeilen die Proben und die Spalten die Variablen darstellen. Aus der Datenmatrix wird eine Kovarianzmatrix berechnet, welche die Varianz der Proben beschreibt. Im nächsten Schritt wird die Eigenwertberechnung durchgeführt, woraus die Matrizes S , P und T resultieren. Die Matrix S enthält die Eigenwerte, welche der Größe nach sortiert sind und deren Beträge der Varianz entsprechen. Der erste Eigenvektor (PC1, die erste Spalte in P) hat die Richtung der größten Varianz und der zweite Eigenvektor (PC2) die Richtung der zweit größten Varianz. Der letzte Eigenvektor beschreibt den geringsten Teil der Varianz und hat den kleinsten Eigenwert. In der Loadingsmatrix P befinden sich die Eigenvektoren der Kovarianzmatrix, welche das neue Koordinatensystem bilden. Die Faktorenladungen oder auch Loadings beschreiben dabei, wie groß der Einfluss der einzelnen Variablen auf die jeweiligen PC ist und können im Loadingsplot graphisch dargestellt werden. Besitzt eine Variable einen hohen Loadingswert für PC1, so ist diese zu einem großen Teil mitverantwortlich für die Trennung der Klassen in Richtung von PC1. Die Scoresmatrix T enthält die mit den Eigenwerten gewichteten Scoreswerte. Diese sind die Werte für das neue Koordinatensystem. Die Scores können im Scoresplot graphisch dargestellt werden, wodurch die Position der Proben, Gruppierungen dieser und Ausreißer in dem neuen Koordinatensystem ersichtlich sind [185, 198–200, 202].

Graphisch betrachtet können die PCs bestimmt werden, indem bei der ersten PC (PC1) die Richtung der maximalen Varianz in den Ausgangsdaten gesucht wird (Abbildung 3-2(A)). Anschließend wird möglichst viel der verbleibenden Varianz mit der zweiten PC (PC2), welche orthogonal zu PC1 ist, extrahiert (Abbildung 3-2(B)). Diese Vorgehensweise wird schrittweise weitergeführt, bis die maximal mögliche Gesamtvarianz der Ausgangsdaten erklärt ist. Die auf diese Weise gefundenen PCs bilden ein neues Koordinatensystem, wobei der Ursprung des Koordinatensystems der Datenmittelpunkt ist. Die Ursprungsdaten werden in das neue Koordinatensystem transformiert, indem jede Variable auf jede dieser PC projiziert wird (Abbildung 3-2(C)) [198, 199].

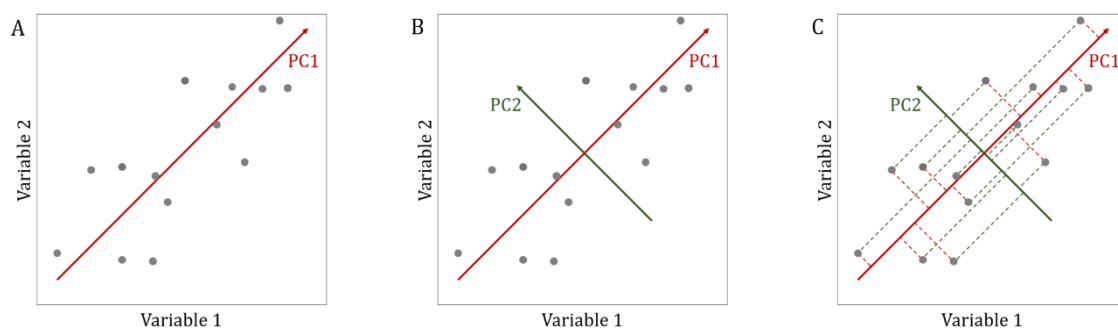


Abbildung 3-2: Graphische Darstellung der Vorgehensweise der Hauptkomponentenanalyse (Principal Component Analysis, PCA). (A) Bestimmung der ersten Hauptkomponente (PC1) in Richtung der maximalen Varianz. (B) Bestimmung der zweiten Hauptkomponente (PC2) in Richtung der maximalen verbleibenden Varianz. PC2 ist orthogonal zu PC1. (C) Projektion der Ursprungsdaten auf das neue Koordinatensystem mit PC1 und PC2 als Achsen (modifiziert nach [198, 199])

Die PCA extrahiert jedoch nur die Hauptvariation des Datensatzes und reduziert die Dimensionalität der Daten. Um eine Klassifizierung durchführen zu können, ist eine Kombination mit beobachteten Methoden wie der Diskriminanzanalyse nötig [185, 200, 203]. Zusätzlich ist es möglich, dass die PCA einige Diskriminanzinformationen vernachlässigt, was die Klassifizierungsleistung beeinträchtigen kann [204]. In solchen Fällen können andere multivariate Datenanalysen wie die PLS (Kapitel 3.5.4) durchgeführt werden.

In der vorliegenden Arbeit wurde mit den durch die Datenvorbehandlung erhaltenen Buckets der ^1H -NMR-Spektren der lipophilen und hydrophilen Metaboliten jeweils eine PCA durchgeführt (Arbeitsvorschrift 8.4.3).

3.5.4 Überwachte (supervised) multivariate Datenanalyse

Im Gegensatz zu den unüberwachten Methoden ist das Ziel der überwachten Methoden die Klassifizierung der Proben und die Vorhersage der Klassenzugehörigkeit neuer Proben [182, 203]. Dabei wird die Beziehung zwischen einer Matrix von Variablen und einer Matrix von Antworten, in diesem Fall die Klassenzugehörigkeit, verwendet, um neue Daten vorherzusagen [166]. Bei den überwachten Methoden wird zwischen der Regression und der Klassifizierung unterschieden. Eine sehr häufig angewandte Regression ist die PLS [198]. Im Gegensatz zur PCA nutzt die PLS bei der Bestimmung der Hauptkomponenten bereits die Struktur der Klassenzugehörigkeit. Die Vorgehensweise der PLS ist die, dass der Zusammenhang zwischen einer oder mehreren Zielgrößen und den Variablen bestimmt wird. Der Ausgangspunkt ist wie bei der PCA die Datenmatrix X bestehend aus den Proben und den Variablen. Zu jeder Probe wird zusätzlich eine Zielgröße gemessen, welche den Vektor y bildet. Bei mehreren Zielgrößen existieren demnach auch mehrere Vektoren. Diese Vektoren ergeben die Matrix Y , welche die Proben und die Zielgrößen enthält. Ziel der PLS ist es nun, sowohl mit der Datenmatrix X , als auch mit der Matrix Y eine PCA durchzuführen [198]. Daraus resultieren jeweils eine Scores- und eine Loadingsmatrix für X und Y . Bei der Berechnung der Hauptkomponenten für die PLS erfolgt ein Informationsaustausch zwischen den Scores und Loadings von X und Y [198].

Zu den klassischen Klassifizierungsverfahren, welche häufig bei der Analyse von Lebensmitteln Anwendung finden, zählt die Diskriminanzanalyse [166, 205]. Dabei wird unter anderem zwischen der linearen Diskriminanzanalyse (LDA) und der quadratischen Diskriminanzanalyse (Quadratic Discriminant Analysis, QDA) unterschieden [185]. Im Fall einer LDA erfolgt die Klassifizierung, indem die Grenzen zwischen benachbarten Klassen oder

Klassen als Geraden oder Hyperebenen berechnet werden, bei der QDA sind dies quadratische Kurven oder gekrümmte Ebenen [185, 206]. Die LDA sucht eine lineare Funktion, welche höher dimensionale Merkmalsvektoren, die zu unterschiedlichen Klassen gehören, in einen niedriger dimensional Raum transformiert [182, 207]. In diesem niedriger dimensional Raum (LDA-Raum) sollen die Merkmalsvektoren anderer Klassen gut voneinander getrennt vorliegen. Mit anderen Worten sucht die LDA also nach einer Funktion und richtet diese so aus, dass die Varianz zwischen den Klassen erhöht und die Varianz innerhalb der Klassen minimiert wird [182, 207]. Die Varianz zwischen den Klassen wird bestimmt, indem der Abstand der Mittelwerte verschiedener Klassen berechnet wird. Durch die Berechnung des Abstands zwischen dem Mittelwert und den Proben jeder Klasse erfolgt die Bestimmung der Varianz innerhalb einer Klasse [204]. Die Zuordnung von Testproben erfolgt über die Diskriminanzfunktion zu der Klasse, von deren Mittelwert die Testprobe den geringsten Abstand aufweist [208]. Die relativen Abstände von Testproben zum jeweiligen Klassenmittelpunkt können als sogenannte *conf*-Werte ausgedrückt werden. Je geringer der *conf*-Wert ist, desto geringer ist auch der relative Abstand der Testprobe zur jeweiligen Klasse.

Voraussetzung der LDA ist, dass für die Durchführung die Anzahl der Stichproben größer ist, als die Anzahl der Variablen. Daher wird die LDA in Kombination mit dimensionsreduzierenden Methoden wie der PCA oder der PLS kombiniert [182, 205–207]. Eine der am häufigsten angewandten Methoden ist die PLS zusammen mit der Diskriminanzanalyse (Partial Least Squares Discriminant Analysis, PLS-DA) [182]. Bei der PLS-DA wird im Vergleich zur PLS das Modell auf Grundlage der Klassifizierungsleistung statt auf Grundlage der Vorhersageleistung optimiert [197]. Das PLS-DA Modell muss durch geeignete Validierungstechniken überprüft werden, um die Klassentrennung zu bestätigen [190, 197].

Eine Alternative zur LDA ist die Soft Independent Modeling of Class Analogy (SIMCA), die angewendet werden kann, wenn die Klassen unterschiedlich definierte Strukturen, sogenannte asymmetrische Datenstrukturen (Abbildung 3-3), aufweisen und daher keine Abstandsberechnungen möglich sind [209].

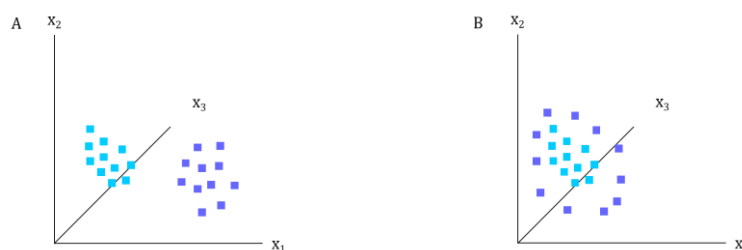


Abbildung 3-3: Schematische Darstellung verschiedener Datenstrukturen. (A) Symmetrische Datenstruktur; (B) asymmetrische Datenstruktur (modifiziert nach [210])

Bei der SIMCA werden PCA Modelle für jede Klasse erstellt, wodurch eine Klassifizierung von Testproben in keine, eine oder mehrere Klassen möglich ist [185]. Dabei kann jede Klasse in einer anderen Dimension zusammengefasst werden. Neue Beobachtungen werden dann anhand ihrer Abweichung zu den verschiedenen PCA Modellen klassifiziert. Diese Abweichung wird in der orthogonalen Distanz einer Beobachtung zum PCA-Raum dargestellt [211].

In der vorliegenden Arbeit wurde die PCA mit einer LDA kombiniert, um Klassifizierungsmodelle nach dem geographischen Ursprung, nach der Anbaumethode, nach dem

Röstverfahren und dem Röstgrad zu erstellen. Diese PCA-LDA stellt ein Hybridverfahren von PCA und LDA dar, welches von der Firma Bruker entwickelt wurde. Der Algorithmus der verwendeten PCA-LDA ist im Anhang ersichtlich. Dabei wurden jeweils zwei-Klassen Modelle erstellt (Arbeitsvorschrift 8.4.3).

3.5.5 Validierung der Klassifizierungsmodelle

Ein wichtiger Aspekt der Klassifizierung ist die Validierung. Die sehr leistungsfähigen überwachten Methoden können Modelle konstruieren, die sehr an die vorgegebenen Daten angelehnt sind, selbst wenn es keine tatsächliche Beziehung zwischen den Daten gibt [166]. Dies wird als Überanpassung (Overfitting) bezeichnet und entsteht durch eine erhöhte Dimensionalität der Daten [166, 190, 212]. Es gibt mehrere Möglichkeiten, ein Modell bezüglich der Vorhersageleistung zu validieren, darunter die interne Kreuzvalidierung (KV). Um diese durchzuführen, wird der Datensatz in einen Trainings- und Testdatensatz geteilt. Die Auswahl der Proben für den Testdatensatz sollte dabei repräsentativ für jede Klasse sein. Mit dem Trainingsdatensatz erfolgt die Erstellung des Modells, welches mit dem Testdatensatz validiert wird [166]. Die Einteilung des Trainings- und Testdatensatzes wird so oft wiederholt, bis jede Probe einmal im Testdatensatz war [213]. Werden beispielsweise 10% der Proben als Testdatensatz verwendet, erfolgt die Einteilung in Trainings- und Testdatensatz und die Erstellung des Klassifizierungsmodells zehn Mal. Der Erfolg eines Klassifizierungsmodells wird anschließend anhand des Prozentsatzes korrekt klassifizierter Proben aus dem Testdatensatz bestimmt. Das Modell ist umso besser, je höher der Prozentsatz ist [213]. Die KV ist besonders für kleine Datensätze geeignet. Jedoch sind die Trainings- und Testsätze nicht komplett unabhängig voneinander, was dazu führt, dass das erstellte Klassifizierungsmodell für die ausgelassenen Stichproben optimiert ist. Um dieses Problem zu minimieren, kann eine Monte-Carlo-Kreuzvalidierung (MCKV) angewendet werden. In der MCKV wird die KV mehrfach durchgeführt, während bei jeder neuen KV die Proben für Trainings- und Testdatensatz nach dem Zufallsprinzip ausgewählt werden. Dadurch werden die Modelle bezüglich ihrer Vorhersagefähigkeit unter Verwendung vieler verschiedener Kombinationen von Validierungsdatensätzen bewertet [214]. Der Erfolg des Modells wird anschließend als durchschnittlicher Prozentsatz korrekt klassifizierter Proben aus den Testsätzen dargestellt [213].

Eine weitere Möglichkeit, die Optimierung der Klassifizierungsmodelle auf die ausgelassenen Stichproben zu umgehen, ist eine externe Validierung. Bei dieser Validierung erfolgt die Verwendung eines nicht zur Modellerstellung verwendeten und somit unabhängigen Probandatensatzes [213, 215, 216]. Dieser Probandatensatz ist ein Teil des originalen Datensatzes, welcher nicht für die Modellerstellung und KV verwendet wurde. Dazu wird der Datensatz in ein Trainings-, Test- und unabhängigen Probandatensatz geteilt. Mit dem Trainings- und Testdatensatz erfolgt die Modellerstellung und KV, während die Vorhersagekraft des optimierten Modells anschließend mit dem unabhängigen Probandatensatz überprüft wird [213].

In der vorliegenden Arbeit wurde eine interne MCKV als auch eine externe Validierung der Klassifizierungsmodelle durchgeführt. Für den unabhängigen Probandatensatz wurden 20% der Originaldaten verwendet als Beurteilungskriterium für die Vorhersagekraft des Modells bezüglich des unabhängigen Probandatensatzes diente der *conf*-Wert, welcher als Maß für den Abstand der Testprobe zum Mittelpunkt der entsprechenden Klasse darstellt. Besitzt eine Testprobe einen *conf*-Wert von 0, so liegt diese im Datenmittelpunkt der entsprechenden Klasse. Beträgt der *conf*-Wert 1, so liegt die Testprobe noch innerhalb des

berechneten 95% Prognoseellipsoids der entsprechenden Klasse. Die interne Validierung erfolgte mit einer zehnfachen MCKV. Der Datensatz wurde zu 90% Trainingsdatensatz und 10% Testdatensatz eingeteilt, sodass es sich um eine zehnfache KV handelte. Diese zehnfache KV wurde zehn Mal wiederholt, während nach jeder KV die Auswahl der Proben für Trainings- und Testdatensatz nach dem Zufallsprinzip erfolgte. Dies führte dazu, dass der Trainingsdatensatz bei jeder KV aus einer anderen Kombination von Proben bestand. Demnach erfolgte die Erstellung und Überprüfung des Modells 100 Mal. Aus den 100 Überprüfungen wurde der Mittelwert der korrekt klassifizierten Proben aus den Testsätzen in einer Konfusionsmatrix dargestellt (Arbeitsvorschrift 8.4.3).

3.5.6 Datenfusionierung der ^1H -NMR-spektroskopischen Daten der lipophilen und hydrophilen Metaboliten

Die Überprüfung der Authentizität von Lebensmitteln erfordert eine ganzheitliche Charakterisierung der Produkte. Allerdings reicht eine einzelne Messung oftmals nicht aus, um eine komplexe Lebensmittelmatrix komplett zu charakterisieren. Daher kann eine verbesserte Charakterisierung des Lebensmittels auch durch Kombination mehrerer Analysemethoden in einer Datenfusionierung erfolgen. Durch die Datenfusionierung wird der Informationsgehalt erhöht und gleichzeitig kann der Einfluss störender Variablen reduziert werden. Wichtig für eine Datenfusionierung ist jedoch, dass die Datensätze, die kombiniert werden, mindestens teilweise unterschiedliche Informationen enthalten [217]. Häufig kombinierte Analysemethoden sind sensorische Daten, NMR, MS, oder Isotopendaten, wobei normalerweise zwei Methoden miteinander kombiniert werden [205, 218]. Die Datenfusionierung kann in die Kategorien der Low-, Mid- und High-Level Datenfusionierung eingeteilt werden. In der Low-Level Datenfusionierung werden die Verfahren zusammengefasst, bei denen die ursprünglichen Datenblöcke verschiedener Methoden zusammengefügt werden. Die ursprünglichen Datenblöcke stellen in dieser Arbeit die Buckets dar und werden mittels Zeilenverkettung in eine neue Matrix überführt. Die Zeilen der neuen Matrix entsprechen der Anzahl der analysierten Proben und die Spalten der Anzahl der Variablen, die aus den verschiedenen Informationsquellen stammen. Mit dieser Matrix erfolgt dann die Erstellung des Klassifizierungsmodells. Low-Level Ansätze sind einfach anzuwenden, jedoch ist ein Skalierungsschritt erforderlich, um einen größeren Einfluss eines Datenblocks mit höherer Varianz zu vermeiden [205, 217]. Nachteil der Low-Level Datenfusionierung ist, dass der Einfluss störender oder irrelevanter Variablen möglicherweise nicht reduziert wird. Bei einer hohen Anzahl von Variablen sollte vor der Erstellung eines Klassifizierungsmodells eine Dimensionsreduktion mittels PCA oder PLS erfolgen [218].

Die hohe Dimensionalität der Daten in einem Metabolomics-Ansatz macht eine Datenreduktion zwingend erforderlich, sodass eine Mid-Level Datenfusionierung oft besser geeignet ist [218–220]. Bei der Mid-Level Datenfusionierung erfolgt zunächst die Berechnung der latenten Variablen für jeden ursprünglichen Datenblock. Die latenten Variablen der Datenblöcke, welche die PCA- oder PLS-Scores darstellen, werden anschließend in einer Matrix verkettet. Das endgültige Klassifizierungsmodell wird anschließend auf Basis dieser Matrix berechnet [205, 217, 218]. Die Mid-Level Datenfusionierung ergibt im Vergleich zur Low-Level Datenfusionierung eine bessere Klassifizierungsfähigkeit [218].

Die Anwendung der High-Level Datenfusionierung basiert auf der Erstellung von Klassifizierungsmodellen für jeden Datenblock, welche anschließend für die Entscheidungsfindung kombiniert werden. Daher wird dieser Ansatz auch als Entscheidungsfusion bezeichnet [205, 217].

In der vorliegenden Arbeit wurde getestet, ob eine Low-Level oder eine Mid-Level Datenfusionierung ^1H -NMR-Spektren der lipophilen und hydrophilen Metaboliten eine bessere Vorhersagefähigkeit als die der jeweiligen Einzelmodelle ergibt (Arbeitsvorschrift 8.4.4).

3.6 Identifizierung der Kaffeemetaboliten mittels NMR-Spektroskopie

Das ^1H -NMR-Spektrum eines Pflanzen-Extraktes liefert Signale vieler verschiedener Primär- und Sekundärverbindungen. Der Vergleich mit Referenzverbindungen kann für viele dieser Verbindungen zu einer erfolgreichen Identifizierung der Signale in einem NMR-Spektrum führen. Jedoch sind für viele pflanzliche Sekundärmetaboliten keine Referenzverbindungen verfügbar. In diesem Fall erfordert die Zuordnung der NMR-Signale und die Identifizierung der entsprechenden Verbindung mehr Spektraldaten [135]. So können zweidimensionale (2D)-NMR-Experimente zur Identifizierung angewandt werden [135]. Die Einführung einer zusätzlichen Zeitspanne, der Evolutionsperiode, zwischen der Vorbereitungs- und Erfassungsperiode des 1D-Experiments ist die Grundlage für 2D-Experimente. Der FID ist nun von der Evolutionszeit t_1 und der Detektionszeit t_2 abhängig. Durch eine FT resultiert ein NMR-Spektrum, das durch die zwei Frequenzen F_1 und F_2 charakterisiert ist. Das 2D-NMR-Spektrum kann anschließend in sogenannten Konturdiagrammen dargestellt werden [119]. Die F_1 -Ebene im 2D-NMR-Spektrum stammt von t_1 und entspricht der Evolutionszeit, wohingegen die F_2 -Ebene aus t_2 stammt und der Erfassungszeit entspricht [121].

Ein 2D-Experiment besteht aus der Vorbereitung, der Evolution, der Mischung und der Erfassung. Die Vorbereitungs- und Mischperiode beinhalten abhängig vom entsprechenden Experiment jeweils ein oder mehrere Pulse, wohingegen die Detektionsperiode analog zu einem 1D-NMR-Experiment ist. Der Puls in der Vorbereitungsperiode erzeugt einen Ungleichgewichtszustand des Spinsystems, welches sich nach dem Puls in der Evolutionsperiode frei entwickelt. Die Art der Wechselwirkung während der Evolutionsperiode ist abhängig von den physikalischen Bedingungen wie der chemischen Verschiebung oder der Spin-Spin-Kopplung und charakterisiert die Art des 2D-Experiments. Die Evolutionsperiode endet mit einem zweiten Puls in der Mischphase, welcher zur Erfassung des FIDs führt. Um die Evolutionszeit während der Aufnahme des FID indirekt miterfassen zu können, wird das NMR-Experiment mehrmals wiederholt, wobei die Evolutionszeit um einen einheitlichen Betrag erhöht wird [119, 121]. Der Betrag der schrittweisen Erhöhung der Evolutionszeit bestimmt wesentlich die Gesamtdauer des 2D-Experiments. Um die Gesamtdauer des 2D-NMR-Experiments zu reduzieren, kann ein Non Uniform Sampling (NUS) eingesetzt werden. Durch NUS wird die Anzahl der gesammelten Datenpunkte reduziert und damit die Gesamtdauer des 2D-NMR-Experiments verkürzt. Dies wird erreicht, indem nur eine Teilmenge der t_1 -Datenpunkte erfasst wird. In einem NUS Experiment können 25-50% der Datenpunkte erfasst werden, was zu einer zwei- bis vierfachen Zeitersparnis führt. Die Erfassung einer zu geringen Anzahl von t_1 -Datenpunkten kann jedoch zur Bildung von Artefakten bis hin zur Verfälschung des NMR-Spektrums führen. Zudem muss berücksichtigt werden, dass vor der FT eine Rekonstruktion der fehlenden Datenpunkte erforderlich ist [121].

2D-NMR-Experimente können in solche unterteilt werden, die homonukleare oder heteronukleare Korrelationen erfassen. Homonukleare Korrelationen beinhalten die gleichen

Kerne und sind meistens ^1H - ^1H -NMR-Experimente. Zur Identifizierung von Metaboliten können Correlation Spectroscopy (COSY) und Total Correlation Spectroscopy (TOCSY) angewendet werden. Das ^1H - ^1H -COSY-Experiment zeigt direkte Kopplungen zwischen benachbarten Protonen über zwei ($^2J_{\text{HH}}$) oder drei Bindungen ($^3J_{\text{HH}}$). Signale, die im COSY-Spektrum auf der Diagonalen liegen, weisen die chemische Verschiebung auf, die auch in einem 1D- ^1H -NMR-Spektrum erkennbar sind. Die von der Diagonalen abweichenden Kreuzsignale zeigen eine skalare Kopplung benachbarter Protonen. Dahingegen zeigt das TOCSY-Experiment skalare Kopplungen innerhalb des gesamten Spinsystems an [121].

Bei heteronuklearen 2D-NMR-Experimenten handelt es sich um die Messung zwei verschiedener Kerne, wie beispielsweise ^1H und ^{13}C . Die beiden weit verbreiteten Experimente, die auf heteronuklearen Korrelationen beruhen, sind das Hetero Nuclear Single Quantum Coherence (HSQC) und das Heteronuclear Multiple Bond Correlation (HMBC) Experiment. Die Anwendung dieser Techniken erfordert die Unterdrückung der störenden Resonanz von Protonen, welche an Kerne gebunden sind, die eine Kernspinzahl $I \neq \frac{1}{2}$ haben, wie es bei ^1H gebunden ^{12}C -Kernen der Fall ist. Dies wird heutzutage mittels gepulster Feldgradienten erreicht. Sowohl im HSQC- als auch im HMBC-Experiment erfolgt die Übertragung der Quermagnetisierung von ^1H auf gekoppelte ^{13}C -Kerne. Durch das ^1H - ^{13}C -HSQC-Experiment werden Korrelationen von Protonen mit direkt benachbarten ^{13}C -Kernen ($^1J_{\text{CH}}$) ersichtlich, wohingegen beim ^1H - ^{13}C -HMBC-Experiment Korrelationen von Protonen und ^{13}C -Kernen über zwei ($^2J_{\text{CH}}$) oder drei Bindungen ($^3J_{\text{CH}}$) aufgezeigt werden [121, 221]. Ein zusätzlicher Informationsgewinn kann durch die Verwendung eines Experimentes erfolgen, welches das Hetero Nuclear Single Quantum Coherence-Experiment mit der Total Correlation Spectroscopy (HSQC-TOCSY) verbindet, da hier neben der Korrelation von Protonen und direkt benachbarten ^{13}C -Kernen auch Spinsysteme eines Moleküls aufgezeigt werden können [129].

Für die Zuordnung der Signale zu Koffeinhaltstoffen wurden in der vorliegenden Arbeit das HSQC- und HSQC-TOCSY-Experiment verwendet. Für die Erfassung der Daten wurde NUS angewendet, wobei 50% der t_1 -Datenpunkte erfasst wurden, was zu einer Zeitersparnis um den Faktor zwei führte (Arbeitsvorschrift 8.5). Zusätzlich erfolgte der Vergleich mit Referenzsubstanzen (Arbeitsvorschrift 8.5).

3.7 Zielgerichtete Analyse zur Quantifizierung verschiedener Kaffeemetaboliten

Die Quantifizierung der Metaboliten in Kaffee kann durch verschiedene Methoden erfolgen. Während die Analyse der Aromastoffe normalerweise mittels GC-MS erfolgt, wird für die Analyse nicht flüchtiger Kaffee-Inhaltsstoffe die Hochleistungsflüssigkeitschromatographie (High Performance Liquid Chromatography, HPLC) eingesetzt. Bei den mittels HPLC untersuchten Inhaltsstoffen handelt es sich dabei um Coffein, CQA, Trigonellin, NMP, HMF und Nicotinsäure [222–226]. Mittels qNMR konnten in der Literatur unter anderem die Koffeinhaltstoffe Trigonellin, NMP, Nicotinsäure, Ameisensäure, 5-CQA, 3-CQA, Koffein und Ameisensäure in einem methanolischen Extrakt bestimmt werden [88]. Del Campo et al. quantifizierten Koffein, Ameisensäure, Trigonellin und HMF in löslichem Kaffee mittels qNMR [227].

Die HPLC ist eine chromatographische Methode, welche die Trennung von Analytgemischen innerhalb kurzer Zeit ermöglicht. Die Trennung erfolgt aufgrund unterschiedlich

starker Wechselwirkungen sowohl zwischen Analyt und der mobilen Phase, als auch zwischen dem Analyten und der stationären Phase. Die stationäre Phase stellt das Füllmaterial gepackter Säulen dar, durch welche die mobile Phase fließt. Je weniger stark die Wechselwirkung von Analyt mit der stationären Phase ist, desto schneller passiert der Analyt die Säule. Bei der Verwendung einer polaren stationären Phase, wie beispielsweise Kieselgel, und einer relativ unpolaren mobilen Phase, handelt es sich um ein Normalphasen-System. Das Trennprinzip beruht auf einer unterschiedlich starken Adsorption der verschiedenen Analyten an der stationären Phase. Dagegen werden bei dem Umkehrphasen-System chemisch modifizierte Trägermaterialien als stationäre Phase verwendet, die relativ unpolar sind. Dies wird erreicht, indem an die polaren, endständigen Silanolgruppen an der Oberfläche des Kieselgels Alkylreste bestimmter Kettenlänge oder Reste mit einer funktionellen Gruppe gebunden werden. Als mobile Phase dient Wasser, welches verschiedene Anteile an organischen Modifiern wie Methanol oder Acetonitril enthalten kann. Das Umkehrphasen-System wird überwiegend zur Auftrennung von Gemischen mittelpolarer Verbindungen eingesetzt [228–230]. Die Analyse der Kaffeeinhaltsstoffe erfolgt in der Literatur meist mit einem Umkehrphasen-System, welches mit einem Diodenarray-Detektor gekoppelt ist (HPLC-DAD) [223, 225]. In der vorliegenden Arbeit wurde eine HPLC-DAD Methode zur Quantifizierung von Trigonellin, Koffein, HMF, 3-CQA, 4-CQA, und 5-CQA eingesetzt und die Ergebnisse mit denen der qNMR-Methode verglichen (Arbeitsvorschrift 8.6.1). Zusätzlich wurde die amtliche Untersuchungsmethode L 46.00-3 zur Bestimmung von Koffein in Kaffee und Kaffee-Erzeugnissen angewendet (Arbeitsvorschrift 8.6.2). Nach dieser Methode erfolgte eine Heißextraktion der Kaffeeproben und eine anschließende Bestimmung des Koffeins mittels HPLC-DAD.

Wie in Kapitel 1.2.2 beschrieben, ermöglicht die qNMR die simultane Bestimmung mehrerer Analyten aus verschiedenen Substanzklassen ohne aufwändige Probenvorbereitung [119]. Für die Quantifizierung der Kaffeeinhaltsstoffe mittels qNMR wurde in der Literatur sowohl die Verwendung eines IStd, als auch die Anwendung der PULCON-Methode beschrieben [88, 227]. In der vorliegenden Arbeit wurden die ^1H -NMR-Spektren der nicht zielgerichteten Methode für die Untersuchung der hydrophilen Metaboliten verwendet, um ausgewählte Kaffeeinhaltsstoffe zu quantifizieren. Aufgrund der vielen hydrophilen Metaboliten, deren Signale sich über das gesamte NMR-Spektrum erstrecken, konnte kein geeigneter IStd gefunden werden, welcher keinerlei Überlappung mit den Analytsignalen aufwies. Aus diesem Grund wurde die PULCON-Methode (Kapitel 1.2.2.1) zur Quantifizierung von HMF, Trigonellin, NMP, Koffein, CQA, Ameisensäure, Essigsäure und Milchsäure angewendet als Grundlage für die Quantifizierung diente die entwickelte PULCON-Methode von Monakhova et al. [127, 231]. Durch einen separat gemessenen externen Std, welcher als Quantification Reference (QuantRef) bezeichnet und bei jeder Messserie mitvermessen wird, erfolgte die Ermittlung eines gerätespezifischen Faktors (ERETIC-Faktor). Dazu wurden die Signale der sich in der QuantRef befindenden Milch- und Zitronensäure integriert und für jedes dieser Signale der ERETIC-Faktor bestimmt (Formel 14) [127, 231]. Es gilt:

$$ERETIC = \frac{I_{Ref} \cdot SW_{Ref} \cdot MG_{Ref}}{SI_{Ref} \cdot c_{Ref} \cdot N_{Ref}} \quad (14)$$

Dabei beschreibt I_{Ref} das absolute Integral des jeweiligen Signals, SW_{Ref} die spektrale Breite in ppm, MG_{Ref} das Molekulargewicht in g/mol der entsprechenden Referenzsubstanz,

SI_{Ref} die Anzahl der Datenpunkte des prozessierten $^1\text{H-NMR}$ -Spektrums, c_{Ref} die Konzentration der Referenzsubstanz in mg/L und N_{Ref} die Anzahl der Protonen, die das entsprechende Signal erzeugen. Nach der Bildung des Mittelwertes der bestimmten ERETIC-Faktoren wurde die Konzentration der Analyten in den Kaffeeproben nach Formel 15 ermittelt [127, 231]. Aufgrund der gleichen Anzahl an Scans für die $^1\text{H-NMR}$ -Messung von QuantRef und Probe, wurde dieser Parameter in Formel 15 vernachlässigt.

$$c_X = \frac{I_X \cdot SW_X \cdot MG_X \cdot P1_X \cdot V}{SI_X \cdot ERETIC \cdot N_X \cdot P1_{Ref} \cdot VF \cdot EW} \quad (15)$$

Das tiefgestellte X bezieht sich auf dem entsprechenden Analyten, c_X ist die Konzentration des Analyten in mg/kg, I_X das absolute Integral des jeweiligen Signals, SW_X die spektrale Breite in ppm, MG_X das Molekulargewicht in g/mol, $P1_X$ die Länge des 90° Pulses der Probe, V das Volumen der Extraktionslösung in L, SI_X die Anzahl der Datenpunkte des prozessierten NMR-Spektrums, N_X die Anzahl der Protonen, die das entsprechende Signal erzeugen, $P1_{Ref}$ die Länge des 90° Pulses der QuantRef, der VF der Verdünnungsfaktor und EW die Einwaage der Kaffeeprobe in kg.

Bei der qNMR-Methode war die separate Bestimmung von 3-, 4- und 5-CQA nicht möglich, sodass diese Isomere als Summenparameter bestimmt wurden. Die Signale von Koffein, Essigsäure, Milchsäure und CQA überlappten im $^1\text{H-NMR}$ -Spektrum, sodass für deren Integration ein Line-Shape-Fitting Verfahren mittels der Software MATLAB® angewendet wurde (Arbeitsvorschrift 8.6.3).

Die entwickelte qNMR Methode wurde anschließend bezüglich Richtigkeit, Präzision, Linearität, Robustheit, Nachweis- (NG) und Bestimmungsgrenze (BG) überprüft (Arbeitsvorschrift 8.6.3). Die Richtigkeit ist ein Maß für die Abweichung des ermittelten Wertes vom wahren Wert aufgrund eines systematischen Fehlers und wird durch die Wiederfindung belegt. Die Präzision ist ein Maß für die zufällige Streuung der Messwerte um den wahren Wert und beinhaltet Wiederhol- und Vergleichspräzision. Die Wiederholpräzision beschreibt Schwankungen, die während der Probenvorbereitung entstehen können und kann durch eine mehrfache Probenvorbereitung einer Probe bestimmt werden. Die Vergleichspräzision oder auch Reproduzierbarkeit beschreibt zu der Wiederholpräzision die Schwankungen durch die Probenaufarbeitung und Messung durch verschiedene Personen oder Labore. Die Reproduzierbarkeit wird durch die Aufarbeitung und Messung einer Probe von mindestens zwei verschiedenen Personen bestimmt, wobei Werte von 10 – 20 % akzeptabel sind. Bei der Linearität wird der Konzentrationsbereich eines Analyten bestimmt, in welchem das Messergebnis proportional zur Analytkonzentration in der Probe ist. Die Bestimmung der Linearität erfolgt durch eine Regression des Signals gegen die Konzentration des Analyten. Die NG ist das Messergebnis, ab welchem mit hinreichender Sicherheit ausgesagt werden kann, dass der Analyt in der Probe tatsächlich vorhanden ist. Dahingegen beschreibt die BG das Messergebnis, bei welchem die Unsicherheit ausreichend klein, und damit eine Quantifizierung möglich ist [232]. NG und BG können nach DIN 32645 bestimmt werden. Die Robustheit gibt Auskunft darüber, ob das Messergebnis durch Änderungen der Bestimmungsbedingungen variiert. Sie kann bestimmt werden, indem bewusst Veränderungen vorgenommen werden, wie beispielsweise die Messung an einem anderen Analysengerät, eine Veränderung des pH-Wertes oder die Verwendung unterschiedlicher Chemikalien [232].

4 Ergebnisse und Diskussion

4.1 Probenmaterial

Für einen Metabolomics-Ansatz ist die Auswahl des Probenmaterials von großer Bedeutung. Der Pflanzenstoffwechsel zeichnet sich neben einer Vielzahl aus Stoffwechselwegen durch eine hohe Anpassungsfähigkeit an die Umgebung der Pflanze aus. Bei der Probenahme von Pflanzen im Rahmen eines Metabolomics-Ansatzes kann jedoch nur ein aktuelles Bild dieses vielfältigen Stoffwechsels erfasst werden. Um die Vielfältigkeit dieses Pflanzenmetaboloms erfassen zu können, ist daher die Messung vieler Proben, welche in unterschiedlichen Zeiträumen gesammelt wurden, erforderlich [135, 233, 234]. In der vorliegenden Arbeit wurden über einen Zeitraum von drei Jahren 688 Arabica-Kaffeeproben unterschiedlicher Varietät untersucht. Dabei wurden möglichst viele Metadaten über geographischen Ursprung sowie Anbau- und Erntebedingungen der Kaffeeproben mit aufgenommen. Neben 603 trommelgerösteten Kaffeeproben wurden 85 Rohkaffee-Proben mittels Heißluft-Röstverfahren geröstet. Es wurden ausschließlich geröstete Kaffeeproben untersucht, wodurch ein zusätzlicher Trocknungsschritt während der Probenvorbereitung nicht nötig war. Bedingt durch verschiedene Röstbedingungen der einzelnen Kaffeeproben, wiesen diese einen unterschiedlichen Feuchtegehalt und einen unterschiedlichen Röstgrad auf. Um die biologischen Variationen zu erfassen, wurde der Feuchtegehalt und der Röstgrad jeder Kaffeeprobe bestimmt. Die in dieser Arbeit verwendeten Kaffeeproben mit den jeweiligen Metadaten sind im Anhang in Tabelle 9-1 aufgelistet.

4.2 Entwicklung einer nicht zielgerichteten Methode für die Untersuchung der lipophilen Metaboliten

4.2.1 Extraktionsmethode

Das Ziel des nicht zielgerichteten Metabolomics-Ansatzes ist es, den Ausschluss von Metaboliten durch den Einsatz geeigneter Probenvorbereitungsverfahren weitgehend zu vermeiden, um so eine umfassende Analyse biologischer Systeme zu ermöglichen [233]. Je mehr Schritte bei der Probenvorbereitung durchgeführt werden, desto geringer wird die chemische Vielfalt der im Extrakt vorhandenen Verbindungen [235]. Für die Extraktion von Pflanzenmaterial wird daher eine schnelle, reproduzierbare und unselektive Extraktionsmethode bevorzugt, die möglichst viele verschiedene Metaboliten der zahlreich vorhandenen Primär- und Sekundärbestandteile herauslöst [187, 236]. Dazu wurden während der Entwicklungsphase für die Untersuchung der lipophilen Metaboliten zunächst verschiedene Extraktionsmethoden mit unterschiedlichen Lösungsmittelkombinationen getestet (Kapitel 3.3.1). Die Methoden nach Folch et al., Bligh und Dyer sowie die Methode nach Ackermann et al. verwenden neben unterschiedlichen Anteilen an Chloroform und Methanol zusätzlich Wasser [153, 154, 156]. Die Methode von Ackermann et al. zeichnet sich im Vergleich zu der Methode nach Folch et al. durch ein deutlich geringeres Verhältnis von Probe zu Extraktionsmittel aus [156]. Neben diesen Methoden wurde zusätzlich eine Extraktion nur mit Chloroform und einer Mischung aus Chloroform und Methanol (1:1, v/v) getestet. Auch wurden Extraktionen mit weniger toxischen Lösungsmitteln durchgeführt. Dazu wurde Aceton und eine Mischung aus Ethanol und Ethylacetat (1:1, v/v) als Extraktionsmittel verwendet [155]. Die Extrakte der verschiedenen Lösungsmittel und -gemische

wurden mittels $^1\text{H-NMR}$ vermessen. Die Extraktionsmethoden wurden anhand der $^1\text{H-NMR}$ -Spektren hinsichtlich Reproduzierbarkeit und Extraktion möglichst vieler Metaboliten miteinander verglichen. Ein Vergleich der $^1\text{H-NMR}$ -Spektren der verschiedenen Extraktionsmethoden ist in Abbildung 4-1 dargestellt.

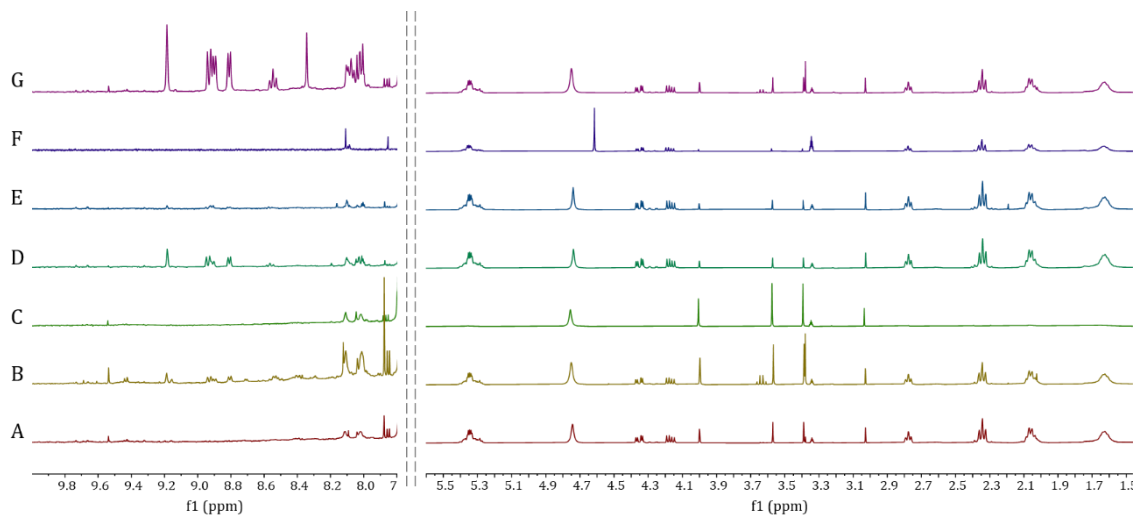


Abbildung 4-1: Ausschnitte aus dem Hoch- und Tieffeld im Bereich $\delta_{\text{H}} = 1,50 - 5,60$ ppm und $\delta_{\text{H}} = 7,80 - 9,80$ ppm der 400 MHz $^1\text{H-NMR}$ -Spektren von einer Kaffeeprobe, die mit verschiedenen Lösungsmitteln und Lösungsmittelgemischen extrahiert wurden. Die Messung erfolgte in $\text{CDCl}_3:\text{MeOH-}d_4$ 1:1 (v/v). Die chemische Verschiebung wurde auf $\delta_{\text{TMS}} = 0,00$ ppm referenziert. Extraktionsmethoden: A: Folch et al. (Chloroform:Methanol: 2:1 (v/v)); B: Ackermann et al. (Chloroform:Methanol:Wasser 10:5:1 (v/v/v)); C: Bligh und Dyer (Chloroform:Methanol 1:2 (v/v)); D: Lin et al. (Ethylacetat:Ethanol 1:1 (v/v)), E: Aceton; F: Chloroform; G: Chloroform:Methanol 1:1 (v/v)

Aus Abbildung 4-1 ist zu erkennen, dass durch die Anwendung der verschiedenen Extraktionsmethoden eine unterschiedliche Anzahl an Kaffeemetaboliten erfasst wurde. Mit Ausnahme der Methode nach Bligh und Dyer konnte eine Extraktion der Lipidbestandteile, deren Signale im Bereich $\delta_{\text{H}} = 0,50 - 5,50$ ppm auftreten, durch alle Extraktionsmethoden erzielt werden. Mit der Methode nach Bligh und Dyer wurden deutlich weniger lipophile Kaffeemetaboliten extrahiert. Durch eine Extraktion mit Chloroform und Methanol im Verhältnis 1:1 (v/v) konnten zusätzlich polarere Metaboliten erfasst werden, welche Signale im Bereich $\delta_{\text{H}} = 8,00 - 9,50$ ppm erzeugen. Dies konnte auch mit der Methode nach Ackermann et al. teilweise erreicht werden. Im Gegensatz dazu sind im $^1\text{H-NMR}$ -Spektrum der mittels Folch et al. extrahierten Kaffeeproben im Tieffeld nur wenige Signale ersichtlich. Bei der Methode nach Folch et al. wurde nach der Extraktion ein zusätzlicher Waschschriff durchgeführt, in welchem ein Verlust der zuvor extrahierten Metaboliten auftreten konnte. Auch durch die Verwendung der toxikologisch weniger bedenklichen Methode nach Lin et al. konnten polarere Metaboliten mit Signalen im Bereich von $\delta_{\text{H}} = 8,00 - 9,50$ ppm extrahiert werden. Eine Extraktion von Kaffee mit dem reinen Lösungsmittel Chloroform bzw. Aceton ergab $^1\text{H-NMR}$ -Spektren, welche deutlich weniger Signale im Tieffeld aufwiesen.

Um die getesteten Methoden bezüglich der Extraktionseffizienz zu vergleichen, wurden zunächst in den $^1\text{H-NMR}$ -Spektren zwölf Signale im Bereich $\delta_{\text{H}} = 0,50 - 7,30$ ppm ausgewählt. Anschließend wurde der Mittelwert der jeweiligen Signalfäche aus allen Extraktionsmethoden gebildet und die prozentuale Abweichung jeder Methode zum Mittelwert berechnet. Die prozentuale Abweichung vom Mittelwert aller getesteten Extraktionsmethoden für die zwölf ausgewählten Signale ist in Abbildung 4-2 dargestellt.

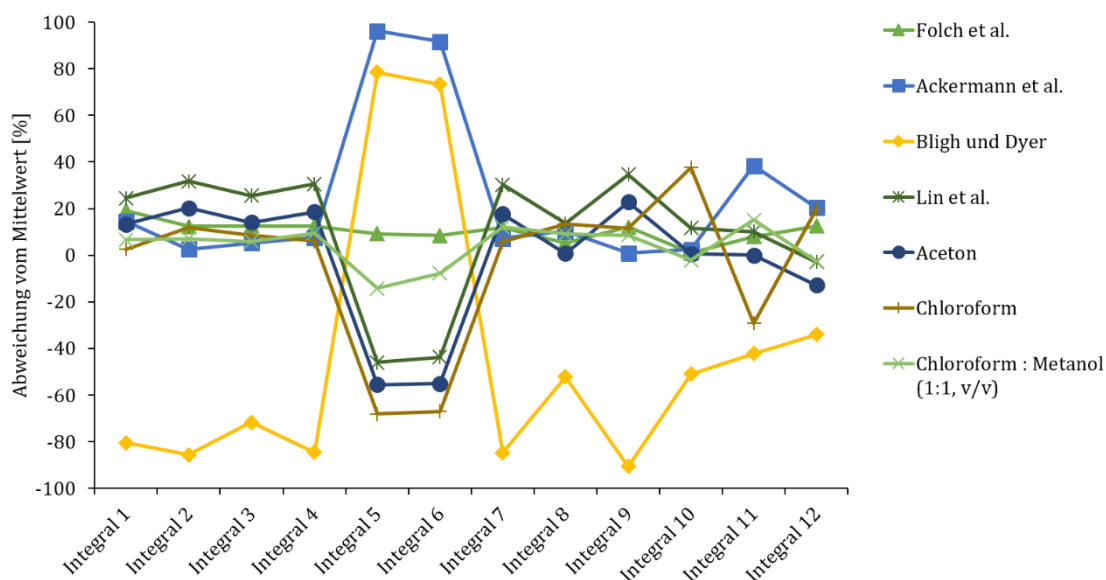


Abbildung 4-2: Prozentuale Abweichungen vom Mittelwert der entsprechenden Integrale ausgewählter Signale im Bereich $\delta_H = 0,5\text{--}7,3$ ppm. Für die Berechnung des Mittelwertes eines Integrals wurden die entsprechenden Integrale aller Extraktionsmethoden herangezogen. Extraktionsmethoden: Folch et al.: Chloroform: Methanol: 2:1 (v/v); Ackermann et al.: Chloroform: Methanol: Wasser 10:5:1 (v/v/v); Bligh und Dyer: Chloroform: Methanol 1:2 (v/v); Lin et al.: Ethylacetat: Ethanol 1:1 (v/v), Aceton, Chloroform, Chloroform, Chloroform: Methanol 1:1 (v/v). Integral 1: $\delta_H = 0,76 - 0,87$ ppm; Integral 2: $\delta_H = 1,22 - 1,34$ ppm; Integral 3: $\delta_H = 1,46 - 1,66$ ppm; Integral 4: $\delta_H = 2,68 - 2,78$ ppm; Integral 5: $\delta_H = 3,53 - 3,56$ ppm; Integral 6: $\delta_H = 3,96 - 4,00$ ppm; Integral 7: $\delta_H = 4,08 - 4,15$ ppm; Integral 8: $\delta_H = 4,21 - 4,26$ ppm; Integral 9: $\delta_H = 5,21 - 5,34$ ppm; Integral 10: $\delta_H = 5,84 - 5,90$ ppm; Integral 11: $\delta_H = 7,21 - 7,23$ ppm; Integral 12: $\delta_H = 7,25 - 7,29$ ppm

Die Methode nach Bligh und Dyer zeigte die geringste Extraktionseffizienz. Lediglich die Integrale der Signale 5 und 6, welche Signale von Koffein darstellen, sind im Vergleich zu den anderen Methoden erhöht. Dies deutet darauf hin, dass mit der Methode nach Bligh und Dyer besser polarere Lipidbestandteile aus der Kaffeematrix herausgelöst werden konnten. Ursprünglich wurde die Methode von Bligh und Dyer zur Extraktion von Fischproben mit einer hohen Konzentration von Phospholipiden verwendet [153]. Wie in Kapitel 1.1.7 beschrieben, besteht die Lipidfraktion von Kaffee zu einem großen Teil aus unpolaren Triacylglyceriden und Diterpenen, welche durch das relativ polare Extraktionsgemisch der Bligh und Dyer Methode nicht vollständig extrahiert werden konnten.

Auch mit der Methode nach Ackermann et al. konnten polarere Bestandteile wie Koffein gut extrahiert werden. Neben Koffein wies auch das Integral von Signal 11, welches der Verbindung Cafestol zugeordnet werden konnte, einen vergleichsweise hohen Integralwert auf. Nach genauerer Analyse der ^1H -NMR-Spektren war dieser erhöhte Wert jedoch auf eine Überlagerung des Signals zurückzuführen. Die Integrale der Signale 1 bis 4, welche gesättigten FA und UFA zugeordnet werden konnten, waren durch die Extraktion von Ackermann et al. vergleichbar zu den anderen getesteten Methoden. Die Anwendung der Methode nach Lin et al. mit toxikologisch weniger bedenklichen Lösungsmitteln führte zu einer effizienten Extraktion der unpolaren Lipidbestandteile. Durch diese Extraktion konnten vergleichsweise höhere Integralwerte der Signale 1 bis 4 sowie 7 bis 9, welche Glycerin und der sn-2 Position der Triacylglyceride zugeordnet werden konnten, erzielt werden. Lin et al. schlugen die Extraktion mit Ethanol und Ethylacetat als Alternative zur Methode nach Folch et al. zur Untersuchung von tierischen Lipiden vor und konnten diesbezüglich vergleichbare Ergebnisse erreichen [155]. Allerdings wurde die Fettsäureverteilung bezüglich der Triacylglyceride, Cholesterinester, Phospholipide und anderer Lipide untersucht. In dem Fettextrakt von Kaffee waren zusätzlich polarere Bestandteile wie beispielsweise Koffein

vorhanden, welche mit der Extraktion nach Lin et al. weniger gut aus der Matrix herausgelöst wurden. Die Integrale der Signale 5 bis 6 wiesen hier vergleichsweise geringere Werte auf. Auch die Lösungsmittel Aceton und Chloroform waren für die Extraktion von unpolaren Kaffeemetaboliten gut geeignet, jedoch bezüglich der Extraktion von polareren Metaboliten limitiert. Die Methode nach Folch et al. und eine Extraktion mit Chloroform und Methanol (1:1, v/v) führten zu einer relativen guten Extraktion sowohl der unpolaren Metaboliten wie den Triacylglyceriden oder Diterpenen, als auch den polareren Bestandteile wie Koffein. Die Integrale der Signale 5 und 6 wiesen mittlere Werte im Vergleich zu den anderen getesteten Methoden auf.

Neben der Extraktion möglichst vieler Metaboliten bei einer guten Extraktionseffizienz ist bei zielgerichteten und nicht zielgerichteten Untersuchungen eine hohe Reproduzierbarkeit wichtig [152]. Um eine Vergleichbarkeit der Reproduzierbarkeit zu erhalten, wurden für jede getestete Methode eine Dreifachbestimmung durchgeführt und der Variationskoeffizient (VK) für einzelne Signale gebildet (Tabelle 4-1). Zusätzlich wurde der mittlere VK über alle Signale für die entsprechende Methode berechnet.

Tabelle 4-1: Verfahrensvariationskoeffizienten (VKs) für die Bestimmung von 12 ausgewählten Signalen im Bereich $\delta_H = 0,5-7,3$ ppm und mittlere VKs für die getesteten Extraktionsmethoden: Folch et al.: Chloroform: Methanol: 2:1 (v/v); Ackermann et al.: Chloroform: Methanol: Wasser 10:5:1 (v/v/v); Bligh und Dyer: Chloroform: Methanol 1:2 (v/v); Lin et al.: Ethylacetat: Ethanol 1:1 (v/v), Aceton, Chloroform, Chloroform: Methanol 1:1 (v/v)

	Folch et al.	Ackermann et al.	Bligh und Dyer	Lin et al.	Aceton	Chloroform	Chloroform: Methanol (1:1, v/v)
Signal 1	0,8	0,9	3,6	0,7	1,4	1,0	0,8
Signal 2	0,7	1,1	3,7	0,7	1,3	1,1	0,7
Signal 3	0,5	1,1	3,3	0,7	1,3	1,0	0,7
Signal 4	0,6	1,2	3,2	0,7	1,3	0,9	0,6
Signal 5	4,9	0,9	2,9	1,2	1,3	7,1	0,9
Signal 6	4,8	0,9	3,1	1,3	1,4	7,3	0,7
Signal 7	0,6	1,2	3,1	0,7	1,3	0,9	0,6
Signal 8	0,5	1,3	3,3	0,7	1,3	7,4	0,6
Signal 9	0,6	1,2	4,1	0,6	1,4	1,0	0,6
Signal 10	0,3	1,4	3,3	0,6	1,4	0,9	0,7
Signal 11	0,4	1,0	2,7	0,8	1,3	0,3	0,7
Signal 12	1,0	1,5	3,1	0,9	1,4	0,8	0,7
Ø VK	1,3	1,1	3,3	0,8	1,3	2,5	0,7

Die VKs lagen für alle getesteten Methoden in einem akzeptablen Bereich. Lediglich bei der Extraktion mit Chloroform waren die VKs für die Signale fünf, sechs und acht relativ hoch. Im Durchschnitt besaß die Methode nach Bligh und Dyer den höchsten VK, wohingegen die Extraktion mit der Mischung aus Chloroform und Methanol (1:1, v/v) den geringsten VK von 0,7% aufwies. Neben dem geringen VK wies eine Extraktion mit der Mischung aus Chloroform und Methanol (1:1, v/v) auch den Vorteil auf, viele Metaboliten aus der Kaffeematrix, bei einer jedoch vergleichsweise mittleren Extraktionseffizienz, herauszulösen. Im Vergleich zu den Methoden nach Folch et al., Ackermann et al. und Bligh und Dyer war zudem die Probenvorbereitung weniger aufwändig, da kein zusätzlicher Waschschritt oder eine Phasentrennung durchgeführt werden musste. Eine Extraktion nach Lin et al. wäre eine gute Alternative zu der Extraktion mit Chloroform und Methanol. Auch die Methode nach Lin et al. wies einen geringen VK von durchschnittlich 0,8% auf und konnte relativ viele Metaboliten aus der Kaffeematrix herauslösen. Die Extraktionseffizienz war für

die unpolaren Metaboliten vergleichsweise höher, jedoch für die polarerer Metaboliten geringer. Nachteil der Methode nach Lin et al. waren die anschließend hohen Abdampfzeiten, welche zur vollständigen Entfernung von Ethanol benötigt wurden.

Nach der Extraktion der Probe und dem Entfernen des Lösungsmittels (Arbeitsvorschrift 8.2.3) waren bei Raumtemperatur im Rückstand der Probe feste sowie flüssige Anteile vorhanden. Dies erschwerte eine genaue und reproduzierbare Einwaage des Kaffee-fetts. Aus diesem Grund wurde nach der Extraktion ein Aliquot des Extraktes von 2 mL abgenommen, membranfiltriert und vom Filtrat das Lösungsmittel entfernt. Der Rückstand wurde anschließend ausgewogen, sodass definierte Mengen der lipophilen Metaboliten für die $^1\text{H-NMR}$ -Messung eingesetzt wurden. Zur Erhöhung der Konzentration der Metaboliten in dem Extrakt wurde die Probeneinwaage der Kaffeeproben von 0,5 g auf 0,575 g erhöht und das Volumen des Extraktionsmittels (Chloroform und Methanol, 1:1, v/v) von 10 mL auf 8 mL herabgesetzt (Arbeitsvorschrift 8.2.3). Zur Überprüfung der Zeitabhängigkeit der Extraktion wurden Versuche mit Extraktionszeiten von 30, 60, 90, 120 und 180 min durchgeführt. Aus den resultierenden $^1\text{H-NMR}$ -Spektren wurden zwölf Signale im Bereich $\delta_{\text{H}} = 0,80 - 9,20$ ppm ausgewählt und integriert. Ein Vergleich dieser Integrale zeigte, dass durch eine längere Extraktionszeit eine höhere Ausbeute dieser Kaffeemetaboliten erzielt wurde (Abbildung 9-1(A)). Auch das Gesamtintegral, welches aus der Summe der Signale eins bis zwölf gebildet wurde, zeigte mit längerer Extraktionszeit einen höheren Wert (Abbildung 9-1(B)). Allerdings wies der mittlere VK bei einer Extraktionszeit von 60 min einen bereits sehr geringen Wert von 0,7% auf (Tabelle 9-4). Dies zeigte, dass bei einer Extraktionszeit von 60 min reproduzierbare Ergebnisse erzielt werden konnten. Eine Extraktion unter zusätzlicher Verwendung von Glasperlen erzielte höhere Integralwerte verschiedener Signale (Abbildung 9-2). Aufgrund der erzielten Ergebnisse wurde eine Extraktion mit einer Mischung aus Chloroform und Methanol (1:1, v/v) und einer Extraktionszeit von 60 min unter Verwendung von Glasperlen festgelegt (Arbeitsvorschrift 8.2.3).

4.2.2 Vorbereitung der Messlösung für die $^1\text{H-NMR}$ -spektroskopische Messung

Zur Optimierung der chemischen Verschiebungen und der Auflösung der Metaboliten-signale wurden verschiedene Kombinationen deuterierter Lösungsmittel für die Messlösung getestet (Kapitel 3.3.2). Dazu wurde der vom Lösungsmittel befreite Rückstand der Extraktion in CDCl_3 , in Gemischen von CDCl_3 und $\text{MeOH-}d_4$ in unterschiedlichen Verhältnissen und in einer Mischung aus CDCl_3 und $\text{DMSO-}d_6$ (5:1, v/v) gelöst. Abbildung 4-3 zeigt einen Ausschnitt der erhaltenen $^1\text{H-NMR}$ -Spektren im Bereich $\delta_{\text{H}} = 7,00 - 11,5$ ppm. Durch die Verwendung des häufig zur Untersuchung von Kaffee eingesetzten CDCl_3 [89, 90, 90, 237, 238] konnten insbesondere die polarerer Kaffeemetaboliten nicht gelöst werden. Aus Abbildung 4-3 ist zu erkennen, dass vergleichsweise weniger Signale im Tieffeld des $^1\text{H-NMR}$ -Spektrums auftraten. Auch durch die Verwendung der Mischung aus CDCl_3 und $\text{DMSO-}d_6$ (5:1, v/v) trat ein Verlust der polarerer Kaffeemetaboliten während des Lösens auf. Eine Mischung aus CDCl_3 und $\text{MeOH-}d_4$ führte dazu, dass alle bei der Extraktion gewonnenen Kaffeemetaboliten wieder vollständig gelöst werden konnten. Die Auflösung der Signale blieb dabei in unabhängig von dem Mischungsverhältnis aus CDCl_3 und $\text{MeOH-}d_4$.

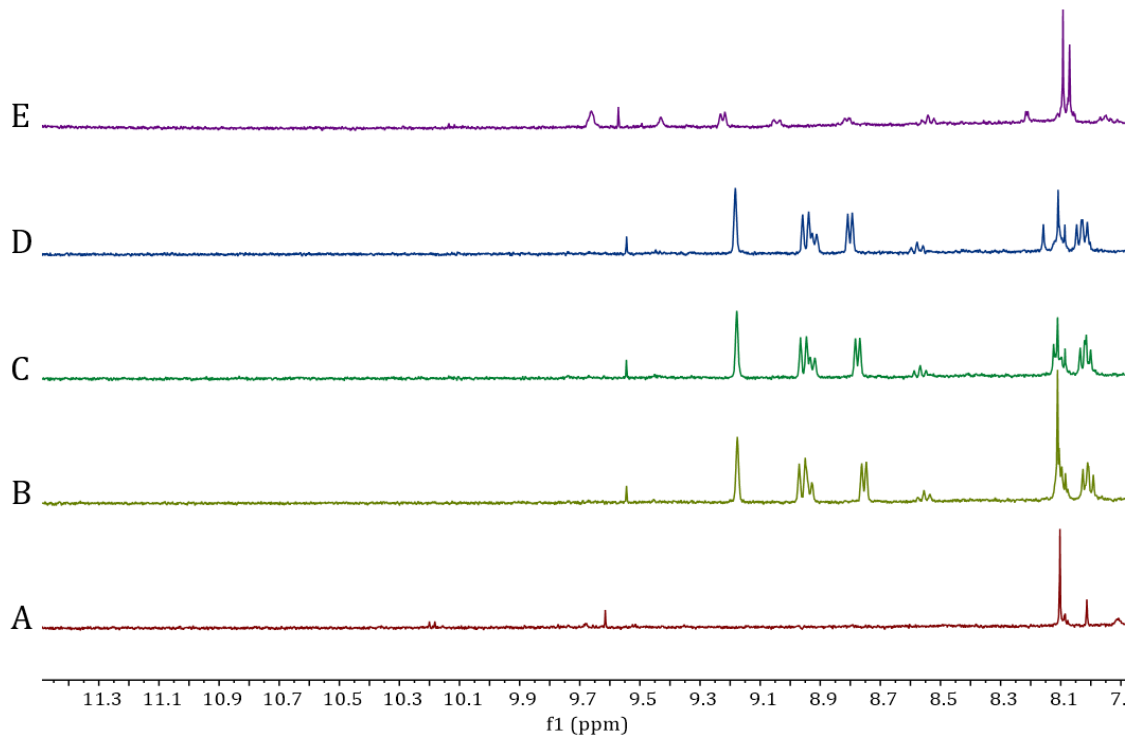


Abbildung 4-3: Ausschnitte verschiedener 400 MHz ^1H -NMR-Spektren im Bereich $\delta_{\text{H}} = 7,00 - 11,5$ ppm der lipophilen Kaffeemetaboliten, gemessen in verschiedenen Lösungsmitteln bzw. Lösungsmittelgemischen. Die chemische Verschiebung wurde auf $\delta_{\text{TMS}} = 0,00$ ppm referenziert. Messlösung: A: CDCl_3 ; B: $\text{CDCl}_3 : \text{MeOH-}d_4$ (2:1, v/v); C: $\text{CDCl}_3 : \text{MeOH-}d_4$ (3:2, v/v); D: $\text{CDCl}_3 : \text{MeOH-}d_4$ (1:1, v/v); E: $\text{CDCl}_3 : \text{DMSO-}d_6$ (5:1, v/v)

Das Lösen des evaporierten Extraktes in einer Mischung aus CDCl_3 und $\text{MeOH-}d_4$ im Verhältnis 2:1 führte im ^1H -NMR-Spektrum zu einer Überlagerung der Analytsignale durch Restwasser (Abbildung 9-3). Eine Erhöhung des Anteils von $\text{MeOH-}d_4$ zu einem Verhältnis von $\text{CDCl}_3 : \text{MeOH-}d_4$ 1:1 (v/v) führte zu einer Tieffeldverschiebung des Restwasser-Signals, wodurch eine Überlagerung mit Analytsignalen vermieden werden konnte.

Wie in Kapitel 3.3.2 beschrieben, besteht die Möglichkeit, dass sich CDCl_3 unter der Bildung von reaktiven Produkten, wie Phosgen oder Salzsäure, zersetzen kann. Unerwünschte Kontaminanten wie Phosgen in CDCl_3 können zur Bildung von Artefakten beitragen [161]. Daher wurde der Einfluss der Abbauprodukte von CDCl_3 auf die Erfassung der Kaffeemetaboliten überprüft. Dazu wurde eine Kaffeeprobe an verschiedenen Tagen aufgearbeitet und vermessen, wobei für die Messlösung jeweils unbehandeltes und entschärftes CDCl_3 (Arbeitsvorschrift 8.2.4 B) verwendet wurde. Ein Vergleich der resultierenden ^1H -NMR-Spektren zeigte, dass die Abbauprodukte des CDCl_3 einen Einfluss auf die Erfassung der Metaboliten hatten (Abbildung 9-4). Durch die Verwendung des unbehandelten CDCl_3 traten in den ^1H -NMR-Spektren vermehrt Signal-Shifts im Tieffeldbereich auf (Abbildung 9-4(A)). Der starke Shift des Ameisensäuresignals bei $\delta_{\text{H}} = 8,34$ ppm ließ darauf schließen, dass das Vorhandensein von Salzsäure eine Änderung des pH-Wertes zur Folge hatte. Durch die Anwendung von entschärftem CDCl_3 konnten die auftretenden Signal-Shifts deutlich reduziert werden (Abbildung 9-4 (B)). Somit wurde der vom Extraktionsmittel befreite Rückstand in einer Mischung aus $\text{CDCl}_3 : \text{MeOH-}d_4$ 1:1 (v/v) gelöst, wobei mögliche Abbauprodukte aus dem CDCl_3 entfernt wurden.

4.2.3 Optimierung der Messparameter für die ^1H -NMR-spektroskopische Messung

Um die optimale Erfassung der Kaffeemetaboliten sicherzustellen, wurden der Pulswinkel und die Messtemperatur überprüft sowie ein Inversion Recovery Experiment zur Bestimmung der T_1 -Zeit durchgeführt. Für einen Pulswinkel von 90° wird eine Repetitionszeit von Faktor fünf der längsten vorhandenen T_1 -Zeit empfohlen. Eine Repetitionszeit von 5-10 s, wie sie in vielen Metabolomics-Ansätzen angewendet wird, ist oft nicht ausreichend, um eine vollständige Relaxation aller Kerne zu gewährleisten [187]. Die Verringerung des Pulswinkels auf 30° führt zu der Verwendung einer kürzeren Repetitionszeit, sodass beispielsweise die Anzahl der Scans heraufgesetzt werden kann, ohne die Gesamtmesszeit zu verlängern. Jedoch resultiert ein geringerer Pulswinkel auch in einer verringerten Intensität der Signale. Mit dem Ziel, eine möglichst hohe Intensität der Metabolitensignale mit einem hohen S/N bei kurzer Messzeit zu erhalten, wurden ein Pulswinkel von 90° und einer von 30° miteinander verglichen. Dazu wurden in den erhaltenen ^1H -NMR-Spektren zwölf Signale integriert und die Flächen miteinander verglichen. Durch die Verwendung eines Pulswinkels von 30° bei gleichbleibender Scanzahl waren die Integrale im Vergleich zu einem Pulswinkel von 90° deutlich reduziert (Abbildung 9-5(A)). Durch eine Verringerung der Repetitionszeit und eine Erhöhung der Scanzahl konnten mit einem Pulswinkel von 30° Integrale erzielt werden, welche im Vergleich zu einem Pulswinkel von 90° nur leicht verringert waren (Abbildung 9-5(A)). Wie für die Integrale wurde auch das höchste S/N bei einem Pulswinkel von 90° erzielt (Abbildung 9-5(B)).

Die Überprüfung der Messtemperatur erfolgte, indem eine Kaffeeprobe bei 280 K, 290 K und 300 K vermessen wurde. Ein Vergleich der erhaltenen ^1H -NMR-Spektren zeigte eine Tieffeldverschiebung des Restwassersignals bei Verringerung der Messtemperatur (Abbildung 4-4).

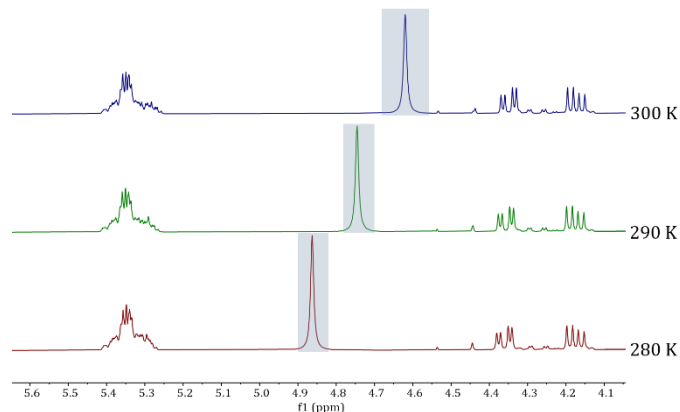


Abbildung 4-4: Ausschnitte verschiedener 400 MHz ^1H -NMR-Spektren im Bereich $\delta_{\text{H}} = 4,10 - 5,60$ ppm der lipophilen Kaffeemetaboliten, gemessen bei verschiedenen Temperaturen. Die chemische Verschiebung wurde auf $\delta_{\text{TMS}} = 0,00$ ppm referenziert. Der grau hinterlegte Bereich stellt das Signal des Restwassers dar.

Durch eine Verringerung der Messtemperatur von 300 K auf 280 K trat eine Tieffeldverschiebung des Restwassersignals um 0,25 ppm auf, durch welche eine Überlagerung des Restwassersignals mit Analytsignalen vermieden werden konnte. Bereits durch die Verringerung der Messtemperatur auf 290 K konnte diese Überlagerung vermieden werden und benötigte im Gegensatz zu einer Messtemperatur von 280 K eine geringere Zeit zum Herunterkühlen des Probenkopfes.

Mittels eines Inversion Recovery Experiments wurden die T_1 -Zeiten der in höheren Konzentrationen vorkommenden Kaffeemetaboliten bestimmt. Anhand der längsten vorhandenen T_1 -Zeit erfolgte die Bestimmung der Repetitionszeit. Die Auswertung des Inversion Recovery Experimentes bei der Analyse der lipophilen Metaboliten ergab T_1 -Zeiten der Protonen von 0,24 – 2,70 s (Tabelle 9-5). Für die CH_2 -Gruppe der FA sowie für die Protonen des GlycerinGrundgerüsts der Triacylglyceride wurden kurze T_1 -Zeiten von 240 – 650 ms ermittelt, wohingegen die CH_2 -Gruppe in α -Position der Carboxylgruppe und die endständige CH_3 -Gruppe der FA längere T_1 -Zeiten von 2,14 s und 1,41 s aufwiesen. Die längste T_1 -Zeit mit 2,70 s wurde für ein Proton des Koffeins bestimmt.

Ein Vergleich verschiedener Pulswinkel und Messtemperaturen sowie eine Bestimmung der T_1 -Zeit der in höheren Konzentrationen vorkommenden Kaffeemetaboliten wurde ergab eine Messung mit einem Pulswinkel von 90° bei einer Temperatur von 290 K und einer D1-Zeit von 15 s.

4.2.4 Überprüfung der Reproduzierbarkeit der Extraktion und der Messung

Um sicherzustellen, dass die Extraktion und nachfolgende Messung der Kaffeeproben für die nicht zielgerichtete Untersuchung der lipophilen Kaffeemetaboliten für die große Probenanzahl reproduzierbar ist, wurde eine Kontrollprobe mitgeführt. Diese Kontrollprobe war ein biologisch angebauter Kaffee ohne Herkunftsangabe. In der Vorperiode wurde diese Kontrollprobe fünfmal an jeweils drei verschiedenen Tagen von zwei unterschiedlichen Personen aufgearbeitet und vermessen. Aus den resultierenden ^1H -NMR-Spektren wurden über einen Bereich von $\delta_{\text{H}} = 0,90 - 9,20$ ppm zehn Signale ausgewählt und diese automatisch integriert. Dabei wurden für die Integration der Signale festgelegte Bereiche verwendet. Die absoluten Integrale der einzelnen Signale wurden aufsummiert und der prozentuale Anteil jedes Signals bezogen auf das Gesamtintegral bestimmt. Anhand der mehrfachen Aufarbeitung und Messung wurden der Mittelwert und die Standardabweichung gebildet. Anschließend wurde eine obere bzw. untere Warngrenze durch Addition bzw. Subtraktion der zweifachen Standardabweichung vom Mittelwert berechnet. Die obere und untere Kontrollgrenze ergab sich aus der Addition und Subtraktion der dreifachen Standardabweichung vom entsprechenden Mittelwert jedes Signals. Anschließend wurde bei jeder Messserie eine Kontrollprobe mit aufgearbeitet und vermessen. Die zehn Signale wurden in jedem Spektrum der Kontrollprobe innerhalb der festgelegten Bereiche integriert. Daraus wurden die prozentualen Anteile bezogen auf das Gesamtintegral gebildet. Die berechneten prozentualen Anteile für jedes Integral wurden in die Kontrollkarte eingetragen. Abbildung 4-5 stellt die Kontrollkarte von Integral 1 und Integral 10 dar. Aus der Abbildung sowie den Kontrollkarten der anderen Integrale ist ersichtlich, dass die unpolaren Metaboliten reproduzierbarer extrahiert und vermessen wurden als die polaren Metaboliten (Abbildung 9-6 und Abbildung 9-7). Die unpolaren Metaboliten wiesen sehr enge Warn- und Kontrollgrenzen auf, wohingegen diese bei den polaren Metaboliten breiter angesetzt waren. Jedoch wurde für keine Kontrollprobe die Kontrollgrenze überschritten.

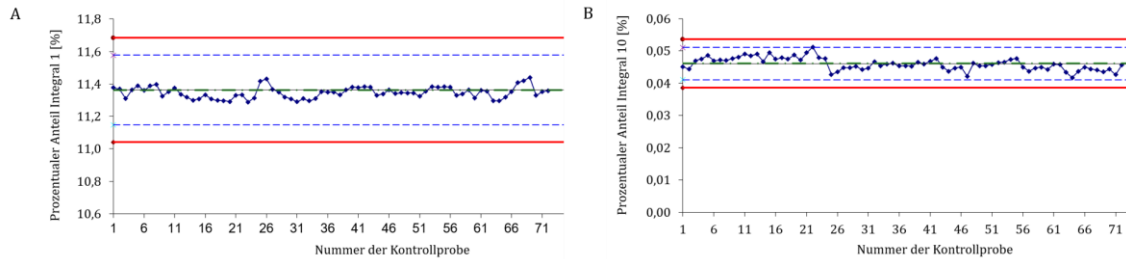


Abbildung 4-5: Kontrollkarten von (A) Integral 1 ($\delta_H = 0,85 - 0,90$ ppm) und (B) Integral 10 ($\delta_H = 9,16 - 9,21$ ppm) der ausgewählten Signale für die lipophilen Kaffeemetaboliten für insgesamt 74 Kontrollproben. Aus den Integralen der zehn Signale wurde das Gesamtintegral berechnet und anschließend der prozentuale Anteil eines jeden Integrals bezogen auf das Gesamtintegral bestimmt. Die grüne Linie zeigt den Mittelwert. Die rote, durchgezogene Linie stellt die Kontrollgrenze dar (Mittelwert ± 3 ·Standardabweichung). Die blaue, gestrichelte Linie beschreibt die Warngrenze (Mittelwert ± 2 ·Standardabweichung).

4.3 Entwicklung einer nicht zielgerichteten Methode für die Untersuchung der hydrophilen Metaboliten

4.3.1 Entwicklung der Extraktionsmethode

Wie in Kapitel 3.4.1 beschrieben, führt die Selbstassoziation von Koffein sowie die Bildung eines π -Komplexes mit CGA zu Shifts der entsprechenden Signale im $^1\text{H-NMR}$ -Spektrum [163, 164]. Um das Ausmaß dieser Shifts zu minimieren, wurden verschiedene Probeneinwaagen bei gleichbleibendem Volumen des Extraktionsmittels getestet. Abbildung 4-6 zeigt den Tieffeldbereich verschiedener $^1\text{H-NMR}$ -Spektren, bei welchen eine unterschiedliche Probeneinwaage einer Kaffeeprobe erfolgte. Deutlich erkennbar sind die Shifts der Koffein- und CGA-Signale im Bereich $\delta_H = 6,10 - 8,00$ ppm, welche aufgrund der unterschiedlichen Probenkonzentration verursacht wurden.

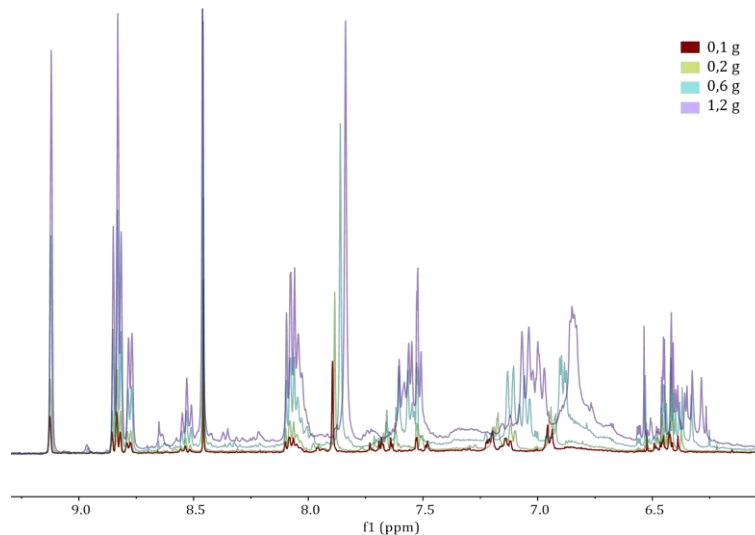


Abbildung 4-6: Ausschnitt verschiedener 400 MHz $^1\text{H-NMR}$ -Spektren im Bereich $\delta_H = 6,00 - 9,30$ ppm einer Kaffeeprobe, gemessen in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die Lösungsmittelunterdrückung erfolgte mit dem NOESY-Pulsprogramm. Die chemische Verschiebung wurde auf $\delta_{\text{TSP}} = 0,00$ ppm referenziert. Getestet wurden die Probeneinwaagen 0,10 g, 0,20 g, 0,60 g und 1,20 g. Die Extraktion erfolgte in jeweils 8,00 mL H_2O .

Auch der Vergleich zweier Kaffeeproben mit unterschiedlichen Röstgraden verdeutlicht insbesondere einen Shift der Koffeinsignale bei einem steigenden Verhältnis von Probe zu Extraktionsmittel (Abbildung 9-8). Da Signalverschiebungen die multivariate Datenauswertung der $^1\text{H-NMR}$ -Spektren erschweren könnten, wurde die Probeneinwaage

möglichst gering gewählt. Eine Probeneinwaage von 0,10 g führte jedoch zu $^1\text{H-NMR}$ -Spektren, in welchen Minorsignale ein zu geringes S/N aufwiesen. Daher wurde eine Probeneinwaage von 0,20 g gewählt und zusätzlich der Receiver Gain während der Erfassung der hydrophilen Metaboliten erhöht. In bisherigen nicht zielgerichteten Untersuchungen zu Kaffee wurde ein deutlich höheres Verhältnis von Probe zu Extraktionsmittel gewählt, wobei auftretende Shifts der Koffeinsignale nicht erwähnt wurden [144, 145, 150, 239]. Toci et al. beobachteten Shifts der Signale von NMP und den CGA und berücksichtigten diese durch Verwendung von breiteren Buckets [150].

Zur Sicherstellung einer reproduzierbaren Extraktion wurden die verschiedenen Extraktionszeiten 20, 30, 60, 90, 120, 150 und 180 min getestet. Aus den resultierenden $^1\text{H-NMR}$ -Spektren wurden 13 Signale im Bereich von $\delta_{\text{H}} = 1,00 - 9,10$ ppm ausgewählt und integriert. Der Vergleich der Integrale zeigte, dass eine Extraktion für 20 Minuten bereits ausreichend war, um die hydrophilen Kaffeemetaboliten vollständig aus der Matrix herauszulösen (Abbildung 9-9(A)). Lediglich bei Signal 2 führte eine längere Extraktion zu höheren Integralwerten. Ein Vergleich der entsprechenden $^1\text{H-NMR}$ -Spektren zeigte jedoch, dass sich durch eine längere Extraktionszeit das breite Signal der Proteine erhöhte, welches Signal 2 überlagert. Der dadurch begründete steigende Integralwert von Signal 2 führte weiterhin dazu, dass auch das Gesamtintegral mit steigender Extraktionszeit zunahm (Abbildung 9-9(B)). Nach 20 min Extraktionszeit wurde ein mittlerer VK von 0,5% erreicht, welcher auch durch eine längere Extraktionszeit nicht weiter reduziert werden konnte. Eine Extraktion unter zusätzlicher Verwendung von Glasperlen ergab keine Verbesserung bezüglich der Ausbeute und Reproduzierbarkeit (Abbildung 9-10). Bei einer Extraktion von 0,2 g Kaffeeprobe mit 8 mL H_2O und einer Extraktionszeit von 20 min wurde daher auf die Verwendung von Glasperlen verzichtet.

Die in den extrahierten Kaffeeproben enthaltenen Kohlenhydrate könnten die nachfolgende zielgerichtete und insbesondere die nicht zielgerichtete Untersuchung negativ beeinflussen. Grund hierfür ist der Abbau der Polysaccharide während der Röstung, welcher von der Varietät und den Röstbedingungen beeinflusst wird und daher im unterschiedlichen Ausmaß stattfinden kann [53]. Bedingt durch die Untersuchung von Kaffeeproben verschiedener Varietät und unterschiedlichen Röstgrads könnte das Vorhandensein unterschiedlicher Polysaccharidstrukturen, auch in den Melanoidinen, die nicht zielgerichtete Untersuchung hinsichtlich des geographischen Ursprungs erschweren. Zusätzlich erzeugen die Protonen der Polysaccharide Signalberge im $^1\text{H-NMR}$ -Spektrum, welche andere Metabolitensignale überlagern. Daher wurden für die Entfernung der Polysaccharide und Melanoidine aus der extrahierten Kaffeeprobe eine UF mit Membranen verschiedener Porengröße getestet. Bekedam et al. verwendeten zur Untersuchung von Melanoidinen eine Diafiltration mit einer 3 kDa Membran [51]. Auch Gniechwitz et al. nutzen zur Isolierung und Charakterisierung der hochmolekularen Bestandteile eines Kaffeeextraktes unter anderem eine UF, wobei die Membranen eine Porengröße von 2, 3, 10, 50 und 100 kDa besaßen [41, 240]. Zur Bestimmung der Gesamtpolysaccharide in Kaffee verwendeten Redgwell et al. eine Dialysemembran mit einer Porengröße von 3,5 kDa [53]. Mit dem Ziel auch kleinere Polysaccharide abzutrennen, wurden in dieser Arbeit Membranen mit einer Porengröße von 1, 2, 3 und 5 kDa getestet. Zusätzlich zu der Porengröße wurden bei den Membranen verschiedene Materialien untersucht. Bei der 1 kDa Membran wurde ein modifiziertes Polyethersulfon eingesetzt, wohingegen bei der Porengröße von 2 kDa eine stabilisierte Membran auf Cellulosebasis (Hydrosart®) eingesetzt wurde. Die 3 und 5 kDa Membranen bestanden jeweils aus Polyethersulfon. Die Membranen verschiedener Porengröße und unterschiedlichen Membranmaterials wurden anschließend hinsichtlich

Effektivität der Kohlenhydratentfernung und der Glättung der Basislinie im resultierenden ^1H -NMR-Spektrum miteinander verglichen. Bei allen getesteten Membranen war jedoch Glycerin vorhanden, welches durch Waschen mit Wasser vor der Filtration der Proben entfernt werden sollte, um dadurch unerwünschte Signale im ^1H -NMR-Spektrum zu vermeiden [166, 167]. Durch fünf Waschschrte konnte das Glycerin bei den Filtern mit einer Porengröße von 1 und 2 kDa vollständig entfernt werden (Abbildung 4-7 (B) und (C)). In ^1H -NMR-Spektren der Kaffeeproben, welche mit den Filtern mit einer Porengröße von 3 und 5 kDa behandelt wurden, traten nach fünf Waschschrten weiter Signale von Glycerin im Bereich $\delta_{\text{H}} = 3,5\text{-}3,7$ ppm auf (Abbildung 4-7 (D) und (E)).

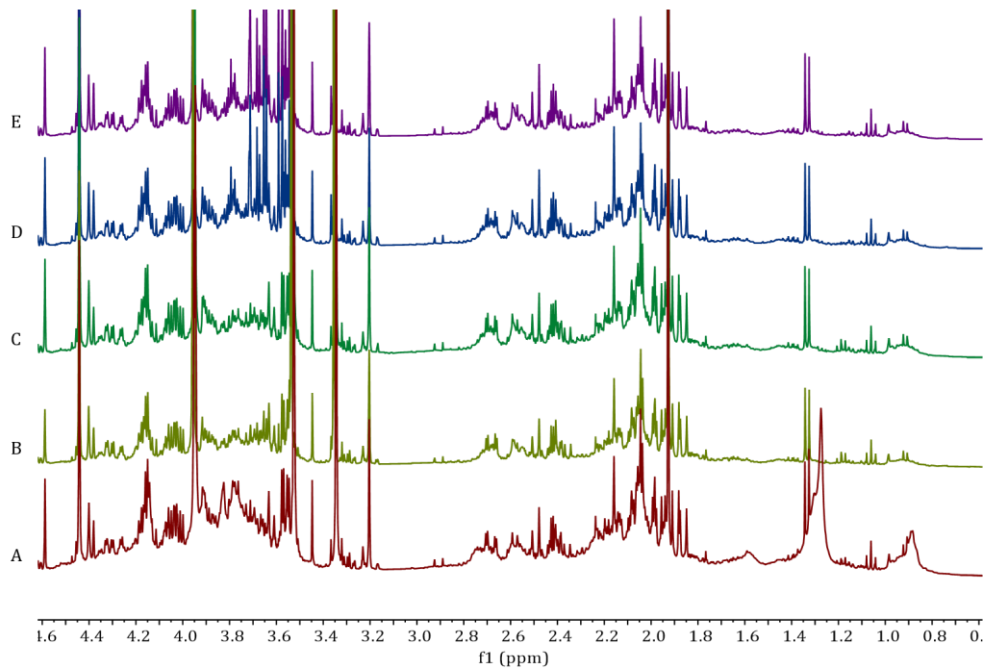


Abbildung 4-7: Ausschnitte verschiedener 400 MHz ^1H -NMR-Spektren im Bereich $\delta_{\text{H}} = 0,60 - 4,60$ ppm einer Kaffeeprobe aufgearbeitet mittels Ultrafiltration unter Verwendung verschiedener Membranen, gemessen in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die Lösungsmittelunterdrückung erfolgte mit dem NOESY-Pulsprogramm. Die chemische Verschiebung wurde auf $\delta_{\text{TSP}} = 0,00$ ppm referenziert. A: Kaffeeextrakt membranfiltriert (GF/PET, 25 mm, $0,2 \mu\text{m}$); B: Verwendung einer Membran mit der Porengröße 1 kDa; C: Verwendung einer Membran mit der Porengröße 2 kDa; D Verwendung einer Membran mit der Porengröße 3 kDa; E: Verwendung einer Membran mit der Porengröße 5 kDa

In Abbildung 4-7 kennzeichnet (A) das membranfiltrierte Kaffeeextrakt. Die Signalberge in den Bereichen $\delta_{\text{H}} = 1,80 - 2,30$ ppm, $\delta_{\text{H}} = 3,50 - 4,40$ ppm und $\delta_{\text{H}} = 1,20 - 1,40$ ppm sind deutlich erkennbar. Die Proteine konnten durch die Anwendung aller Zentrifugenfilter entfernt werden. Im Bereich $\delta_{\text{H}} = 1,20 - 1,40$ ppm ist in den Spektren (B)-(E) der Signalberg, verursacht durch die Proteine, nicht mehr erkennbar. Die effektivste Abtrennung der Kohlenhydrate konnte durch eine Porengröße von 1 kDa erreicht werden. Spektrum (B), welches die Anwendung der 1 kDa Membran darstellt, weist die beste Glättung der Basislinie auf. Durch die Verwendung einer 2 kDa Membran (Spektrum (C)) konnte eine geringfügig schlechtere Abtrennung der Kohlenhydrate erreicht werden. Dies deutet an, dass durch eine UF mit Zentrifugenfilter mit einer Porengröße von 1 und 2 kDa hochmolekulare Melanoidine entfernt werden konnten. Durch die Anwendung der Zentrifugenfilter mit dieser Porengröße konnten die Signalberge insbesondere im Bereich $\delta_{\text{H}} = 3,50 - 4,40$ ppm reduziert, jedoch nicht vollständig entfernt werden. Grund hierfür könnten die teilweise abgebauten Polysaccharide und niedermolekulare Melanoidine sein, welche nach der Röstung vorlagen und durch die UF nicht vollständig entfernt werden konnten. In Spektrum (D) und (E), welche jeweils durch die Anwendung der 3 und 5 kDa

Membran entstanden sind, sind neben den Signalen der Kohlenhydrate zusätzlich Signale von Glycerin erkennbar. Bei diesen beiden Filtern wären zusätzliche Waschschriffe nötig, um das Glycerin vollständig aus der Membran zu entfernen.

Die Auswirkung der Entfernung der hochmolekularen Melanoidine und die damit verbundene teilweise Entfernung der Kohlenhydrate auf die nicht zielgerichtete Untersuchung der hydrophilen Kaffeemetaboliten bezüglich des geographischen Ursprungs wurde anhand von 75 Kaffeeproben untersucht. Dazu wurden die Proben jeweils mit einer Membranfiltration und der UF unter Verwendung von Filter mit einer Porengröße von 1 und 2 kDa aufgearbeitet und vermessen. Die resultierenden ^1H -NMR-Spektren wurden anschließend multivariat mittels PCA-LDA ausgewertet und die erhaltenen Klassifizierungsmodelle miteinander verglichen (Abbildung 9-11). Die Anwendung des Membranfilters und der UF mit einem 2 kDa Zentrifugenfilter resultierten in sehr ähnlichen Klassifizierungsmodellen. Durch die Anwendung der UF mit einer 1 kDa Membran sank die richtige Zuordnung der Testproben in der internen Validierung. Wie in Kapitel 1.1.7.4 beschrieben, werden neben Kohlenhydrate auch eine Vielzahl weitere Verbindungen in Melanoidine eingebaut. Durch die Abtrennung der Melanoidine wurden folglich auch weitere Verbindungen aus dem Kaffeeextrakt entfernt. Möglicherweise trug ein Teil dieser Verbindungen zur Abgrenzung von Kaffeeproben nach dem geographischen Ursprung von Kaffee bei und die Entfernung dieser führte zu einer Verschlechterung der Klassifizierungsgüte. Da durch die Anwendung der UF ein zusätzlicher Schritt in der Probenvorbereitung nötig war und dieser nicht zu einer Verbesserung der Klassifizierungsgüte der PCA-LDA führte, wurde auf eine UF verzichtet und nur eine Membranfiltration der wässrigen Extrakte durchgeführt.

4.3.2 Vorbereitung der Messlösung für die ^1H -NMR-spektroskopische Messung

Um die durch einen variierenden pH-Wert entstehenden Signal-Shifts in den ^1H -NMR-Spektren der hydrophilen Metaboliten zu reduzieren oder zu vermeiden, sollte zu der Messlösung ein geeigneter Puffer zugegeben werden. Dabei sollten bevorzugt anorganische Puffer verwendet werden, welche keine zusätzlichen Signale im ^1H -NMR-Spektrum erzeugen [133, 134, 167]. In dieser Arbeit wurden Natriumphosphat-Puffer verschiedener Molarität und mit unterschiedlichen pH-Werten getestet. Nach der wässrigen Extraktion der Kaffeeproben besaßen diese einen pH-Wert zwischen 5 und 6. Daher wurden bei einem pH-Wert von 6 zunächst unterschiedliche Molaritäten des Phosphatpuffers getestet, welche die Konzentrationen 0,10, 0,50, 1,00, 2,00 und 3,00 mol/L aufwiesen. Als Proben dienten drei Kaffeeproben mit unterschiedlichem Röstgrad. Eine Konzentration des Natriumphosphat-Puffers von 0,10 mol/L reichte nicht aus, um die Signal-Shifts der organischen Säuren zu verhindern. Insbesondere das Signal der Essigsäure wies in den verschiedenen Kaffeeproben unter Verwendung des Puffers dieser Konzentration unterschiedliche chemische Verschiebungen auf (Abbildung 9-12(A)). Diese Shifts konnten durch die Verwendung des Puffers mit einer Konzentration von 0,50 mol/L minimiert und durch die Verwendung des Puffers mit der Konzentration von 1,00 mol/L vermieden werden (Abbildung 9-12(B) und (C)). Auch durch die Verwendung des Puffers mit der Konzentration von 2,00 und 3,00 mol/L traten diese Signal-Shift nicht mehr auf (Abbildung 9-12(D) und (E)). Nachteil einer hohen Salzkonzentration sind jedoch die möglichen Schwierigkeiten bei der Abstimmung und Anpassung der Spektrometerfrequenz [123]. Da durch einen Puffer mit der

Konzentration von 1,00 mol/L pH-abhängige Signal-Shifts vermieden werden konnten, wurde daher auf einen Puffer höher Konzentration verzichtet.

Im nächsten Schritt wurde ein 1 M Natriumphosphat-Puffer mit unterschiedlichen pH-Werten im Bereich pH = 2,00 – 6,00 getestet. Als Proben dienten wieder drei Kaffeeproben mit unterschiedlichem Röstgrad. Aus Abbildung 9-13 wird ersichtlich, dass nur durch die Verwendung eines 1 M Natriumphosphat-Puffers mit einem pH-Wert von 2,00 oder 6,00 pH-abhängige Signal-Shifts vermieden werden konnten. Ein zu basischer pH-Wert wird nicht empfohlen, da dies zur Bildung von Artefakten oder zu Isomerisierungsreaktionen führen kann [135]. Beispielsweise wird 5-CQA leicht in 4- oder 3-CQA umgewandelt oder CGAL kann gebildet werden [135]. Aus diesem Grund wurde mit einem Natriumphosphat-Puffer mit einer Konzentration von 1,00 mol/L und einem pH-Wert von 6,00 gearbeitet.

4.3.3 Optimierung der Messparameter für die ^1H -NMR-spektroskopische Messung

Aufgrund der Extraktion der Kaffeemetaboliten mit H_2O statt mit D_2O wurde im ^1H -NMR-Spektrum ein dominierendes Wassersignal erzeugt. Daher sollte die Intensität der Wasserresonanz durch eine geeignete Pulssequenz verringert werden. Es wurden verschiedenen Pulssequenzen getestet und bezüglich der Effektivität der Lösungsmittelunterdrückung und Beeinflussung der Metabolitensignale miteinander verglichen. Alle getesteten Pulsprogramme wurden an der gleichen Probe und unter Verwendung der gleichen Mess- und Prozessierungsparameter (Tabelle 8-3) durchgeführt.

Neben einem einfachen Vorsättigungspuls wurde auch eine Pulssequenz getestet, bei welcher nach der Vorsättigung ein Composite Pulse erfolgte. Dazu wurden die Pulsprogramme von Bruker namens „zgpr“ und „zgcppr“ verwendet. Für die 1D-NOESY Pulssequenz wurde das „noesygppr1d“ von Bruker ausgewählt. Daneben wurde das Pulsprogramm „noesygppr1d_d7“ getestet, welches vor der D_1 -Relaxationsverzögerung eine weitere Wartezeit (D_7) verwendete. Neben der „wet“ Pulssequenz wurden die WATERGATE Unterdrückungsmethoden „p3919“ und „zggpwg“, sowie die Excitation-Sculpting-Experimente „zggpw5“ und „zgesgp“ getestet.

Im Vergleich zu dem Spektrum ohne Lösungsmittelunterdrückung (Abbildung 9-14(A)) konnte mit allen getesteten Pulssequenzen eine Unterdrückung der Wasserresonanz erreicht werden (Abbildung 9-14 bis Abbildung 9-18). Durch einen Vorsättigungspuls konnte die Wasserresonanz soweit unterdrückt werden, dass die Metabolitensignale gut sichtbar waren (Abbildung 9-15(A)). Auch wies das ^1H -NMR-Spektrum keine Phasen- oder Basislinienfehler auf. Allerdings war die Wasserresonanz auch nach der Vorsättigung noch sehr dominant, sodass Metabolitensignale nahe dem Wassersignal von diesem beeinflusst wurden. Die Verwendung eines Vorsättigungspulses in Verbindung mit einem Composite Pulse konnte die Breite des Wassersignals weiter reduzieren (Abbildung 9-15(B)), wodurch Metabolitensignale nahe dem Wassersignal nicht mehr von diesem überlagert wurden. Die Anwendung der 1D-NOESY-Pulssequenzen führte zu einer zufriedenstellenden Unterdrückung der Wasser-Resonanz. Das Wassersignal wies in beiden entsprechenden ^1H -NMR-Spektren eine geringe Breite und keine Überlagerung mit den Metabolitsignalen auf (Abbildung 9-16). Bei der Verwendung des WET-Experimentes musste aufgrund des stark negativen Wassersignals der Receiver Gain herabgesetzt werden. In dem resultierenden ^1H -NMR-Spektrum erfolgte eine gute Unterdrückung der Wasserresonanz ohne Phasenfehler (Abbildung 9-14(B)). Allerdings war durch die Verwendung eines geringeren Receiver Gain

die Intensität der Metabolitensignale stark herabgesetzt. Für die WATERGATE Experimente wurden die Pulssequenzen von Bruker „p3919“ und „zggpwg“ eingesetzt. Bei der Pulssequenz „p3919“ schlossen zwei Gradientenpulse eine Spin-Echo-Sequenz, bestehend aus drei Pulspaaren, ein. Dahingegen verwendete die „zggpwg“ neben den Gradientenpulsen zwei selektive 90° Pulse zur Unterdrückung der Wasserresonanz. Jedoch kam es durch die Anwendung insbesondere der Pulssequenz „p3919“ zu Phasenfehlern in den entsprechenden ^1H -NMR-Spektren (Abbildung 9-17(A)). Durch die Pulssequenz „zggpwg“ reduzierten sich diese Phasenfehler, jedoch war die Intensität der Metabolitensignale im Vergleich zu der zuvor beschriebenen WATERGATE Pulssequenz reduziert (Abbildung 9-17(B)). Durch die Anwendung der Excitation-Sculpting-Experimente konnte die Wasserresonanz vollständig unterdrückt werden (Abbildung 9-18). Allerdings waren im ^1H -NMR-Spektrum der „zggpw5“ Pulssequenz Phasenfehler in den zuvor beschriebenen Bereichen erkennbar (Abbildung 9-18(A)). Bei einer sehr guten Unterdrückung der Wasserresonanz traten diese Fehler im Spektrum der „zgesgp“ Pulssequenz nicht auf (Abbildung 9-18(B)). Während sich die Pulssequenz „zggpw5“ durch eine Kombination aus harten Pulsen und Verzögerungen auszeichnete, verwendete die „zgesgp“ Pulssequenz selektive 180° Pulse zur Unterdrückung des Wassersignals.

Um die getesteten Methoden zur Unterdrückung der Wasserresonanz bezüglich der Beeinflussung der Kaffeemetaboliten zu vergleichen, wurden zunächst in den entsprechenden ^1H -NMR-Spektren 13 Signale im Bereich $\delta_{\text{H}} = 0,50 - 9,30$ ppm ausgewählt. Anschließend wurde der Mittelwert der jeweiligen Signalfäche aus allen getesteten Pulssequenzen gebildet und die prozentuale Abweichung jeder Methode zum Mittelwert berechnet. Die prozentuale Abweichung der Integrale vom Mittelwert aller getesteten Pulssequenzen für die 13 ausgewählten Signale ist in Abbildung 4-8 dargestellt.

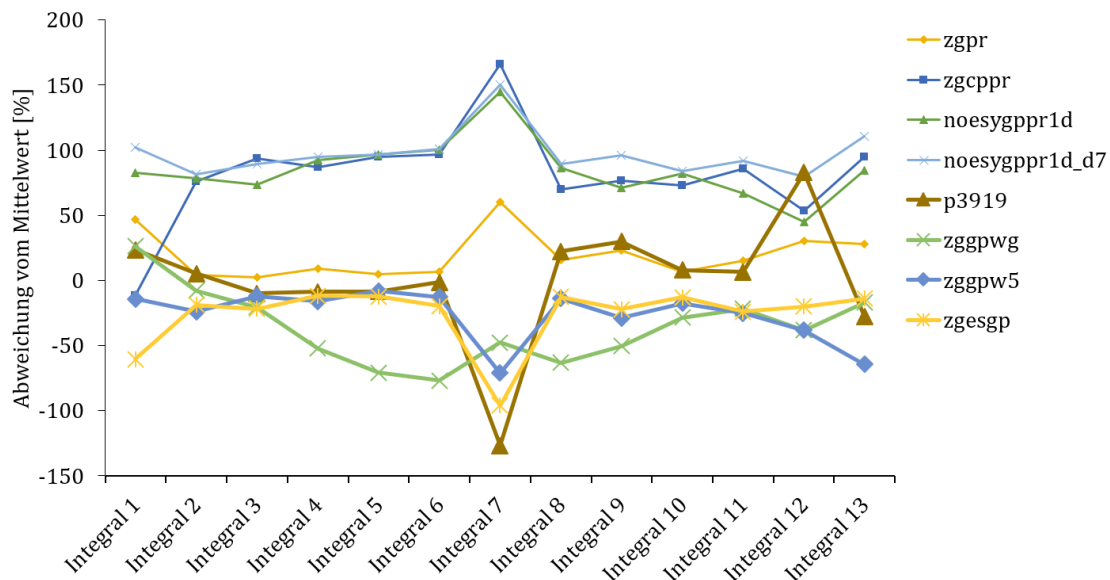


Abbildung 4-8: Prozentuale Abweichungen von den Mittelwerten ausgewählter Signale im Bereich $\delta_{\text{H}} = 0,50 - 9,30$ ppm. Für die Berechnung des Mittelwertes eines Integrals wurden die entsprechenden Integrale aller getesteten Unterdrückungsmethoden herangezogen. Unterdrückungsmethoden: Vorsättigungspuls (zgpr), Vorsättigungspuls kombiniert mit Composite Pulse (zgcppr), 1D-NOESY (noesygppr1d und noesygppr1d_d7), WATERGATE (p3919, zggpwg), Excitation-Sculpting (zggpw5 und zgesgp). Integral 1: $\delta_{\text{H}} = 1,05 - 1,06$ ppm; Integral 2: $\delta_{\text{H}} = 1,31 - 1,33$ ppm; Integral 3: $\delta_{\text{H}} = 1,92 - 1,93$ ppm; Integral 4: $\delta_{\text{H}} = 3,19 - 3,20$ ppm; Integral 5: $\delta_{\text{H}} = 3,33 - 3,35$ ppm; Integral 6: $\delta_{\text{H}} = 3,50 - 3,52$ ppm; Integral 7: $\delta_{\text{H}} = 4,42 - 4,44$ ppm; Integral 8: $\delta_{\text{H}} = 7,07 - 7,12$ ppm; Integral 9: $\delta_{\text{H}} = 7,52 - 7,53$ ppm; Integral 10: $\delta_{\text{H}} = 7,85 - 7,88$ ppm; Integral 11: $\delta_{\text{H}} = 8,41 - 8,45$ ppm; Integral 12: $\delta_{\text{H}} = 8,50 - 8,54$ ppm; Integral 13: $\delta_{\text{H}} = 9,09 - 9,13$ ppm. NOESY: Nuclear Overhauser Effect Spectroscopy; WATERGATE: Water Suppression by Gradient Tailored Excitation

Die Lösungsmittelunterdrückung mittels WET wurde in Abbildung 4-8 nicht mit aufgenommen. Aufgrund der Herabsetzung des Receiver Gain bei diesem Experiment ist die Vergleichbarkeit zu den anderen getesteten Methoden zur Lösungsmittelunterdrückung nicht mehr gegeben. Die Pulssequenzen der WATERGATE-Experimente beeinflussten die Resonanzen der Metaboliten vergleichsweise stark (Abbildung 4-8). Beide Pulssequenzen des Excitation-Sculpting zeigten eine ähnliche Beeinflussung der Integrale 1 bis 13. Auch durch die Verwendung eines Vorsättigungspulses in der „zgpr“ Sequenz konnten keine zufriedenstellenden Ergebnisse erzielt werden. Lediglich die Verwendung eines Vorsättigungspulses mit der Kombination eines Composite Pulse sowie durch die Verwendung der beiden 1D-NOESY-Experimente konnte eine geringe bis keine Beeinflussung der Metabolitenresonanzen bei gleichzeitig guter Unterdrückung der Wasserresonanz erreicht werden. Bei Verwendung des 1D-NOESY-Experimentes mit zusätzlicher D_7 -Zeit (noesygppr1d_d7) wiesen die Integrale vergleichsweise die höchsten Werte auf (Abbildung 4-8).

Durch die Verwendung einer zu hohen Leistung des Sättigungspulses bei dem 1D-NOESY-Experimentes kann es zu einem dispersiven oder negativem Wassersignal kommen. Um die optimale Leistung des Sättigungspulses zu ermitteln, wurden daher die Werte 10, 15, 20, 25, 30 und 50 Hz für die Leistung des Unterdrückungspulses getestet. Die Leistung sollte so gewählt werden, dass die Wasserresonanz vollständig unterdrückt wird und gleichzeitig benachbarte Metabolitensignale möglichst wenig vom Sättigungspuls beeinflusst werden. Durch eine Leistung von 10 Hz konnte keine ausreichende Unterdrückung der Wasserresonanz sichergestellt werden (Abbildung 9-19(A)). Die Anwendung einer Leistung von 15 Hz reduzierte die Intensität des Wassersignals (Abbildung 9-19(B)). Durch eine Leistung des Sättigungspulses von 20 Hz und 25 Hz konnte die Wasserresonanz sehr gut unterdrückt werden, während keine erkennbare Beeinflussung der benachbarten Metabolitensignale auftrat (Abbildung 9-19(C) und (D)). Die Verwendung einer höheren Pulsleistung von 30 Hz und 50 Hz führte resultierte in einer zu starken Unterdrückung und beeinflusste die benachbarten Metabolitensignale (Abbildung 9-19(E) und (F)).

Um die Relaxationsprozesse von Kernen ausgewählter Metaboliten zu untersuchen, wurde das 1D-NOESY Experiment mit verschiedenen Wartezeiten durchgeführt. Dazu wurde die Pulssequenz „noesygppr1d“ mehrfach hintereinander durchgeführt, wobei für die D_1 -Zeit Werte von 0,10 – 40,0 s verwendet wurden. Aus den resultierenden ^1H -NMR-Spektren wurden die entsprechenden Signale von Milchsäure, Essigsäure, Koffein, CQA, Ameisensäure, NMP, Trigonellin und HMF integriert und die Integrale gegen die D_1 -Zeit aufgetragen (Abbildung 9-20). Die Ergebnisse zeigten, dass die untersuchten Kerne von Milchsäure, Koffein, CQA und NMP bei $D_1 = 8,00$ s vollständig relaxiert waren. Für den untersuchten Kern von Trigonellin war eine Zunahme des Integrals ab einer D_1 -Zeit von 15 s nicht mehr beobachtbar. Für die untersuchten Kerne von Essigsäure und Ameisensäure war dies nach 20 s und 30 s der Fall. Die Integrale in Abhängigkeit der D_1 -Zeit für den untersuchten Kern von HMF ergab eine Kurve, welche keine genaue Aussage der Wartezeit zuließ. Grund hierfür könnte das geringe S/N dieses Signals im ^1H -NMR-Spektrum gewesen sein. Eine Relaxationsverzögerung von 30 s würde zu sehr hohen Messzeiten führen. Daher wurde für die Erfassung der hydrophilen Metaboliten eine Relaxationsverzögerung von 8 s gewählt. Für die nicht vollständig ausgeschwungenen Kerne von Ameisensäure, Essigsäure und Trigonellin Korrekturfaktoren ermittelt, indem die Integrale der Signale bei einer Relaxationsverzögerung von 8 s und 40 s verglichen wurden. Für Ameisensäure ergab sich dadurch ein Korrekturfaktor von 1,2 und für Essigsäure und Trigonellin einer von 1,1.

Ein weiterer Punkt bei der Optimierung der Messparameter war die Überprüfung der Messtemperatur. Diese erfolgte, indem eine Kaffeeprobe bei einer Messtemperatur von 280 K, 290 K und 300 K vermessen und die resultierenden ^1H -NMR-Spektren anschließend miteinander verglichen wurden (Abbildung 9-21). Die Veränderung der Messtemperatur führte sowohl zu einem Shift des Wassersignals, als auch zu einem der Metabolitensignale. Durch eine Messtemperatur von 300 K lag das Wassersignal im ^1H -NMR-Spektrum weit genug von den Metabolitsignalen entfernt, sodass es zu keiner Beeinflussung kam.

Aufgrund der erzielten Ergebnisse wurde zur Erfassung der hydrophilen Kaffeemetaboliten die NOESY-Pulssequenz unter zusätzlicher Verwendung der D7-Deit „noesygppr1d_d7.eba“ mit einer Leistung des Sättigungspulses von 25 Hz, einer Relaxationsverzögerung von 8 s und einer Messtemperatur von 300 K verwendet.

4.3.4 Überprüfung der Reproduzierbarkeit der Extraktion und der Messung

Für die Sicherstellung einer reproduzierbaren Extraktion und Messung für die Untersuchung der hydrophilen Kaffeemetaboliten wurde auch hier mit jeder Messserie eine Kontrollprobe aufgearbeitet und vermessen. Wie für die Untersuchung der lipophilen Kaffeemetaboliten war diese Kontrollprobe ein biologisch angebauter Kaffee ohne Herkunftsangabe. In der Vorperiode wurde die Kontrollprobe fünfmal an jeweils drei verschiedenen Tagen von zwei unterschiedlichen Personen aufgearbeitet und vermessen. Aus den resultierenden ^1H -NMR-Spektren wurden über einen Bereich von $\delta_{\text{H}} = 1,0\text{-}9,2$ ppm zehn Signale ausgewählt und diese automatisch integriert. Für die Integration der Signale wurden zuvor festgelegte Bereiche verwendet. Die Integrale der einzelnen Signale wurden aufsummiert und der prozentuale Anteil jedes Signals bezogen auf das Gesamtintegral bestimmt. Anhand der mehrfachen Aufarbeitung und Messung wurde der Mittelwert und die Standardabweichung gebildet. Die obere und untere Warn- sowie Kontrollgrenze wurde wie in Kapitel 4.3.4 beschrieben aus Addition und Subtraktion der zweifachen und dreifachen Standardabweichung vom Mittelwert gebildet. Aus dem ^1H -NMR-Spektrum jeder anschließend aufgearbeiteten Kontrollprobe wurden anhand der festgelegten Bereiche die zehn Signale integriert und der prozentuale Anteil bezogen auf das Gesamtintegral gebildet. Die berechneten prozentualen Anteile für jedes Integral wurden in die Kontrollkarte eingetragen. Abbildung 4-9 stellt die Kontrollkarte von Integral 1 und Integral 10 dar. Die zehn Integrale für die hydrophilen Metaboliten wiesen enge Warn- und Kontrollgrenzen auf, was auf eine reproduzierbare Probenvorbereitung und Messung hindeutete. Die Kontrollgrenze wurde für die Integrale 1 und 10 nicht überschritten (Abbildung 4-9). Dies war jedoch für die Integrale 2, 5, 6 und nicht der Fall (Abbildung 9-22 und Abbildung 9-23), bei welchen die Kontrollgrenze jeweils zweimal überschritten wurde. In diesen Fällen wurde die Extraktion sowie die Messung der gesamten Messserie wiederholt. Die Wiederholung führte dazu, dass die Kontrollprobe für alle zehn Integrale wieder innerhalb der Grenzen lag.

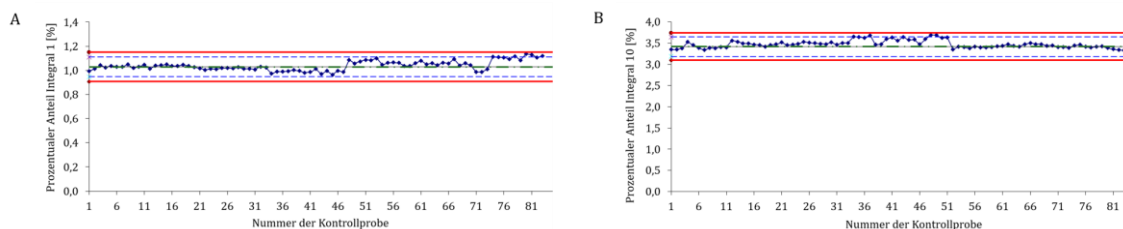


Abbildung 4-9: Kontrollkarten von (A) Integral 1 ($\delta_H = 1,05 - 1,07$ ppm) und (B) Integral 10 ($\delta_H = 9,10 - 9,16$ ppm) der ausgewählten Signale für die hydrophilen Kaffeemetaboliten für insgesamt 84 Kontrollproben. Aus den Integralen der zehn Signale wurde das Gesamtintegral berechnet und anschließend der prozentuale Anteil jedes Integrals bezogen auf das Gesamtintegral bestimmt. Die rote, durchgezogene Linie stellt die Kontrollgrenze dar (Mittelwert ± 3 -Standardabweichung). Die blaue, gestrichelte Linie beschreibt die Warngrenze (Mittelwert ± 2 -Standardabweichung).

4.4 Nicht zielgerichtete Untersuchung durch Verfahren der uni- und multivariaten Statistik

4.4.1 Datenvorbehandlung der $^1\text{H-NMR}$ -Spektren

Vor der Erstellung der Klassifizierungsmodelle wurden verschiedene Methoden der Datenvorbehandlung auf die $^1\text{H-NMR}$ -Spektren angewendet. Zunächst wurde jedes $^1\text{H-NMR}$ -Spektrum bezüglich eines Standardsignals normalisiert. In den $^1\text{H-NMR}$ -Spektren der lipophilen Kaffeemetaboliten wurde dazu das Signal von DMS bei $\delta_H = 3,00 - 3,05$ ppm und in den $^1\text{H-NMR}$ -Spektren der hydrophilen Kaffeemetaboliten das TSP-Signal im Bereich $\delta_H = -0,10 - 0,10$ ppm verwendet. Im nächsten Schritt wurde ein Alignment auf das TMS-Signal für die NMR-Spektren der lipophilen und auf das TSP-Signal für NMR-Spektren der hydrophilen Metaboliten bei jeweils $\delta_H = 0,00$ ppm durchgeführt. Für die Bucketierung der $^1\text{H-NMR}$ -Spektren der lipophilen Metaboliten wurden 2000 äquidistante Buckets im Bereich von $\delta_H = 0,50 - 9,80$ ppm verwendet, was einer Bucketbreite von 0,005 ppm entsprach. Nach Ausschluss der Lösungsmittelsignale standen noch 1775 Buckets zur Erstellung der Klassifizierungsmodelle zur Verfügung. Die $^1\text{H-NMR}$ -Spektren der hydrophilen Metaboliten wurden im Bereich von $\delta_H = 0,50 - 9,80$ ppm in 1000 äquidistante Buckets eingeteilt, wodurch eine Bucketbreite von 0,009 ppm gegeben war. Nach Ausschluss des Wassersignals resultierten 983 Buckets. Zuletzt wurde mit den bucketierten Daten eine log-Transformation durchgeführt, wodurch ein pseudo-Skalierungseffekt erreicht wurde. Durch die nichtlineare Umwandlung der Daten wurden große Integrale stärker gestaucht als kleine.

Die Kaffeeproben wurden im bereits gerösteten Zustand vermessen und wiesen daher Variationen wie unterschiedliche Feuchtegehalte auf. Abbildung 9-24(C) zeigt die Variationen des Feuchtegehaltes der Kaffeeproben in den untersuchten Klassen. Die Variation sowie der Mittelwert der Feuchtegehalte waren zwischen den Klassen vergleichbar. Eine Ausnahme bildeten hier die heißluftgerösteten Kaffeeproben, welche einen deutlich geringeren Feuchtegehalt aufwiesen als die trommelgerösteten Kaffeeproben. Um den Einfluss dieser Variation auf die Klassifizierung bezüglich des geographischen Ursprungs zu überprüfen, wurden die $^1\text{H-NMR}$ -Spektren neben der Normalisierung auf ein Standardsignal zusätzlich auf einen einheitlichen Feuchtegehalt von 3,00% normalisiert. Dazu wurden zwei Klassifizierungsmodelle basierend auf der Analyse der hydrophilen bzw. lipophilen Metaboliten erstellt. Für das erste Klassifizierungsmodell wurden die $^1\text{H-NMR}$ -Spektren wie im oberen

Abschnitt beschrieben vorbehandelt. In der Datenvorbehandlung für das zweite Klassifizierungsmodell fand zusätzlich eine Normalisierung bezüglich eines einheitlichen Feuchtegehaltes von 3% statt. Aus Abbildung 9-24(C) geht hervor, dass die Kaffeeprobe aus Brasilien und Kolumbien die größten Variationen bezüglich des Feuchtegehaltes aufwiesen. Aus diesem Grund wurden die Klassifizierungsmodelle mit diesen beiden Probenklassen gebildet. Abbildung 9-25 (A)-(D) zeigt, dass durch eine Normalisierung bezüglich eines einheitlichen Feuchtegehaltes Klassifizierungsmodelle mit geringfügig schlechteren Werten für die richtige Zuordnung in der internen Validierung entstanden. Daher wurde in der weiteren Arbeit keine Normalisierung bezüglich eines einheitlichen Feuchtegehaltes durchgeführt. Da bei der nicht zielgerichteten Untersuchung der lipophilen Metaboliten ein Aliquot des Extraktes verwendet wurde, wurden unterschiedliche Mengen des evaporierten Rückstands in der Messlösung gelöst (Abbildung 9-24(B)). Um den Einfluss dieser Variation zu untersuchen, wurde neben der Normalisierung der $^1\text{H-NMR}$ -Spektren bezüglich eines Standardsignals zusätzlich eine Normalisierung bezüglich eines einheitlichen Gewichtes des evaporierten Rückstands von 20,0 mg durchgeführt. Auch hier wurden zwei Klassifizierungsmodelle basierend auf der Analyse der lipophilen Metaboliten erstellt. Während im ersten Modell die Datenvorbehandlung wie beschrieben durchgeführt wurde, wurden für das zweite Klassifizierungsmodell die $^1\text{H-NMR}$ -Spektren zusätzlich bezüglich einer einheitlichen Menge des evaporierten Rückstands von 20 mg normalisiert. Als Klassen wurden Kaffeeprobe aus Brasilien und Kolumbien gewählt. Ein Vergleich der erhaltenen Klassifizierungsmodelle zeigte, dass sich die Ergebnisse der internen Validierung durch eine Normalisierung bezüglich eines einheitlichen Gewichtes des evaporierten Rückstands nicht veränderten (Abbildung 9-25(E) und (F)). Aus diesem Grund wurde auf eine zusätzliche Normalisierung bezüglich eines einheitlichen Gewichtes des evaporierten Rückstands verzichtet.

4.4.2 Erstellung von Klassifizierungsmodellen basierend auf der Analyse der lipophilen und hydrophilen Kaffeemetaboliten zur Abgrenzung von Kaffee nach Kontinenten

4.4.2.1 Modellerstellung und interne Validierung

Die Abgrenzung von Kaffeeprobe aus den Kontinenten Afrika, Asien, Nordamerika und Südamerika erfolgte jeweils in zwei-Klassen Modellen. Für die Erstellung der Klassifizierungsmodelle wurde zu Beginn 20% des Probandensatzes entfernt. Dieser Teil diente als Probandensatz für die externe Validierung. Somit wurden 71 Proben aus Afrika, 53 Proben aus Asien, 141 Proben aus Nordamerika und 152 Proben aus Südamerika für die Erstellung der Klassifizierungsmodelle mit interner Validierung mittels MCKV verwendet. Mit den log-transformierten Buckets des verbleibenden Probandensatzes wurde zunächst eine PCA zur Dimensionsreduktion durchgeführt. Anschließend wurde die LDA anhand der Scores der PCA berechnet. Für die LDA wurden die Scores der ersten 29 PCs der PCA verwendet, mit welchen mindestens 97,8% der Gesamtvarianz beschrieben werden konnte. Als interne Validierung diente eine MCKV, wobei für die KV jeweils 10% des Probandensatzes als Testdatensatz verwendet wurde. Das Ergebnis der MCKV wurde in einer Konfusionsmatrix dargestellt, aus der hervorgeht, wie viele Proben aus dem Testdatensatz der „richtigen“ oder „falschen“ Klasse zugeordnet wurden. Abbildung 4-10 zeigt beispielhaft das Ergebnis der MCKV der erhaltenen Klassifizierungsmodelle, erstellt aus den $^1\text{H-NMR}$ Spektren der lipophilen und hydrophilen Metaboliten für die zwei Klassen Afrika und Nordamerika.

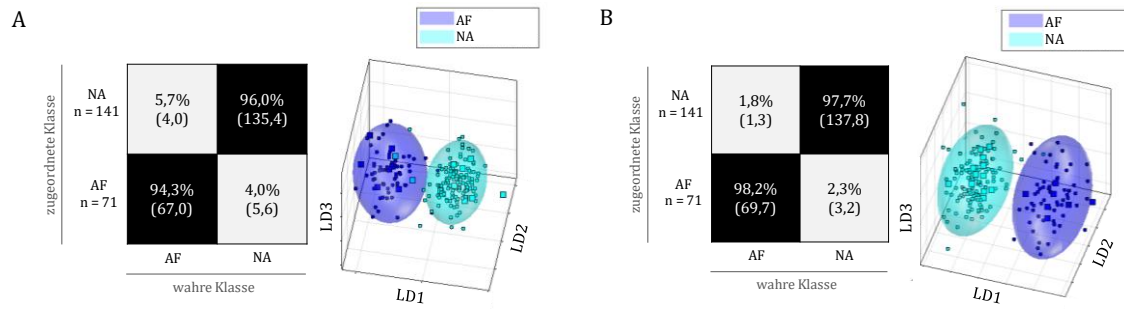


Abbildung 4-10: Ergebnisse der MCKV der PCA-LDA zur Vorhersage des geographischen Ursprungs von Kaffee bezüglich der Kontinente Afrika und Nordamerika aus den $^1\text{H-NMR}$ -Spektren der (A) lipophilen und (B) hydrophilen Kaffeemetaboliten. Jeweils rechts in der Abbildung (A) und (B) ist der LDA-Raum eines Kreuzvalidierungsschrittes dargestellt. Das 95%ige Prognoseellipsoid des Trainingsdatensatzes ist als Wolke dargestellt, die Proben des Testdatensatzes werden durch Rechtecke symbolisiert. Links in der Abbildung (A) und (B) ist die Konfusionsmatrix der MCKV dargestellt. Die x-Achse und die y-Achse zeigen die zugeordnete Klasse bzw. die wahre Klasse. In der Konfusionsmatrix ist die Wahrscheinlichkeit des Vorhersageergebnisses in Prozent angegeben. Anzahl der für die LDA verwendeten PC: 29, Erklärte Varianz: (A) 99,6% und (B) 98,8%. MCKV: Monte-Carlo Kreuzvalidierung; PCA-LDA: Hauptkomponentenanalyse mit linearer Diskriminanzanalyse; LD: Lineare Diskriminanzfunktion; AF: Afrika; NA: Nordamerika

Eine Abgrenzung von Kaffee aus den beiden Kontinenten Afrika und Nordamerika mittels PCA-LDA war sowohl durch die Analyse der lipophilen (Abbildung 4-10(A)) als auch der hydrophilen (Abbildung 4-10(B)) Metaboliten möglich. Von den Testproben der Klasse Afrika konnten 94,1% bzw. 94,5% der richtigen Klasse zugeordnet werden. Für die Testproben der Klasse Nordamerika erfolgte bei der Analyse der hydrophilen Metaboliten mit 95,7% eine geringfügig bessere Zuordnung, als bei der Analyse der lipophilen Metaboliten. Bei Verwendung des Modells basierend auf den Daten der lipophilen Metaboliten wurden die Proben AF30 und AF150 sowie die Proben N180, N200 und N401 über die gesamte MCKV der falschen Klasse zugeordnet. Bei den Proben AF30 und AF150 wurden im Vergleich zu den $^1\text{H-NMR}$ -Spektren der anderen afrikanischen Kaffeeproben erhöhte Intensitäten der Cafestolsignale beobachtet. Dies könnte ein möglicher Grund für die falsche Zuordnung der beiden Proben sein. Auch ein Vergleich der $^1\text{H-NMR}$ -Spektren der lipophilen Metaboliten der Proben aus Nordamerika zeigte für die Proben N180, N200 und N401 eine geringere Intensität der Cafestolsignale, was ein möglicher Grund für die falsche Zuordnung gewesen sein könnte. In dem Modell erstellt aus der Analyse der hydrophilen Metaboliten wurden die Proben AF150, 20HA009 und N401 während der gesamten MCKV der falschen Klasse zugeordnet. Im Gegensatz zu der Untersuchung der lipophilen Metaboliten konnten für diese Proben in den $^1\text{H-NMR}$ -Spektren der hydrophilen Metaboliten jedoch keine unmittelbaren Auffälligkeiten beobachtet werden.

Neben dem Modell Afrika-Nordamerika wurden Klassifizierungsmodelle für die Klassen Afrika-Asien, Afrika-Südamerika, Asien-Nordamerika, Asien-Südamerika und Nordamerika-Südamerika erstellt. Tabelle 4-2 zeigt eine Übersicht der Ergebnisse der MCKV der Klassifizierungsmodelle zur Vorhersage des geographischen Ursprungs bezüglich aller Kontinente.

Tabelle 4-2: Übersicht über die Ergebnisse der MCKV der PCA-LDA zur Vorhersage des geographischen Ursprungs von Kaffee bezüglich der Kontinente aus den $^1\text{H-NMR}$ -Spektren der lipophilen und hydrophilen Kaffeemetaboliten. Anzahl der für die LDA verwendeten PC: 29. MCKV: Monte-Carlo Kreuzvalidierung; PCA-LDA: Hauptkomponentenanalyse mit linearer Diskriminanzanalyse; PC: Hauptkomponente; AF: Afrika, AS: Asien, NA: Nordamerika, SA: Südamerika, n: Anzahl richtig zugeordneter Proben aus dem Testdatensatz

		Modell erstellt aus der Analyse der lipophilen Metaboliten		Modell erstellt aus der Analyse der hydrophilen Metaboliten	
Klasse 1	Klasse 2	Richtige Zuordnung Klasse 1	Richtige Zuordnung Klasse 2	Richtige Zuordnung Klasse 1	Richtige Zuordnung Klasse 2
AF	AS	94,1% n = 66,8	92,1% n = 65,4	94,5% n = 67,1	95,7% n = 51,7
AF	NA	94,3% n = 67,0	96,0% n = 135,4	98,2% n = 69,7	97,7% n = 137,8
AF	SA	92,4% n = 65,6	98,0% n = 149,0	95,2% n = 67,6	97,0% n = 147,7
AS	NA	87,9% n = 46,5	96,2% n = 136,6	83,2% n = 44,1	93,2% n = 131,4
AS	SA	89,1% n = 47,2	94,2% n = 143,2	82,6% n = 43,8	93,2% n = 141,7
NA	SA	82,1% n = 115,8	85,4% n = 129,8	90,0% n = 126,9	90,8% n = 138,0

Aus Tabelle 4-2 wird ersichtlich, dass die Vorhersagekraft der Klassifizierungsmodelle Afrika-Asien, Afrika-Nordamerika, Afrika-Südamerika und Nordamerika-Südamerika in der Analyse der hydrophilen Metaboliten geringfügig bessere Ergebnisse ergaben. Dahingegen war die richtige Zuordnung der Testproben in den Klassifizierungsmodellen Asien-Nordamerika und Asien-Südamerika bei der Analyse der lipophilen Metaboliten leicht höher. Die interne Validierung aller Modelle zeigte für Kaffeeproben aus Afrika sowohl bei der Analyse der lipophilen als der hydrophilen Metaboliten die besten Ergebnisse. Neben dem geographischen Ursprung könnten auch die angepflanzten Varietäten oder eine unterschiedliche Verarbeitung der Grund für die bessere Abgrenzung von Kaffee aus Afrika sein. Auf diesem Kontinent wird eine große Vielfalt an Arabica-Varietäten in kleinen Plantagen mit Schattenanbau angepflanzt [16, 25]. Schattenanbau wird auch auf vielen Plantagen in Zentralamerika und Asien betrieben [3]. Aus diesem Grund kann dieser Parameter als Grund für die Abgrenzung von Kaffeeproben aus Afrika ausgeschlossen werden. Zusätzlich könnte die nasse Verarbeitung der Rohkaffeebohnen, welche vorwiegend in Afrika angewendet wird [7, 15], zu der Abgrenzung beitragen. Die geringste Vorhersagekraft bei der Verwendung der Daten der lipophilen Metaboliten zeigte das Modell Nordamerika-Südamerika. Dies könnte durch die geographische Nähe der Kaffeeproben aus beiden Kontinenten erklärt werden. Bei der Analyse der hydrophilen Metaboliten zeigte das Modell Asien-Nordamerika die geringste Vorhersagefähigkeit.

Auch Consonni et al. führten Untersuchungen zur Trennung von Kaffee nach Kontinenten durch. In ihrer Studie führten sie Proben aus Nord- und Südamerika zusammen als Klasse Amerika auf. So konnten sie durch die Analyse der der hydrophilen Metaboliten Kaffeeproben erfolgreich nach den Kontinenten Amerika, Asien und Afrika auftrennen [145]. Allerdings wurden bei dieser Arbeit nur 40 Kaffeeproben untersucht und eine spezielle Form der PLS-DA, die Orthogonal Partial Least Square Discriminant Analysis (OPLS-DA),

verwendet [145]. Aufgrund der geringen Anzahl von Proben und der Verwendung einer anderen multivariaten Datenanalyse in der Studie von Consonni et al. ist allerdings kein direkter Vergleich zu den in dieser Arbeit erzielten Ergebnissen möglich.

Bei allen auf den Daten der lipophilen Metaboliten beruhenden Modellen wurden die Proben AF150, AS166, AS200, AS250 und N180 über die gesamte MCKV der falschen Klasse zugeordnet. Für die Probe AF150 könnte wie bereits beschrieben, eine erhöhte Intensität der Cafestolsignale ein Grund für die falsche Zuordnung sein. Im Gegensatz dazu wies die Probe N180 eine höhere Intensität der Kahweolsignale im Vergleich zu den anderen Proben aus Nordamerika auf. Auch bei den Proben aus Asien könnte ein abweichendes Verhältnis der Signalintensitäten von Kahweol und Cafestol eine mögliche Ursache für die falsche Zuordnung sein. Bei den auf den Daten der hydrophilen Metaboliten beruhenden Modellen wurden die Proben AF150, 20HA009 und 20DU070 in der MCKV immer der falschen Klasse zugeordnet. Für diese Proben konnten jedoch durch einen Vergleich der entsprechenden $^1\text{H-NMR}$ -Spektren keine Auffälligkeiten beobachtet werden.

4.4.2.2 Externe Validierung

Um die Vorhersage über zukünftige Beobachtungen besser beurteilen zu können, wurden die Klassifizierungsmodelle auf einen externen Datensatz angewendet. Voraussetzung für eine externe Validierung ist ein ausreichend großer Datensatz [198]. Die Zusammensetzung der Proben bezüglich der Klassen im externen Testdatensatz sollte dabei vergleichbar mit der des Originaldatensatzes sein [198]. Beide Anforderungen konnten in der vorliegenden Arbeit umgesetzt werden. Für den externen Testdatensatz wurden 20% der Proben vor der Erstellung der Klassifizierungsmodelle aus dem originalen Probensatz entfernt. Dabei handelte es sich um 18 Proben aus Afrika, 13 Proben aus Asien, 35 Proben aus Nordamerika und 38 Proben aus Südamerika. Diese Proben wurden nicht zur Erstellung der Klassifizierungsmodelle und der MCKV verwendet. Mit den Proben des externen Datensatzes wurde die gleiche Datenvorbehandlung (Kapitel 8.4.1) durchgeführt, wie mit den Proben, welche zur Modellerstellung dienten. Anschließend wurden die Klassifizierungsmodelle auf den externen Probendatensatz angewendet und dadurch die Vorhersagefähigkeit beurteilt. Als Beurteilungskriterium diente der *conf*-Wert, welcher den Abstand der Testprobe zum Mittelpunkt der Datenwolke der entsprechenden Klasse im LDA-Raum angibt. Bei einem *conf*-Wert von 0 lag die Testprobe im Datenmittelpunkt der einer Klasse, bei einem Wert von 1 innerhalb des berechneten 95%igem Prognoseellipsoids der Klasse. Für eine richtige Zuordnung einer externen Testprobe sollte diese demnach einen möglichst kleinen *conf*-Wert für die entsprechende Klasse aufweisen.

Die *conf*-Werte des Probendatensatzes der externen Validierung für die Modelle Afrika-Nordamerika basierend auf der Analyse der lipophilen bzw. hydrophilen Metaboliten sind in Tabelle 9-6 dargestellt. Für das Modell Afrika-Nordamerika bei Verwendung der Daten der lipophilen Metaboliten konnten 51 von insgesamt 53 Testproben der richtigen Klasse zugeordnet werden, was 96,2% entspricht. Das auf den Daten der hydrophilen Metaboliten basierende Modell Afrika-Nordamerika zeigte im Gegensatz zu der internen Validierung mittels MCKV bei der externen Validierung eine schlechtere Vorhersagefähigkeit. Hier wurden 49 der 53 Proben der richtigen Klasse zugeordnet, was 92,5% entspricht.

Bei der externen Validierung des Klassifizierungsmodells basierend auf der Analyse der lipophilen Metaboliten wurde die Probe AF3 nicht der Klasse Afrika zugeordnet. Diese Probe lag mit einem *conf*-Wert für Afrika von 1,80 und einem *conf*-Werten für Nordamerika von 0,20 außerhalb des Prognoseellipsoids der Klasse Afrika und innerhalb dem der Klasse

Nordamerika. Die externe Testprobe N5 wurde in dem Modell bei der Verwendung der Daten der lipophilen Metaboliten nicht der Klasse Nordamerika zugeordnet. Hier betrug der *conf*-Wert für Afrika 0,10 und der für Nordamerika 2,10. In einem Vergleich der ¹H-NMR-Spektren der afrikanischen Kaffeeproben konnten bei der Probe AF3 geringere Intensitäten der TAG- und FA-Signale beobachtet werden. Die Probe N5 zeigte bei einem Vergleich der ¹H-NMR-Spektren mit anderen Proben aus Nordamerika eine leicht geringere Intensität der Cafestolsignale. Möglicherweise könnte eine geringere Signalintensität der Grund für eine falsche Zuordnung sein.

In dem mit den Daten der hydrophilen Metaboliten erstellten Modell Afrika-Nordamerika wurden die externen Testproben AF3 und AF22a, sowie N5 und N450 nicht der richtigen Klasse zugeordnet (Tabelle 9-6). Die *conf*-Werte der Proben AF3 und AF22a waren für Afrika 1,89 und 1,84 und für Nordamerika 0,53 und 0,59. Somit lagen diese außerhalb des Prognoseellipsoids der Klasse Afrika. Die Proben N5 und N450 besaßen einen *conf*-Wert für Afrika von 0,35 und 0,66, wohingegen dieser für Nordamerika 2,14 und 1,81 betrug. Die Proben AF3 und AF22a wiesen bei einem Vergleich der ¹H-NMR-Spektren mit anderen Proben aus Afrika eine geringere Intensität des Signalberges bei $\delta_{\text{H}} = 1,28$ ppm auf. Dagegen war bei den Proben N5 und N450 durch einen Spektrenvergleich mit anderen Proben aus Nordamerika eine geringere Intensität der aromatischen Signale der CGA zu beobachten, was einen möglichen Grund für die falsche Zuordnung darstellt.

Die Testproben AF3, AF22a, AS251, AS450, N5 und N450 konnten in den weiteren Modellen basierend auf der Analyse der lipophilen Metaboliten nicht der richtigen Klasse zugeordnet werden. In den Modellen basierend auf der Analyse der hydrophilen Metaboliten wurden die Testproben AF3, AF22a N5 und N450 der falschen Klasse zugeordnet. Bei Verwendung der Daten der lipophilen Metaboliten zeigte das Modell Afrika-Südamerika die beste Vorhersagefähigkeit auf die Testproben. Hier wurden 98,2% der Proben der richtigen Klasse zugeordnet. Die beste Vorhersagefähigkeit bei Verwendung der Daten der hydrophilen Metaboliten zeigte das Modell Afrika-Nordamerika, bei welchem 92,5% der Testproben richtig zugeordnet wurden. Die geringste Vorhersagefähigkeit ist durch das Modell Nordamerika-Südamerika sowohl basierend auf der Analyse der lipophilen als auch der hydrophilen Metaboliten gegeben. Hier wurden 89,0% der Testproben der richtigen Klasse zugeordnet.

4.4.3 Erstellung von Klassifizierungsmodellen basierend auf der Analyse der lipophilen und hydrophilen Kaffeemetaboliten zur Abgrenzung von Kaffee nach ausgewählten Ländern

4.4.3.1 Modellerstellung und interne Validierung

Neben den Kontinenten wurden Klassifizierungsmodelle zur Abgrenzung von Kaffeeproben aus den Ländern Brasilien, Äthiopien und Kolumbien erstellt. Brasilien ist weltweit der größte Kaffeeproduzent, in welchem Kaffee sowohl auf großen Plantagen, als auch auf kleinen Farmen angebaut wird [241]. Im Gegensatz dazu wird in Äthiopien über 95% des Kaffees auf Farmen produziert, welche eine Größe kleiner als einen Hektar haben und auf welchen traditionelle Landwirtschaftstechniken angewendet werden [25]. Aufgrund seiner unverwechselbaren feinen Qualität wird Kaffee aus Äthiopien, speziell aus den Regionen Yirgacheffe und Harrar, auf dem internationalen Markt zu hohen Preisen verkauft [242]. Kaffee aus Kolumbien ist der teuerste weltweit [15]. Ein Verbund von kolumbianische

Kaffeebauern erschuf bereits in den 1930er Jahre mit dem „Café de Colombia“ ein Gütesiegel, welches für qualitativ hochwertigen Kaffee aus Kolumbien steht [243]. Die Registrierung von „Café du Colombia“ als Ursprungsbezeichnung wurde 2005 in Kolumbien erreicht und dessen Registrierung als geschützte geographische Angabe (g.g.A.) wurde in der Europäischen Union 2007 abgeschlossen [243]. Aus diesem Grund ist es wichtig, Methoden zur Überprüfung der Authentizität von Kaffee dieses geographischen Ursprunges zu implementieren.

Die Abgrenzung von Kaffeeproben aus den drei Ländern erfolgte in zwei-Klassen-Modellen mittels PCA-LDA, wobei jeweils die Länder und die entsprechenden Kontinente als Klassen dienten. Auch hier wurden 20% der Proben als externer Testdatensatz verwendet, welcher nicht zur Modellerstellung und zur internen Validierung mittels MCKV verwendet wurde. Für die Modellerstellung standen somit 45 Kaffeeproben aus Äthiopien, 93 Kaffeeproben aus Brasilien und 28 Proben aus Kolumbien zur Verfügung. Die PCA diente zunächst zur Dimensionsreduktion, wobei die Scores der ersten 29 PC der PCA für die LDA zur Abgrenzung von Kaffeeproben aus den Ländern Äthiopien und Brasilien verwendet wurden. Da bei einer hohen Anzahl von Dimensionen die Gefahr des Overfittings besteht [190], und die Anzahl der Kaffeeproben aus Kolumbien sehr gering war, wurden bei dieser Klasse nur die Scores der ersten 19 PC der PCA für die LDA verwendet. Bei der Analyse der lipophilen bzw. der hydrophilen Metaboliten zur Abgrenzung von Kaffeeproben aus Äthiopien und Brasilien konnten unter Verwendung von 29 PC bei der PCA eine Gesamtvarianz von mindestens 98,1% bzw. 98,4% beschrieben werden. Durch die Verwendung von 19 PC bei der Modellerstellung mit den Kaffeeproben aus Kolumbien lagen diese Prozentsätze nur bei 96,4% bzw. 97,3%. Als interne Validierung diente wie bereits bei der Untersuchung der Kontinente eine MCKV, wobei für die KV jeweils 10% des Probandatensatzes als Testdatensatz verwendet wurde. Abbildung 4-11 zeigt das Ergebnis der MCKV der enthaltenen Klassifizierungsmodelle basierend auf der Analyse der lipophilen und hydrophilen Metaboliten für die zwei Klassen Brasilien und Afrika.

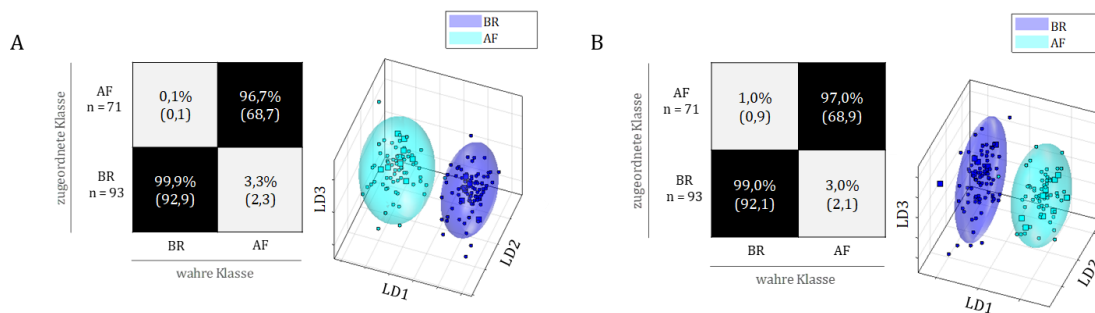


Abbildung 4-11: Ergebnisse der MCKV der PCA-LDA zur Vorhersage des geographischen Ursprungs von Kaffee bezüglich des Landes Brasilien gegen und den Kontinent Afrika aus den $^1\text{H-NMR}$ -Spektren der (A) lipophilen und (B) hydrophilen Kaffeemetaboliten. Jeweils rechts in der Abbildung (A) und (B) ist der LDA-Raum eines Kreuzvalidierungsschrittes dargestellt. Das 95%ige Prognoseellipsoid des Trainingsdatensatzes ist als Wolke dargestellt, die Proben des Testdatensatzes werden durch Rechtecke symbolisiert. Links in der Abbildung (A) und (B) ist die Konfusionsmatrix der MCKV dargestellt. Die x-Achse und die y-Achse zeigen die zugeordnete Klasse bzw. die wahre Klasse. In der Konfusionsmatrix ist die Wahrscheinlichkeit des Vorhersageergebnisses in Prozent angegeben. Anzahl der für die LDA verwendeten PC: 29, Erklärte Varianz: (A) 98,4% und (B) 98,7%. MCKV: Monte-Carlo Kreuzvalidierung; PCA-LDA: Hauptkomponentenanalyse mit linearer Diskriminanzanalyse; LD: Lineare Diskriminanzfunktion; AF: Afrika; BR: Brasilien.

Das Ergebnis der MCKV zeigt, dass das Modell basierend auf der Analyse der lipophilen bzw. der hydrophilen Metaboliten zur Abgrenzung von Kaffeeproben aus Brasilien und Afrika geeignet war. Von dem internen Testdatensatz konnten bei der Verwendung der Daten

der lipophilen Metaboliten jeweils 99,9% und 96,7% der richtigen Klasse zugeordnet werden (Abbildung 4-11(A)). Für das auf den Daten der hydrophilen Metaboliten basierende Modell konnte ein ähnlich gutes Ergebnis erzielt werden (Abbildung 4-11(B)). Hier wurden 99,0% und 97,0% der Proben des internen Testdatensatzes der richtigen Klasse zugeordnet.

Bei Klassifizierungsmodell erstellt durch die Analyse der lipophilen Metaboliten konnten bis auf die Probe S8 alle Proben aus Brasilien der richtigen Klasse zugeordnet werden. Wie bereits bei der MCKV der Klassifizierungsmodelle zur Abgrenzung von Kaffeeproben aus verschiedenen Kontinenten wurden die Proben AF30 und AF150 immer der falschen Klasse zugeordnet als möglicher Grund für die falsche Zuordnung dieser Proben könnte eine im Vergleich zu den anderen Kaffeeproben aus Afrika geringere Intensität der Cafestolsignale sein. Die Testproben 20DU070, AF153 und AF30 konnten bei dem auf den Daten der hydrophilen Metabolite basierenden Modell in der MCKV kein einziges Mal richtig zugeordnet werden. Dies war bei der Probe 20DU070 bereits bei den Modellen zur Abgrenzung von Kaffee aus verschiedenen Kontinenten der Fall. Durch einen Vergleich der ^1H -NMR-Spektren dieser Probe mit anderen Kaffeeproben aus Brasilien konnte eine geringfügig erhöhte Intensität der Signale der aromatischen Protonen der CGA festgestellt werden, was ein Grund für die falsche Zuordnung sein könnte.

Neben dem Modell Brasilien-Afrika wurden für die Abgrenzung von Kaffee aus Brasilien zusätzlich die Klassifizierungsmodelle Brasilien-Asien sowie Brasilien-Nordamerika erstellt. Für Kolumbien waren dies die Klassifizierungsmodelle Kolumbien-Afrika, Kolumbien-Asien und Kolumbien-Nordamerika. Zudem wurden die Klassifizierungsmodelle Äthiopien-Asien, Äthiopien-Nordamerika und Äthiopien-Südamerika erstellt. Eine Übersicht der Ergebnisse der MCKV dieser Klassifizierungsmodelle ist in Tabelle 4-3 dargestellt.

Tabelle 4-3: Übersicht über die Ergebnisse der MCKV der PCA-LDA zur Vorhersage des geographischen Ursprungs von Kaffee bezüglich der Länder Äthiopien, Brasilien und Kolumbien aus den ¹H-NMR-Spektren der lipophilen und hydrophilen Kaffeemetaboliten. Anzahl der für die LDA verwendeten PC: 29 für die Modelle mit Äthiopien und Brasilien; 19 für die Modelle mit Kolumbien. MCKV: Monte-Carlo Kreuzvalidierung; PCA-LDA: Hauptkomponentenanalyse mit linearer Diskriminanzanalyse; PC: Hauptkomponente; AF: Afrika, AS: Asien, NA: Nordamerika, SA: Südamerika, AT: Äthiopien, BR: Brasilien, K: Kolumbien; n: Anzahl richtig zugeordneter Proben aus dem Testdatensatz.

		Modell erstellt aus der Analyse der lipophilen Metaboliten		Modell erstellt aus der Analyse der lipophilen Metaboliten	
Klasse 1	Klasse 2	Richtige Zuordnung Klasse 1	Richtige Zuordnung Klasse 2	Richtige Zuordnung Klasse 1	Richtige Zuordnung Klasse 2
AT	AS	95,6% n = 43,0	94,0% n = 49,8	97,8% n = 44,0	96,8% n = 51,3
AT	NA	97,8% n = 44,0	97,8% n = 137,9	99,8% n = 44,9	99,1% n = 139,7
AT	SA	97,6% n = 43,9	98,4% n = 149,6	99,3% n = 44,7	99,0% n = 150,5
BR	AF	99,9% n = 92,9	96,7% n = 68,7	99,0% n = 92,1	97,0% n = 68,9
BR	AS	96,3% n = 89,6	91,1% n = 48,3	96,0% n = 89,3	94,5% n = 50,1
BR	NA	90,6% n = 84,3	94,3% n = 133,0	93,8% n = 87,2	97,5% n = 137,5
K	AF	89,7% n = 25,1	90,3% n = 64,1	96,8% n = 27,1	96,5% n = 68,5
K	AS	83,6% n = 23,4	88,1% n = 46,7	96,1% n = 26,9	87,5% n = 46,3
K	NA	70,4% n = 19,7	77,7% n = 109,6	89,6% n = 25,1	92,8% n = 130,8

Bis auf das Modell Brasilien-Afrika konnten durch die Verwendung der Daten der hydrophilen Metaboliten bei der MCKV geringfügig bessere Ergebnisse erzielt werden. In den Modellen erstellt durch die Analyse der hydrophilen Metaboliten konnten vergleichsweise mehr Proben aus dem internen Testdatensatz der richtigen Klasse zugeordnet werden. Die Probe 20DU070 konnte nur in dem Modell Brasilien-Afrika bei Verwendung der Daten der lipophilen Metaboliten der richtigen Klasse zugeordnet werden. In allen anderen Modellen basierend auf der Analyse der lipophilen und hydrophilen Metaboliten zur Abgrenzung von Kaffeeproben aus Brasilien wurde diese Probe der falschen Klasse zugeordnet. Bei dieser Probe war in einem Vergleich der ¹H-NMR-Spektren der lipophilen Metaboliten mit anderen Proben aus Brasilien eine geringere Intensität der Kahweol- und Cafestolsignale zu beobachten.

In den Klassifizierungsmodellen zur Abgrenzung von Kaffeeproben aus Äthiopien konnten bei der Analyse der lipophilen Metaboliten 94,0-96,8% und der Analyse der hydrophilen Metaboliten 96,8-99,8% aller Proben der richtigen Klasse zugeordnet werden. Dies zeigt, dass insbesondere die PCA-LDA-Modelle basierend auf den Daten der hydrophilen Metaboliten dazu geeignet waren, Kaffee aus Äthiopien von solchen aus anderen Kontinenten abzutrennen. In den Klassifizierungsmodellen für die Abgrenzung von Kaffee aus Brasilien konnten durch die Verwendung der Daten der lipophilen bzw. hydrophilen Metaboliten

ähnlich gute Ergebnisse erzielt werden. Hier konnten 90,6-99,9% bzw. 93,8-99,0% aller Proben der richtigen Klasse zugeordnet werden. Bei der MCKV konnte die Probe AF 30 in den Modellen basierend auf der Analyse der lipophilen Metaboliten und der hydrophilen Metaboliten nicht der richtigen Klasse zugeordnet werden. In den $^1\text{H-NMR}$ -Spektren der lipophilen Metaboliten konnte eine im Vergleich zu den restlichen afrikanischen Kaffeeproben geringfügig erhöhte Intensität der Cafestolsignale als mögliche Ursache festgestellt werden. Dagegen konnte in den $^1\text{H-NMR}$ -Spektren der hydrophilen Metaboliten kein Grund für die falsche Zuordnung beobachtet werden.

Die Abgrenzung von Kaffeeproben aus Kolumbien mittels PCA-LDA konnte durch die Verwendung der Daten der lipophilen Metaboliten nicht vollständig erreicht werden. Insbesondere bei dem Klassifizierungsmodell Kolumbien-Nordamerika war keine eindeutige Abgrenzung der Kaffeeproben mehr möglich. Mit einer richtigen Zuordnung der Proben von 70,4% und 77,7% war nur eine Tendenz zur Trennung der Klassen erkennbar. Für diese beiden Klassen konnten jedoch durch das Modell basierend auf der Analyse der hydrophilen Metaboliten mit 89,6% und 92,8% richtiger Zuordnung deutlich bessere Ergebnisse erzielt werden. Allerdings kam es in den verschiedenen Modellen auch bei unterschiedlichen Proben aus Kolumbien bei der MCKV zur falschen Zuordnung. Es konnte keine Probe beobachtet werden, die in allen Modellen der falschen Klasse zugeordnet wurde. Ein Grund für die schlechteren Ergebnisse der MCKV bei der Abgrenzung von Kaffee aus Kolumbien könnte hier eine zu geringe Probenanzahl sein. Für die Modellerstellung standen nur 28 Proben aus Kolumbien zur Verfügung. Möglicherweise könnte durch eine deutliche höhere Probenanzahl eine bessere Abgrenzung von Kaffee aus diesem Land erreicht werden.

Neben den falsch zugeordneten Proben aus Brasilien und Äthiopien konnten weiter die Proben 20HA009, AF150, AS200 und AS250 in vielen Modellen bei der MCKV nicht der richtigen Klasse zugeordnet werden. Diese Proben waren bereits bei der Abgrenzung von Kaffeeproben aus den verschiedenen Kontinenten auffällig. Als mögliche Gründe für eine falsche Zuordnung wurden im Vergleich zu den anderen Kaffeeproben der jeweils gleichen Klasse abweichende Verhältnisse der Signalintensitäten von Kahweol und Cafestol bei der Untersuchung der lipophilen Metaboliten genannt.

Außer dem geographischen Ursprung könnte die gute Abgrenzung von Kaffeeproben aus Äthiopien und Brasilien auch durch unterschiedlich angebaute Varietäten sowie durch unterschiedliche Anbau-, Ernte- und Rohkaffeeverarbeitungsverfahren begründet sein. In Brasilien werden hauptsächlich die Varietäten Bourbon, Mundo Novo, Catuai, Icatu, Catucaí angebaut [24, 244]. Dort gibt es Kaffeeplantagen mit einer Größe von kleiner als zehn Hektar und große Plantagen mit über 2000 Hektar [244]. Diese Plantagen werden hauptsächlich im Sonnenanbau betrieben. Die Ernte der Kaffeebohnen erfolgt in Brasilien vermehrt maschinell und die Rohkaffeebohnen werden mit der halb-trockenen oder trockenen Methode verarbeitet [3, 244, 245]. In Kolumbien dagegen werden die Rohkaffeebohnen mit der nassen Methode verarbeitet [15]. Äthiopien verwendet zu 70% die trockene und zu 30% die nasse Methode zur Rohkaffeeverarbeitung [246]. Auch wird in Äthiopien der Kaffee zu 90% in Farmen mit einer Größe von kleiner als einem Hektar mit Schattenanbau produziert [19, 25, 246]. Aufgrund der unterschiedlichen agronomischen Eigenschaften sind in Äthiopien viele ursprüngliche Varietäten und die weltweit größte genetische Variabilität zu finden [247]. Viele Farmer in Äthiopien haben nicht den Zugang zu den Mitteln und der Infrastruktur für Bewässerung oder vergleichsweise teure Pestizide und verzichten auf Düngemittel, Pestizide und Herbizide [19, 246, 247]. Zur Entfernung von Unkraut werden in Äthiopien größtenteils mechanische Verfahren angewendet. Nur 13% der Kaffeeplantagen verwenden

Chemikalien zur Unkrautbekämpfung [246]. Dahingegen werden zersetzte Kaffeeschalen und die Pulpe eingesetzt, um den Wachstum und den Ertrag von Kaffee zu erhöhen [246]. Insbesondere aufgrund der sehr traditionellen Kaffeeproduktion in Äthiopien könnte die gute Abgrenzung von Kaffeeproben aus Äthiopien gegenüber den anderen Klassen möglich sein.

Arana et al. verwendeten die $^1\text{H-NMR}$ -Spektren methanolischer Kaffeeextrakte, um ein Klassifizierungsmodell zur Abgrenzung von Kaffeeproben aus Kolumbien zu erstellen [138]. Dazu verwendeten sie 84 Proben aus Kolumbien, 94 Proben aus Amerika, 22 Proben aus Afrika und zwölf Proben aus Asien. In einem PLS-DA Modell differenzierten Arana et al. kolumbianischen Kaffee gegen Kaffee eines anderen Ursprungs. Durch eine interne Kreuzvalidierung konnten sie 95% und 97% richtige Zuordnung für die Kaffeeproben aus Kolumbien solchen aus anderer geographischer Herkunft erreichen. Sie stellten fest, dass der Grund für eine falsche Zuordnung mancher Proben eine gleich durchgeführt Rohkaffeeverarbeitung sein könnte [138]. Medina et al. konnten bei ihren Untersuchungen sowohl mittels NMR, als auch mittels Nahinfrarot-Spektroskopie kombiniert mit einer PLS-DA eine Trennung von Arabica aus Kolumbien und Arabica aus anderen Ländern erreichen [248]. Die Ergebnisse von Arana et al. und Medina et al. deuten an, dass eine Abgrenzung von Kaffee aus Kolumbien zu Kaffee aus anderen Ländern möglich ist, was im Rahmen dieser Arbeit nicht erreicht werden konnte. Wie bereits erwähnt, könnte eine deutlich höhere Probenanzahl von Kaffee aus Kolumbien zu einer besseren Abgrenzung mittels PCA-LDA führen. Daneben könnte überprüft werden, ob die Verwendung einer anderen multivariaten Methode zur Datenanalyse wie die PLS-DA zu einer besseren Abgrenzung von Kaffeeproben aus Kolumbien führt.

Auch Nunez et al. konnten in ihren nicht zielgerichteten Untersuchungen zum geographischen Ursprung von Kaffee eine Abgrenzung der Erzeugerländer Brasilien, Kolumbien, Äthiopien, Indien, Indonesien, Nicaragua und Uganda erzielen [249]. Sie führten eine PLS-DA mit den HPLC-DAD-Daten wässriger Kaffeeextrakte durch [249]. Rodrigues et al. verwendeten eine Isotopen- und Elementanalyse zur Abgrenzung von Kaffee aus verschiedenen Ländern und konnten durch eine PCA der Daten eine Trennung erreichen [250]. Allerdings wurden im Gegensatz zu dieser Arbeit Rohkaffeebohnen in geringer Anzahl untersucht, was einen Vergleich der Ergebnisse erschwert.

4.4.3.2 Externe Validierung

Analog zu Kapitel 4.4.2.2 wurden die Klassifizierungsmodelle zur Abgrenzung von Kaffeeproben aus den Ländern Brasilien, Äthiopien und Kolumbien extern validiert. Für den externen Testdatensatz standen 18 Proben aus Afrika zur Verfügung, von welchen 10 Proben aus Äthiopien stammten. Von 38 Proben aus Südamerika kamen 23 Proben aus Brasilien und 7 Proben aus Kolumbien. Weiter wurden zur externen Validierung 13 Proben aus Asien und 35 Proben aus Nordamerika verwendet.

Die Ergebnisse der externen Validierung sind beispielhaft für das Modell Brasilien-Afrika basierend auf der Analyse der lipophilen und der hydrophilen Metaboliten in Tabelle 9-7 aufgelistet. Anhand der Tabelle ist zu erkennen, dass alle externen Testproben aus Brasilien in den Modellen basierend auf der Analyse der lipophilen und der hydrophilen Metaboliten der richtigen Klasse zugeordnet wurden. Die *conf*-Werte der Testproben für die Klasse Brasilien bzw. für die Klasse Afrika waren bei der Verwendung der Daten der lipophilen Metaboliten ausnahmslos $<1,00$ bzw. $>1,40$. Demnach lagen alle Testproben

innerhalb des 0,95 Prognoseellipsoides der Klasse Brasilien und außerhalb des 0,95 Prognoseellipsoides der Klasse Afrika. Bis auf die Testprobe S16 war dies auch bei der Verwendung der Daten der hydrophilen Metaboliten der Fall. Testprobe S16 besaß für beide Klassen einen *conf*-Wert $>1,30$ und lag außerhalb der 0,95 Prognoseellipsoide für beide Klassen. Allerdings war der *conf*-Wert für die Klasse Brasilien geringer, wodurch eine richtige Zuordnung dieser Probe erfolgte.

Auch für alle afrikanischen externen Testproben erfolgte im Modell basierend auf der Analyse der lipophilen und der hydrophilen Metaboliten eine richtige Zuordnung. In beiden Modellen wiesen die Proben AF10 und AF3 für beide Klassen einen *conf*-Wert $>1,00$ auf, konnten jedoch aufgrund des geringeren Wertes für die Klasse Afrika dieser richtig zugeordnet werden. Aus den Ergebnissen der externen Validierung für das Modell Brasilien-Afrika konnte geschlossen werden, dass dieses sowohl bei der Verwendung der Daten der lipophilen als auch der hydrophilen Metaboliten eine sehr gute Vorhersagefähigkeit besaß. Von 41 externen Testproben konnten alle der richtigen Klasse zugeordnet werden.

Bei der externen Validierung der anderen Modelle basierend auf der Analyse der lipophilen Metaboliten waren die Proben AF3, AF22a und N5 auffällig. Diese wurden in über der Hälfte der Modelle der falschen Klasse zugeordnet. Diese Proben konnten auch bei den Modellen basierend auf die Analyse der lipophilen Metaboliten zur Abgrenzung von Kaffeeproben aus den verschiedenen Kontinenten in der externen Validierung nicht der richtigen Klasse zugeordnet werden. Auch zeigte sich, dass die Modelle zur Abgrenzung von Kaffeeproben aus Kolumbien im Vergleich zu denen mit Brasilien oder Äthiopien eine schlechtere Vorhersagefähigkeit aufwiesen. Bei dem Modell Kolumbien-Asien sowie Kolumbien-Nordamerika wurden nur 80,0% der externen Testproben der richtigen Klasse zugeordnet. Die vergleichsweise schlechtere Vorhersagefähigkeit wurde bereits durch die Ergebnisse der internen Validierung angedeutet.

Die gleichen externen Testproben wurden auch in den Modellen erstellt durch die Analyse der hydrophilen Metaboliten wiederholt der falschen Klasse zugeordnet. Insbesondere die Proben AF22a und AF3 konnten in 80,0% der Modelle nicht der richtigen Klasse zugeordnet werden. Auch bei den Modellen basierend auf der Analyse der hydrophilen Metaboliten zur Abgrenzung von Kaffeeproben aus den verschiedenen Kontinenten waren diese Proben auffällig. Die Modelle zur Abgrenzung von Kaffeeproben aus Brasilien zeigten die beste Vorhersagefähigkeit. In der externen Validierung dieser Klassifizierungsmodelle konnten nahezu alle Proben der richtigen Klasse zugeordnet werden. Auch die Anwendung der Modelle zur Abgrenzung von Kaffeeproben aus Äthiopien auf externe Proben zeigte einen Vorhersagefehler von $<10,0\%$. Diese Vorhersagefähigkeit konnte bei den Modellen zur Abgrenzung von Kaffeeproben aus Kolumbien nicht erreicht werden. In den Modellen Kolumbien-Afrika und Kolumbien-Asien konnten nur 80,0% der externen Proben der richtigen Klasse zugeordnet werden.

4.4.4 Univariate Analyse der $^1\text{H-NMR}$ -spektroskopischen Daten mittels Kruskal-Wallis Test

Die Untersuchung der $^1\text{H-NMR}$ -Spektren mittels Kruskal-Wallis Test diente zur Erkennung von spektralen Unterschieden zwischen den Klassen. Zur Vermeidung falsch-positiver Ergebnisse bei der Überprüfung aller Buckets wurde das Signifikanzniveau einer Šidák-Korrektur unterzogen, welche die Anzahl der Buckets berücksichtigt. Zur Analyse der Spektren

der lipophilen und hydrophilen Metaboliten wurde ein Signifikanzniveau von $\alpha = 0,05$ gewählt. Mit der Verwendung von 1775 Buckets bei den Spektren der lipophilen und 983 Buckets bei den Spektren der hydrophilen Metaboliten ergab sich ein korrigiertes Signifikanzniveau von $\alpha_{\text{korrr}} = 2,89 \cdot 10^{-5}$ und $\alpha_{\text{korrr}} = 5,22 \cdot 10^{-5}$. Für die Untersuchung der univariaten Unterschiede wurden jeweils Kaffeeproben aus den Ländern Brasilien und Äthiopien gegen Kaffeeproben aus anderen Kontinenten aufgetragen. Aufgrund der unzureichenden Abtrennung von Kaffee aus Kolumbien wurde für diese Klasse kein Kruskal-Wallis Test durchgeführt. Abbildung 4-12 zeigt das Ergebnis des Kruskal-Wallis Tests von den $^1\text{H-NMR}$ -Daten der lipophilen Metaboliten der Klassen Afrika, Asien, Nordamerika und Brasilien. Pink hinterlegt sind die Buckets, bei welchen der p -Wert kleiner war als das korrigierte Signifikanzniveau α_{korrr} . Die spektralen Daten dieser Buckets wiesen somit Unterschiede zwischen den Klassen auf.

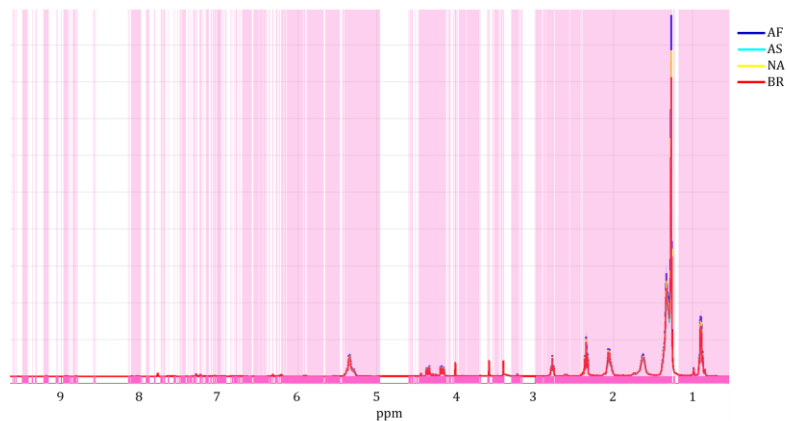


Abbildung 4-12: Datenanalyse zur Untersuchung univariater Unterschiede zwischen $^1\text{H-NMR}$ -Spektren der lipophilen Kaffee-metaboliten der Klassen Afrika, Asien, Nordamerika und Brasilien mittels Kruskal-Wallis Test. Pink hinterlegt sind die Buckets mit einem kleineren p -Wert als das Signifikanzniveau. $\alpha = 0,05$, Anzahl Buckets: 1775, nach Šidák korrigiertes Signifikanzniveau $\alpha_{\text{korrr}} = 2,89 \cdot 10^{-5}$. AF: Afrika; AS: Asien; NA: Nordamerika; BR: Brasilien

Anhand Abbildung 4-12 ist zu erkennen, dass nahezu alle Buckets der $^1\text{H-NMR}$ -Spektren der lipophilen Metaboliten einen p -Wert kleiner als $2,89 \cdot 10^{-5}$ und somit Informationen zu Unterschieden besaßen. Bei näherer Betrachtung der p -Werte einzelner Buckets wurde ersichtlich, dass Buckets mit Signalen von TAG und gesättigten FA sowie UFA sehr geringe Werte $< 1,00 \cdot 10^{-10}$ aufwiesen. Daraus ließ sich schließen, dass in diesen Buckets signifikante Unterschiede zwischen den Klassen vorhanden waren. Auch der Kruskal-Wallis Test der spektralen Daten der Klassen Äthiopien, Asien, Nordamerika und Südamerika ergab sehr kleine p -Werte für diese Buckets.

Auch nahezu alle Buckets der $^1\text{H-NMR}$ -Spektren der hydrophilen Metaboliten wiesen einem p -Wert $< 5,22 \cdot 10^{-5}$ auf, wodurch in den Buckets der $^1\text{H-NMR}$ -Spektren signifikante Unterschiede zwischen den Klassen Afrika, Asien, Nordamerika und Brasilien vorhanden waren (Abbildung 9-26). Die Buckets mit den Signalen von aliphatischen organischen Säuren wie Propionsäure, Äpfelsäure und Maleinsäure sowie von CGA wiesen dabei sehr geringe Werte $< 1,00 \cdot 10^{-10}$ auf. Die gleiche Feststellung ergab sich bei dem Kruskal-Wallis Test der spektralen Daten der hydrophilen Metaboliten der Klassen Äthiopien, Asien, Nordamerika und Südamerika.

Die Ergebnisse der Kruskal-Wallis Tests zeigten, dass sowohl für die NMR-Spektren der lipophilen als auch der hydrophilen Metaboliten nahezu alle Buckets Unterschiede aufwiesen. Die deutlichsten Unterschiede waren in den $^1\text{H-NMR}$ -Spektren der lipophilen

Metaboliten bei den Signalen der TAG und gesättigten FA und UFA zu beobachten. Die deutlichsten Unterschiede in den ^1H -NMR-Spektren der hydrophilen Metaboliten traten bei den Signalen von Propionsäure, Äpfelsäure, Maleinsäure und CGA auf. Allerdings konnte keine Aussage getroffen werden, ob diese Unterschiede, welche durch den Kruskal-Wallis Test angezeigt wurden, auch der Grund für eine Abgrenzung der Kaffeeproben in den Klassifizierungsmodellen waren. Jedoch konnte anhand des Kruskal-Wallis Tests gezeigt werden, dass die Abgrenzung der Klassen nicht durch Anwesenheit oder Abwesenheit von konkreten Signalen definierter Verbindungen verursacht wurde. Die Abgrenzung der Kaffeeproben in den Klassifizierungsmodellen wurde durch ein Zusammenspiel unterschiedlicher Verbindungen verursacht, welche in höheren oder niedrigeren Konzentrationen vorhanden waren.

4.4.5 Identifizierung möglicher lipophiler Metaboliten als Ursache für die Abgrenzung von Kaffee nach ausgewählten Ländern

Eine Identifizierung lipophiler Metaboliten, die einen Einfluss auf die Abgrenzung der Klassen in den Modellen haben, kann über den Loadingsplot der PCA-LDA erfolgen. Wie in Kapitel 3.5.3 erwähnt, zeigen die Loadings, wie groß die Beeinflussung der einzelnen Variablen auf die Trennung der Proben im LDA-Raum ist. Mit einem hohen oder geringen Loadingswert befinden sich die Variablen im Loadingsplot weit vom Datenmittelpunkt entfernt und kennzeichnen so spektrale Bereiche, die eine Mitursache für die Trennung der Klassen darstellen. Die Analyse des Loadingsplot erfolgte für die Abgrenzung von Kaffeeproben aus den Ländern Brasilien und Äthiopien gegen die entsprechenden Kontinente. Die Klassifizierungsmodelle für die Abgrenzung von Kaffeeproben aus Kolumbien wurden außen vorgelassen, da die Trennung der Klassen in diesen Modellen unzureichend war (vgl. Kapitel 4.4.3).

Die Abgrenzung von Kaffeeproben aus Brasilien erfolgte anhand der Klassifizierungsmodelle Brasilien-Afrika, Brasilien-Asien und Brasilien-Nordamerika. Abbildung 4-13 zeigt den LDA-Raum und dazugehörigen Loadingsplot zur Abgrenzung von Kaffeeproben aus Brasilien und Afrika. Die Proben beider Klassen zeigten im zweidimensionalen LDA-Raum zwei Cluster, welche entlang der ersten linearen Diskriminanzfunktion (LD1) vollständig getrennt waren. Dabei besaß die Klasse Brasilien positive und die Klasse Afrika negative Werte für LD1. Für den Loadingsplot bedeutete dies, dass Variablen mit hohen Loadingswerten für LD1 einen Einfluss auf die Abgrenzung der Klasse Brasilien hatten. Dahingegen stieg der Einfluss einer Variable auf die Abgrenzung der Klasse Afrika, je geringer der Loadingswert für LD1 war.

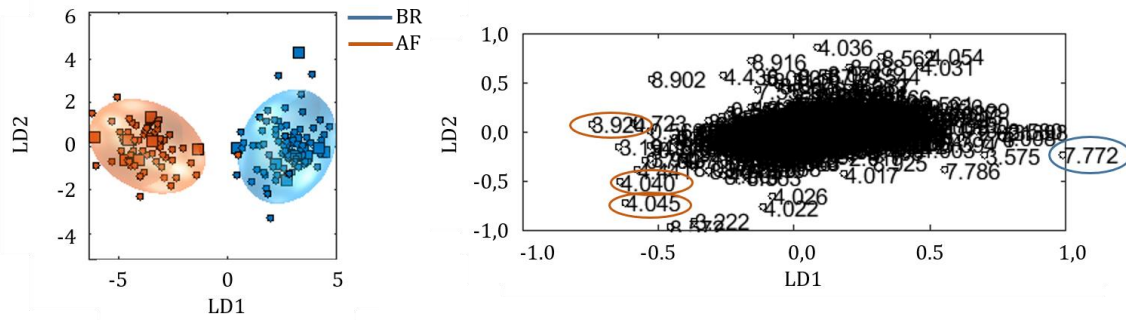


Abbildung 4-13: Links: zweidimensionaler LDA-Raum für die Klasse Brasilien (blau) und Afrika (orange). Rechts: Zugehöriger Loadingsplot der 1775 verwendeten Buckets der ¹H-NMR-Spektren der lipophilen Metaboliten als Variablen. Die Zahlenwerte kennzeichnen die spektrale Region im ¹H-NMR-Spektrum. Die blau/orange umkreisten Variablen haben für die Trennung der Klasse Brasilien/Afrika den größten Einfluss. BR: Brasilien; AF: Afrika; LD: Lineare Diskriminanzfunktion; PCA-LDA: Hauptkomponentenanalyse mit linearer Diskriminanzanalyse

Anhand des Loadingsplots in Abbildung 4-13 ist zu erkennen, dass sich sehr viele Variablen im Datenmittelpunkt befanden. Diese hatten auf die Trennung der Klassen keinen Einfluss. Die Variablen $\delta_H = 7,77$ ppm und $\delta_H = 4,01$ ppm besaßen einen vergleichsweise hohen Loadingswert für LD1 und standen in gleichgerichteter Beziehung mit den positiven Werten im LDA-Raum für die Kaffeeprobe aus Brasilien. Somit hatten diese Variablen einen Einfluss auf die Abgrenzung der Klasse Brasilien. Die Variablen $\delta_H = 3,92$ ppm, $\delta_H = 4,04$ ppm, $\delta_H = 4,05$ ppm besaßen relativ geringe Loadingswerte für LD1 und wiesen einen gleichgerichteten Zusammenhang mit den negativen Werten der afrikanischen Kaffeeprobe im LDA-Raum auf. Diese Variablen hatten einen Einfluss auf die Trennung der Klasse Afrika. Für einen direkten Vergleich wurden die log-transformierten Integrale der Variablen (Buckets) $\delta_H = 7,77$ ppm, $4,01$ ppm, $3,92$ ppm und $4,04$ ppm für die Klasse Brasilien und Afrika in jeweils in einem Box-Whisker-Plot dargestellt (Abbildung 4-14).

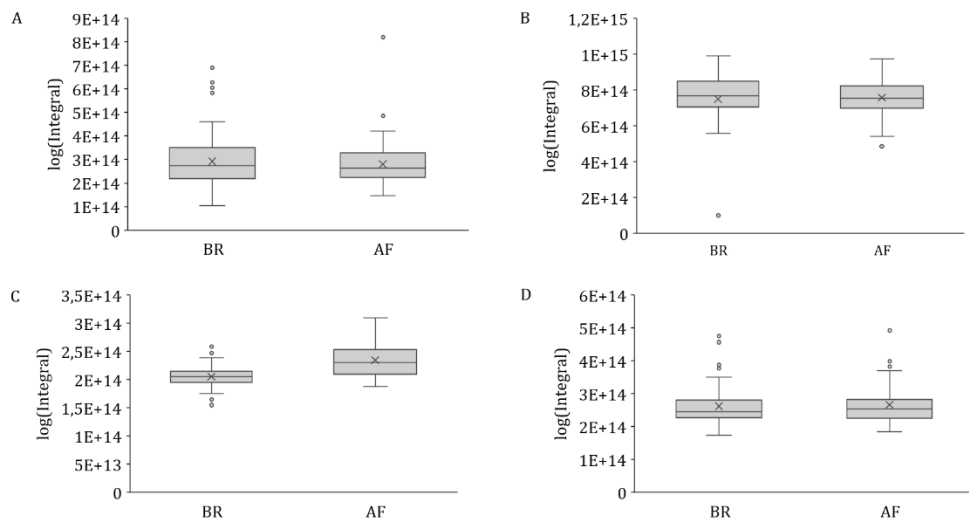


Abbildung 4-14: Darstellung der log-transformierten Integrale der für die Unterscheidung der Klassen Brasilien und Afrika relevanten Variablen bei der Analyse der lipophilen Metaboliten in einem Box-Whisker-Plot. (A) Variable bei $\delta_H = 7,77$ ppm; (B) Variable bei $\delta_H = 4,01$ ppm; (C) Variable bei $\delta_H = 3,92$ ppm; (D): Variable bei $\delta_H = 4,04$ ppm. BR: Brasilien; AF: Afrika

Die Box-Whisker-Plots der einzelnen Variablen zeigen, dass zwischen den Klassen nur minimale Unterschiede in den Integralen vorlagen. Die Variablen $\delta_H = 7,77$ ppm und

4,01 ppm, welche für die Abgrenzung der Kaffeeproben aus Brasilien mitverantwortlich waren, wiesen sehr ähnliche Integralwerte für beiden Klassen auf. Lediglich der Median dieser Variablen war bei den Kaffeeproben aus Brasilien geringfügig erhöht (Abbildung 4-14 (A) und (B)). Bei der Variable $\delta_H = 4,04$ ppm, welche zur Unterscheidung von afrikanischen Kaffeeproben beitrug, war der Median der log-transformierten Integrale in der Klasse Afrika leicht höher (Abbildung 4-14 (C)). Für die Variable $\delta_H = 3,92$ ppm zeigte sich zwischen den Klassen ein größerer Unterschied. Die log-transformierten Integrale der Klasse Brasilien wiesen eine geringere Streuung auf und besaßen kleinere Werte (Abbildung 4-14 (D)).

In Abbildung 4-15 sind die spektralen Bereiche der vier Variablen $\delta_H = 3,92$ ppm, 4,01 ppm, 4,05 ppm und 7,77 ppm im $^1\text{H-NMR}$ -Spektrum der lipophilen Metaboliten einer afrikanischen Kaffeeprobe farblich hinterlegt. Die für die Trennung der brasilianischen Kaffeeproben mitverantwortlichen Variablen $\delta_H = 7,77$ ppm und 4,01 ppm sind im Spektrum blau hinterlegt und konnten den Signalen von Koffein zugeordnet werden. Dahingegen konnten die Variablen $\delta_H = 3,92$ ppm und 4,01 ppm, welche zur Trennung der Klasse Afrika beitrugen und im Spektrum orange hinterlegt sind, nicht identifiziert werden. Aus Abbildung 4-15 ist zu erkennen, dass sich in diesen spektralen Bereichen sehr kleine Signale befanden, welche teilweise überlagert waren und ein sehr geringes S/N aufwiesen. Dies erschwerte eine Identifizierung der Signale in diesen spektralen Bereichen.

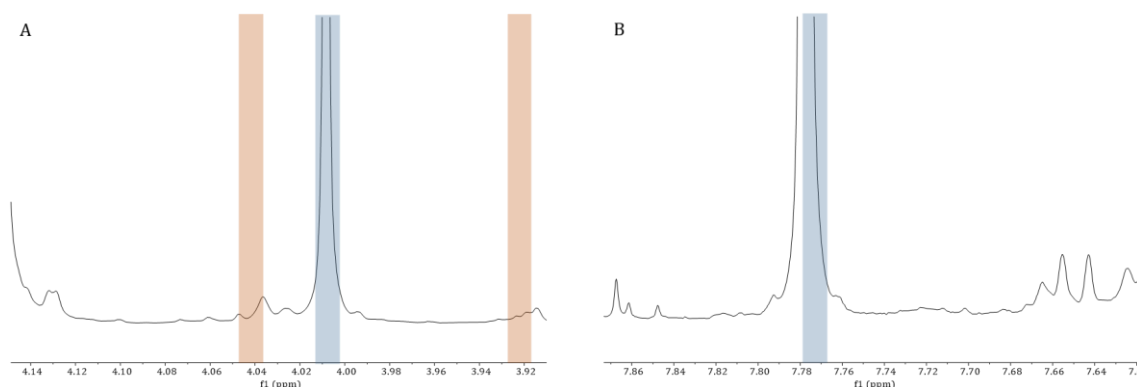


Abbildung 4-15: Ausschnitte der 400 MHz $^1\text{H-NMR}$ -Spektren der lipophilen Metaboliten in der Probe AF15 gemessen in $\text{CDCl}_3:\text{MeOH-}d_4$ 1:1 (v/v). (A) $\delta_H = 3,93 - 4,15$ ppm; (B) $\delta_H = 7,61 - 7,88$ ppm. Orange hinterlegt sind die spektralen Bereiche der Variablen $\delta_H = 3,92$ ppm und 4,04 ppm, welche für die Abgrenzung der Klasse Afrika mitverantwortlich waren. Blau hinterlegt sind die spektralen Bereiche der Variablen $\delta_H = 7,77$ ppm und 4,01 ppm, welche die Abgrenzung der Klasse Brasilien beeinflussten.

Für die Unterscheidung der Klassen Brasilien und Asien waren erneut die Variablen mit den spektralen Bereichen der Koffeinsignale für die Abgrenzung der brasilianischen Kaffeeproben mitverantwortlich. Ein Teil der Variablen, welche mit den Werten der Kaffeeproben aus Asien im LDA-Raum gleichgerichtet zusammenhingen, konnten bereits für die Klasse Afrika nicht identifiziert werden. Die Variablen $\delta_H = 4,44$ ppm und 9,18 ppm wiesen hohe Loadingswerte für LD1 auf und zeigten einen gleichgerichteten Zusammenhang mit den positiven LD1-Werten der Kaffeeproben aus Asien im LDA-Raum. Diese Variablen konnten der Verbindung Trigonellin zugeordnet werden. Bei dem Klassifizierungsmodell Brasilien-Nordamerika zeigte sich, dass die spektralen Bereiche der Koffeinsignale die Abgrenzung von Kaffeeproben aus Brasilien nicht beeinflussten. Für die Trennung von Brasilien waren unter anderem die Variablen $\delta_H = 2,02$ ppm 3,29 ppm und 3,47 ppm ursächlich. Allerdings konnten diesen Signalen keine entsprechenden Verbindungen zugeordnet werden. Beispielsweise war in dem spektralen Bereich der Variable $\delta_H = 2,02$ ppm eine angedeutete Schulter bei dem Signal der UFA erkennbar (Abbildung 9-27). In dem spektralen Bereich der Variable $\delta_H = 3,47$ ppm können die Protonen der Galactose Signale erzeugen. In diesem

Bereich waren viele Signale mit einem geringen S/N zu finden, die überlagert waren (Abbildung 9-27). Aus diesem Grund konnte auch hier keine Verbindung eindeutig identifiziert werden. Der Loadingswert der Variable $\delta_H = 9,54$ ppm zeigte einen Zusammenhang mit den Werten der Kaffeeproben aus Asien im LDA-Raum und könnte daher einen Einfluss auf die Trennung haben. Dieser Variable konnte die Verbindung HMF zugeordnet werden. Allerdings wird HMF stark von der Röstung beeinflusst. Dies deutet an, dass sich der Einfluss der Abgrenzung asiatischer Kaffeeproben nicht auf den geographischen Ursprung, sondern auf unterschiedliche Röstgrade beziehen könnte.

Bei den Klassifizierungsmodellen für die Abgrenzung von Kaffeeproben aus Äthiopien zeigte die Variable $\delta_H = 3,92$ ppm einen Einfluss auf die Trennung der Kaffeeproben von Äthiopien. Diese Variable kennzeichnete einen spektralen Bereich mit keiner bekannten Verbindung, sodass eine direkte Identifizierung nicht möglich war. Dieser spektrale Bereich zeigte viele Signale mit geringer Intensität und starker Überlagerung (Abbildung 9-28). Durch den Box-Whisker-Plot ist jedoch ersichtlich, dass die äthiopischen Kaffeeproben im Vergleich zu den Kaffeeproben anderer Klassen ein höheres log-transformiertes Integral aufwiesen (Abbildung 4-16). Auch die Variable $\delta_H = 0,71$ ppm zeigte in allen Modellen einen Einfluss auf die Trennung der Klasse Äthiopien. Der spektrale Bereich dieser Variable kennzeichnet Signale von Sterolen wie beispielsweise Stigmasterin. Aufgrund der auch hier geringen Signalintensitäten konnte jedoch keine eindeutige Aussage über die Identifizierung der Variable getroffen werden.

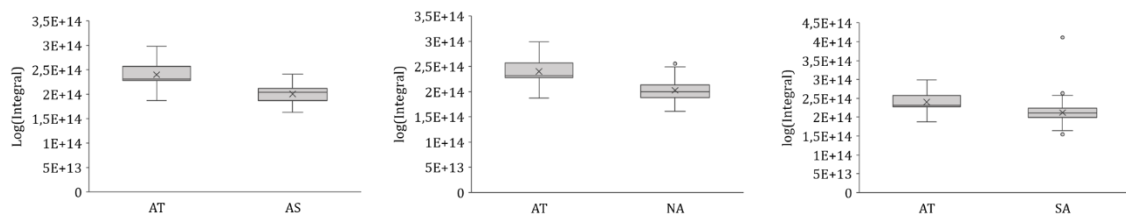


Abbildung 4-16: Darstellung der log-transformierten Integrale für die Variable $\delta_H = 3,92$ ppm in den Klassifizierungsmodellen (A) Äthiopien-Asien, (B) Äthiopien-Nordamerika und (C) Äthiopien-Südamerika. Die Variable $\delta_H = 3,92$ zeigte in allen Klassifizierungsmodellen eine Beeinflussung der Trennung von Äthiopien. AT: Äthiopien; AS: Asien; NA: Nordamerika; SA: Südamerika

In den Klassifizierungsmodellen Äthiopien-Nordamerika und Äthiopien-Südamerika zeigten die Variablen $\delta_H = 3,40$ ppm, $3,58$ ppm, $4,01$ ppm und $7,78$ ppm jeweils für die Trennung der beiden Klassen Nord- und Südamerika einen Einfluss im Loadingsplot. In diesen spektralen Bereichen traten Signale von Koffein auf (Abbildung 9-28). Bei der Untersuchung des Loadingsplots für das Klassifizierungsmodell Äthiopien-Asien konnte keine Verbindung als Grund für die Trennung der Klasse Asien identifiziert werden. Die Variablen $\delta_H = 6,86$ ppm und $7,11$ ppm, welche die Trennung von Asien beeinflussten, lagen in einem spektralen Bereich, in welchem auch die aromatischen Protonen der CGA-Signale erzeugen. Allerdings befanden sich auch in diesem Bereich viele überlappende Signale mit geringer Intensität, sodass keine gesicherte Aussage über die Identifizierung dieser Variablen getroffen werden konnte.

4.4.6 Identifizierung möglicher hydrophiler Metaboliten als Ursache für die Abgrenzung von Kaffee nach ausgewählten Ländern

Auch die Identifizierung möglicher hydrophiler Metaboliten, die einen Einfluss auf die Abgrenzung von Kaffeeproben aus den Ländern Brasilien und Äthiopien hatten, erfolgte zunächst über den Loadingsplot. Dazu wurden die Klassifizierungsmodelle Brasilien-Afrika, Brasilien-Asien, Brasilien-Nordamerika, Äthiopien-Asien, Äthiopien-Nordamerika und Äthiopien-Südamerika basierend auf der Analyse der hydrophilen Metaboliten verwendet. Die Klassifizierungsmodelle zur Abgrenzung von Kaffeeproben aus Kolumbien wurden wie bei der Analyse der lipophilen Metaboliten aufgrund der unzureichenden Trennung und Vorhersagefähigkeit außen vorge lassen. In Abbildung 4-17 ist der zweidimensionale LDA-Raum und der dazugehörige Loadingsplot für das Klassifizierungsmodell Brasilien-Afrika unter der Verwendung der Daten der hydrophilen Metaboliten dargestellt. Die Proben der Klassen Brasilien und Afrika zeigten im zweidimensionalen LDA-Raum eine Trennung entlang der LD1, wobei die Klasse Brasilien negative und die Klasse Afrika positive LD1-Werte im LDA-Raum aufwies. Im Loadingsplot hatten daher die Variablen mit einem geringen Loadingswert für LD1 einen Einfluss auf die Trennung der brasilianischen Kaffeeproben. Die Variablen mit einem hohen Loadingswert für LD1 hatten einen Einfluss auf die Trennung der afrikanischen Kaffeeproben.

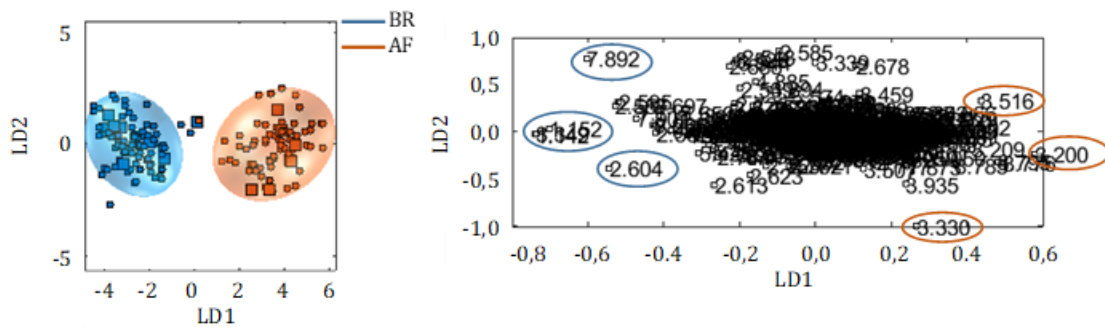


Abbildung 4-17: Links: zweidimensionaler LDA-Raum für die Klasse Brasilien (blau) und Afrika (orange). Rechts: Zugehöriger Loadingsplot der 983 verwendeten Buckets der $^1\text{H-NMR}$ -Spektren der hydrophilen Metaboliten als Variablen. Die Zahlenwerte kennzeichnen die spektrale Region im $^1\text{H-NMR}$ -Spektrum. Die blau/orange umkreisten Variablen hatten für die Trennung der Klasse Brasilien/Afrika den größten Einfluss. BR: Brasilien; AF: Afrika; LD: Lineare Diskriminanzfunktion; PCA-LDA: Hauptkomponentenanalyse mit linearer Diskriminanzanalyse

Wie bei der Analyse der lipophilen Metaboliten befanden sich sehr viele Variablen im Datenmittelpunkt des Loadingsplots und hatten somit keinen Einfluss auf die Trennung der Klassen Brasilien und Afrika. Die Variablen $\delta_{\text{H}} = 1,14 \text{ ppm}$, $1,15 \text{ ppm}$, $2,60 \text{ ppm}$ und $7,89 \text{ ppm}$ waren weiter vom Datenmittelpunkt entfernt und besaßen vergleichsweise geringe Loadingswerte für LD1. Die negativen Loadingswerte dieser Variablen standen im Zusammenhang mit den negativen Werten der Kaffeeproben aus Brasilien im LDA-Raum und hatten somit einen Einfluss auf die Trennung dieser Klasse. Im Gegensatz dazu besaßen die Variablen $\delta_{\text{H}} = 3,20 \text{ ppm}$, $3,30 \text{ ppm}$, $3,37 \text{ ppm}$ und $3,52 \text{ ppm}$ hohe Loadingswerte für LD1 und standen in gleichgerichteter Beziehung mit den positiven LD1-Werten für die Kaffeeproben aus Afrika im LDA-Raum. Die Abgrenzung der Klasse Afrika wurde somit von diesen Variablen beeinflusst. Um einen Unterschied der beiden Klassen in den spektralen Bereichen besser beobachten zu können, wurden die log-transformierten Integrale der Variablen $\delta_{\text{H}} = 1,14 \text{ ppm}$, $1,15 \text{ ppm}$, $2,06 \text{ ppm}$, $3,20 \text{ ppm}$, $3,30 \text{ ppm}$ und $3,37 \text{ ppm}$ für die Klassen Brasilien und Afrika in Box-Whisker-Plots dargestellt.

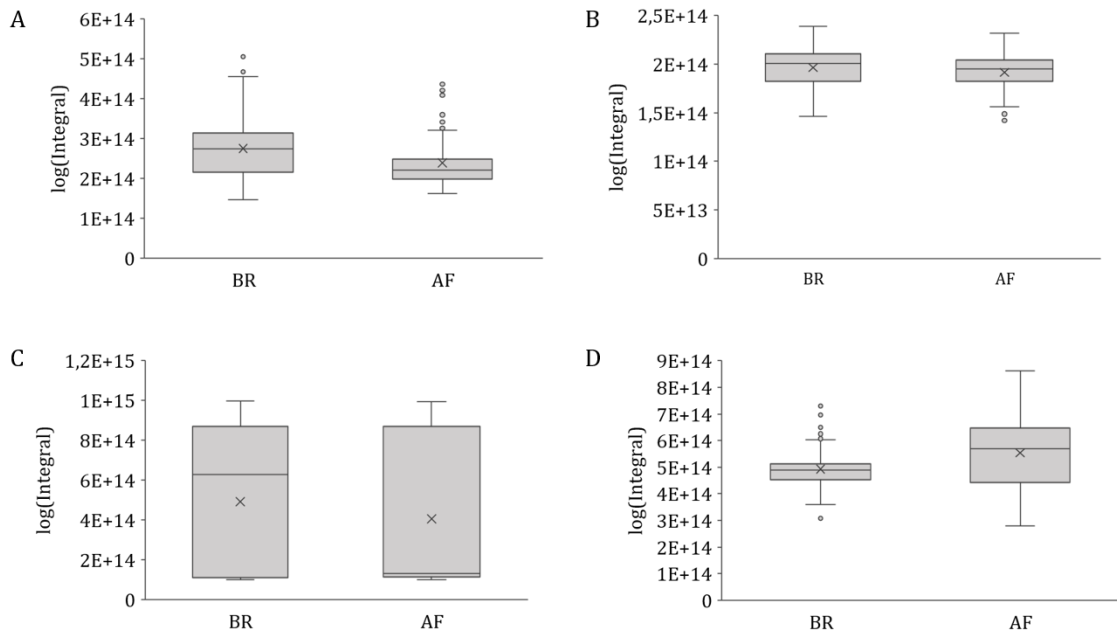


Abbildung 4-18: Darstellung der log-transformierten Integrale der für die Unterscheidung der Klassen Brasilien und Afrika relevanten Variablen bei der Analyse der hydrophilen Metaboliten in einem Box-Whisker-Plot. (A) Variable bei $\delta_H = 1,14$ ppm; (B) Variable bei $\delta_H = 2,06$ ppm; (C) Variable bei $\delta_H = 3,20$ ppm; (D) Variable bei $\delta_H = 3,37$ ppm. BR: Brasilien; AF: Afrika

Durch die Box-Whisker-Plots ist zu erkennen, dass bei den log-transformierten Integralen der Variablen, welche die Abgrenzung der Kaffeeproben beeinflussten, zwischen den Klassen nur minimale Unterschiede vorhanden waren. Die Klasse Afrika wies bei der Variablen $\delta_H = 1,14$ ppm geringere log-transformierte Integrale als die Klasse Brasilien auf (Abbildung 4-18(A)). Geringere Integralwerte bei der Klasse Afrika konnten auch bei der Variablen $\delta_H = 1,15$ ppm beobachtet werden. Für die Variable $\delta_H = 2,06$ ppm waren die Mittelwerte sowie die Streuung der Werte für beide Klassen nahezu identisch (Abbildung 4-18(B)). Bei der Variablen $\delta_H = 3,20$ ppm dagegen war der Mittelwert der log-transformierten Integrale sehr ähnlich, jedoch war die Verteilung beider Werte in den Klassen unterschiedlich. Während bei der Klasse Brasilien der Median leicht oberhalb des Mittelwertes lag, war dieser bei der Klasse Afrika deutlich unterhalb des entsprechenden Mittelwertes (Abbildung 4-18(C)). Auch bei der Variablen $\delta_H = 3,37$ ppm waren die Mittelwerte der log-transformierten Daten sehr ähnlich. Allerdings wies die Klasse Brasilien eine deutlich geringere Streuung der Werte auf als die Klasse Afrika (Abbildung 4-18(D)).

Auch für die Analyse der hydrophilen Metaboliten konnte nicht jeder Variablen, welche einen Einfluss auf die Abgrenzung der Klassen hatte, eine Verbindung zugeordnet werden. So konnte bei den Variablen $\delta_H = 1,14$ ppm und $\delta_H = 1,15$ ppm, welche im $^1\text{H-NMR}$ -Spektrum direkt nebeneinanderlagen, keine Identifizierung erfolgen. In diesem spektralen Bereich waren im $^1\text{H-NMR}$ -Spektrum Signale mit sehr geringer Intensität ersichtlich (Abbildung 4-19). Die Variable $\delta_H = 2,06$ ppm bildete einen spektralen Bereich ab, in welchem die Protonen der Chinasäure Signale erzeugen. Allerdings konnte Chinasäure als Grund für eine Trennung der Klassen Brasilien und Afrika ausgeschlossen werden. Durch den Vergleich der log-transformierten Integralwerte konnte gezeigt werden, dass diese zwischen den beiden Klassen keinen Unterschied aufwiesen (Abbildung 4-18(B)). Zusätzlich erzeugt Chinasäure weitere Signale in den Bereichen $\delta_H = 1,80 - 1,90$ ppm, $1,90 - 2,10$ ppm, $3,50 - 3,60$ ppm und $4,00 - 4,20$ ppm. Die entsprechenden Variablen lagen im Loadingsplot

im Datenmittelpunkt, sodass diese im Gegensatz zur Variablen $\delta_H = 2,06$ ppm keinen Einfluss auf die Trennung der Klasse Brasilien hatten.

Die Variable $\delta_H = 3,20$ ppm, welche die Trennung der afrikanischen Kaffeeproben beeinflusste, konnte der Verbindung Cholin zugeordnet werden. Die Variablen $\delta_H = 3,30$ ppm und $3,52$ ppm kennzeichneten spektrale Bereiche, in welchen Signale von Koffein ersichtlich waren (Abbildung 4-19). Das Signal, welches in dem spektralen Bereich der Variablen $\delta_H = 3,37$ ppm zu beobachten war, konnte nicht durch Aufstockungen mit Standardsubstanzen oder 2D-NMR-Experimenten identifiziert werden. Zur vollständigen Identifizierung schlagen Kim et al. die Kopplung von NMR mit HPLC vor [133]. Dies ermöglicht die Einsicht von spektralen Daten einzelner Metaboliten nach der Abtrennung durch die HPLC [133]. Durch chromatographischen Trennung könnte die Verbindung, welche das Signal bei $\delta_H = 3,37$ ppm erzeugt, im Anschluss einzeln mit der NMR analysiert und so identifiziert werden. Weiter könnten andere Detektionen wie die MS zusätzliche Informationen liefern, die eine bessere Identifizierung ermöglichen könnten [135].

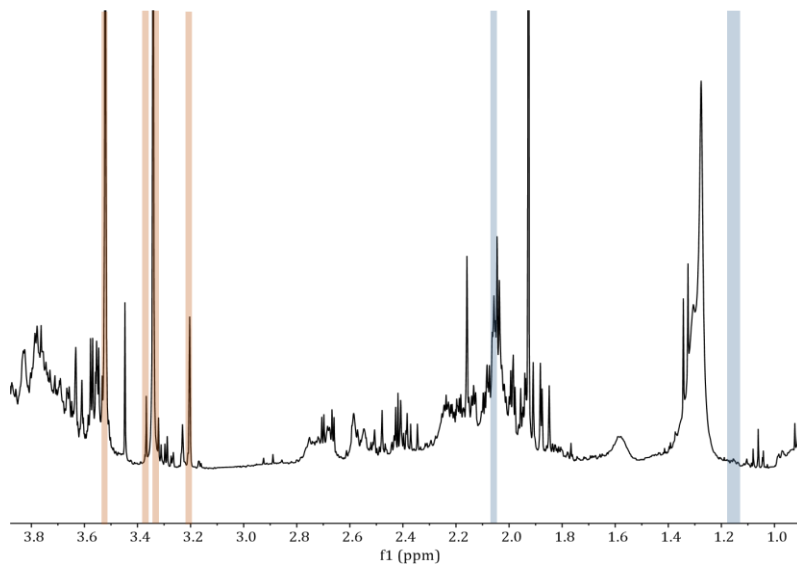


Abbildung 4-19: Ausschnitt des 400 MHz ^1H -NMR-Spektrums der hydrophilen Metaboliten in der Probe AF15 im Bereich $\delta_H = 0,80 - 3,90$ ppm gemessen in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Orange hinterlegt sind die spektralen Bereiche der Variablen $\delta_H = 3,20$ ppm, $3,30$ ppm, $3,37$ ppm und $3,52$ ppm, welche für die Abgrenzung der Klasse Afrika mitverantwortlich waren. Blau hinterlegt sind die spektralen Bereiche der Variablen $\delta_H = 1,14$ ppm, $1,15$ ppm und $2,06$ ppm, welche die Abgrenzung der Klasse Brasilien beeinflussten.

Auch für die Modelle Brasilien-Asien und Brasilien-Nordamerika zeigten die Loadingswerte von LD1 der Variablen $\delta_H = 1,14$ ppm und $\delta_H = 1,15$ ppm einen gleichgerichteten Zusammenhang mit den Werten für LD1 im zweidimensionalen LDA-Raum der Kaffeeproben aus Brasilien. Daneben hatten die Variablen $\delta_H = 3,33$ ppm und $\delta_H = 3,34$ ppm einen Einfluss auf die Trennung der Klasse Asien. Diese Variablen kennzeichneten wieder spektrale Bereiche mit Koffeinsignalen. Für die Klasse Nordamerika zeigten die Variablen $\delta_H = 3,20$ ppm und $\delta_H = 3,21$ ppm einen Einfluss auf die Trennung, welche den spektralen Bereich der Cholinsignale kennzeichneten.

Bei dem Klassifizierungsmodell Äthiopien-Asien konnten keine relevanten Verbindungen eindeutig identifiziert werden. Auf die Trennung der Klasse Äthiopien hatten die Variablen $\delta_H = 3,37$ ppm, $3,45$ ppm und $5,15$ ppm einen Einfluss. In diesen spektralen Bereichen waren überwiegend überlappende Signale mit geringer Signalintensität vorhanden, sodass diesen Signalen keine Verbindung zugeordnet werden konnte. In den Klassifizierungsmodellen Äthiopien-Nordamerika und Äthiopien-Südamerika hatten die Variablen

$\delta_H = 8,78$ ppm und $8,80$ ppm einen Einfluss auf die Trennung der äthiopischen Kaffeeproben. Diesen Variablen konnten Signale von NMP zugeordnet werden. Einen Einfluss auf die Abgrenzung der Klasse Asien zeigten die Variablen $\delta_H = 2,60$ ppm, $2,67$ ppm und $2,70$ ppm. In den spektralen Bereichen dieser Variablen waren Signale von Zitronensäure und Äpfelsäure zu finden. Es konnte jedoch keine Aussage darüber getroffen werden, welche der beiden Verbindungen durch die Variablen gekennzeichnet war. Grund hierfür war die starke Überlappung der Signale beider Verbindungen im $^1\text{H-NMR}$ -Spektrum. Auch war die Konzentration dieser organischen Säuren zu gering, sodass keine Signalidentifizierung über 2D-NMR-Spektren erfolgen konnte. Da in dem Bereich $\delta_H = 2,50 - 2,60$ ppm spezifische Signale der Zitronensäure zu finden sind, wäre es möglich, dass die Variable $\delta_H = 2,60$ ppm durch Zitronensäure gekennzeichnet war. Auch bei den Klassen Nord- und Südamerika hatten die Variablen im Bereich $\delta_H = 2,60 - 2,70$ ppm, in welchem Signale von Äpfel- und Zitronensäure vorhanden waren, einen Einfluss auf die Trennung.

Consonni et al. untersuchten die $^1\text{H-NMR}$ -Spektren wässriger Kaffeeextrakte und konnten durch eine OPLS-DA eine vollständige Trennung von Kaffee aus den Kontinenten Afrika, Asien und Amerika erreichen [145]. Einen Einfluss auf die Trennung der Klasse Afrika hatten CGA, Trigonellin und Koffein, wohingegen die Verbindungen Acetat, Chinasäure und myo-Inositol die Trennung von Asien beeinflusste [145]. Auch in den in dieser Arbeit erstellten Modellen konnte gezeigt werden, dass Signale von Koffein einen möglichen Einfluss auf die Trennung bestimmter Klassen haben. Einen Einfluss auf die Trennung durch die Metaboliten Trigonellin, Chinasäure, Acetat und myo-Inositol konnte jedoch in dieser Arbeit jedoch nicht beobachtet werden.

Toledo et al. nutzten GC-MS, um Kaffee aus Brasilien von Kaffee aus Afrika, Asien und Zentralamerika zu differenzieren. Mittels Diskriminanzanalyse erreichten sie eine vollständige Trennung der Klassen und konnten die Analyten 2-Methylbutanal, 3-Methylbutanal, 2,3-Pentadion, 2,3-Butadion, 1-Methylpyrrol, Pyrazin, 2-Methylpyrazin und 2,5-Dimethylpyrazin als Ursache für die Trennung identifizieren [251].

4.5 Datenfusionierung der $^1\text{H-NMR}$ -spektroskopischen Daten der lipophilen und hydrophilen Metaboliten für die Abgrenzung von Kaffee nach ausgewählten Ländern

Durch eine Datenfusionierung kann der Informationsgehalt erhöht, und gleichzeitig der Einfluss von Variablen, die keine relevanten Informationen enthalten, reduziert werden [217]. Daher wurde getestet, ob durch eine Datenfusionierung eine bessere Trennung und Vorhersagekraft der Klassifizierungsmodelle zur Abgrenzung von Kaffeeproben aus ausgewählten Ländern erreicht werden konnte. Dazu wurde eine Low-Level und eine Mid-Level Datenfusionierung durchgeführt. Bei der Low-Level Datafusion wurden die Buckets der $^1\text{H-NMR}$ -Spektren der lipophilen und hydrophilen Metaboliten kombiniert, wohingegen bei der Mid-Level Datenfusionierung die PCA-Scores vereint wurden. Als Klassifizierungsmodell wurde beispielhaft das Modell Brasilien-Afrika ausgewählt, durch welches eine gute Abgrenzung der beiden Klassen bei der Verwendung der Daten der lipophilen bzw. hydrophilen Metaboliten erreicht werden konnte. Daneben wurde bei der unzureichenden Abgrenzung der Klassen Kolumbien und Asien bei der Analyse der lipophilen bzw. hydrophilen Metaboliten getestet, ob eine Datenfusionierung zu besseren Ergebnissen führte.

Abbildung 4-20 zeigt das Ergebnis der MCKV der Klassifizierungsmodelle basierend auf der Analyse der (A) lipophilen und (B) hydrophilen Metaboliten sowie das Ergebnis der MCKV der (C) Low-Level und (D) Mid-Level Datenfusionierung für die Klassen Brasilien und Afrika. Die Modellerstellung basierend auf der Analyse der lipophilen und hydrophilen Metaboliten erfolgte analog zu Kapitel 4.4.3 unter Verwendung der Scores der ersten 29 PC der PCA für die LDA. Für die Low-Level Datenfusionierung wurden die 1775 Buckets der NMR-Spektren der lipophilen Metaboliten mit den 983 Buckets der NMR-Spektren der hydrophilen Metaboliten in einer Matrix vereint. Somit waren nach der Datenfusionierung 2785 Buckets als Variablen vorhanden. Zur Dimensionsreduktion wurde nach der Datenfusionierung eine PCA durchgeführt, von welcher die ersten 29 PC für die LDA verwendet wurden. Dahingegen wurden bei der Mid-Level Datenfusionierung jeweils die Scores der ersten 29 PC der PCA für die LDA vereinigt.

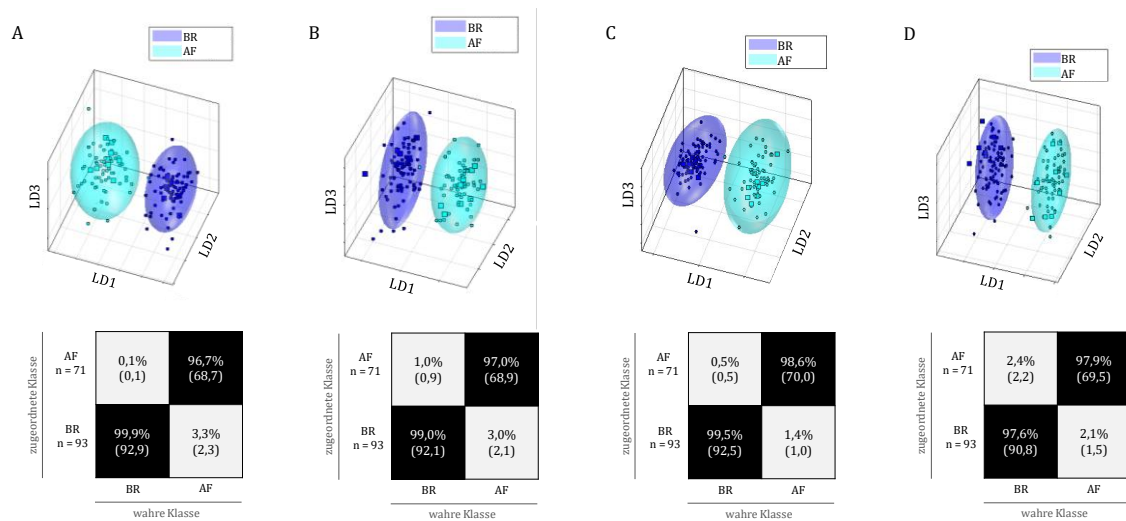


Abbildung 4-20: Ergebnisse der MCKV des Klassifizierungsmodells zur Vorhersage des geographischen Ursprungs von Kaffee bezüglich des Landes Brasilien und den Kontinent Afrika. (A) Klassifizierungsmodell durch die Analyse der ^1H -NMR-Spektren der lipophilen Metaboliten; (B) Klassifizierungsmodell durch die Analyse der ^1H -NMR-Spektren der hydrophilen Kaffeemetaboliten; (C) Klassifizierungsmodell durch die Low-Level Datenfusionierung der ^1H -NMR-Spektren der lipophilen und hydrophilen Metaboliten; (D) Klassifizierungsmodell durch die Mid-Level Datenfusionierung der ^1H -NMR-Spektren der lipophilen und hydrophilen Metaboliten. In (A)-(D) ist oben der LDA-Raum eines Kreuzvalidierungsschrittes dargestellt. Die Wolken kennzeichnen das 95%ige Prognoseellipsoid des Trainingsdatensatzes. Unten in (A)-(D) ist die Konfusionsmatrix der MCKV dargestellt. Die x-Achse und die y-Achse zeigen die zugeordnete Klasse bzw. die wahre Klasse. In der Konfusionsmatrix ist die Wahrscheinlichkeit des Vorhersageergebnisses in Prozent angegeben. Anzahl der für die LDA verwendeten PC: (A)-(C) 29; PC: (D) 58. MCKV: Monte-Carlo Kreuzvalidierung; LD: Lineare Diskriminanzfunktion; AF: Afrika; BR: Brasilien

Die Ergebnisse in Abbildung 4-20 zeigen, dass jeweils bereits die Klassifizierungsmodelle basierend auf jeweils der Analyse der lipophilen und hydrophilen Metaboliten eine gute Abgrenzung der Klassen Brasilien und Afrika ermöglichten. Eine Low-Level Datenfusionierung erzielt eine geringfügig bessere Zuordnung der Testproben aus Afrika. Hier wurde in der MCKV eine richtige Zuordnung des Testdatensatzes von 98,6% im Gegensatz zu 96,7% bei Verwendung der Daten der lipophilen und 97,0% bei Verwendung der Daten der hydrophilen Metaboliten erreicht. Zudem ist zu erkennen, dass sich die Streuung der Proben in dem LDA-Raum durch die Low-Level Datenfusionierung geringfügig verbesserte. Auch die Mid-Level Datenfusionierung erzielte mit 97,9% richtiger Zuordnung der Testproben aus Afrika in der MCKV ein geringfügig besseres Ergebnis als die Einzelmodelle. Allerdings verringerte sich die richtige Zuordnung der brasilianischen Proben in der Mid-Level Datenfusionierung. Bei der Mid-Level Datenfusionierung wurden in der MCKV 97,6%

der brasilianischen Testproben richtig zugeordnet. Dahingegen konnten bei dem Model basieren auf der Analyse der lipophilen Metaboliten 99,9% und bei dem Model basieren auf der Analyse der hydrophilen Metaboliten 99,0% richtige Zuordnung der Testproben aus Brasilien erreicht werden. Im Vergleich zur Low-Level Datenfusionierung war die richtige Zuordnung der Testproben beider Klassen bei der Mid-Level Datenfusionierung in der MCKV geringer. Allerdings konnte durch die Mid-Level Datenfusionierung die Streuung der Proben im LDA-Raum noch weiter reduziert werden. Dies führte zu geringeren Grenzen des Prognoseellipsoids, wodurch die Cluster beider Klassen kleiner wurden und besser voneinander abgetrennt waren. Die geringeren Grenzen der Prognoseellipsoide bei einem schlechteren Ergebnis der MCKV deuteten allerdings an, dass bei diesem Klassifizierungsmodell ein leichtes Overfitting vorlag. Eine weitere Analyse hatte ergeben, dass bereits die ersten 15 Dimensionen der fusionierten PCA-Scores der Daten der lipophilen bzw. hydrophilen Metaboliten in der Mid-Level Datenfusionierung ausreichend war, um eine sehr gute Abgrenzung der beiden Klassen zu ermöglichen (Abbildung 9-30(A)). Durch die Verwendung einer geringeren Anzahl an Dimensionen konnte auch die Gefahr der Overfittings verringert werden.

Die Fusionierung der $^1\text{H-NMR}$ -Daten der lipophilen und hydrophilen Metaboliten für das Klassifizierungsmodell Kolumbien-Asien führte weder bei der Low-Level noch bei der Mid-Level Datenfusionierung zu einer besseren Zuordnung des Testdatensatzes bei der MCKV (Abbildung 9-29). Das Klassifizierungsmodell basierend auf der Analyse der hydrophilen Metaboliten zeigte bessere Ergebnisse der MCKV (Abbildung 9-29(A)) als das Model basierend auf der Analyse der lipophilen Metaboliten (Abbildung 9-29(B)). Durch die Low-Level und Mid-Level Datenfusionierung (Abbildung 9-29(C) und (D)) sank die richtige Zuordnung der Testproben beider Klassen im Vergleich zum Model basierend auf der Analyse der hydrophilen Metaboliten. Durch die Mid-Level Datenfusionierung konnte die Streuung der Proben im LDA-Raum verringert werden, was dazu führte, dass die Cluster besser voneinander getrennt waren. Auch hier deuteten die Ergebnisse an, dass bei der Verwendung von 38 Dimensionen bei der Mid-Level Datenfusionierung ein leichtes Overfitting vorlag. So zeigte eine weitere Analyse, dass der Mid-Level Datenfusionierung bereits 11 Dimensionen der fusionierten PCA-Scores der $^1\text{H-NMR}$ -Daten der lipophilen und hydrophilen Metaboliten ausreichend waren, um eine sehr gute Abgrenzung der beiden Klassen zu ermöglichen (Abbildung 9-30(B)).

Um eine gesicherte Aussage Vorhersagefähigkeit der Low-Level und Mid-Level fusionierten $^1\text{H-NMR}$ -Daten zu ermöglichen, ist die Untersuchung von mehr Kaffeeproben nötig. Mit einer deutlich höheren Probenanzahl könnte die Vorhersagefähigkeit der Klassifizierungsmodelle der fusionierten Daten erhöht und eine bessere Bewertung der Güte der Zuordnung getroffen werden. Neben den $^1\text{H-NMR}$ Daten könnten auch Daten aus anderen Analysenmethoden verwendet werden, um die Vorhersagefähigkeit der Klassifizierungsmodelle zu erhöhen. Beispielsweise kombinierten Choi et al. Daten basierend auf der Analyse mittels LC/MS- und GC mit einem Flammenionisationsdetektor (GC-FID) und Daten aus Protein-, Monosaccharid- und Aminosäureanalyse, um Kaffeeproben den Kontinenten Asien, Südamerika und Afrika zuzuordnen [143].

4.6 Erstellung eines Klassifizierungsmodells basierend auf der Analyse der lipophilen und hydrophilen Kaffeemetaboliten zur Abgrenzung von Kaffee nach verschiedenen Anbaumethoden

Das wachsende Bewusstsein für die weitreichenden Auswirkungen durch die Anwendung chemisch-synthetischer eingesetzter Agrarchemikalien führt zu neuen Vorgehensweisen mit dem Ziel, auf diese zu verzichten [15]. Der Einsatz von synthetischen Chemikalien und Praktiken wie die Bestrahlung zur Schädlingsbekämpfung sind in der ökologischen Landwirtschaft verboten [32]. Allerdings kann dadurch der Ertrag biologisch erzeugter Ernten im Vergleich zum konventionellen Anbau um bis zu 50% verringert sein [15]. Um den reduzierten Ertrag ausgleichen zu können, wird demnach für biologisch erzeugte Produkte ein höherer Preis verlangt. Die unterschiedlichen Preise für biologisch und konventionell erzeugte Produkten erfordern die Überprüfung der Authentizität von biologisch erzeugten Produkten. Daher wurde überprüft, ob anhand der lipophilen und hydrophilen Metaboliten von biologisch und konventionell angebautem Kaffee eine Abgrenzung dieser beiden Anbaumethoden möglich ist. Dazu wurde erneut eine PCA-LDA basierend auf der Analyse der $^1\text{H-NMR}$ -Spektren der lipophilen und hydrophilen Metaboliten durchgeführt (Abbildung 4-21). Für die Modellerstellung und interne Validierung mittels MCKV wurden 143 Kaffeeproben aus biologischem Anbau und 214 Kaffeeproben aus konventionellem Anbau verwendet. Der Testdatensatz für die externe Validierung bestand aus 37 biologischen und 54 konventionellen Kaffeeproben.

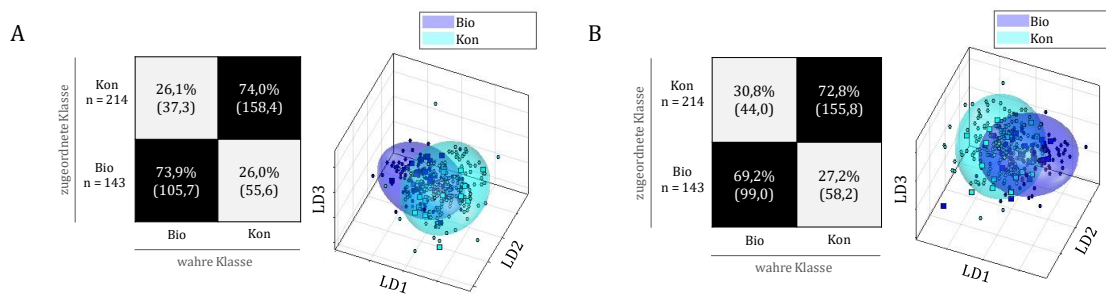


Abbildung 4-21: Ergebnisse der MCKV der PCA-LDA zur Vorhersage der Anbaumethode aus den $^1\text{H-NMR}$ -Spektren der (A) lipophilen und (B) hydrophilen Kaffeemetaboliten. Jeweils rechts in der Abbildung (A) und (B) ist der LDA-Raum eines Kreuzvalidierungsschrittes dargestellt. Das 95%ige Prognoseellipsoid des Trainingsdatensatzes ist als Wolke dargestellt, die Proben des Testdatensatzes werden durch Rechtecke symbolisiert. Links in der Abbildung (A) und (B) ist die Confusionsmatrix der MCKV dargestellt. Die x-Achse und die y-Achse zeigen die zugeordnete Klasse bzw. die wahre Klasse. In der Confusionsmatrix ist die Wahrscheinlichkeit des Vorhersageergebnisses in Prozent angegeben. Anzahl der für die LDA verwendeten PC: 23. Erklärte Varianz: (A) 98,9% und (B) 97,2%. MCKV: Monte-Carlo Kreuzvalidierung; PCA-LDA: Hauptkomponentenanalyse mit linearer Diskriminanzanalyse; LD: Lineare Diskriminanzfunktion; Bio: biologisch zertifizierter Anbau; Kon: Konventioneller Anbau

Die Ergebnisse der MCKV der PCA-LDA mit 69,2 – 74,0% richtiger Zuordnung der Testproben zeigen, dass die Kaffeeproben nicht zuverlässig nach biologischem und konventionellem Anbau getrennt werden konnten. Es war lediglich eine leichte Tendenz zur Trennung dieser beiden Klassen gegeben. In der MCKV des Klassifizierungsmodells basierend auf der Analyse der lipophilen Metaboliten konnten geringfügig mehr Testproben der richtigen Klasse zugeordnet werden (Abbildung 4-21(A)) als im Modell basierend auf der Analyse der der hydrophilen Metaboliten (Abbildung 4-21(B)). Das gleiche Muster zeigte die externe Validierung. Durchschnittlich konnten wie bereits bei der internen Validierung nur 2/3 des externen Testdatensatzes der jeweils richtigen Klasse zugeordnet werden. In

dem Klassifizierungsmodell basierend auf der Analyse der lipophilen Metaboliten wurden von den 54 konventionellen Kaffeeproben 36 Proben der Klasse „konventionell“ zugeordnet. Von 37 biologischen Kaffeeproben wurden 30 der richtigen Klasse „bio“ zugeordnet. In der externen Validierung des Klassifizierungsmodells basierend auf der Analyse der hydrophilen Metaboliten wurden 25 der 37 biologischen und 34 der 54 konventionellen Proben jeweils der richtigen Klasse zugeordnet.

Die Ergebnisse in dieser Arbeit entsprechen somit nicht denen, die Consonni et al. in ihrer Studie erzielen konnten. In ihren Untersuchungen verwendeten sie 42 konventionell und 26 biologisch angebaute Kaffeeproben und erstellten mit den Daten der $^1\text{H-NMR}$ -Spektren der hydrophilen Metaboliten ein OPLS-DA-Modell [144]. Durch eine OPLS-DA konnten Consonni et al. eine Abgrenzung der beiden Klassen erzielen [144]. Biologisch erzeugter Kaffee unterschied sich insbesondere durch Signale der Galactose, Mannose und CGAL, wohingegen konventionell angebaute Kaffee durch Signale von Trigonellin, CGA, Koffein und Acetat abgegrenzt wurde [144]. Consonni et al. verwendeten in ihren Untersuchungen Kaffee unterschiedlicher geographischer Herkunft und Herstellungszeiträume. Ein Grund für die unterschiedlichen Ergebnisse könnte die Verwendung unterschiedlicher Verfahren der multivariaten Datenanalyse sein. Consonni et al. führten mit den Kaffeeproben zunächst eine PCA durch, mit welcher sie jedoch auch eine unzureichende Trennung erreichten [144]. Erst der Einsatz einer überwachten Methode wie der OPLS-DA verbesserte die Trennung der biologisch angebauten Kaffeeproben von den konventionell angebauten. Jedoch stellten auch do Carmo Carvalho et al. bei ihren Untersuchungen von biologisch und konventionell angebautem Kaffee mittels HPLC Unterschiede zwischen biologisch und konventionell angebautem Kaffee fest und konnten erhöhte Gehalte von CGA, Koffein und Trigonellin in biologischem Kaffee ermitteln [252]. Die Kaffeeproben aus ihren Untersuchungen waren Red Catuaí und stammten von Brasilien [252]. Barbosa et al. konnten durch eine Elementanalyse kombiniert mit einer multivariaten Datenanalyse erfolgreich biologisch von konventionell angebauten Kaffeeproben aus Brasilien trennen [141]. Sie stellten fest, dass organisch angebaute Kaffee im Allgemeinen einen höheren Elementgehalt aufwies [141]. Auch Gordillo-Delgado et al. untersuchten Rohkaffeebohnen aus Kolumbien und konnten mittels FT Infrarotspektroskopie kombiniert mit einer PCA biologisch angebauten Kaffee abgrenzen [253]. Sie entdeckten, dass Fructose, Pyrazine, Brenztraubensäure und Oxalsäure einen Einfluss auf die Trennung hatten [253]. Auch hier ist ein direkter Vergleich mit den Ergebnissen dieser Arbeit nicht möglich. Insbesondere die Verbindungen, welche die Trennung von biologisch und konventionell angebautem Rohkaffee beeinflussten, werden von dem Röstprozess stark verändert. Die Ergebnisse von do Carmo Carvalho et al., Barbosa et al. und Gordillo-Delgado et al. deuten an, dass eine Abgrenzung von biologisch und konventionell erzeugten Kaffee möglich ist. Alle Arbeitsgruppen verwendeten Kaffeeproben aus jeweils einem Land. In der vorliegenden Arbeit wurden Kaffeeproben unterschiedlicher geographischer Herkunft verwendet, um die Trennung verschiedener Anbaumethoden zu untersuchen. Die Analyse der in dieser Arbeit erhaltenen $^1\text{H-NMR}$ Daten von Kaffeeproben nur aus Brasilien könnte somit möglicherweise zu einer besseren Abtrennung biologisch erzeugten Kaffee führen.

Der Grund für die unzureichende Abgrenzung der beiden Klassen in dieser Arbeit könnte in den länderübergreifenden Zertifizierungsverfahren liegen. Wie in Kapitel 1.1.3.1 beschrieben, kann biologisch erzeugter Kaffee nur dann als solcher ausgelobt werden, wenn dieser auf biologisch zertifizierten Kaffeefarmen produziert wird. Diese Zertifizierungsverfahren erzeugen bei den Kaffeebauern jedoch erhebliche Kosten, die insbesondere von kleinen Kaffeefarmen nicht aufgebracht werden können. Viele Kaffeefarmen haben nicht

den Zugang und die Mittel, um teure Pestizide und Fungizide zu verwenden und betreiben daher Schattenanbau. In dieser Art des Anbaus ist die biologische Produktion oft mit inbegriffen, jedoch aufgrund der hohen Kosten nicht zwingend mit einer bio-Zertifizierung verbunden [15, 27, 28]. Dadurch ist es möglich, dass keine klare Grenze zwischen biologisch und konventionell angebautem Kaffee vorhanden ist, was die unzureichende Trennung der Klassifizierungsmodelle in dieser Arbeit erklären könnte.

4.7 Erstellung eines Klassifizierungsmodells basierend auf der Analyse der lipophilen und hydrophilen Kaffeemetaboliten zur Abgrenzung von Kaffee nach dem Röstgrad und dem Röstverfahren

Neben der Abgrenzung von Kaffee aus unterschiedlichem geographischen Ursprungs und von unterschiedlichen Anbaumethoden wurde weiter getestet, inwieweit sich die PCA-LDA auf die ^1H -NMR-Spektren der lipophilen und hydrophilen Kaffeemetaboliten anwenden lässt, um die Röstung zu untersuchen. Dazu wurde getestet, ob eine PCA-LDA der ^1H -NMR-Spektren der lipophilen und hydrophilen Metaboliten eine Abgrenzung von Kaffee nach dem Röstverfahren sowie eine Abgrenzung nach hellem und dunklem Röstgrad ermöglicht. In dieser Arbeit wurden Kaffeeproben neben dem klassischen Trommelröstverfahren mit einem Heißluftverfahren geröstet. Zudem wurden geröstete Kaffeeproben mit teilweise sehr unterschiedlichen Röstgraden untersucht. Für die Modellerstellung zur Abgrenzung von Kaffeeproben nach dem Röstverfahren und für die interne Validierung mittels MCKV wurden 481 trommelgeröstete und 69 heißluftgeröstete Kaffeeproben verwendet. Der Testdatensatz für die externe Validierung bestand aus 120 trommelgerösteten und 16 heißluftgerösteten Kaffeeproben. Die Modellerstellung und die interne Validierung zur Abgrenzung von Kaffeeproben nach dem Röstgrade erfolgte nur mit trommelgerösteten Kaffeeproben. Unter diesen besaßen 84 Kaffeeproben einen dunklen Röstgrad mit einem Tonino-Farbwert <80 und 242 Kaffeeproben einen hellen Röstgrad mit einem Tonino-Farbwert von >100 . Für die externe Validierung wurden 21 dunkel und 57 hell geröstete Kaffeeproben verwendet.

In Abbildung 4-22 ist das Ergebnis der MCKV der Klassifizierungsmodelle basierend auf der Analyse der lipophilen und hydrophilen Metaboliten zur Abgrenzung von Kaffeeproben nach dem Röstverfahren dargestellt. Es war sowohl durch die Analyse der lipophilen als durch die Analyse der hydrophilen Metaboliten eine Abgrenzung der trommelgerösteten Kaffeeproben von den heißluftgerösteten Kaffeeproben möglich.

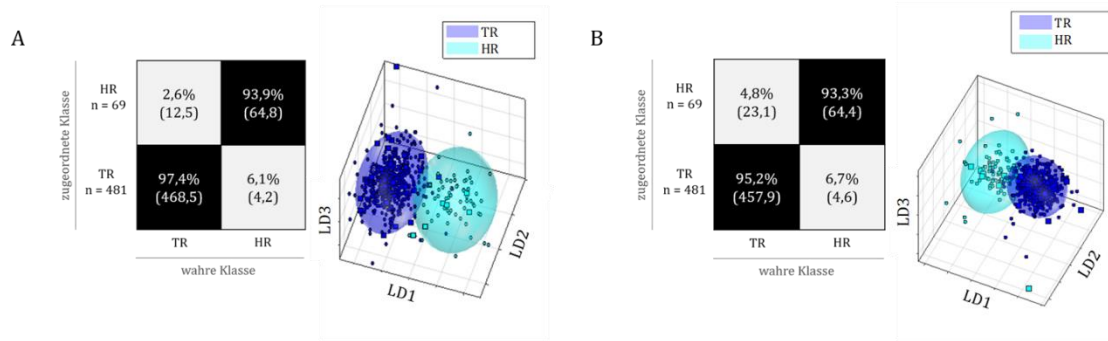


Abbildung 4-22: Ergebnisse der MCKV der PCA-LDA zur Vorhersage des Röstverfahrens aus den ¹H-NMR-Spektren der (A) lipophilen und (B) hydrophilen Kaffeemetaboliten. Jeweils rechts der Abbildung (A) und (B) ist der LDA-Raum eines Kreuzvalidierungsschrittes dargestellt. Das 95%ige Prognoseellipsoid des Trainingsdatensatzes ist als Wolke dargestellt, die Proben des Testdatensatzes werden durch Rechtecke symbolisiert. Links in der Abbildung (A) und (B) ist die Konfusionsmatrix der MCKV dargestellt. Die x-Achse und die y-Achse zeigen die zugeordnete Klasse bzw. die wahre Klasse. In der Konfusionsmatrix ist die Wahrscheinlichkeit des Vorhersageergebnisses in Prozent angegeben. Anzahl der für die LDA verwendeten PC: 39, erklärte Varianz: (A) 99,3% und (B) 98,6%. MCKV: Monte-Carlo Kreuzvalidierung; PCA-LDA: Hauptkomponentenanalyse mit linearer Diskriminanzanalyse; LD: Lineare Diskriminanzfunktion; TR: Trommelröstung; HR: Heißluftröstung

In der internen Validierung konnten 97,4% der trommelgerösteten Kaffeeproben und 93,9% der heißluftgerösteten Kaffeeproben bei dem Modell basierend auf der Analyse der lipophilen Metaboliten der richtigen Klasse zugeordnet werden (Abbildung 4-22(A)). Das auf der Analyse der hydrophilen Metaboliten basierende Modell erzielte geringfügig schlechtere Ergebnisse. Hier wurden bei der internen Validierung 95,2% der trommelgerösteten Kaffeeproben und 93,3% der heißluftgerösteten Kaffeeproben der richtigen Klasse zugeordnet (Abbildung 4-22(B)). In der externen Validierung konnten alle heißluftgerösteten Kaffeeproben der richtigen Klasse zugeordnet werden. In dem Modell basierend auf der Analyse der lipophilen Metaboliten wurden 112 von 120 der trommelgerösteten Kaffeeproben der richtigen Klasse zugeordnet, was einem Prozentsatz von 93,3% entsprach. Von 120 der trommelgerösteten Kaffeeproben konnten dagegen im Modell basierend auf der Analyse der hydrophilen Metaboliten 117 Proben der richtigen Klasse zugeordnet werden, was 97,5% entsprach. Somit zeigte sowohl die interne als auch die externe Validierung, dass eine gute Vorhersagefähigkeit für beide Modelle gegeben war. Allerdings waren die Cluster beider Klassen in dem LDA-Raum in beiden Modellen nicht vollständig voneinander getrennt. Zudem lag bei der Klasse der heißluftgerösteten Kaffeeproben eine große Streuung vor. Für den Modellaufbau waren 84 heißluftgeröstete Proben beteiligt. Möglicherweise könnte die Streuung der Proben dieser Klasse durch die Untersuchung einer größeren Probenzahl minimiert werden.

Im wissenschaftlichen Interesse wurde in einer Machbarkeitsstudie überprüft, ob die PCA-LDA basierend auf der Analyse der lipophilen und hydrophilen Metaboliten dazu verwendet werden kann, Kaffee nach hellem und dunklem Röstgrad abzugrenzen. Abbildung 4-23 zeigt das Ergebnis der MCKV der Klassifizierungsmodelle basierend auf der Analyse der lipophilen und hydrophilen Metaboliten zur Abgrenzung von Kaffeeproben nach dem Röstgrad. Die Abbildung zeigt, dass durch die PCA-LDA basierend auf den spektralen Daten der lipophilen und hydrophilen Metaboliten eine Trennung der Kaffeeproben nach hellem und dunklem Röstgrad möglich ist. In der Praxis hat der Parameter Röstgrad keine rechtliche Relevanz. Das erstellte Klassifizierungsmodell könnte beispielsweise in größeren Röstereien zur Nachverfolgung des Röstprozesses eingesetzt werden. Da aber die Farbwertmessung von Kaffee eine geringere Probenvorbereitung erfordert und daher leichter

umzusetzen ist, wird das erstellte Klassifizierungsmodell in der Praxis wahrscheinlich keine Anwendung finden.

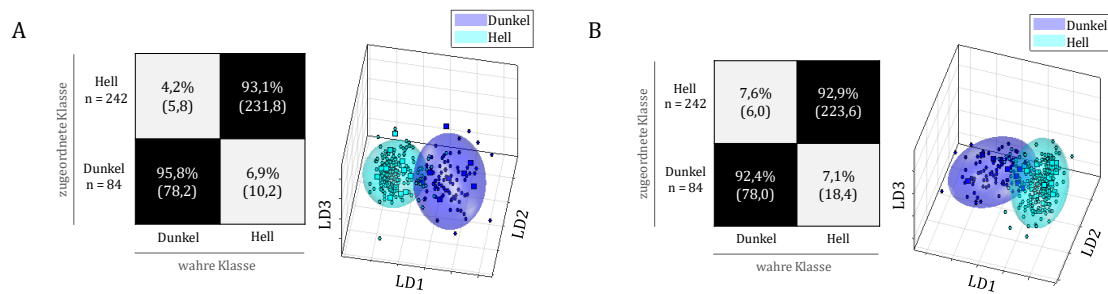


Abbildung 4-23: Ergebnisse der MCKV der PCA-LDA zur Vorhersage des Röstgrades aus den $^1\text{H-NMR}$ -Spektren der (A) lipophilen und (B) hydrophilen Kaffeemetaboliten. Jeweils rechts in der Abbildung (A) und (B) ist der LDA-Raum eines Kreuzvalidierungsschrittes dargestellt. Das 95%ige Prognoseellipsoid des Trainingsdatensatzes ist als Wolke dargestellt, die Proben des Testdatensatzes werden durch Rechtecke symbolisiert. Links in der Abbildung (A) und (B) ist die Konfusionsmatrix der MCKV dargestellt. Die x-Achse und die y-Achse zeigen die zugeordnete Klasse bzw. die wahre Klasse. In der Konfusionsmatrix ist die Wahrscheinlichkeit des Vorhersageergebnisses in Prozent angegeben. Anzahl der für die LDA verwendeten PC: 27, Erklärte Varianz: (A) 98,5% und (B) 98,1%. MCKV: Monte-Carlo Kreuzvalidierung; PCA-LDA: Hauptkomponentenanalyse mit linearer Diskriminanzanalyse; LD: Lineare Diskriminanzfunktion

Die interne Validierung zeigte eine geringfügig bessere Zuordnung der Kaffeeproben bei dem Modell basierend auf der Analyse der lipophilen Metaboliten. Von den dunkel gerösteten Kaffeeproben konnten 95,8% und von den hell gerösteten Kaffeeproben 93,1% der richtigen Klasse zugeordnet werden. In dem Modell basierend auf der Analyse der hydrophilen Metaboliten waren dies 92,4% und 92,9%. In der externen Validierung konnten in beiden Modellen jeweils eine dunkle und eine hell geröstete Probe nicht der richtigen Klasse zugeordnet werden. Somit wurden bei der externen Validierung 95,2% der dunkel und 98,3% der hell gerösteten Testproben der richtigen Klasse zugeordnet. Somit konnte anhand der internen und externen Validierung gezeigt werden, dass auch die erstellten Klassifizierungsmodelle zur Abgrenzung von dunkel und hell geröstetem Kaffee eine gute Vorhersagefähigkeit besaßen. In dem LDA-Raum streuten die Proben der Klasse des dunkel gerösteten Kaffees mehr als die Proben des hell gerösteten Kaffees. Dies führte zu einem breiteren Cluster und zu einem größeren Prognoseellipsoid dieser Klasse. Durch die Untersuchung einer größeren Probenanzahl könnte diese Streuung reduziert werden und gegebenenfalls zu einer besseren Abtrennung der Cluster beider Klassen führen.

Eine erfolgreiche Abgrenzung von Kaffeeproben nach dem Röstgrad konnte bereits in Untersuchungen von Febvay et al., Nunez et al. und Ruosi et al. erreicht werden [72, 249, 254]. Febvay et al. konnten mittels OPLS-DA der $^1\text{H-NMR}$ -Spektren von hydrophilen Metaboliten Kaffee trennen, welcher bei unterschiedlichen Rösttemperaturen geröstet wurde. Ihre Untersuchungen ergaben, dass die Verbindungen HMF, NMP, Methylnicotinat, CGA, CGAL und Milchsäure einen Einfluss auf die Trennung hatten [72]. Nunez et al. konnten durch eine PLS-DA der HPLC-DAD-Daten wässriger Kaffeeextrakte kommerziell erworbener Kaffee nach hellem und dunklem Röstgrad trennen [249]. Ruosi et al. führten eine PCA mit GC-MS-Daten durch und konnten so eine Trennung von Kaffee nach dunklem, mittlerem und hellem Röstgrad erreichen [254].

Es konnte gezeigt werden, dass durch die Erstellung eines Klassifizierungsmodells basierend auf der Analyse der lipophilen und hydrophilen $^1\text{H-NMR}$ -Daten eine Abgrenzung

nach Röstverfahren und Röstgrad möglich ist. Insbesondere die Abgrenzung von Kaffeeproben nach dem Röstverfahren könnte in Zukunft dazu beitragen, die Authentizität von trommelgerösteten Kaffee zu untersuchen. Im Vergleich zu der Wirbelschichtröstung, welches ein Röstverfahren mittels Heißluft darstellt und auf eine Röstzeit von bis zu zwei Minuten verkürzt werden kann, benötigt die Trommelröstung eine Röstzeit von 8,00-20,0 Minuten [5]. Durch eine längere Röstzeit könnten erhöhte Kosten entstehen, wodurch für Kaffeeprodukte mit der Auslobung „schonend gerösteter Kaffee im Langzeit-Trommelröstverfahren“ ein höherer Preis verlangt werden könnte.

4.8 Identifizierung lipophiler Kaffeemetaboliten mittels NMR-Spektroskopie

Die Identifizierung der in den Kaffeeextrakten vorhandenen lipophilen Primär- und Sekundärverbindungen erfolgte durch Aufstockung mit Referenzsubstanzen und ^1H -NMR-Messungen sowie durch die Anwendung von ^1H - ^{13}C -korrelierten 2D-Experimenten. In Abbildung 4-24 ist ein ^1H - ^{13}C -korreliertes HSQC-NMR-Spektrum der lipophilen Kaffeemetaboliten im Bereich $\delta_{\text{H}} = 0,40 - 5,60$ ppm und $\delta_{\text{C}} = 0,00 - 200$ ppm dargestellt. Dieser Bereich des HSQC-Spektrums wurde von Signalen verschiedener FA (lila-blau), welche in Form der TAG an Glycerin gebunden sind, dominiert. Auch erschienen Signale der endständigen CH_3 -Gruppen der FA, Signale der CH_2 -Gruppen sowie die Signale von Protonen in α - oder β -Position einer Doppelbindung oder der Carboxylgruppe. Zudem waren in diesem Bereich Signale der ^1H -Kerne von Kahweol und Cafestol (pink) ersichtlich. Bei einer chemischen Verschiebung von $\delta_{\text{H}} = 5,30$ ppm und $\delta_{\text{C}} = 79,0$ ppm erschienen die Signale des Glycerinrundgerüsts der TAG (hellblau) und bei $\delta_{\text{H}} = 5,40$ ppm und $\delta_{\text{C}} = 128 - 130$ ppm die ^1H -Kerne in direkter Umgebung einer Doppelbindung (graublau). Daneben waren Signale von Koffein (grün), Trigonellin (dunkelgrün), Cholin (hellblau) und die des Istd DMS (grau) erkennbar. Die Strukturen dieser Verbindungen sind in Abbildung 9-31 dargestellt.

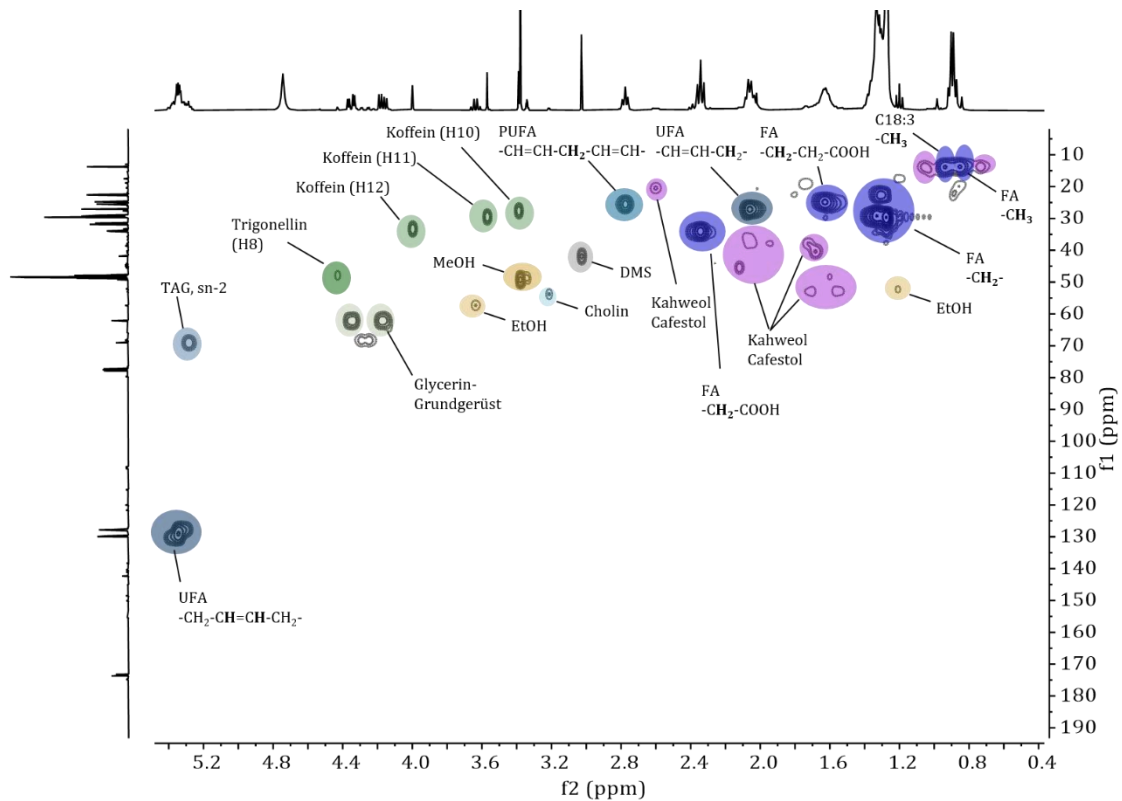


Abbildung 4-24: Ausschnitt eines 400 MHz ^1H - ^{13}C -korrelierten HSQC-NMR-Spektrums der lipophilen Kaffeemetaboliten im Bereich $\delta_{\text{H}} = 0,40 - 5,60$ ppm und $\delta_{\text{C}} = 30,0 - 190$ ppm, gemessen in $\text{CDCl}_3:\text{MeOH}-d_4$ 1:1 (v/v). Die Signale der identifizierten Metaboliten sind farbig hinterlegt. Die Nummern H8-H11 stellen die ^1H -Kerne der entsprechenden Verbindung dar. FA: Fettsäuren; EtOH: Ethanol; UFA: ungesättigte Fettsäuren; PUFA: mehrfach ungesättigte Fettsäuren; DMS: Dimethylsulfon

Im ^1H - ^{13}C -korrelierten HSQC-TOCSY-NMR-Spektrum der lipophilen Kaffeemetaboliten im Bereich $\delta_{\text{H}} = 0,20 - 5,00$ ppm und $\delta_{\text{C}} = 0,00 - 190$ ppm waren die Spinsysteme der FA sehr gut erkennbar. Daneben konnten die Spinsysteme von Kahweol und Cafestol betrachtet werden (Abbildung 9-32).

Abbildung 4-25 zeigt das ^1H - ^{13}C -korrelierte HSQC-NMR-Spektrum der lipophilen Kaffeemetaboliten im Bereich $\delta_{\text{H}} = 5,40 - 10,0$ ppm und $\delta_{\text{C}} = 30,0 - 190$ ppm. In diesem Bereich erschienen erneut Signale von Kahweol und Cafestol (pink). Diese wurden durch die ^1H -Kerne in direkter Umgebung einer Doppelbindung erzeugt. In der Struktur von Kahweol befindet sich im Vergleich zu Cafestol eine zusätzliche Doppelbindung, wodurch in diesem Bereich mehr Signale von Kahweol als von Cafestol vorhanden waren. Auch waren erneut Signale von Koffein (hellgrün) und Trigonellin (grün) erkennbar. Zusätzlich traten Signale der Hydroxyzimtsäure-Isomere Kaffeesäure, Ferulasäure und *p*-Cumarsäure auf (alle rot), welche als CGA oder CGAL an Chinasäure gebunden vorlagen. Aufstockungsversuche haben gezeigt, dass sich die Signale dieser Hydroxyzimtsäure-Isomere im 400 MHz ^1H -NMR-Spektrum überlagerten und bedingt durch eine zu geringe Auflösung auch nicht durch die zweite Ebene des HSQC-Experiments getrennt werden konnten. Wie in Kapitel 1.1.7.6 beschrieben, sind CGA mit Ferulasäure und *p*-Cumarsäure nur in geringen Mengen in geröstetem Kaffee vertreten [7]. Daher konnte davon ausgegangen werden, dass die Signale der Hydroxyzimtsäure-Isomere in diesem ^1H - ^{13}C -HSQC-NMR-Spektrum hauptsächlich von Kaffeesäure erzeugt wurden. Daneben war das Signal der Ameisensäure (lachsfarben), welche nicht zu den lipophilen Metaboliten zählt, bei $\delta_{\text{H/C}} = 8,34/167$ ppm erkennbar. Womöglich wurde diese Verbindung aufgrund des Anteils an Methanol im Extraktionsmittel mitextrahiert.

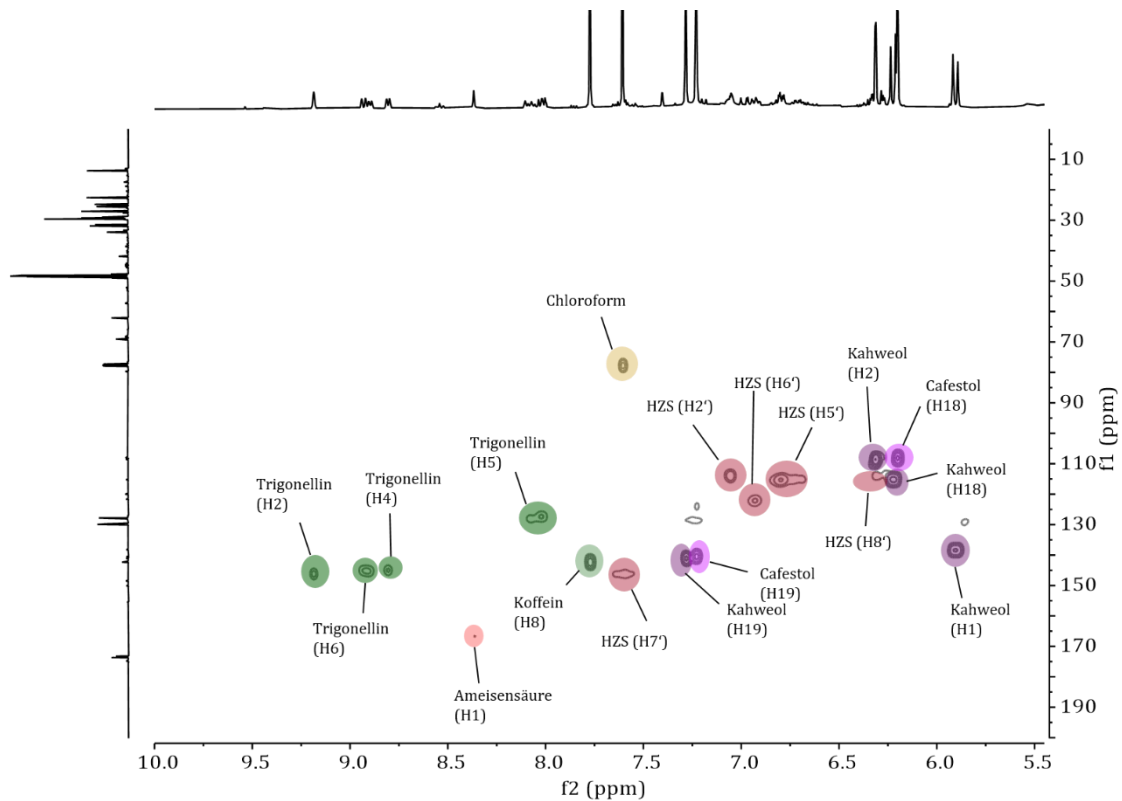


Abbildung 4-25: Ausschnitt eines 400 MHz ^1H - ^{13}C -korrelierten HSQC-NMR-Spektrums der lipophilen Kaffeemetaboliten im Bereich $\delta_{\text{H}} = 5,40 - 10,0$ ppm und $\delta_{\text{C}} = 30,0 - 190$ ppm, gemessen in $\text{CDCl}_3:\text{MeOH}-d_4$ 1:1 (v/v). Die Signale der identifizierten Metaboliten sind farbig hinterlegt. Die Nummern H2-H20 stellen die ^1H -Kerne der entsprechenden Verbindung dar. HZS: Hydroxyzimtsäure-Isomere (Kaffeensäure, Ferulasäure, p-Cumarsäure).

Die im ^1H - ^{13}C -korrelierten HSQC-TOCSY-NMR-Spektrum der lipophilen Kaffeemetaboliten im Bereich $\delta_{\text{H}} = 5,00 - 10,0$ ppm und $\delta_{\text{C}} = 0,00 - 190$ ppm aufgezeigten Spinsysteme von Kahweol und Cafestol trugen weiter zur Signalidentifizierung bei. Daneben war in diesem Spektrum das Spinsystem von Trigonellin ersichtlich (Abbildung 9-33).

Durch die ^1H - ^{13}C -korrelierten HSQC- und ^1H - ^{13}C -korrelierten HSQC-TOCSY-NMR-Experimente sowie durch Aufstockung einer Kaffeeprobe mit Referenzsubstanzen konnten nahezu alle dominanten Signale des ^1H -NMR-Spektrums der lipophilen Metaboliten identifiziert werden. Die Signalzuordnung des Hoch-, Mittel- und Tieffeldbereichs des ^1H -NMR-Spektrums ist in Abbildung 9-34-Abbildung 9-36 dargestellt.

4.9 Identifizierung hydrophiler Kaffeemetaboliten mittels NMR-Spektroskopie

Auch die Identifizierung der Signale der in den Kaffeeextrakten vorhandenen hydrophilen Metaboliten erfolgte durch Aufstockung mit Referenzsubstanzen und ^1H -NMR-Messungen sowie durch die Anwendung von ^1H - ^{13}C -korrelierten 2D-Experimenten. Zusätzlich wurden Literaturdaten zur Identifizierung herangezogen. Abbildung 4-26 zeigt ein ^1H - ^{13}C -korreliertes HSQC-NMR-Spektrum der hydrophilen Kaffeemetaboliten im Bereich $\delta_{\text{H}} = 0,20 - 5,00$ ppm und $\delta_{\text{C}} = 30,0 - 190$ ppm. Wie im Spektrum der lipophilen Metaboliten waren auch im HSQC-Spektrum der hydrophilen Metaboliten Signale von Trigonellin (grün), Koffein (hellgrün) und Cholin (hellblau) zu erkennen. Die Strukturen dieser Verbindungen sind in Abbildung 9-31 abgebildet. Direkt neben dem H8-Signal des Trigonellins erschien ein Signal von NMP (khaki). Zusätzlich waren Signale von Kaffeeproteinen (rosa),

welche teilweise in Form der AGPs in die Melanoidine eingebaut sein können, zu sehen. Daneben waren Signale der aliphatischen organischen Säuren Milchsäure (hellrot), Essigsäure (dunkelrot) und Glycolsäure (pink) erkennbar. Die Kerne der Chinasäure, welche hauptsächlich als CGA oder CGAL an die Hydroxyzimtsäure-Isomere gebunden vorliegt, erzeugten Signale im Bereich von $\delta_H = 1,80 - 4,60$ ppm und $\delta_C = 25,0 - 70,0$ ppm des HSQC-Spektrums (orange). Die zu einem Spinsystem gehörigen ^1H -Kerne der Chinasäure konnten im ^1H - ^{13}C -korrelierten HSQC-TOCSY-NMR-Spektrum beobachtet werden (Abbildung 9-37).

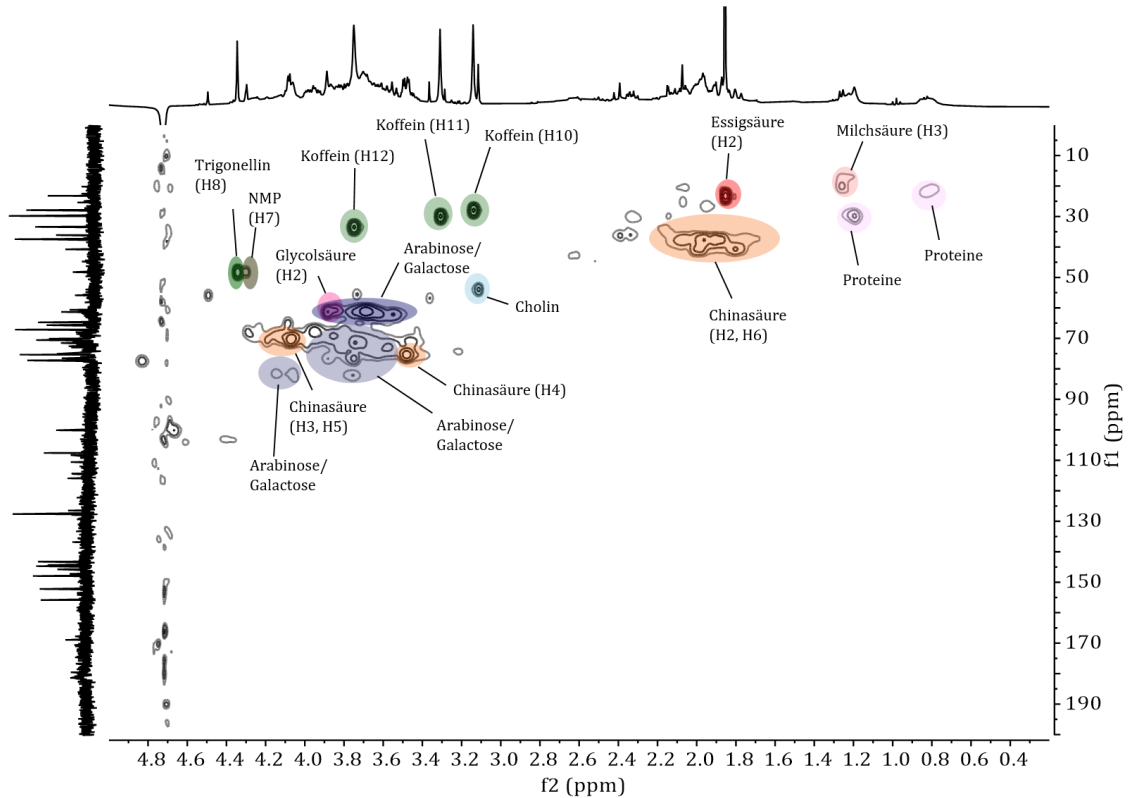


Abbildung 4-26: Ausschnitt eines 400 MHz ^1H - ^{13}C -korrelierten HSQC-NMR-Spektrums der hydrophilen Kaffeemetaboliten im Bereich $\delta_H = 0,20 - 5,00$ ppm und $\delta_C = 30,0 - 190$ ppm gemessen in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die Signale der identifizierten Metaboliten sind farblich hinterlegt. Die Nummern H2-H8' stellen die ^1H -Kerne der entsprechenden Verbindung dar. NMP: N-Methylpyridinium

Die Signalzuordnung der in AGPs und Melanoidine eingebundenen Monosaccharide Galactose und Arabinose erfolgte aufgrund der in der Literatur beschriebenen Ergebnisse der ^1H -NMR, ^{13}C -NMR und ^1H - ^{13}C -korrelierten HSQC-Experimente von Capeck et al., Nunes et al. und Redgwell et al. [255–257]. Capeck et al. untersuchten AGPs, welche aus Instantkaffee isoliert wurden, mittels 2D-NMR. Die untersuchten APGs bestanden aus Galactose (85%), Arabinose (8,2%), Mannose (2,7%) und Glucose (1,4%). Die in dem Bereich $\delta_H = 3,50 - 5,30$ ppm und $\delta_C = 62,0 - 110$ ppm auftretenden Signale konnten sie Arabinfuranose und Galactopyranose der AGPs zuordnen [255]. Ähnliche Ergebnisse konnten Nunes et al. durch ihre Untersuchungen von AGPs aus der hochmolekularen Fraktion von grünen Kaffeebohnen [256] sowie Redgwell et al. bei ihren Untersuchungen zu AGPs aus der Pflanze *Lycium barbarum* [257] erzielen. Ein direkter Vergleich mit den Literaturdaten von Capeck et al., Nunes et al. und Redgwell et al. ist leider nicht möglich, da die Untersuchungen mit jeweils unterschiedlichen experimentellen Parameter durchgeführt wurden. Zudem untersuchten Capeck et al. Instantkaffee, Nunes et al. grüne Kaffeebohnen und Redgwell et al. Proben der Pflanze *Lycium barbarum*, wohingegen in dieser Arbeit geröstete Kaffeebohnen untersucht wurden. Dies erschwert zusätzlich ein Vergleich mit der Literatur.

Die Spinsysteme der Monosaccharide, die im ^1H - ^{13}C -korrelierten HSQC-TOCSY-NMR-Spektrum im Bereich $\delta_{\text{H}} = 3,30 - 4,20$ ppm $\delta_{\text{C}} = 55,0 - 75,0$ ppm beobachtet wurden (Abbildung 9-37), konnten weiter zur Signalidentifizierung beitragen. Allerdings sind für eine genaue Signalidentifizierung in diesem Bereich des in Abbildung 4-27 dargestellten ^1H - ^{13}C -korrelierten HSQC-NMR-Spektrums noch weitere 2D-NMR Experimente nötig. Beispielsweise könnte zusätzlich ein ^1H - ^1H korreliertes COSY zur Signalidentifizierung beitragen.

Im Bereich $\delta_{\text{H}} = 5,00 - 10,0$ ppm und $\delta_{\text{C}} = 30,0 - 190$ ppm des ^1H - ^{13}C -korrelierten HSQC-NMR-Spektrums der hydrophilen Kaffeemetaboliten erschienen erneut Signale von Trigonellin, NMP und Koffein (Abbildung 4-27). Auch war wie im HSQC-Spektrum der lipophilen Metaboliten das Signal der Ameisensäure (hellrot) und Signale der Hydroxyzimtsäure-Isomere (rot) erkennbar. Bei $\delta_{\text{H}/\text{C}} = 5,33/69,4$ ppm und $\delta_{\text{H}/\text{C}} = 5,40/71,1$ ppm erschienen Signale der CQA-Isomere 3-CQA und 5-CQA, die sich durch die Bindungsstelle der Kaffeesäure an der Chinasäure unterscheiden. Der Bereich $\delta_{\text{H}} = 5,00 - 10,0$ ppm und $\delta_{\text{C}} = 0,00 - 190$ ppm des ^1H - ^{13}C -korrelierten HSQC-TOCSY-NMR-Spektrums zeigte die Spinsysteme der aromatischen Protonen Hydroxyzimtsäure-Isomere sowie die von Trigonellin und NMP (Abbildung 9-38).

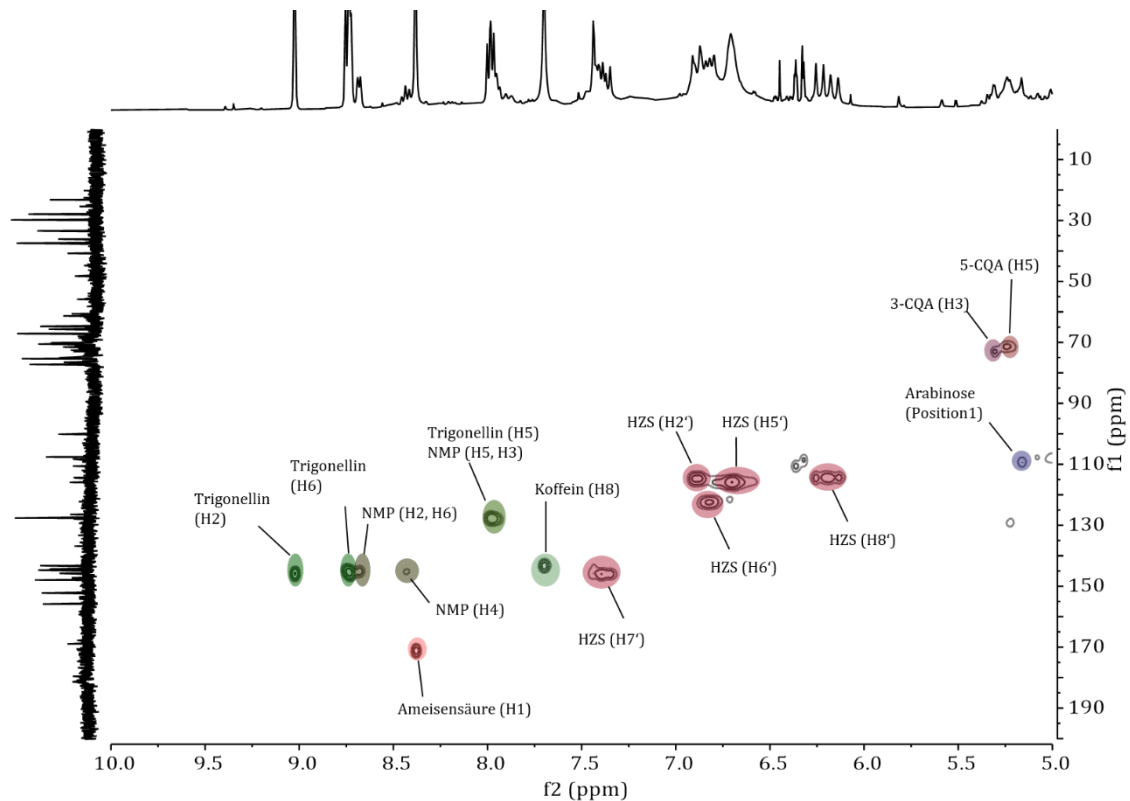


Abbildung 4-27: Ausschnitt eines 400 MHz ^1H - ^{13}C -korrelierten HSQC-NMR-Spektrums der hydrophilen Kaffeemetaboliten im Bereich $\delta_{\text{H}} = 5,00 - 10,0$ ppm und $\delta_{\text{C}} = 30,0 - 190$ ppm, gemessen in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die Signale der identifizierten Metaboliten sind farbig hinterlegt. Die Nummern H2-H8' stellen die ^1H -Kerne der entsprechenden Verbindung dar. HZS: Hydroxyzimtsäure-Isomere (Kaffeesäure, Ferulasäure und p-Cumarsäure); 5-CQA: 5-Caffeoylchinasäure; 3-CQA: 3-Caffeoylchinasäure

Zusätzlich konnten durch Aufstockung mit Reinsubstanzen die Signale von Propion-, Äpfel-, Zitronen- und Maleinsäure identifiziert werden. Diese organischen aliphatischen Säuren lagen jedoch in einer zu geringen Konzentration vor, um die Signale dieser Metaboliten in den ^1H - ^{13}C -korrelierten HSQC- und ^1H - ^{13}C -korrelierten HSQC-TOCSY-NMR-Spektren zu beobachten. Somit konnten anhand der ^1H - ^{13}C -korrelierten HSQC- und ^1H - ^{13}C -korrelierten HSQC-TOCSY-NMR-Experimente sowie durch Aufstockung einer Kaffeeprobe mit

Referenzsubstanzen nahezu alle dominanten Signale des $^1\text{H-NMR}$ -Spektrums der hydrophilen Metaboliten zugeordnet werden (Abbildung 9-39-Abbildung 9-41).

4.10 Zielgerichtete Analyse zur Quantifizierung ausgewählter Kaffeemetaboliten

4.10.1 Entwicklung einer Methode zur Quantifizierung ausgewählter Kaffeemetaboliten durch $^1\text{H-NMR}$ -Spektroskopie und PULCON

Neben der nicht zielgerichteten Analyse wurden die $^1\text{H-NMR}$ -Spektren der hydrophilen Metaboliten im Rahmen einer zielgerichteten Analyse verwendet, um die Metaboliten HMF, Trigonellin, NMP, Koffein, CQA, Ameisensäure, Milchsäure und Essigsäure zu quantifizieren. Abbildung 4-28 zeigt das $^1\text{H-NMR}$ -Spektrum der hydrophilen Metaboliten, in welchem die zur Quantifizierung verwendeten Signale gekennzeichnet wurden. Die entsprechenden ^1H -Kerne, welche die Signale erzeugten, sind in Abbildung 9-42 markiert. Über das Proton H5' konnten die CQA-Isomere 5-CQA, 4-CQA und 3-CQA als Summenparameter (CQA) quantifiziert werden.

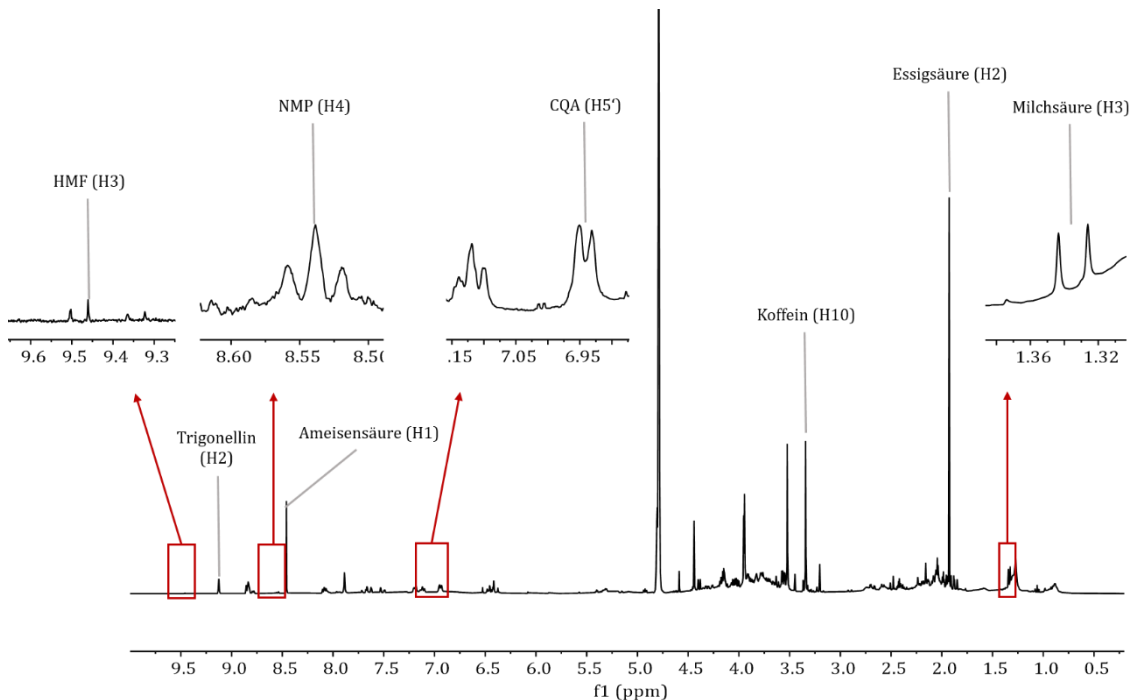


Abbildung 4-28: Ausschnitt eines 400 MHz $^1\text{H-NMR}$ -Spektrums der hydrophilen Kaffeemetaboliten im Bereich $\delta_{\text{H}} = -0,20 - 10,0$ ppm, gemessen in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die Lösungsmittelunterdrückung erfolgte mit einem NOESY-Pulsprogramm. Die chemische Verschiebung wurde auf $\delta_{\text{TSP}} = 0,00$ ppm referenziert. Die Bereiche, in denen die Signale von Milchsäure, NMP, HMF und CQA erscheinen, wurden entsprechend vergrößert. CQA: Isomere der Caffeoylchinasäure; NMP: N-Methylpyridinium; HMF: 5-Hydroxymethylfurfural

Die zur Quantifizierung herangezogenen Signale der Milchsäure, Essigsäure, CQA und des Koffeins lag im $^1\text{H-NMR}$ -Spektrum überlagert vor. Zur Integration dieser Signale wurde eine Kurvenanpassungsmethode verwendet. In dieser Methode wurde eine Voigt-Funktion an die Signale angepasst [258]. Durch Parameter wie die chemische Verschiebung und Breite des Signals, Intensitätsverhältnisse, Multiplettstruktur oder Kopplungskonstante kann diese Funktion an das Signal angepasst und anschließend das Integral unter der gebildeten Kurve berechnet werden [258]. In Abbildung 9-43 ist die Darstellung der

Kurvenanpassungsmethode an die Signale von Milchsäure, Essigsäure, Koffein und CQA ersichtlich. Weiter wurden zur Berechnung der Konzentrationen die in Kapitel 4.3.3 bestimmten Korrekturfaktoren für Ameisensäure (1,2), Essigsäure (1,1) und Trigonellin (1,1) verwendet. Diese Korrekturfaktoren berücksichtigen die nicht vollständige Relaxation dieser ^1H -Kerne bei der verwendeten Repetitionszeit.

Zunächst wurde die entwickelte qNMR-Methode hinsichtlich der Validierungsparameter Richtigkeit, Linearität, NG, BG, Präzision und Robustheit für die acht Analyten überprüft. Die Richtigkeit wurde anhand der Wiederfindung belegt. Dazu wurden die Reinsubstanzen der Analyten in acht verschiedenen Konzentrationen einer Kaffeematrix zugesetzt. Die Wiederfindung wurde als Quotient des bestimmten gegenüber dem tatsächlich vorhandenen Gehalt bestimmt. Die Wiederfindung der Analyten ist in Tabelle 9-8 aufgelistet. Die Wiederfindungen für die Analyten Milchsäure (96,3-103%), Essigsäure (96,2-109%), Ameisensäure (97,4-104%) und Trigonellin (102-105%) lagen in einem akzeptablen Bereich und belegten, dass die Richtigkeit für diese Analyten gegeben war. Für CQA konnte lediglich eine Wiederfindung von 76,2-80,6% und für Koffein eine Wiederfindung von 84,0-86,1% ermittelt werden. Dies zeigte, dass CQA und Koffein mit der qNMR-Methode unterbestimmt wurden. Ein Grund für diese Unterbestimmung könnte eine nicht quantitative Extraktion dieser Verbindungen aus der Kaffeematrix sein. Daneben könnte die Voigt-Funktion die Struktur der Koffein- und CQA-Signale nicht vollständig beschreiben, was weiter zu einer Unterbestimmung dieser Analyten führen könnte. Dahingegen deuteten die Wiederfindungen für HMF von 106-116% und NMP von 103-110% eine Überbestimmung an. Grund für diese Überbestimmung könnte das geringe S/N der Signale sein, über welche die beiden Analyten quantifiziert wurden. Bharti und Roy schlagen für genaue quantitative Ergebnisse ein S/N von 250:1 vor [123], wohingegen in der vorliegenden Arbeit für die Signale von HMF und NMP nur ein S/N von 242:1 bzw. 142:1 für die Aufstockung mit der jeweils höchsten Analytkonzentration berechnet wurde.

Zur Überprüfung der Linearität wurde erneut die Aufstockung mit Reinsubstanzen der Analyten in acht verschiedenen Konzentrationen herangezogen. Die bestimmte Konzentration wurde in einem Diagramm gegen die tatsächlich vorhandene Konzentration aufgetragen und der lineare Zusammenhang bewertet. Zusätzlich wurde anhand der Residuen überprüft, ob die gewählte lineare Regression die Ergebnisse hinreichend beschreiben konnte. Dabei stellten die Residuen die Abweichung des durch die lineare Regression bestimmten Wertes vom tatsächlich beobachteten Wert dar. Die lineare Regression sowie das Ergebnis der Residuenanalyse für die Analyten Milchsäure und Essigsäure ist in Abbildung 4-29 dargestellt. Die Ergebnisse zeigten, dass für beide Analyten die Linearität über den getesteten Konzentrationsbereich gegeben war. Das Bestimmtheitsmaß betrug $R^2 = 0,9997$ für Milchsäure und $R^2 = 0,9933$ für Essigsäure. Zudem streuten die Residuen ungerichtet um die y -Achse, wodurch weiter von einem linearen Zusammenhang ausgegangen werden konnte.

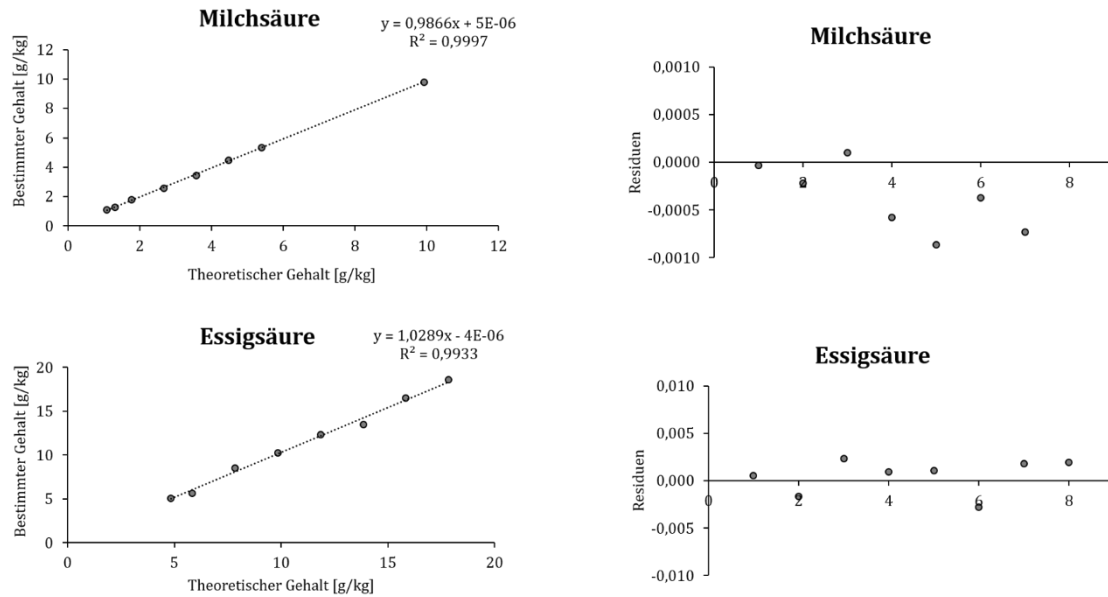


Abbildung 4-29: Ergebnisse der linearen Regression und Residuenanalyse zur Überprüfung der Linearität für die Analyten Milchsäure und Essigsäure. Links dargestellt ist die Auftragung der bestimmten Konzentration gegenüber der tatsächlich vorhandenen Konzentration für die acht Aufstockungen der jeweiligen Reinsubstanz auf die Kaffeematrix unter Angabe der Geradengleichung und des Bestimmtheitsmaßes R^2 . Rechts dargestellt ist die Auftragung der Residuen gegen die acht Aufstockungen.

Auch die Auftragung der bestimmten gegenüber der tatsächlich vorhandenen Konzentration sowie die Ergebnisse der Residuenanalyse zeigten für die Analyten Koffein, CQA, Ameisensäure, Trigonellin, NMP und HMF einen linearen Zusammenhang in dem getesteten Konzentrationsbereich (Abbildung 9-44). Das Bestimmtheitsmaß der Geradengleichung aller Analyten war $R^2 > 0,998$ und auch die Residuen wiesen eine gleichmäßige Streuung um die y-Achse auf. Lediglich für CQA war eine nicht-zufällige Verteilung der Residuen zu erkennen. Da das Residuum allerdings auch für die höchste CQA-Aufstockung einen Wert nahe Null aufwies, konnte für den getesteten Konzentrationsbereich auch bei CQA ein linearer Zusammenhang angenommen werden. Dieser könnte allerdings bei einem höheren Konzentrationsbereich nicht mehr gegeben sein.

Die Bestimmung der NG und BG erfolgte nach DIN 32645 durch die Aufstockung mit den Reinsubstanzen der Analyten in acht verschiedenen Konzentrationen. Aus den bestimmten und tatsächlich vorhandenen Konzentrationen der Aufstockung wurde eine Regressionsgerade gebildet, aus welcher die Verfahrensstandardabweichung berechnet wurde. Die Bestimmung der NG und BG erfolgte anschließend anhand der Verfahrensstandardabweichung [259]. In Tabelle 4-4 sind die NG und BG für die untersuchten Analyten aufgelistet.

Tabelle 4-4: Nachweisgrenzen (NG) und Bestimmungsgrenzen (BG) (angegeben in g/kg ermittelt nach DIN 32645) für die Analyten Milchsäure, Essigsäure, Koffein, Caffeylchinasäure-Isomere (CQA), Ameisensäure, N-Methylpyridinium (NMP), Trigonellin und 5-Hydroxymethylfurfural (HMF) in der Kaffeematrix

	Milchsäure [g/kg]	Essigsäure [g/kg]	Koffein [g/kg]	CQA [g/kg]	Ameisensäure [g/kg]	NMP [g/kg]	Trigonellin [g/kg]	HMF [g/kg]
NG	0,12	0,97	1,35	1,35	0,36	0,21	0,45	0,06
BG	0,43	3,33	4,80	4,87	1,27	0,77	1,64	0,22

Der Validierungsparameter Präzision beinhaltet die Wiederhol- und Vergleichspräzision. Für diese Parameter wurde die in Kapitel 4.3.4 verwendete Kontrollprobe jeweils fünf Mal an verschiedenen Tagen von zwei unterschiedlichen Personen aufgearbeitet und vermessen. Anhand der daraus resultierenden ^1H -NMR-Spektren wurden die Konzentrationen der acht Analyten und daraus der VK für jeden dieser Analyten berechnet. Die Ergebnisse der ermittelten VKs sind in Tabelle 9-9 aufgelistet. Die VKs für die Wiederholpräzision, für welche die Kontrollprobe jeweils fünf Mal an zwei unterschiedlichen Tagen aufgearbeitet wurde, besaßen für die Analyten Milchsäure, Koffein, CQA, Ameisensäure, NMP und Trigonellin Werte von je $<2,50\%$. Daraus resultierte eine akzeptable Wiederholpräzision für die Analyten. Auch zeigten diese Analyten eine akzeptable Vergleichspräzision, für welche die Kontrollprobe jeweils fünf Mal von zwei unterschiedlichen Personen aufgearbeitet wurde. Die VKs der Analyten Milchsäure, Koffein, CQA, Ameisensäure, NMP und Trigonellin lagen bei je $<3,00\%$. Mit einem VK für die Wiederholpräzision von $5,20\%$ und einem VK für die Vergleichspräzision von $5,40\%$ zeigte Essigsäure schlechtere Werte als die anderen Analyten, wobei diese Werte in einem noch akzeptablen Bereich lagen. Lediglich für HMF konnte eine unzureichende Präzision festgestellt werden. Der VK für die Wiederhol- und Vergleichspräzision betrug $>10,0\%$. Der Grund hierfür könnte der geringe bestimmte Gehalt von HMF in der Kaffeekontrollprobe gewesen sein. Für diesen Analyten wurde ein Gehalt unterhalb der BG bestimmt. Dadurch konnte es zu einer größeren Streuung der Messwerte kommen, wodurch der hohe VK erklärt werden konnte.

Zur Überprüfung der Robustheit wurde die Kontrollprobe fünf Mal aufgearbeitet und an unterschiedlichen 400 MHz NMR-Spektrometern vermessen und die Konzentrationen der acht Analyten bestimmt. Aus den bestimmten Konzentrationen wurden erneut die VKs für jeden Analyten gebildet (Tabelle 9-9). Wie bereits für die Wiederhol- und Vergleichspräzision nachgewiesen, zeigten auch die VKs für die Robustheit akzeptable Werte für die Analyten Milchsäure, Essigsäure, Koffein, CQA, Ameisensäure, NMP und Trigonellin. Dahingegen wurde für HMF erneut ein höherer VK von $7,50\%$ ermittelt.

Die Validierungsdaten zeigten, dass die entwickelte ^1H -qNMR-Methode für die Quantifizierung der Analyten Milchsäure, Essigsäure, Koffein, CQA, Ameisensäure, NMP, HMF und Trigonellin geeignet war. Um die entwickelte ^1H -qNMR-Methode weiter zu überprüfen und so eine gesicherte Aussage über die ermittelten Konzentrationen treffen zu können, wurde eine zusätzliche Quantifizierungsmethode mittels HPLC-DAD entwickelt. und die Ergebnisse beider Methoden wurden miteinander verglichen.

4.10.2 Entwicklung einer HPLC-DAD-Referenzmethode zur Quantifizierung ausgewählter Kaffeemetaboliten

Angelehnt an die Methode nach Gant et al. [225] erfolgte eine Bestimmung der Analyten 3-CQA, 4-CQA, 5-CQA, Koffein, Trigonellin, NMP und HMF mittels HPLC-DAD. Da das erhaltene Kaffeeextrakt der hydrophilen Metaboliten für die Messung mittels ^1H -NMR-Spektroskopie nicht für die Messung mittels HPLC geeignet war, wurde die Probenvorbereitung in dieser Hinsicht optimiert. Um eine geringere Konzentration der Analyten zu erhalten, wurde das Probe/Lösungsmittelverhältnis herabgesetzt, sodass eine Probeneinwaage von $50,0\text{ mg}$ mit $8\text{ mL H}_2\text{O}$ extrahiert wurde. Für eine vollständig quantitative Extraktion wurde die Extraktionszeit von 20 min auf 60 min erhöht. Die Probenvorbereitung und Messung erfolgte nach Arbeitsvorschrift 8.6.1. Das HPLC-DAD Chromatogramm mit den Peaks von Trigonellin, HMF, NMP, 3-CQA, 4-CQA, 5-CQA und Koffein ist beispielhaft

für eine Kaffeeprobe in Abbildung 9-45 dargestellt. Anhand Abbildung 9-45 war zu erkennen, dass der Peak von NMP im Chromatogramm eine geringe Intensität besaß und dadurch sehr verbreitert vorlag. Daraus resultierte ein geringes S/N, sodass keine Quantifizierung möglich war. Aus diesem Grund konnte NMP mit der HPLC-DAD-Quantifizierungsmethode nicht bestimmt werden.

Die Analyten wurden zunächst hinsichtlich der Validierungsparameter Richtigkeit, Linearität, NG, BG und Präzision überprüft. Die Richtigkeit wurde anhand der Wiederfindung belegt. Dazu wurden wie für die ^1H -qNMR-Methode die Reinsubstanzen in acht verschiedenen Konzentrationen zu einer Kaffeematrix gegeben. Die Wiederfindung wurde als Quotient der bestimmten Konzentration und der tatsächlich vorhandenen Konzentration berechnet und ist für die Analyten in Tabelle 9-10 aufgelistet. Die aus den Aufstockungsversuchen ermittelten NG und BG sind in Tabelle 9-11 gezeigt.

Die Auftragung der bestimmten gegenüber der tatsächlich vorhandenen Konzentration zeigte für die Analyten einen linearen Zusammenhang in dem getesteten Konzentrationsbereich (Abbildung 9-46). Das Bestimmtheitsmaß der Geradengleichung aller Analyten bis auf das von 4-CQA lag bei $R^2 > 0,996$. Das Bestimmtheitsmaß der 4-CQA zeigte mit einem R^2 von 0,989 einen geringfügig schlechteren Wert, welcher allerdings noch in einem akzeptablen Bereich lag. Zudem zeigten die Residuen von 4-CQA eine zufällige Streuung um die y-Achse, was der Annahme eines linearen Zusammenhangs entsprach. Auch die Residuen für Trigonellin und Koffein zeigten eine zufällige Streuung um die y-Achse. Dahingegen deuteten die Residuen von 3-CQA, 5-CQA und HMF eine nicht-zufällige Verteilung an. Für diese Analyten war eine Abnahme der Residuen in eine negative Richtung bei einer steigenden Konzentration erkennbar. Diese Tendenz deutete an, dass für die Analyten 3-CQA, 5-CQA und HMF die Daten durch eine polynomische Funktion 2. Grades besser approximiert werden könnten. Allerdings fielen die Werte der Residuen im Vergleich zu den Werten der Konzentrationen sehr klein aus, wodurch die Verbesserung der Ergebnisse bei Verwendung einer quadratischen Kurve nur minimal wäre. Weiter würde eine Beschreibung der Daten mit einer quadratischen Kurve in einem Verlust von Freiheitsgraden resultieren. Durch den Verlust von Freiheitsgraden steigt der T-Wert und damit die Prognosegrenzen. Dies würde in einer größeren Messunsicherheit durch die Verwendung einer Funktion 2. Grades resultieren. Aufgrund dessen wurde für diesen Fall für die Abschätzung der Leistungsfähigkeit ein linearer Zusammenhang der Daten angenommen. Die VKs bestimmt aus einer fünffachen Aufarbeitung und Messung einer Probe an zwei unterschiedlichen Tagen von zwei verschiedenen Personen betrug für alle Analyten 1,47-5,06%, was eine akzeptable Wiederhol- und Vergleichspräzision zeigte.

Die Validierungsdaten belegten, dass die entwickelte HPLC-DAD-Quantifizierungsmethode für die Bestimmung Analyten Trigonellin, HMF, 3-CQA, 4-CQA, 5-CQA und Koffein grundsätzlich geeignet war. Anhand der Validierungsdaten von NMP konnte gezeigt werden, dass eine Bestimmung von NMP mit der HPLC-DAD-Quantifizierungsmethode nicht möglich war.

4.10.3 Vergleich der ^1H -NMR-spektroskopischen und HPLC-DAD-Quantifizierungsmethode

Um die ^1H -qNMR-Methode mit der HPLC-DAD-Quantifizierungsmethode zu vergleichen, wurden aus dem in Kapitel 3.2 beschriebenen Probandensatz 78 Proben mittels ^1H -NMR und HPLC-DAD vermessen (Tabelle 9-12).

Die Gehalte von Trigonellin, Koffein und CQA wurden anschließend miteinander verglichen. Die Gehalte von HMF konnten nicht miteinander verglichen werden, da die mittels HPLC-DAD-Quantifizierungsmethode bestimmten Konzentrationen unterhalb der BG waren. Da durch die ^1H -qNMR-Methode die CQA-Isomere nicht einzeln, sondern nur als Summenparameter bestimmt werden konnten, wurde auch bei der HPLC-DAD-Quantifizierungsmethode die Summe aus den Isomeren 3-CQA, 4-CQA und 5-CQA gebildet. Um die in Kapitel 4.10.1 beschriebene geringe Wiederfindung von Koffein von 84,0-86,1% weiter zu untersuchen, wurde zusätzlich Koffein in 39 Proben mit der amtlichen Untersuchungsmethode L 46.00-3, welche auch auf eine HPLC-DAD-Messung beruht, quantifiziert (Tabelle 9-12). Abbildung 4-30 zeigt die mittels der verschiedenen Quantifizierungsmethoden bestimmten Gehalte von HMF, Trigonellin, CQA und Koffein.

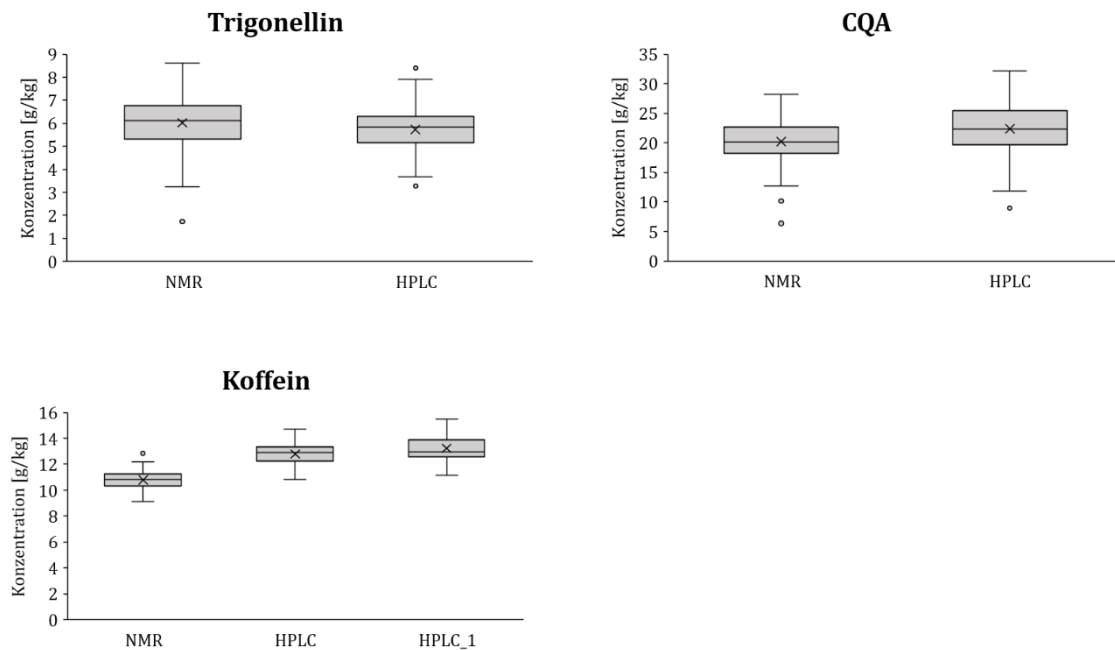


Abbildung 4-30: Darstellung der mittels ^1H -qNMR-Methode (NMR), HPLC-DAD-Quantifizierungsmethode (HPLC) und amtlicher Untersuchungsmethode L 46.00-3 (HPLC_1) bestimmten Konzentrationen von, Trigonellin, Caffeoyl-chinasäure-Isomere (CQA) und Koffein in Box-Whisker-Plots

Die mit den beiden Quantifizierungsmethoden bestimmten Gehalte von Trigonellin waren sehr ähnlich. Aus Abbildung 4-30 wird jedoch ersichtlich, dass die mittels ^1H -NMR bestimmten Gehalte eine geringfügig größere Streuung aufwiesen. Für CQA lagen die Gehalte, welche mit HPLC-DAD ermittelt wurden, höher im Vergleich zu den mittels ^1H -NMR erzielten Ergebnissen. Grund hierfür könnte die Kurvenanpassungsmethode der CQA sein, mit welcher die ^1H -NMR-Signale der CQA integriert wurden. Aus Abbildung 9-43(C) wurde ersichtlich, dass der Algorithmus der Kurvenanpassung das Signal von CQA nicht vollständig beschreibt. Weiter könnte hier der Algorithmus der Kurvenanpassung verbessert werden. Beispielsweise könnte das Duplett der CQA durch zwei Kurvenanpassungen ausgedrückt werden, welche wiederum jeweils nur ein Singulett beschreiben. So könnten aufgrund der unterschiedlichen CQA-Isomere entstehende Schultern besser mitbeschrieben werden. Die mittels HPLC-DAD bestimmten Gehalte der CQA waren im Durchschnitt jedoch nur 11,0% höher, was zeigte, dass die Abweichung der mittels ^1H -NMR bestimmten Gehalte in einem akzeptablen Bereich lagen. Aus Abbildung 4-30 wird ersichtlich, dass die mit der ^1H -qNMR-Methode ermittelten Koffeingehalte geringer waren als die mittels HPLC-DAD quantifizierten Konzentrationen. Dabei lagen die mittels entwickelter HPLC-DAD-

Quantifizierungsmethode und die mittels amtlicher Untersuchungsmethode quantifizierte Gehalte von Koffein in einem ähnlichen Bereich. Die Ursache für die durch die ^1H -qNMR bestimmten geringeren Gehalte könnte in der Probenvorbereitung gelegen haben. Bei der amtlichen Untersuchungsmethode erfolgte eine Heißextraktion von Koffein für 60 min. Bei der entwickelten HPLC-DAD-Quantifizierungsmethode wurde eine Extraktion bei Raumtemperatur für 60 min durchgeführt. Dahingegen erfolgte die Extraktion der Kaffeemetaboliten bei der ^1H -qNMR-Methode für lediglich 20 min. Letztere Extraktionszeit war ausreichend lang, um die Metaboliten reproduzierbar aus der Matrix herauszulösen (vgl. Kapitel 4.3.4). Allerdings zeigte die in der Validierung der ^1H -qNMR-Methode bestimmte Wiederfindung für Koffein von 84,0-86,1% eine nicht quantitative Extraktion dieser Verbindung. Da die Extraktion bei einer Zeit von 20 min jedoch reproduzierbar war, bestand die Möglichkeit, für die Konzentrationsbestimmung für Koffein mittels ^1H -qNMR ein Korrekturfaktor zu ermitteln. Bei einem direkten Vergleich der entwickelten ^1H -NMR und HPLC-DAD anhand der gemessenen Proben zeigte sich, dass die mit HPLC-DAD quantifizierten Werte um $18,8 \pm 7,80\%$ höher waren. Die mittels amtlicher Untersuchungsmethode bestimmten Gehalte waren $22,6 \pm 3,10\%$ höher im Vergleich zu den Koffein-Gehalten ermittelt durch ^1H -qNMR. Anhand des Vergleiches mit der amtlichen Untersuchungsmethode konnte ein Korrekturfaktor für Koffein von 1,2 ermittelt werden, welcher weiter bei der Quantifizierung von Koffein mittels ^1H -qNMR eingesetzt wurde.

4.10.4 Anwendung der ^1H -qNMR-spektroskopischen Methode auf Kaffeeproben

Die in Kapitel 3.2 beschriebenen Kaffeeproben wurden mit der Methode für die nicht zielgerichtete Untersuchung der hydrophilen Kaffeemetaboliten aufgearbeitet und vermessen. Anschließend wurden die resultierenden ^1H -NMR-Spektren mit der entwickelten Quantifizierungsmethode ausgewertet. Für die Quantifizierung wurden die bereits beschriebenen Korrekturfaktoren für Ameisensäure (1,2), Essigsäure (1,1), Trigonellin (1,1) und Koffein (1,2) eingesetzt.

Zunächst wurde die Kontrollprobe aus Kapitel 4.3.4, welche bei jeder Messserie mitvermessen wurde, ausgewertet. Für die 83 Messungen der Kontrollprobe trat eine geringe Schwankung der Gehalte von Trigonellin (3,00%), NMP (5,40%), Koffein (3,40%), Essigsäure (4,90%), Milchsäure (3,20%) und CQA (2,70%) auf, wodurch weiter gezeigt werden konnte, dass die entwickelte ^1H -qNMR-Methode für diese Analyten eine gute Reproduzierbarkeit aufwies. Für Ameisensäure trat eine erhöhte Schwankung der bestimmten Gehalte von 9,20% auf. Grund hierfür könnte möglicherweise die Flüchtigkeit dieser Verbindung sein. Für HMF konnte keine Schwankung der bestimmten Gehalte ermittelt werden, da diese unterhalb der BG von 0,19 g/kg.

Die mittels ^1H -NMR quantifizierten Gehalte der Analyten Koffein, CQA, Trigonellin, Ameisensäure, Essigsäure, HMF, NMP und Milchsäure in den in der in dieser Arbeit untersuchten Proben sind in Abbildung 4-31 dargestellt.

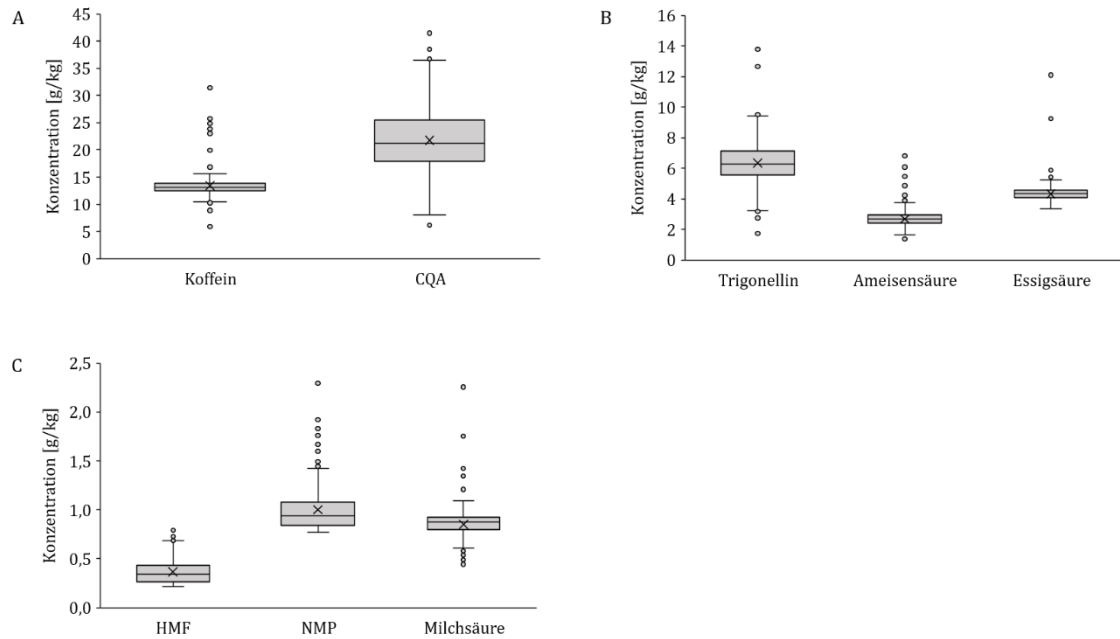


Abbildung 4-31: Die mittels der ^1H -qNMR-Methode ermittelten Gehalte von (A) Koffein und Caffeyolchinasäure-Isomere (CQA), (B) Trigonellin, Ameisensäure und Essigsäure, (C) 5-Hydroxymethylfurfural (HMF), N-Methylpyridinium (NMP) und Milchsäure in den in dieser Arbeit untersuchten Kaffeeproben, dargestellt in Box-Whisker-Plots

Anhand Abbildung 4-31 ist zu erkennen, dass die Gehalte von Milchsäure, Ameisensäure, Essigsäure und Koffein über alle gemessenen Proben eine sehr geringe Streuung aufwiesen, wohingegen die Gehalte von HMF, NMP und Trigonellin weiter streuten. Die größte Streuung wiesen die Gehalte der CQA auf. Unter Einbeziehung der untersuchten Kaffeeproben wurden Gehalte von HMF, Trigonellin, NMP, Ameisensäure, Koffein, Essigsäure, Milchsäure und CQA im Bereich von 0,22-0,81 g/kg, 1,73-13,8 g/kg, 0,77-2,29 g/kg, 1,37-6,81 g/kg, 5,93-31,5 g/kg, 3,34-12,1 g/kg, 0,44-2,26 g/kg und 6,17-41,5 g/kg bestimmt. Von allen analysierten Kaffeeproben waren die ermittelten Konzentrationen von 476 Kaffeeproben unterhalb der BG von NMP, 276 unterhalb der BG von HMF, 28 unterhalb der von Essigsäure, zwei unterhalb der BG von Milchsäure und jeweils eine unterhalb der BG von Ameisensäure und Trigonellin und wurden somit nicht berücksichtigt. Burton et al. quantifizierten unter anderem Trigonellin, Koffein, CQA, NMP, Ameisensäure und Essigsäure in methanolischen Kaffeeextrakten [88]. Sie untersuchten 292 Kaffeeproben und konnten Gehalte von Koffein und NMP quantifizieren, welche in einem gleichen Bereich wie die in dieser Arbeit ermittelten Konzentrationen waren. Allerdings quantifizierten Burton et al. einen deutlich geringeren Gehalt von Trigonellin, CQA, Essigsäure und Ameisensäure im Vergleich zu den Ergebnissen in dieser Arbeit. Aufgrund unterschiedlicher verwendeter Kaffeeproben sind die Ergebnisse in dieser Arbeit mit denen von Burton et al. allerdings nicht direkt vergleichbar.

Galli et al. quantifizierten verschiedene organische Säuren mittels Kapillarelektrophorese und konnten Konzentrationen von Ameisensäure, Essigsäure und Milchsäure von 2,70 g/kg, 7,30 g/kg und 0,72 g/kg bestimmen [260], welche den in dieser Arbeit ermittelten Konzentrationen ähnelte.

4.10.5 Anwendung der ^1H -qNMR-spektroskopischen Methode auf Kaffeeproben mit verschiedenen Defekten

Im nächsten Schritt wurde die ^1H -qNMR-Methode auf Kaffeeproben angewendet, welche verschiedene Defekte aufwiesen (Tabelle 9-2). Es wurden die Defekte „Black Beans“, „Eaten Beans“, „Well Eaten Beans“, „Sour Beans“, „Dark Green Beans“, „Faded Beans“, „Bleached Beans“, „Quaker“, „Malformed“, „Unripe Beans“, „Brocken Beans“, „Pods“, „Pulper Cuts“, „Shells“, „Twigs“, „Husks“ und „Parchment“ untersucht. Die Defekte wurden nach der New Yorker Methode klassifiziert, bei welcher die Defekte in Bezug auf ein Äquivalent des Defektes „Black Beans“ eingeteilt werden. Von den Klassifizierungsstufen NY2-NY8 wurden die Klassen NY2, NY4 und NY6 untersucht (Abbildung 9-47-Abbildung 9-51). Als Vergleich diente der entsprechende Arabica-Kaffee ohne Defekt.

Die Abbildung 9-47 bis Abbildung 9-51 zeigen die quantifizierten Analyten in den Kaffeeproben mit verschiedenen Defekten. Bei den Defekten „Eaten Beans“ und „Well Eaten Beans“, die durch den Kaffeekirschenbohrer verursacht werden, war im Vergleich zur Probe ohne Defekt eine Zunahme der aliphatischen organischen Säuren zu beobachten, wohingegen die Konzentration der Essigsäure mit einem steigenden Anteil der Defekte abnahm. Auch für Koffein war eine leichte Zunahme der Konzentration erkennbar. Für die „Black Beans“ war mit Ausnahme der Ameisensäure eine Abnahme aller Analyten im Vergleich zur Probe ohne Defekte zu beobachten. Auch Franca et al. verglichen in ihren Untersuchungen die Defekte „Black Beans“, „Unripe Beans“ und „Sour Beans“ mit einer Kaffeeprobe ohne Defekte [118]. Für „Black Beans“ konnten sie einen reduzierten Gehalt von Trigonellin und CQA feststellen. Allerdings wurden in dieser Studie die Defekte nicht nach der New Yorker Methode vermischt, wodurch ein direkter Vergleich erschwert wurde [118]. Bei einem Vergleich der „Sour Beans“ mit einer Kaffeeprobe ohne Defekt konnten verringerte Gehalte von HMF, CQA und Trigonellin festgestellt werden.

Bei den Defekten „Quaker“, „Malformed“ und „Unripe Beans“ zeigte Ameisensäure keine Veränderung, wohingegen die Konzentration der anderen Analyten abnahm. Eine ähnliche Tendenz war auch bei den Defekten „Dark Green Beans“, „Faded Beans“ und „Bleached Beans“ zu beobachten. Bei den „Unripe Beans“, bei welchen es sich um den Zusatz von Kaffeebohnen unreifer Kirschen handelt, konnte insbesondere ein geringerer Koffeingehalt ermittelt werden. Diese Beobachtung stimmt mit den Ergebnissen von Kidist et al. überein, welche auch einen geringeren Koffeingehalt in Kaffeebohnen unreifer Früchte ermitteln konnten [242].

Für den Defekt „Brocken Beans“ wurden dahingegen erhöhte Gehalte von HMF, NMP, Ameisensäure, Essigsäure und Milchsäure bestimmt. Der Gehalt von Koffein blieb unverändert und die Konzentrationen von Trigonellin und CQA sanken. Diese Veränderungen könnten durch den Röstprozess erklärbar sein. Bei dem Defekt „Brocken Beans“ handelte es sich um eine Beimischung von Bruchstücken von Rohkaffeebohnen. Während des Röstprozesses könnte es zur stärkeren Röstung dieser kleinen Bruchstücke gekommen sein, wodurch eine erhöhte Konzentration von Röstprodukten resultieren könnte. Auch bei dem Defekt „Pulper Cuts“ konnten diese Veränderungen beobachtet werden. Bei diesem Defekt wurden die Rohkaffeebohnen während der nassen Rohkaffeeverarbeitung vom Pulper gequetscht oder angeschnitten. Auch hier wäre eine stärkere Röstung der beschädigten Kaffeebohnen und damit eine höhere Konzentration der Röstprodukte möglich.

Bei dem Defekt „Pods“ blieb die Konzentration von Ameisensäure unverändert, wohingegen sich Konzentration der anderen Analyten verringerte. Dies war auch für den Defekt

„Parchment“ zu erkennen. Grund für diese Veränderungen könnte ein Verdünnungseffekt durch den Zusatz der Hülse und der Pergamenthaut zu den Kaffeebohnen sein.

Für den Defekt „Shells“ zeigte sich eine leicht erhöhte Konzentration von HMF und Ameisensäure und unveränderte Konzentrationen der anderen Analyten. Bei dem Defekt „Husks“ war zusätzlich eine erhöhte Konzentration für Milchsäure zu beobachten. Die Veränderungen in den Konzentrationen der Analyten bei dem Zusatz der Defekte „Shells“ und „Husks“ waren nur minimal. Möglicherweise war die chemische Zusammensetzung der beiden Bestandteile der Kaffeekirsche ähnlich der der Rohkaffeebohnen. Dadurch würde ein Zusatz dieser zu den Kaffeebohnen in keinen oder nur minimalen Veränderungen resultieren.

Der Defekt „Twigs“ wies eine höhere Konzentration von HMF und Ameisensäure auf. Für die Analyten Trigonellin, NMP, Koffein, Essigsäure und Milchsäure konnten keine Veränderungen der Konzentrationen festgestellt werden und der Gehalt von CQA sank leicht. Dieser Defekt kommt durch den Zusatz kleiner Zweige zustande kommt. Diese könnten durch den Röstprozess stärker geröstet werden, wodurch eine erhöhte Konzentration der Röstprodukte HMF und Ameisensäure erklärbar wäre.

Für alle untersuchten Defekte gilt jedoch, dass die beobachteten Veränderungen in den Konzentrationen der untersuchten Analyten nur minimal waren. Auch Franca et al. deuten an, dass die Unterschiede in der Zusammensetzung der Inhaltstoffe von defekten und nicht defekten Kaffee nur minimal sind [261].

4.11 Zusammenfassende Diskussion

Für die Untersuchung des geographischen Ursprungs, der Produktionsmethode und des Röstverfahrens von Kaffee wurden bereits Metabolomics-Ansätze in der Literatur beschrieben. Der Metabolomics-Ansatz bezeichnet dabei die Analyse möglichst vieler niedermolekularer Stoffwechselverbindungen, die in einem biologischen System vorhanden sind. Dazu wurden bisher verschiedene Analysetechniken wie chromatographische Methoden gekoppelt mit unterschiedlichen Detektoren, spektroskopische Methoden sowie Isotopen- und Elementanalysen [137, 139, 140, 142, 143, 147, 248, 250, 262] eingesetzt. Dabei fand jeweils eine anschließende Analyse der erhaltenen Daten mit Verfahren der multivariaten Statistik statt. Neben diesen Analysenmethoden wurde die ^1H -NMR kombiniert mit einer PCA oder PLS-DA bereits dazu verwendet, um den geographischen Ursprung von Kaffee zu untersuchen [138, 144, 146, 150].

Im Rahmen dieser Arbeit wurde ein Metabolomics-Ansatz dazu verwendet, um den geographischen Ursprung und die Anbaumethode, sowie den Röstgrad und das verwendete Röstverfahren von gerösteten Kaffeeproben zu untersuchen. Insgesamt wurden 603 trommelgeröstete Kaffeeproben der Art *Coffea arabica* L. (Arabica) untersucht. Davon stammten 89 Kaffeeproben aus Afrika, 66 aus Asien, 176 aus Nordamerika und 190 aus Südamerika, wovon wiederum 116 aus Brasilien, 35 aus Kolumbien und 51 Kaffeeproben aus Äthiopien kamen. Weiter wurden von diesen Proben 180 biologisch angebaut. Neben den 603 trommelgerösteten Kaffeeproben wurden zusätzlich 82 Kaffeeproben mit Heißluft geröstet. Bis auf Kaffee aus Kolumbien war somit eine ausreichend hohe Probenanzahl für jede Gruppe gegeben, sodass hier multivariate statistische Verfahren zur Untersuchung des geographischen Ursprungs, der Produktionsmethode und des Röstverhaltens angewendet werden konnten.

Um die ^1H -NMR-Daten mittels PCA-LDA auswerten zu können, wurden zu Beginn der Arbeit zwei Methoden entwickelt, welche möglichst viele lipophile und hydrophile Metaboliten aus der Kaffeematrix extrahieren sollten. Ein Lösungsmittelgemisch, bestehend aus Chloroform und Methanol im Verhältnis 1:1 (v/v), konnte zur umfassenden Extraktion der lipophilen Metaboliten aus der Kaffeematrix verwendet werden. Um eine Sättigung des Empfängers und eine Unterdrückung der vorkommenden Lösungsmittelsignale während der ^1H -NMR-Messung zu vermeiden, wurde von einem Aliquot des Extraktes das Lösungsmittel nach der Extraktion entfernt und der Rückstand in einer Mischung aus CDCl_3 und $\text{MeOH-}d_4$ im Verhältnis 1:1 (v/v) gelöst. Dazu mussten zunächst mögliche reaktive Abbauprodukte des CDCl_3 aus diesem entfernt werden [161]. Abbauprodukte wie beispielsweise Salzsäure könnten zu einer Änderung des pH-Wertes des im CDCl_3 vorhandenen Restwassers führen. Dies könnte in Signal-Shifts von polarerer Metaboliten resultieren, was eine Auswertung mittels PCA-LDA erschwerte. Durch die Entfernung der Abbauprodukte konnten Signal-Shifts von Ameisensäure und Trigonellin im ^1H -NMR-Spektrum der lipophilen Metaboliten deutlich reduziert werden.

Für die Extraktion der hydrophilen Metaboliten wurde als Lösungsmittel H_2O verwendet. Bei einem hohen Verhältnis von Probe zu Lösungsmittel traten Shifts der Koffein- und CGA-Signale auf, was auf die Selbstassoziation von Koffein bzw. die Bildung eines π -Komplexes mit CGA zurückgeführt werden kann. Da dies die anschließende multivariate Datenanalyse mittels PCA-LDA negativ beeinflusste, wurde das Verhältnis möglichst gering gewählt. Auftretende pH-Wert abhängige Signal-Shifts der polaren Kaffeemetaboliten konnten durch den Zusatz eines Natriumphosphat-Puffers mit einem pH-Wert = 6,00 und einer

Konzentration von 1,00 mol/L zur Messlösung minimiert werden. Zur Unterdrückung des Wassersignals während der Erfassung der Metaboliten mittels $^1\text{H-NMR}$ eignete sich die NO-ESY-Pulssequenz. Es wurde überprüft, ob eine Entfernung der polymeren Kohlenhydratstrukturen aus dem Kaffeeextrakt positive Auswirkungen auf die Ergebnisse der nachfolgenden PCA-LDA haben könnte. Eine Abtrennung der hochmolekularen Kohlenhydrate durch Ultrazentrifugation resultierte in einer schlechteren Abgrenzung von Kaffeeproben aus verschiedenen Kontinenten. Eine mögliche Erklärung ist, dass Kohlenhydrate in Röstkaffee hauptsächlich in Melanoidinen eingebunden vorliegen. Neben den Kohlenhydraten sind andere niedermolekulare Bestandteile wie beispielsweise CQA in Melanoidinen vorhanden. Durch die Abtrennung der Melanoidine kam es demnach auch zum Verlust anderer Kaffeebestandteile. Auch bei anderen Untersuchungen zum geographischen Ursprung von Kaffee mittels $^1\text{H-NMR}$ und multivariater statistischer Datenanalyse wurde auf eine Entfernung hochmolekularer Kohlenhydrate verzichtet [145, 150].

Nach der Erfassung der lipophilen und hydrophilen Metaboliten erfolgte eine Datenvorbehandlung der $^1\text{H-NMR}$ -Spektren um experimentelle, messtechnische und auswertetechnische Variationen zu minimieren. Dazu wurde zunächst eine Referenzierung auf das TMS-Signal bei den $^1\text{H-NMR}$ -Spektren der lipophilen Metaboliten und auf das TSP-Signal bei den $^1\text{H-NMR}$ -Spektren der hydrophilen Metaboliten durchgeführt. Anschließend erfolgte eine Normalisierung bezüglich eines Standardsignals. Bei den $^1\text{H-NMR}$ -Spektren der lipophilen Metaboliten wurde das Signal des DMS ($\delta_{\text{H}} = 3,00 - 3,05$ ppm) gewählt, wohingegen die Normalisierung bei den $^1\text{H-NMR}$ -Spektren der hydrophilen Metaboliten auf das Signal von TSP ($\delta_{\text{H}} = -0,10 - 0,10$ ppm) stattfand. Zur Reduktion der Anzahl der Datenpunkte und zum Ausgleich geringfügiger Signal-Shifts erfolgte eine Bucketierung der $^1\text{H-NMR}$ -Spektren. Dazu wurden für die Spektren der lipophilen Metaboliten 2000 und für die Spektren der hydrophilen Metaboliten 1000 äquidistante Buckets gewählt. Nach Ausschluss der Lösungsmittelsignale wurden die Integrale der Buckets log-transformiert und als Variablen für die PCA-LDA verwendet. Auf eine weitere Skalierungsmethode wie dem Autoscaling wurde verzichtet, da dadurch der Einfluss von Variablen ohne Information für die Abgrenzung der Klassen erhöht wurde. Auch auf eine zusätzliche Normalisierung der $^1\text{H-NMR}$ -Spektren auf einen einheitlichen Feuchtegehalt der Kaffeeproben wurde verzichtet, da die Variation der Feuchtegehalte der Kaffeeproben zwischen den Klassen nur gering ausfiel. Demnach führte eine PCA-LDA basierend auf der Analyse der lipophilen bzw. hydrophilen Metaboliten nach der Normalisierung auf einen einheitlichen Feuchtegehalt zu keiner Veränderung. Auch andere Arbeitsgruppen, welche die Authentizität von geröstetem Kaffee basierend auf $^1\text{H-NMR}$ in Kombination mit multivariaten statistischen Verfahren untersuchten, verzichteten auf diese Normalisierung [138, 139, 144, 145, 170].

Nach der Datenvorbehandlung erfolgte die Erstellung verschiedener Klassifizierungsmodelle. Es wurde eine PCA-LDA mit den vorbehandelten Daten durchgeführt, wobei die PCA zur weiteren Dimensionsreduktion angewendet wurde und mit der LDA der Aufbau der Klassifizierungsmodelle stattfand. Zur Überprüfung der Klassifizierungsmodelle diente eine interne Validierung mittels MCKV und eine externe Validierung. Für die externe Validierung wurden 20% der Proben vor der Erstellung der Modelle entfernt. Mit den Daten der restlichen 80% der Proben erfolgte die Modellerstellung und interne Validierung. Dazu wurde der Datensatz für die KV erneut in 10% Testdatensatz und 90% Trainingsdatensatz unterteilt. Mit dem Trainingsdatensatz erfolgte die Erstellung des Klassifizierungsmodells, welches mit dem Testdatensatz intern validiert wurde. Für die KV erfolgte die Einteilung des Probandatensatzes in Trainings- und Testdatensatz zehn Mal, sodass jede Probe jeweils einmal bei der Modellerstellung und internen Validierung beteiligt war. Da Trainings- und

Testdatensatz nicht vollständig unabhängig voneinander waren, erfolgte zusätzlich ein Monte-Carlo Ansatz, bei welchem die KV zehn Mal wiederholt wurde. Die Einteilung der Proben in Trainings- und Testdatensatz erfolgte bei jeder neuen KV nach dem Zufallsprinzip, sodass der Testdatensatz aus unterschiedlichen Kombinationen von Proben bestand. Das Ergebnis der internen Validierung war eine Konfusionsmatrix, aus welcher hervorging, welche Anzahl der Proben aus dem Testdatensatz jeweils der richtigen Klasse zugeordnet wurde. Nach der Modellerstellung erfolgte die externe Validierung, in der das Klassifizierungsmodell auf die zuvor entfernten Proben angewendet wurde. Die Beurteilung der Vorhersagefähigkeit des Klassifizierungsmodells erfolgte anhand des *conf*-Wertes, der den relativen Abstand der Testprobe zum Datenmittelpunkt der entsprechenden Klasse angibt. Für eine richtige Zuordnung der Probe sollte der *conf*-Wert für die entsprechende Klasse < 1 sein, wodurch die Probe innerhalb des 95%igen Prognoseellipsoids der Klasse liegen würde.

Für die Untersuchung des geographischen Ursprungs von Kaffee auf der Ebene der Kontinente wurden zwei-Klassen-Modelle erstellt, in welchen jeweils Kaffeeproben der Kontinente Afrika, Asien, Nordamerika und Südamerika gegeneinander aufgetragen wurden. Für die Abgrenzung von Kaffeeproben aus Afrika von Kaffee aus einem anderen Kontinent konnte basierend auf der Analyse der lipophilen und hydrophilen Metaboliten vergleichbare Klassifizierungsmodelle mit hoher Vorhersagekraft erstellt werden. In der externen Validierung dieser Klassifizierungsmodelle basierend auf der Analyse der lipophilen Metaboliten konnten mit 92,8% geringfügig mehr Testproben der richtigen Klasse zugeordnet werden. Dies deutet eine geringfügig höhere Robustheit der Klassifizierungsmodelle basierend auf der Analyse der lipophilen Metaboliten an. Im Gegensatz zur erfolgreichen Abtrennung von Kaffeeproben aus Afrika wurden mit den Modellen zur Zuordnung von Kaffeeproben aus Asien leicht schlechtere Ergebnisse erzielt. Der Prozentsatz für die richtige Zuordnung der Proben aus dem internen und externen Testdatensatz waren auch für diese Modelle basierend auf der Analyse der lipophilen und hydrophilen Metabolite vergleichbar. In der internen Validierung konnten geringfügig mehr Proben aus dem Testdatensatz in dem Modell basierend auf der Analyse der lipophilen Metaboliten zugeordnet werden. Dahingegen konnte in der externen Validierung dieser Klassifizierungsmodelle mit 91,4% mehr Kaffeeproben aus dem Testdatensatz bei der Analyse der hydrophilen Metaboliten der richtigen Klasse zugeordnet werden. Für das PCA-LDA-Modell zur Abgrenzung von Kaffeeproben aus Nord- von Kaffee aus Südamerika basierend auf der Analyse der hydrophilen Metaboliten wurden in der internen und externen Validierung durchschnittlich 87,1% und 89,0% der Testproben der richtigen Klasse zugeordnet. Auch Consonni et al. und Choi et al. konnten erfolgreich Kaffeeproben aus verschiedenen Kontinenten voneinander abgrenzen [143, 145]. Gründe für die unterschiedlich guten Abgrenzungen von Kaffeeproben aus den verschiedenen Kontinenten könnten neben dem Mikrobiom des Bodens auch unterschiedliche Anbau-, Ernte- und Rohkaffeeverarbeitungstechniken sein, welche in den einzelnen Kontinenten angewendet werden. Wie in dieser Arbeit gezeigt werden konnte, scheinen sich diese Faktoren unterschiedlich stark auf das Metabolom von Kaffee auszuwirken. Ein Grund für die schlechtere Abgrenzung von Kaffeeproben aus Nordamerika von Kaffeeproben aus Südamerika könnte die im Vergleich zu den anderen Kontinenten größere geographische Nähe sein. Für die Abgrenzung von Kaffeeproben aus verschiedenen Kontinenten eignete sich sowohl die Analyse der lipophilen Metaboliten als auch die Analyse der hydrophilen Metaboliten. Die externe Validierung zeigte, dass die erstellten Klassifizierungsmodelle robust gegenüber neuen und unbekanntem Kaffeeproben sind.

Im nächsten Schritt wurden Klassifizierungsmodelle zur Abgrenzung von Kaffeeproben aus den Ländern Äthiopien, Brasilien und Kolumbien erstellt. Dazu wurden erneut zwei-Klassen-Modelle gebildet, in welchen Kaffeeproben eines der Länder und Kaffeeproben aus einem anderen Kontinent gegenübergestellt wurden. Die Vorhersagefähigkeit der erstellten Modelle wurde anhand der internen und externen Validierung beurteilt. Die Klassifizierungsmodelle basierend auf der Analyse der hydrophilen Metaboliten zeigten bei der internen Validierung eine geringfügig bessere Vorhersagekraft als die Modelle basierend auf der Analyse der lipophilen Metaboliten. Dahingegen war die Vorhersagefähigkeit bei der externen Validierung für die Klassifizierungsmodelle basierend auf der Analyse der lipophilen und hydrophilen Metaboliten vergleichbar. Bei den Klassifizierungsmodellen zur Abgrenzung von Kaffeeproben aus Äthiopien bzw. Brasilien von Kaffeeproben aus anderen Kontinenten konnten mindestens 91,3% der Proben aus dem internen und externen Testdatensatz der richtigen Klasse zugeordnet werden. Die Ergebnisse zeigen, dass es bereits durch die Anwendung der Klassifizierungsmodelle basierend auf der Analyse der hydrophilen Metaboliten möglich ist, Kaffeeproben den Ländern Äthiopien und Brasilien zuzuordnen. Dadurch könnte auf die aufwändigere Probenvorbereitung mit toxischen Lösungsmitteln, welche für die Analyse der lipophilen Metaboliten benötigt werden, verzichtet werden. Allerdings ist es notwendig, noch Daten weiterer Kaffeeproben in die Modelle aufzunehmen. Mit einer größeren Anzahl an Proben steigt die Robustheit der Modelle und es kann eine sicherere Abschätzung der Klassifizierungsgüte erfolgen. Zu bedenken ist auch, dass klimatische Veränderungen zu einer Veränderung des Metaboloms der Kaffeeproben führen können, weshalb es wichtig ist, auch zukünftig Daten von Kaffeeproben in die Klassifizierungsmodelle aufzunehmen.

Bei der bei der internen und externen Validierung der Klassifizierungsmodelle basierend auf der Analyse der hydrophilen Metaboliten zur Abgrenzung von Kaffeeproben aus Kolumbien von Kaffeeproben aus anderen Kontinenten konnten mindestens 80,0% der Testproben der richtigen Klasse zugeordnet werden. Eine vollständige Abtrennung von Kaffeeproben aus Kolumbien von Kaffeeproben aus anderen Ländern konnte daher nicht erreicht werden. Ein Grund für die vergleichsweise schlechtere Abgrenzung von Kaffeeproben aus Kolumbien könnte die geringe Probenanzahl sein. Für die Modellerstellung standen lediglich 28 Kaffeeproben aus Kolumbien zur Verfügung. Durch eine Erhöhung der Probenanzahl auf 50 – 100 könnte eine bessere Abgrenzung von Kaffeeproben aus Kolumbien von denen aus anderen Kontinenten erreicht werden. Weiter könnte die Anwendung eines anderen Verfahrens der multivariaten Datenanalyse wie eine PLS-DA oder SIMCA auf die ¹H-NMR-Daten zu einer besseren Abgrenzung von Kaffeeproben aus Kolumbien führen. So zeigt die Studie von Arana et al., dass durch die Analyse methanolischer Extrakte mittels ¹H-NMR und PLS-DA eine Abtrennung von 84 Kaffeeproben aus Kolumbien von denen anderer Kontinente möglich ist [138]. Allerdings ist insbesondere für die Anwendung einer SIMCA oder PLS-DA eine höhere Probenanzahl unabdingbar. Das Ziel, die Authentizität von Kaffeeproben aus Kolumbien zu überprüfen, sollte in Zukunft weiterverfolgt werden. Hochwertiger Kaffee aus Kolumbien darf mit der Angabe g.g.A. ausgelobt werden, wodurch eine Fälschung dieser Kaffeeprodukte attraktiv werden kann.

Der Grund für die Abgrenzung von Kaffeeproben aus den Ländern Brasilien und Äthiopien von Kaffeeproben aus anderen Kontinenten könnte insbesondere in unterschiedlichen Anbau-, Ernte- und Rohkaffeeverarbeitungstechniken liegen. So werden beispielsweise in Brasilien die Kaffeeplantagen hauptsächlich im Sonnenanbau betrieben, die Ernte erfolgt überwiegend maschinell und die Rohkaffeebohnen werden mit der halbtrockenen Methode verarbeitet [3, 15, 16, 244, 245]. Auf Plantagen in Äthiopien wird Kaffee hauptsächlich im

Schattenanbau produziert und zu 70% die trockene und zu 30% die nasse Methode zur Rohkaffeeverarbeitung verwendet [19, 25, 242, 246]. In Kolumbien dagegen wird die nasse Rohkaffeeverarbeitung angewendet [15].

Aufgrund der guten Vorhersagefähigkeit der Klassifizierungsmodelle zur Abgrenzung von Kaffeeproben aus den Ländern Brasilien bzw. Äthiopien von Kaffeeproben aus anderen Kontinenten wurden für diese Modelle der Loadingsplot analysiert. Dazu wurden die Daten der lipophilen und hydrophilen Metaboliten herangezogen. Ziel war die Identifizierung möglicher Metaboliten als Ursache für die Abgrenzung von Kaffeeproben aus diesen Ländern. Anhand des Loadingsplots basierend auf der Analyse der lipophilen Metaboliten konnten nur wenige Variablen erkannt werden, welche die Abgrenzung der Klassen beeinflussten. Die Untersuchung des Loadingsplots deutete an, dass die Signale von Koffein die Abgrenzung von Kaffeeproben aus Brasilien beeinflussten. Daneben konnten keine Metaboliten identifiziert werden, welche die Abgrenzung von Kaffeeproben aus Äthiopien beeinflussten. In den spektralen Bereichen der Variablen, welche einen Einfluss auf die Abgrenzung von Kaffeeproben aus Äthiopien hatten, befanden sich viele kleine Signale, die ein geringes S/N aufwiesen und überlagert waren. Für diese Signale war eine Identifizierung der zugrundeliegenden Verbindung nicht möglich. Zudem zeigte ein direkter Vergleich der log-transformierten Integrale dieser Variablen in Box-Whisker-Plots nur geringfügige Unterschiede zwischen den Klassen. Durch die Untersuchung des Loadingsplots basierend auf der Analyse der hydrophilen Metaboliten konnte keine Verbindung identifiziert werden, welche die Abgrenzung von Kaffeeproben aus Äthiopien bzw. Brasilien von Kaffeeproben aus anderen Kontinenten beeinflussten. Variablen, welche einen Einfluss auf die Abgrenzung von Kaffeeproben aus Äthiopien hatten, wiesen beispielsweise spektrale Bereiche auf, in denen Äpfel- und Zitronensäure Signale erzeugten. Dieser spektrale Bereich war erneut durch viele überlappende Signale gekennzeichnet, sodass keine Aussage darüber getroffen werden konnte, ob und welche dieser beiden Verbindungen einen Einfluss auf die Trennung hatte. Die Analyse der Loadingsplots deuten allerdings an, dass die Abgrenzung der Kaffeeproben aus Brasilien bzw. Äthiopien von Kaffeeproben aus anderen Kontinenten nicht auf die An- oder Abwesenheit einzelner Metaboliten zurückzuführen ist. Die Abgrenzung der Klassen könnte vielmehr durch ein Zusammenspiel mehrerer Metaboliten verursacht werden, welche in unterschiedlichen Konzentrationen vorliegen.

Mit dem Ziel, den Informationsgehalt zu erhöhen und gleichzeitig den Einfluss Variablen ohne relevante Informationen zu reduzieren, wurden die $^1\text{H-NMR}$ -Daten der lipophilen und hydrophilen Metaboliten in einer Low-Level und Mid-Level Datenfusionierung zusammengeführt. Dies wurde für die Daten der Kaffeeproben aus Brasilien und Afrika sowie Kolumbien und Nordamerika durchgeführt. Durch die Low-Level Datenfusionierung konnten in beiden Fällen in der internen Validierung mittels MCKV Werte für die richtige Zuordnung der Testproben erreicht werden, die jeweils zwischen denen der Einzelmodelle lagen. Die Mid-Level Datenfusionierung führte zu geringfügig schlechteren Werten für die richtige Zuordnung der Testproben in der internen Validierung verglichen mit den einzelnen Klassifizierungsmodellen auf Basis der Daten der hydrophilen und lipophilen Metaboliten. Ein möglicher Grund hierfür könnte ein leichtes Overfitting der Modelle aufgrund der hohen Anzahl der verwendeten Dimensionen für die LDA sein. Dadurch könnten die Klassifizierungsmodelle zu sehr an die Proben des Trainingssets angepasst worden sein, wodurch sich die Streuung der Proben des Trainingsdatensatz im LDA-Raum verringerte, was zu geringeren Grenzen des 95%igen Prognoseellipsoids führte. Durch die zu genaue Anpassung der Klassifizierungsmodelle an die Proben des Trainingsdatensatz konnten die Proben aus dem internen Testdatensatz nicht mehr vollständig beschrieben werden und es

kam bei der MCKV vermehrt zu einer falschen Zuordnung. Eine Reduktion der Dimensionen für die LDA führte bei der Mid-Level Datenfusionierung zu einer größeren Streuung der Proben des Trainingsdatensatzes im LDA-Raum, allerdings konnten bei der MCKV geringfügig höhere Prozentsätze für die richtige Zuordnung erreicht werden. Die Ergebnisse zeigten, dass durch eine Datenfusionierung der lipophilen und hydrophilen ^1H -NMR-Daten keine bessere Abgrenzung der Kaffeeproben aus den verschiedenen Klassen erreicht werden konnte. Möglicherweise könnte eine höhere Anzahl von Proben dazu beitragen, eine bessere Abgrenzung der Klassen mittels Low-Level oder Mid-Level Datenfusionierung zu erreichen. Auch könnte eine Fusionierung der ^1H -NMR-Daten mit Daten aus einer anderen Analysenmethode zu einer besseren Abgrenzung der Kaffeeproben aus den verschiedenen Klassen führen. So kombinierten Choi et al. LC/MS- und GC-FID-Daten sowie Daten aus Protein-, Monosaccharid- und Aminosäurenanalyse um Kaffeeproben dem geographischen Ursprung zuzuordnen [143].

Die entwickelten Methoden wurden zusätzlich zur Untersuchung der Anbaumethode und des Röstverfahrens angewendet. Dazu wurden erneut Klassifizierungsmodelle basierend auf der Analyse der lipophilen bzw. hydrophilen Metaboliten erstellt und intern sowie extern validiert. Zu der Unterscheidung der Anbaumethode wurden Kaffeeproben aus biologischem und konventionellem Anbau verwendet. Kaffeeproben, die mittels Trommel- oder Heißluftverfahrens geröstet wurden, dienten zur Untersuchung des Röstverfahrens. Eine verlässliche Abgrenzung von Kaffeeproben aus biologischem und konventionellem Anbau war mit dem untersuchten Probandensatz nicht möglich. Das Klassifizierungsmodell basierend auf der Analyse der lipophilen Metaboliten zeigte in der internen und externen Validierung eine geringfügig bessere Zuordnung der Testproben als das Modell basierend auf der Analyse der hydrophilen Metaboliten. Die unzureichende Abgrenzung der Kaffeeproben dieser Klassen könnte darin begründet sein, dass es in der Praxis keine klare Grenze zwischen biologisch und konventionell erzeugtem Kaffee gibt. Viele insbesondere kleine Kaffeeplantagen haben nicht den Zugang und die Mittel, um Agrarchemikalien zu verwenden. Diese Plantagen betreiben traditionellen Kaffeeanbau, in welchem auf diese Chemikalien verzichtet wird. Dadurch erfolgt der Kaffeeanbau biologisch, was nicht zwingend mit einer Zertifizierung verbunden ist. Ohne Zertifizierung darf dieser Kaffee jedoch nicht als biologisch ausgelobt werden [32, 263]. Allerdings konnten do Carmo Carvalho et al. durch die Verwendung von HPLC, Barbosa et al. mittels Elementaranalyse und Gordillo-Delgado et al. mittels FT Infrarotspektroskopie kombiniert mit multivariaten statistischen Verfahren eine Abgrenzung von Kaffeeproben erreichen, welche biologisch bzw. konventionell angebaut wurden [141, 252, 253]. Im Gegensatz zu dieser Arbeit stammten die von diesen Arbeitsgruppen untersuchten Kaffeeproben aus jeweils dem gleichen Land. In dieser Arbeit könnte sich der Einfluss des geographischen Ursprungs nachteilig auf die Abgrenzung von biologisch und konventionell angebauten Kaffeeproben auswirken.

Dahingegen konnten die Kaffeeproben entsprechend dem Parameter Röstverfahren differenziert werden. Das Klassifizierungsmodell basierend auf der Analyse der lipophilen bzw. hydrophilen Metaboliten zeigte jeweils ähnlich gute Ergebnisse. In der internen und externen Validierung konnten mindestens 93,3% der Testproben der richtigen Klasse zugeordnet werden. Die Klassifizierungsmodelle zur Abgrenzung von Kaffeeproben nach dem Röstverfahren könnten in Zukunft z.B. dazu beitragen, die Authentizität von Kaffee zu beurteilen, welcher als „schonend gerösteter Kaffee im Langzeit-Trommelröstverfahren“ ausgelobt wird.

Neben dem nicht zielgerichteten Metabolomics-Ansatz wurde für ausgewählte hydrophile Kaffeemetaboliten eine zielgerichtete Analyse durchgeführt. Die $^1\text{H-NMR}$ -Spektren der hydrophilen Metaboliten wurden dazu verwendet, die Analyten Milchsäure, Essigsäure, Koffein, CQA, Ameisensäure, NMP, Trigonellin und HMF mittels PULCON-Methode zu quantifizieren. Die Isomere der CQA konnten nicht einzeln quantifiziert werden, da sich die Signale der Isomere überlagerten. So wurden Summenparameter für die Isomere 5-CQA, 4-CQA und 3-CQA bestimmt. Eine Quantifizierung der einzelnen CQA-Isomere könnte durch die Verwendung eines anderen Lösungsmittels möglich sein. Burton et al. quantifizierten verschiedene Kaffeemetaboliten in einem methanolischem Extrakt und konnten so die 3-CQA, 4-CQA und 5-CQA einzeln erfassen [88]. Die Signale der Milchsäure, der Essigsäure, des Koffeins und der CQA-Isomere lagen im $^1\text{H-NMR}$ -Spektrum überlagert vor, sodass für die Integration dieser Signale eine Kurvenanpassungsmethode verwendet wurde. Hierbei wurde eine Voigt-Funktion unter Berücksichtigung von Parametern wie Breite des Signals, chemische Verschiebung, Multiplett-Struktur oder Kopplungskonstante an das entsprechende Signal angepasst. Da die Protonen der Analyten Ameisensäure, Essigsäure und Trigonellin bei der verwendeten Repetitionszeit des $^1\text{H-NMR}$ -Experimentes nicht vollständig relaxiert vorlagen, wurden zur Berechnung der Konzentrationen Korrekturfaktoren eingesetzt. Diese waren für Ameisensäure 1,2 und für Essigsäure und Trigonellin jeweils 1,1. Die Validierung zeigte, dass die $^1\text{H-qNMR}$ -Methode zur Bestimmung des Gehaltes von Milchsäure, Essigsäure, Koffein, CQA, Ameisensäure, NMP, Trigonellin und HMF geeignet war. Lediglich Koffein und CQA wiesen bei der Untersuchung der Richtigkeit Wiederfindungen zwischen 84,0-86,1% und 76,2-80,6% auf. Durch einen Vergleich mit einer entwickelten HPLC-DAD-Referenzmethode und mit der amtlichen Untersuchungsmethode zur Bestimmung von Koffein konnte festgestellt werden, dass die Ursache der geringen Wiederfindung von Koffein in der nicht quantitativen Extraktion dieses Analyten während der Probenvorbereitung für die $^1\text{H-NMR}$ -Messung war. Anhand des Vergleiches der $^1\text{H-qNMR}$ -Methode mit der amtlichen Untersuchungsmethode zur Bestimmung von Koffein konnte ein Korrekturfaktor von 1,2 für die Berechnung des Gehaltes von Koffein mittels $^1\text{H-NMR}$ ermittelt werden. Auch wurden mit der entwickelten HPLC-DAD-Referenzmethode im Vergleich zur $^1\text{H-qNMR}$ -Methode durchschnittlich 11,0% höhere CQA-Gehalte bestimmt. Der Grund dafür könnte die Voigt-Funktion sein, welche das CQA-Signal nicht vollständig beschrieb. Eine bessere Anpassung dieses Algorithmus an das Signal könnte zu höheren Gehalten der CQA führen. Beispielsweise könnte das Duplett der CQA, welches zur Quantifizierung verwendet wurde, durch zwei Singulets beschrieben werden. In diesem Fall würde das Signal durch zwei Kurvenanpassungen ausgedrückt werden, in welchen die aufgrund der unterschiedlichen CQA-Isomere entstehende Schultern besser mitbeschrieben werden könnten.

Die Anwendung der $^1\text{H-qNMR}$ -Methode auf die in dieser Arbeit untersuchten Proben resultierte in Gehalten von 0,22-0,81 g/kg für HMF, 1,73-13,8g/kg für Trigonellin, 0,77-2,29 g/kg für NMP, 1,37-6,81 g/kg für Ameisensäure, 5,39-31,5 g/kg für Koffein, 3,34-12,1 g/kg für Essigsäure, 0,44-2,26 g/kg für Milchsäure und 6,17-41,5 g/kg für CQA. Die $^1\text{H-qNMR}$ -Methode konnte erfolgreich zur Untersuchung von Kaffeeproben angewendet werden. Eine Weiterentwicklung dieser Methode könnte dazu beitragen, den Zusatz von defekten Kaffeebohnen und die Zugabe von beispielsweise Getreidekaffee oder Bestandteilen der Kaffeekirsche zu untersuchen. De Moura Ribeiro et al. fanden bereits Unterschiede in den $^1\text{H-NMR}$ -Spektren von Kaffee, Gerste, Sojabohne, Mais und Kaffeeschalen [169]. Diese Spektrale Unterschiede könnten dazu verwendet werden, um den Zusatz von beispielsweise Getreide festzustellen und die zugesetzte Menge zu quantifizieren. Weiter wurden

Kaffeeproben untersucht, die mit verschiedenen Anteilen von unterschiedlichen Kaffeedefekten vermischt wurden. Die Defekte wirkten sich unterschiedlich auf die Veränderungen der Konzentration der untersuchten Analyten aus. Allerdings waren die beobachteten Veränderungen nur minimal. Dass die Veränderungen der Inhaltsstoffe durch den Zusatz von Kaffeebohnen mit Defekten nur minimal ausfallen, stellten auch Franca et al. in ihrer Studie fest [261].

5 Zusammenfassung

Die Protonen-Kernspinresonanz-Spektroskopie ($^1\text{H-NMR}$) ist prinzipiell eine geeignete Methode, um die Authentizität von Lebensmitteln zu überprüfen. Für Kaffee konnte in der Vergangenheit die zielgerichtete $^1\text{H-NMR}$ bereits erfolgreich angewendet werden, um die Authentizität bezüglich der Art zu überprüfen. Als Beurteilungskriterium dient der Gehalt der Verbindung 16-O-Methylcafestol, welche in der Art *Coffea canephora* Pierre ex A. vorkommt und in *Coffea arabica* L. nur in Spuren enthalten ist. Insbesondere nicht zielgerichtete Ansätze, in welchen die $^1\text{H-NMR}$ mit multivariater Datenanalyse kombiniert wird, können dazu verwendet werden, um Unterschiede zwischen verschiedenen Produktgruppen zu erkennen. In der amtlichen Lebensmittelüberwachung fehlen jedoch zuverlässige Analysemethoden für die Überprüfung der Authentizität von Kaffee bezüglich weiterer wertgebender Eigenschaften, wie den geographischen Ursprung, biologischen Anbau oder das angewandte Röstverfahren. Ziel der vorliegenden Arbeit war daher die Entwicklung eines zuverlässigen, nicht zielgerichteten Ansatzes zur Überprüfung der Authentizität von Kaffee der Art *Coffea arabica* L. bezüglich des geographischen Ursprungs, der Anbaumethode und des Röstverfahrens. Darüber hinaus sollten weitere Kaffeeinhaltsstoffe in einem zielgerichteten, auf $^1\text{H-NMR}$ basierenden Verfahren untersucht werden.

Für die Erfassung des Metaboloms mittels $^1\text{H-NMR}$ wurden zunächst Methoden entwickelt, um die lipophilen bzw. hydrophilen Kaffeemetaboliten aus der Matrix zu extrahieren. Anschließend wurden mit den Methoden 688 geröstete Kaffeeproben der Art *Coffea arabica* L. extrahiert und mittels $^1\text{H-NMR}$ analysiert. Die Messung der Kaffeeproben mittels $^1\text{H-NMR}$ erfolgte nach Optimierung der Messparameter an einem 400 MHz Spektrometer. Um experimentelle, messtechnische und auswertetechnische Variationen zu minimieren, wurde mit den resultierenden $^1\text{H-NMR}$ -Spektren eine geeignete Datenvorbehandlung durchgeführt. Diese beinhaltete ein Alignment und eine Normalisierung der $^1\text{H-NMR}$ -Spektren bezüglich eines Standardsignals, die Einteilung der Spektren in äquidistante Buckets, den Ausschluss der Lösungsmittelsignale sowie eine log-Transformation der Integrale der erhaltenen Buckets. Mit den vorbehandelten Daten wurde eine Hauptkomponentenanalyse (PCA) durchgeführt, um die Dimensionalität der Daten zu reduzieren. Die erhaltenen Scores der PCA wurden anschließend für die lineare Diskriminanzanalyse (LDA) verwendet und so Klassifizierungsmodelle erstellt. Die Validierung der Klassifizierungsmodelle erfolgte intern mittels Monte-Carlo-Kreuzvalidierung und extern, indem ein Teil des Originaldatensatzes vor Erstellung der Klassifizierungsmodelle entfernt wurde. Die mittels PCA-LDA erstellten Klassifizierungsmodelle wurden anschließend auf den externen Probandatensatz angewendet.

Zur Abgrenzung der Kaffeeproben bezüglich der Kontinente wurden zwei-Klassen-Modelle gebildet. Die interne und externe Validierung zeigten, dass Kaffeeproben aus Afrika gut gegenüber Kaffeeproben aus anderen Kontinenten abgetrennt werden konnten. Dabei zeigten die PCA-LDA-Modelle basierend auf der Analyse der hydrophilen Metaboliten geringfügig bessere Klassifizierungsgenauigkeiten als die PCA-LDA-Modelle basierend auf der Analyse der lipophilen Metaboliten. Die Proben des internen und externen Validierungsdatensatzes konnten durchschnittlich zu 95,4% der richtigen Klasse zugeordnet werden. Für Klassifizierungsmodelle zur Abgrenzung von Kaffeeproben aus Asien von Kaffeeproben aus anderen Kontinenten konnten in der internen und externen Validierung zufriedenstellende Prozentsätze für die richtige Zuordnung erzielt werden. Mit den Klassifizierungsmodellen

zur Unterscheidung von Kaffee aus Nord- und Südamerika konnte keine eindeutige Trennung erreicht werden. Die richtige Zuordnung erfolgte in der internen bzw. externen Validierung nur zu durchschnittlich 88,2%. Grund für die vergleichsweise schlechtere Trennung könnte hier die geographische Nähe dieser beiden Kontinente sein.

Auch für die Abgrenzung von Kaffeeproben aus Äthiopien, Brasilien und Kolumbien von Kaffeeproben aus anderen Kontinenten wurden zwei-Klassen-Modelle erstellt. Kaffeeproben aus Brasilien und Äthiopien konnten sehr gut von Kaffeeproben aus einem anderen Kontinent abgegrenzt werden. Die Klassifizierungsmodelle basierend auf der Analyse der hydrophilen Kaffeemetaboliten zeigten geringfügig bessere Klassifizierungsgenauigkeiten als die Modelle basierend auf der Analyse der lipophilen Metaboliten. Für die Klassifizierungsmodelle zur Abgrenzung von Kaffeeproben aus Äthiopien von Kaffeeproben aus anderen Kontinenten wurden durchschnittlich 97,8% der Testproben der richtigen Klasse zugeordnet. In den Klassifizierungsmodellen zur Abgrenzung von Kaffeeproben aus Brasilien von Kaffee aus anderen Kontinenten waren dies 95,6%. Allerdings konnte mit der Analyse des vorhandenen Probandensatzes keine vollständige Abgrenzung von Kaffeeproben aus Kolumbien erreicht werden. Durch die Analyse der Loadingsplots der Klassifizierungsmodelle zur Herkunftsbestimmung von Kaffee aus Äthiopien und Brasilien konnten keine Metaboliten identifiziert werden, welche die Trennung beeinflussten. In den spektralen Bereichen der Variablen, welche in den Loadingsplots einen Einfluss auf die Abgrenzung der Klassen andeuteten, traten überwiegend Signale mit geringer Intensität auf, die überlagert vorlagen. Dies erschwerte die Identifizierung dieser Signale. Durch eine Low-Level und eine Mid-Level Datenfusionierung wurden die $^1\text{H-NMR}$ -Daten der lipophilen und hydrophilen Metaboliten kombiniert. Sowohl durch die Low-Level, als auch durch die Mid-Level Datenfusionierung konnte keine bessere Zuordnung der Testproben in der internen Validierung erreicht werden.

Ebenfalls wurden Klassifizierungsmodelle zur Abgrenzung von Kaffeeproben nach der Anbaumethode und des verwendeten Röstverfahrens erstellt und validiert. Mittels PCA-LDA konnte biologisch erzeugter Kaffee nicht von konventionell erzeugtem Kaffee abgetrennt werden. Dahingegen basierend auf der Analyse der lipophilen bzw. hydrophilen Metaboliten eine Abgrenzung von trommelgeröstetem und heißluftgeröstetem Kaffee erreicht werden. In der internen und externen Validierung der Klassifizierungsmodelle konnten mindesten 92,9% der Testproben der richtigen Klasse zugeordnet werden.

Im Rahmen des zielgerichteten Ansatzes wurde eine Methode zur Quantifizierung der hydrophilen Metaboliten Milchsäure, Essigsäure, Koffein, Caffeoylchinasäure-Isomere (CQA), Ameisensäure, N-Methylpyridinium (NMP), Trigonellin und 5-Hydroxymethylfurfural (HMF) entwickelt und erfolgreich validiert. Die Quantifizierung der CQA erfolgte als Summenparameter der Isomere 3-Caffeoylchinasäure, 4-Caffeoylchinasäure und 5-Caffeoylchinasäure. Ein Vergleich der bestimmten Gehalte mit denen einer entwickelten Referenzmethode, basierend auf Hochleistungsflüssigkeitschromatographie und Diodearray-Detektion (HPLC-DAD), zur Quantifizierung von Koffein, Trigonellin und CQA zeigte eine gute Übereinstimmung.

Somit konnten in dieser Arbeit Klassifizierungsmodelle zur Abgrenzung von Kaffeeproben aus Afrika, Äthiopien und Brasilien von Kaffeeproben aus anderen Kontinenten erstellt werden, welche eine gute Vorhersagefähigkeit aufwiesen. Weiter konnten mit den $^1\text{H-NMR}$ -Daten Klassifizierungsmodelle zur Abgrenzung von Kaffee nach dem Röstverfahren erstellt werden. Diese Methoden sind somit zur Überprüfung der Authentizität von Kaffee bezüglich dieser Eigenschaften geeignet.

6 Literaturverzeichnis

- [1] **Bundesministerium für Verbraucherschutz, Ernährung und Landwirtschaft (2001).** *Verordnung über Kaffee, Kaffee- und Zichorien-Extrakte. KaffeeV 2001.*
- [2] **Belitz, H. D., Grosch, W. & Schieberle, P. (2007).** *Lehrbuch der Lebensmittelchemie.* Springer-Lehrbuch. Springer Berlin Heidelberg, Berlin, Heidelberg.
- [3] **Souza, R. M. (2008).** *Plant-parasitic nematodes of coffee.* Springer, Dordrecht.
- [4] **Waller, J. M., Bigger, M. & Hillocks, R. J. (2007).** *Coffee pests, diseases & their management.* Biddles Ltd., Norfol.
- [5] **Folmer, B. (2017).** *The craft and science of coffee.* Elsevier, Amsterdam, Boston MA.
- [6] **Rotzoll, F. & Müller-Henniges, H.-G. (1997).** *Kaffee-Bibliothek.* Deutscher Kaffee-Verband e.V.
- [7] **Farah, A. & Ferreira dos Santos, T. (2015).** *Coffee. In health and disease prevention.* Elsevier, London.
- [8] **Runge, F. F. (1820).** *Neuste pytochemische Entdeckungen zur Begründung einer wissenschaftlichen Phytochemie.* Reimer, Berlin.
- [9] **International Coffee Organization (2021).** *World coffee consumption,* London.
- [10] **International Coffee Organization (2020).** *Total production by all exporting countries,* London.
- [11] **Adepoju, A. F., Adenuga, O. O., Mapayi, E. F., Olaniyi O. O. & Adepoju F. A. (2017).** Coffee: botany, distribution, diversity, chemical composition and its management. *Journal of Agriculture and Veterinary Science* **10**, 57–62.
- [12] **Mussatto, S. I., Machado, E. M. S., Martins, S. & Teixeira, J. A. (2011).** Production, composition, and application of coffee and its industrial residues. *Food and Bioprocess Technology* **4**, 661–672.
- [13] **Bacon, C. (2005).** Confronting the coffee crisis: can fair trade, organic, and specialty coffees reduce small-scale farmer vulnerability in northern Nicaragua? *World Development* **33**, 497–511.
- [14] **International Coffee Organization (2019).** *Prices paid to growers in exporting countries,* London.
- [15] **Wintgens, J. N. (2012).** *Coffee: growing, processing, sustainable production. A guide-book for growers, processors, traders and researchers.* WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.
- [16] **Hofmann, J. (2018).** *The world atlas of coffee. From beans to brewing - coffees explored, explained and enjoyed.* Firefly Books, New York.

- [17] **Anthony, F., Combes, M. C., Astorga, C., Bertrand, B., Graziosi, G. & Lashermes, P. (2002)**. The origin of cultivated *Coffea arabica* L. varieties revealed by AFLP and SSR markers. *Theoretical and Applied Genetics* **104**, 894–900.
- [18] **van der Vossen, H., Bertrand, B. & Charrier, A. (2015)**. Next generation variety development for sustainable production of arabica coffee (*Coffea arabica* L.): a review. *Euphytica* **204**, 243–256.
- [19] **Adhikari, M., Isaac, E. L., Paterson, R. R. M. & Maslin, M. A. (2020)**. A review of potential impacts of climate change on coffee cultivation and mycotoxigenic fungi. *Microorganisms* **8**, 1–12.
- [20] **Klingel, T., Kremer, J. I., Gottstein, V., Rajcic de Rezende, T., Schwarz, S. & Lachenmeier, D. W. (2020)**. A review of coffee by-products including leaf, flower, cherry, husk, silver skin, and spent grounds as novel foods within the european union. *Foods* **9**, 665.
- [21] **Bunn, C., Läderach, P., Ovalle Rivera, O. & Kirschke, D. (2015)**. A bitter cup: climate change profile of global production of Arabica and Robusta coffee. *Climatic Change* **129**, 89–101.
- [22] **Gonzalez-Perez, M.-A. & Gutierrez-Viana, S. (2012)**. Cooperation in coffee markets: the case of Vietnam and Colombia. *Journal of Agribusiness in Developing and Emerging Economies* **2**, 57–73.
- [23] **Perfecto, I., Vandermeer, J., Mas, A. & Pinto, L. S. (2005)**. Biodiversity, yield, and shade coffee certification. *Ecological Economics* **54**, 435–446.
- [24] **Coltro, L., Mourad, A., Oliveira, P., Baddini, J. & Kletecke, R. (2006)**. Environmental profile of brazilian green coffee. *Int J Life Cycle Assessment* **11**, 16–21.
- [25] **Kufa, T., Ashenafi, A., Yilma, A., Kumela, T. & Tefera, W. (2011)**. The contribution of coffee research for coffee seed development in Ethiopia. *Journal of Agricultural Research and Development* **1**, 9–16.
- [26] **Beer, J., Muschler, R., Kass, D. & Somarriba, E. (1997)**. Shade management in coffee and cacao plantations. *Agroforestry Systems* **38**, 139–164.
- [27] **Bote, A. D. & Struik, P. C. (2011)**. Effects of shade growht, production and quality of coffee (*Coffea Arabica*) in Ethiopia. *Journal of Horticulture and Forestry* **3**, 336–341.
- [28] **Mutersbaugh, T. (2002)**. The number is the neast: a political economy of organic-coffee certification and producer unionism. *Environment and Planning A: Economy and Space* **34**, 1165–1184.
- [29] **Monteiro, M. C. & Farah, A. (2012)**. Chlorogenic acids in Brazilian *Coffea Arabica* cultivars from various consecutive crops. *Food Chemistry* **134**, 611–614.
- [30] **Baca, M., Läderach, P., Hagggar, J., Schroth, G. & Ovalle, O. (2014)**. An integrated framework for assessing vulnerability to climate change and developing adaptation strategies for coffee growing families in Mesoamerica. *Plos One* **9**, e88463.

- [31] **Zullo, J., Pinto, H. S., Assad, E. D. & Heuminski de Ávila, A. M. (2011)**. Potential for growing Arabica coffee in the extreme south of Brazil in a warmer world. *Climatic Change* **109**, 535–548.
- [32] **Europäisches Parlament und der Rat der europäischen Union (2018)**. *Verordnung (EU) 2018/848 des Europäischen Parlaments und des Rates vom 30. Mai 2018 über die ökologische/biologische Produktion und die Kennzeichnung von ökologischen/biologischen Erzeugnissen sowie zur Aufhebung der Verordnung (EG) Nr. 834/2007 des Rates. VO(EU) 2018/848*.
- [33] **Vincent, J. C. (1978)**. *Green coffee processing*. Springer, Dordrecht.
- [34] **Bytof, G., Knopp, S. E., Schieberle, P., Teutsch, I. & Selmar, D. (2005)**. Influence of processing on the generation of γ -aminobutyric acid in green coffee beans. *European Food Research and Technology* **220**, 245–250.
- [35] **Selmar, D., Kleinwächter, M. & Bytof, G. (2015)**. *Cocoa and coffee fermentations. Chapter 12: Metabolic responses of coffee beans during processing and their impact on coffee flavor*. CRC Press, Boca Raton.
- [36] **Joët, T., Laffargue, A., Descroix, F., Doubeau, S., Bertrand, B., Kochko, A. de & Dussert, S. (2010)**. Influence of environmental factors, wet processing and their interactions on the biochemical composition of green Arabica coffee beans. *Food Chemistry* **118**, 693–701.
- [37] **Clarke, R. J. & Vitzthum, O. G. (2001)**. *Coffee. Recent developments*. Blackwell Science, Oxford, Malden MA.
- [38] **Spiller, G. A. (1998)**. *Caffeine*. CRC Press LLC, Los Altos.
- [39] **Silva, C. F., Vilela, D. M., Souza C., Duarte, W. F., Dias, D. R. & Schwan, R. F. (2013)**. Evaluation of a potential starter culture for enhance quality of coffee fermentation. *World journal of microbiology & biotechnology* **29**, 235–247.
- [40] **Ledl, F. & Schleicher, E. (1990)**. New aspects of the maillard reaction in foods and in the human body. *Angewandte Chemie International Edition in English* **29**, 565–594.
- [41] **Gniechwitz, D., Reichardt, N., Ralph, J., Blaut, M., Steinhart, H. & Bunzel, M. (2008)**. Isolation and characterisation of a coffee melanoidin fraction. *Journal of the Science of Food and Agriculture* **88**, 2153–2160.
- [42] **Namiki, M. (1988)**. Chemistry of Maillard reactions: recent studies on the browning reaction mechanism and the development of antioxidants and mutagens. *Advanced Food Research* **32**, 115–184.
- [43] **Hodge, J. E. (1953)**. Dehydrated foods, chemistry of browning reactions in model systems. *Journal of Agricultural and Food Chemistry* **46**, 2599–2600.
- [44] **Martins, S. I. F. S., Jongen, W. M. F. & van Boekel, M. A. . S. (2000)**. A review of Maillard reaction in food and implications to kinetic modelling. *Trends in Food Science & Technology* **11**, 364–373.

- [45] **Ginz, M., Balzer, H. H., Bradbury, A. G. W. & Maier, H. G. (2000).** Formation of aliphatic acids by carbohydrate degradation during roasting of coffee. *European Food Research and Technology* **211**, 404–410.
- [46] **Spreng, S., Schaerer, A., Poisson, L., Chaumonteuil, M., Mestdagh, F. & Davidek, T. (2021).** Discovery of polyhydroxyalkyl pyrazine generation upon coffee roasting by in-bean labeling experiments. *Journal of Agricultural and Food Chemistry* **69**, 6636–6649.
- [47] **Moreira, A. S. P., Nunes, F. M., Domingues, M. R. & Coimbra, M. A. (2012).** Coffee melanoidins: structures, mechanisms of formation and potential health impacts. *Food & Function* **3**, 903–915.
- [48] **Bekedam, E. K., Schols, H. A., van Boekel, M. A. J. S. & Smit, G. (2006).** High molecular weight melanoidins from coffee brew. *Journal of Agricultural and Food Chemistry* **54**, 7658–7666.
- [49] **Bekedam, E. K., Schols, H. A., van Boekel, M. A. J. S. & Smit, G. (2008).** Incorporation of chlorogenic acids in coffee brew melanoidins. *Journal of Agricultural and Food Chemistry* **56**, 2055–2063.
- [50] **Bekedam, E. K., Laat, M. P. F. C. de, Schols, H. A., van Boekel, M. A. J. S. & Smit, G. (2007).** Arabinogalactan proteins are incorporated in negatively charged coffee brew melanoidins. *Journal of Agricultural and Food Chemistry* **55**, 761–768.
- [51] **Bekedam, E. K., Loots, M. J., Schols, H. A., van Boekel, M. A. J. S. & Smit, G. (2008).** Roasting effects on formation mechanisms of coffee brew melanoidins. *Journal of Agricultural and Food Chemistry* **56**, 7138–7145.
- [52] **Oosterveld, A., Harmsen, J. S., Voragen, A. & Schols, H. A. (2003).** Extraction and characterization of polysaccharides from green and roasted *Coffea arabica* beans. *Carbohydrate Polymers* **52**, 285–296.
- [53] **Redgwell, R. J., Trovato, V., Curti, D. & Fischer, M. (2002).** Effect of roasting on degradation and structural features of polysaccharides in arabica coffee beans. *Carbohydrate Research* **337**, 421–431.
- [54] **Nunes, F. M., Cruz, A. C. S. & Coimbra, M. A. (2012).** Insight into the mechanism of coffee melanoidin formation using modified in bean models. *Journal of Agricultural and Food Chemistry* **60**, 8710–8719.
- [55] **Maria, C. A. B. de, Trugo, L. C., Aquino Neto, F. R., Moreira, R. F. A. & Alviano, C. S. (1996).** Composition of green coffee water-soluble fractions and identification of volatiles formed during roasting. *Food Chemistry* **55**, 203–207.
- [56] **Nunes, F. M. & Coimbra, M. A. (2002).** Chemical characterization of the high-molecular-weight material extracted with hot water from green and roasted robusta coffees as affected by the degree of roast. *Journal of Agricultural and Food Chemistry* **50**, 7046–7052.

- [57] **Redgwell, R. J., Curti, D., Fischer, M., Nicolas, P. & Fay, L. B. (2002)**. Coffee bean arabinogalactans: acidic polymers covalently linked to protein. *Carbohydrate Research* **337**, 239–253.
- [58] **Redgwell, R. & Fischer, M. (2006)**. Coffee carbohydrates. *Brazilian Journal of Plant Physiology* **18**, 165–174.
- [59] **Nunes, F. M., Coimbra, M. A., Duarte, A. C. & Delgadillo, I. (1997)**. Foamability, foam stability, and chemical composition of espresso coffee as affected by the degree of roast. *Journal of Agricultural and Food Chemistry* **45**, 3238–3243.
- [60] **Redgwell, R. J. & Fischer, M. (2006)**. Coffee carbohydrates. *Brazilian Journal of Plant Physiology* **18**, 165–174.
- [61] **IARC Working Group (2019)**. IARC Working Group on the Evaluation of Carcinogenic Risks to Humans. Furfuryl alcohol. *Monogr. Eval. Carcinog. Risks Hum* **119**, 83–113.
- [62] **IARC Working Group (1994)**. IARC working group on the evaluation of carcinogenic risks to humans. Acrylamide. *IARC Monogr. Eval. Carcinog. Risks Hum.*, 389–433.
- [63] **National Toxicology Program (2010)**. NTP technical report on the toxicology and carcinogenesis studies of 5-(hydroxymethyl)-2-furfural (CAS No. 67-47-0) in F344/N rats and B6C3F1 mice (gavage studies). *National Toxicological Program Technical Report Series* **554**, 7-13; 15-9; 21-31.
- [64] **CIAMPA, A., RENZI, G., TAGLIANTI, A., SEQUI, P. & VALENTINI, M. (2010)**. Studies on coffee roasting process by means of nuclear magnetic resonance spectroscopy. *Journal of Food Quality* **33**, 199–211.
- [65] **Endeshawa, H. & Belay, A. (2020)**. Optimization of the roasting conditions to lower acrylamide content and improve the nutrient composition and antioxidant properties of *Coffea arabica*. *Plos One* **15**, e0237265.
- [66] **Lachenmeier, D. W., Schwarz, S., Teipel, J., Hegmanns, M., Kuballa, T., Walch, S. G. & Breitling-Utzmann, C. M. (2018)**. Potential antagonistic effects of acrylamide mitigation during coffee roasting on furfuryl alcohol, furan and 5-hydroxymethylfurfural. *Toxics* **7**, 1.
- [67] **Farah, A., Paulis, T. de, Trugo, L. C. & Martin, P. R. (2005)**. Effect of roasting on the formation of chlorogenic acid lactones in coffee. *Journal of Agricultural and Food Chemistry* **53**, 1505–1513.
- [68] **Bosco, M., Toffanin, R., Palo, D. de, Zatti, L. & Segre, A. (1999)**. High-resolution ¹H NMR investigation of coffee. *Journal of the Science of Food and Agriculture* **79**, 869–878.
- [69] **Frank, O., Blumberg, S., Kunert, C., Zehentbauer, G. & Hofmann, T. (2007)**. Structure determination and sensory analysis of bitter-tasting 4-vinylcatechol oligomers and their identification in roasted coffee by means of LC-MS/MS. *Journal of Agricultural and Food Chemistry* **55**, 1945–1954.

- [70] **Nunes, F. M. & Coimbra, M. A. (2010)**. Role of hydroxycinnamates in coffee melanoidin formation. *Phytochemistry Reviews* **9**, 171–185.
- [71] **Müller, C., Lang, R. & Hofmann, T. (2006)**. Quantitative precursor studies on di- and trihydroxybenzene formation during coffee roasting using "in bean" model experiments and stable isotope dilution analysis. *Journal of Agricultural and Food Chemistry* **54**, 10086–10091.
- [72] **Febvay, L., Hamon, E., Recht, R., Andres, N., Vincent, M., Aoudé-Werner, D. & This, H. (2019)**. Identification of markers of thermal processing ("roasting") in aqueous extracts of *Coffea Arabica* L. seeds through NMR fingerprinting and chemometrics. *Magnetic Resonance in Chemistry* **57**, 589–602.
- [73] **Wei, F., Furihata, K., Hu, F., Miyakawa, T. & Tanokurra, T. (2011)**. Two-dimensional ¹H-¹³C nuclear magnetic resonance (NMR)-based comprehensive analysis of roasted coffee bean extract. *Journal of Agricultural and Food Chemistry* **59**, 9065–9073.
- [74] **Casal, S., Mendes, E., Oliveira, M. & Ferreira, M. A. (2005)**. Roast effects on coffee amino acid enantiomers. *Food Chemistry* **89**, 333–340.
- [75] **Clifford, M. N. & Willson, K. C. (1985)**. *Coffee: botany, biochemistry and production of beans an beverage. Chapter 13: Chemical and physical aspects of green coffee and coffee products*. Springer-Lehrbuch. Springer, London.
- [76] **Casal, S., Beatriz Oliveira, M. & Ferreira, M. A. (2000)**. HPLC/diode-array applied to the thermal degradation of trigonelline, nicotinic acid and caffeine in coffee. *Food Chemistry* **68**, 481–485.
- [77] **Stadler, R. H., Varga, N., Hau, J., Vera, F. A. & Welti, D. H. (2002)**. Alkylpyridiniums. 1. Formation in model systems via thermal degradation of trigonelline. *Journal of Agricultural and Food Chemistry* **50**, 1192–1199.
- [78] **Wei, F., Furihata, K., Koda, M., Hu, F., Miyakawa, T. & Tanokura, M. (2012)**. Roasting process of coffee beans as studied by nuclear magnetic resonance: time course of changes in composition. *Journal of Agricultural and Food Chemistry* **60**, 1005–1012.
- [79] **Alves, R. P., Antoniosi Filho, N. R., Lião, L. M. & Flores, I. S. (2021)**. Evaluation of the metabolic profile of arabica coffee via NMR in relation to the time and temperature of the roasting procedure. *Journal of the Brazilian Chemical Society* **31**, 123–136.
- [80] **Speer, K. & Kölling-Speer, I. (2006)**. The lipid fraction of the coffee bean. *Brazilian Journal of Plant Physiology* **18**, 201–216.
- [81] **Zhou, L., Khalil, A., Bindler, F., Zhao, M., Marcic, C., Ennahar, S. & Marchioni, E. (2013)**. Effect of heat treatment on the content of individual phospholipids in coffee beans. *Food Chemistry* **141**, 3846–3850.
- [82] **Geiger, R., Perren, R., Kuenzli, R. & Escher, F. (2005)**. Carbon dioxide evolution and moisture evaporation during roasting of coffee beans. *Journal of Food Science* **70**, E124-E130.

- [83] **Fabbri, A., Cevoli, C., Alessandrini, L. & Romani, S. (2011)**. Numerical modeling of heat and mass transfer during coffee roasting process. *Journal of Food Engineering* **105**, 264–269.
- [84] **Bolka, M. & Emire, S. (2020)**. Effects of coffee roasting technologies on cup quality and bioactive compounds of specialty coffee beans. *Food Science and Nutrition* **8**, 6120–6130.
- [85] **Cadena, A. (2019)**. *A heated comparison: Ikawa vs Probat*.
- [86] **Kurzrock, T. & Speer, K. (2001)**. Diterpenes and diterpene esters in coffee. *Food Reviews International* **17**, 433–450.
- [87] **Nikolova-Damyanova, B., Velikova, R. & Jham, G. N. (1998)**. Lipid classes, fatty acid composition and triacylglycerol molecular species in crude coffee beans harvested in Brazil. *Food Research International* **31**, 479–486.
- [88] **Burton, I. W., Martinez Farina, C. F., Ragupathy, S., Arunachalam, T., Newmaster, S. & Berru , F. (2020)**. Quantitative NMR methodology for the authentication of roasted coffee and prediction of blends. *Journal of Agricultural and Food Chemistry* **68**, 14643–14651.
- [89] **Monakhova, Y. B., Ruge, W., Kuballa, T., Ilse, M., Winkelmann, O., Diehl, B., Thomas, F. & Lachenmeier, D. W. (2015)**. Rapid approach to identify the presence of arabica and robusta species in coffee using ¹H NMR spectroscopy. *Food Chemistry* **182**, 178–184.
- [90] **Schievano, E., Finotello, C., Angelis, E. de, Mammi, S. & Navarini, L. (2014)**. Rapid authentication of coffee blends and quantification of 16-O-methylcafestol in roasted coffee beans by nuclear magnetic resonance. *Journal of Agricultural and Food Chemistry* **62**, 12309–12314.
- [91] **Sridevi, V. & Giridhar, P. (2016)**. Variations in diterpenes-cafestol and kahweol content in beans of robusta coffee grown at different altitudes. *Proceedings of the National Academy of Sciences, India Section B: Biological Sciences* **86**, 291–297.
- [92] **Scharnhop, H. & Winterhalter, P. (2009)**. Isolation of coffee diterpenes by means of high-speed countercurrent chromatography. *Journal of Food Composition and Analysis* **22**, 233–237.
- [93] **Gunning, Y., Defernez, M., Watson, A. D., Beadman, N., Colquhoun, I. J., Le Gall, G., Philo, M., Garwood, H., Williamson, D., Davis, A. P. & Kemsley, E. K. (2018)**. 16-O-methylcafestol is present in ground roast arabica coffees: implications for authenticity testing. *Food Chemistry* **248**, 52–60.
- [94] **Guercia, E., Colomban, S. & Navarini, L. (2020)**. 16-O-Methylated diterpenes in green *Coffea arabica*: ultra-high-performance liquid chromatography-tandem mass spectrometry method optimization and validation. *Journal of Mass Spectrometry* **55**, e4636.

- [95] **Ahsan F. & Bashir S. (2019)**. Coffee consumption: health perspectives and drawbacks. *Journal of Nutrition and Obesity* **2**, 1–4.
- [96] **Nunes, F. M. & Coimbra, M. A. (2002)**. Chemical characterization of galactomannans and arabinogalactans from two arabica coffee infusions as affected by the degree of roast. *Journal of Agricultural and Food Chemistry* **50**, 1429–1434.
- [97] **Hiroshi A. & Takeo S. (2004)**. Distribution and biosynthesis of caffeine in plants. *Frontiers in Bioscience* **9**, 1864–1876.
- [98] **Pourshahidi, L. K., Navarini, L., Petracco, M. & Strain, J. J. (2016)**. A comprehensive overview of the risks and benefits of coffee consumption. *Comprehensive Reviews in Food Science and Food Safety* **15**, 671–684.
- [99] **Ashihara, H. (2006)**. Metabolism of alkaloids in coffee plants. *Brazilian Journal of Plant Physiology* **18**, 1–8.
- [100] **Zheng, X.-Q. & Ashihara, H. (2004)**. Distribution, biosynthesis and function of purine and pyridine alkaloids in *Coffea arabica* seedlings. *Plant Science* **166**, 807–813.
- [101] **Lire W. H. (2017)**. Review on health benefit and risk of coffee consumption. *Medicinal & Aromatic Plants* **06**, 1000301.
- [102] **Ashihara, H. & Crozier, A. (2001)**. Caffeine: a well known but little mentioned compound in plant science. *Trends in Plant Science* **6**, 407–413.
- [103] **Mohamadi, N., Sharififar, F., Pournamdari, M. & Ansari, M. (2018)**. A review on biosynthesis, analytical techniques, and pharmacological activities of trigonelline as a plant alkaloid. *Journal of Dietary Supplements* **15**, 207–222.
- [104] **Ashihara, H., Ludwig, I. A., Katahira, R., Yokota, T., Fujimura, T. & Crozier, A. (2015)**. Trigonelline and related nicotinic acid metabolites: occurrence, biosynthesis, taxonomic considerations, and their roles in planta and in human health. *Phytochemistry Reviews* **14**, 765–798.
- [105] **Riedel, A., Hochkogler, C. M., Lang, R., Bytof, G., Lantz, I., Hofmann, T. & Somoza, V. (2014)**. N-methylpyridinium, a degradation product of trigonelline upon coffee roasting, stimulates respiratory activity and promotes glucose utilization in HepG2 cells. *Food & Function* **5**, 454–462.
- [106] **Nunes, F. M., Reis, A., Domingues, M. R. M. & Coimbra, M. A. (2006)**. Characterization of galactomannan derivatives in roasted coffee beverages. *Journal of Agricultural and Food Chemistry* **54**, 3428–3439.
- [107] **Gniechwitz, D., Brueckel, B., Reichardt, N., Blaut, M., Steinhart, H. & Bunzel, M. (2007)**. Coffee dietary fiber contents and structural characteristics as influenced by coffee type and technological and brewing procedures. *Journal of Agricultural and Food Chemistry* **55**, 11027–11034.

- [108] **Jham, G. N., Fernandes, S. A., Garcia, C. F. & Da Silva, A. A. (2002)**. Comparison of GC and HPLC for the quantification of organic acids in coffee. *Phytochemical Analysis* **13**, 99–104.
- [109] **Koshiro, Y., Jackson, M. C., Jackson, C. & Ashihara, H. (2015)**. Changes in the content of sugars and organic acids during ripening of *Coffea Arabica* and *Coffea Canephora* fruits. *European Chemical Bulletin* **4**, 378–383.
- [110] **Clifford, M. N., Jaganath, I. B., Ludwig, I. A. & Crozier, A. (2017)**. Chlorogenic acids and the acyl-quinic acids: discovery, biosynthesis, bioavailability and bioactivity. *Natural product reports* **34**, 1391–1421.
- [111] **Koshiro, Y., Jackson, M. C., Katahira, R., Wang, M.-L., Nagai, C. & Ashihara, H. (2007)**. Biosynthesis of chlorogenic acids in growing and ripening fruits of *Coffea arabica* and *Coffea canephora* plants. *Journal of Biosciences* **62**, 731–742.
- [112] **Kamiyama, M., Moon, J.-K., Jang, H. W. & Shibamoto, T. (2015)**. Role of degradation products of chlorogenic acid in the antioxidant activity of roasted coffee. *Journal of Agricultural and Food Chemistry* **63**, 1996–2005.
- [113] **Gogoasa, I., Pirvu, A., Alda, L. M., Velciov, A., Rada, M., Bordean Despina, M., Moigradean, D., Simion, A. & Gergen, I. (2013)**. The mineral content of different brands of coffee. *Journal of Horticulture, Forestry and Biotechnology* **17**, 68–71.
- [114] **Schwarz, S.** *Defekte bei Rohkaffees*. Crema Magazin, München.
- [115] **International Organization for Standardization (2004)**. *Green coffee - defect reference chart*. ISO10470.
- [116] **International Coffee Organization (2006)**. *Quality: Harmonization of ICC resolution number 420 with ISO green coffee defects chart*. International Coffee Organization, London.
- [117] **Santos, J. R., Sarraguça, M. C., Rangel, A. O. S. S. & Lopes, J. A. (2012)**. Evaluation of green coffee beans quality using near infrared spectroscopy: a quantitative approach. *Food Chemistry* **135**, 1828–1835.
- [118] **Franca, A. S., Oliveira, L. S., Mendonça, J. C. F. & Silva, X. A. (2005)**. Physical and chemical attributes of defective crude and roasted coffee beans. *Food Chemistry* **90**, 89–94.
- [119] **Spyros, A. (2012)**. *NMR spectroscopy in food analysis*. ISSN. Royal Society of Chemistry, Cambridge.
- [120] **Friebolin, H. (2013)**. *Ein- und zweidimensionale NMR-Spektroskopie. Eine Einführung*. WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim.
- [121] **Claridge, T. D. W. (2016)**. *High-resolution NMR techniques in organic chemistry*. Elsevier, Amsterdam, Boston, Heidelberg London, New York, Oxford, Paris.

- [122] **Pauli, G. F., Jaki, B. U. & Lankin, D. C. (2005)**. Quantitative ¹H NMR: development and potential of a method for natural products analysis. *Journal of Natural Products* **68**, 133–149.
- [123] **Bharti, S. K. & Roy, R. (2012)**. Quantitative ¹H NMR spectroscopy. *Trends in Analytical Chemistry* **35**, 5–26.
- [124] **Giraudeau, P. (2017)**. Challenges and perspectives in quantitative NMR. *Magnetic Resonance in Chemistry* **55**, 61–69.
- [125] **Holzgrabe, U. (2010)**. Quantitative NMR spectroscopy in pharmaceutical applications. *Progress in Nuclear Magnetic Resonance Spectroscopy* **57**, 229–240.
- [126] **Wider, G. & Dreier, L. (2006)**. Measuring protein concentrations by NMR spectroscopy. *Journal of the American Chemical Society* **128**, 2571–2576.
- [127] **Monakhova, Y. B., Lachenmeier, D. W., Kuballa, T. & Mushtakova, S. P. (2015)**. Standardless multicomponent qNMR analysis of compounds with overlapped resonances based on the combination of ICA and PULCON. *Magnetic Resonance in Chemistry* **53**, 821–828.
- [128] **Williamson, K. & Hatzakis, E. (2019)**. NMR analysis of roasted coffee lipids and development of a spent ground coffee application for the production of bioplastic precursors. *Food Research International* **119**, 683–692.
- [129] **Pontes, J. G. M., Brasil, A. J. M., Cruz, G. C. F., Souza, R. N. de & Tasic, L. (2017)**. NMR-based metabolomics strategies: plants, animals and humans. *Analytical Methods* **9**, 1078–1096.
- [130] **Vignoli, A., Ghini, V., Meoni, G., Licari, C., Takis, P. G., Tenori, L., Turano, P. & Luchinat, C. (2019)**. High-throughput metabolomics by 1D NMR. *Angewandte Chemie* **131**, 980–1007.
- [131] **Cevallos-Cevallos, J. M., Reyes-De-Corcuera, J. I., Etxeberria, E., Danyluk, M. D. & Rodrick, G. E. (2009)**. Metabolomic analysis in food science: a review. *Trends in Food Science & Technology* **20**, 557–566.
- [132] **Fraga-Corral, M., Carpena, M., Garcia-Oliveira, P., Pereira, A. G., Prieto, M. A. & Simal-Gandara, J. (2020)**. Analytical metabolomics and applications in health, environmental and food science. *Critical Reviews in Analytical Chemistry*, 1–23.
- [133] **Kim, H. K., Choi, Y. H. & Verpoorte, R. (2010)**. NMR-based metabolomic analysis of plants. *Nature Protocols* **5**, 536–549.
- [134] **Kim, H. & Verpoorte, R. (2010)**. Sample preparation for plant metabolomics. *Phytochemical Analysis* **21**, 4–13.
- [135] **Verpoorte, R., Choi, Y. H., Mustafa, N. R. & Kim, H. K. (2008)**. Metabolomics: back to basics. *Phytochemistry Reviews* **7**, 525–537.

- [136] Legat, M. Legat, B., Leenders, J., Vanwinsberghe, J., Rousseau, R., Boulanger, B., Eilers, P., Tullio, P. de & Govaerts, B. (2018). PepsNMR for 1H NMR metabolomic data pre-processing. *Analytica Chimica Acta* **1019**, 1–13.
- [137] Anderson, K. A. & Smith, B. W. (2002). Chemical profiling to differentiate geographic growing origins of coffee. *Journal of Agricultural and Food Chemistry* **50**, 2068–2075.
- [138] Arana, V. A., Medina, J., Alarcon, R., Moreno, E., Heintz, L., Schäfer, H. & Wist, J. (2015). Coffee's country of origin determined by NMR: the Colombian case. *Food Chemistry* **175**, 500–506.
- [139] Arana, V. A., Medina, J., Esseiva, P., Pazos, D. & Wist, J. (2016). Classification of coffee beans by GC-C-IRMS, GC-MS, and 1H-NMR. *Journal of Analytical Methods in Chemistry* **2016**, 8564584.
- [140] Badmos, S., Fu, M., Granato, D. & Kuhnert, N. (2020). Classification of Brazilian roasted coffees from different geographical origins and farming practices based on chlorogenic acid profiles. *Food Research International* **134**, 109218.
- [141] Barbosa, R. M., Batista, B. L., Varrique, R. M., Coelho, V. A., Campiglia, A. D. & Barbosa, F. (2014). The use of advanced chemometric techniques and trace element levels for controlling the authenticity of organic coffee. *Food Research International* **61**, 246–251.
- [142] Bitter, N. Q., Fernandez, D. P., Driscoll, A. W., Howa, J. D. & Ehleringer, J. R. (2020). Distinguishing the region-of-origin of roasted coffee beans with trace element ratios. *Food Chemistry* **320**, 126602.
- [143] Choi, M.-Y., Choi, W., Park, J. H., Lim, J. & Kwon, S. W. (2010). Determination of coffee origins by integrated metabolomic approach of combining multiple analytical data. *Food Chemistry* **121**, 1260–1268.
- [144] Consonni, R., Polla, D. & Cagliani, L. R. (2018). Organic and conventional coffee differentiation by NMR spectroscopy. *Food Control* **94**, 284–288.
- [145] Consonni, R., Cagliani, L. R. & Cogliati, C. (2012). NMR based geographical characterization of roasted coffee. *Talanta* **88**, 420–426.
- [146] dos Santos, M. H., Alvarenga, Y. A. & Boffo, E. F. (2020). 1H NMR metabolic fingerprinting of Chapada Diamantina/Bahia (Brazil) coffees as a tool to assessing their qualities. *Microchemical Journal* **152**, 104293.
- [147] Freitas, A. M. & Mosca, A. I. (1999). Coffee geographic origin - an aid to coffee differentiation. *Food Research International* **32**, 565–573.
- [148] Putri, S. P., Irifune, T., Yusianto & Fukusaki, E. (2019). GC/MS based metabolite profiling of Indonesian specialty coffee from different species and geographical origin. *Metabolomics* **15**, 126.

- [149] **Thorburn Burns, D., Tweed, L. & Walker, M. J. (2017)**. Ground roast coffee: review of analytical strategies to estimate geographic origin, species authenticity and adulteration by dilution. *Food Analytical Methods* **10**, 2302–2310.
- [150] **Toci, A. T., Moura Ribeiro, M. V. de, Toledo, P. R. A. B. de, Boralle, N., Pezza, H. R. & Pezza, L. (2018)**. Fingerprint and authenticity roasted coffees by ¹H-NMR: the Brazilian coffee case. *Food Science and Biotechnology* **27**, 19–26.
- [151] **Marko Luther und Paul Holleis GbR**. *Artisan+*. <https://shop.artisan.plus/de/hardware/tonino>. Accessed 3 May 2022.
- [152] **Verpoorte, R., Choi, Y. H. & Kim, H. K. (2007)**. NMR-based metabolomics at work in phytochemistry. *Phytochemistry Reviews* **6**, 3–14.
- [153] **Bligh, E. G. & Dyer, W. J. (1959)**. A rapid method of total lipid extraction and purification. *Canadian Journal of Biochemistry and Physiology* **37**, 911–917.
- [154] **Folch, J., Lees, M. & Stanley, G. H. (1957)**. A simple method for the isolation and purification of total lipides from animal tissue. *Journal of Biological Chemistry* **226**, 497–509.
- [155] **Lin, J.-H., Liu, L.-Y., Yang, M.-H. & Lee, M.-H. (2004)**. Ethyl acetate/ethyl alcohol mixtures as an alternative to folch reagent for extracting animal lipids. *Journal of Agricultural and Food Chemistry* **52**, 4984–4986.
- [156] **Ackermann, S. M., Lachenmeier, D. W., Kuballa, T., Schütz, B., Spraul, M. & Bunzel, M. (2019)**. NMR-based differentiation of conventionally from organically produced chicken eggs in Germany. *Magnetic Resonance in Chemistry* **57**, 579–588.
- [157] **Fishwick, M. J. & Wright, A. J. (1977)**. Comparison of methods for the extraction of plant lipids. *Phytochemistry* **16**, 1507–1510.
- [158] **Lakshi P. Bhuyan/Pradip Tamuly/Pradip K. Mahanta (1991)**. Lipid content and fatty acid composition of tea shoot and manufactured tea. *Journal of Agricultural and Food Chemistry* **39**, 1159–1182.
- [159] **Post, D. M., Layman, C. A., Arrington, D. A., Takimoto, Quattrochi, J. & Montaña, C. G. (2007)**. Getting to the fat of the matter: models, methods and assumptions for dealing with lipids in stable isotope analyses. *Oecologia* **152**, 179–189.
- [160] **Axelsson, M. & Gentili, F. (2014)**. A single-step method for rapid extraction of total lipids from green microalge. *Plos One* **9**, e89643.
- [161] **Teipel, J., Gottstein, V., Hölzle, E., Kaltenbach, K., Lachenmeier, D. & Kuballa (2022)**. An easy and reliable method for the mitigation of deuterated chloroform decomposition to stabilise susceptible NMR samples. *Chemistry* **4**, 776–785.
- [162] **Skiera, C., Steliopoulos, P., Kuballa, T., Holzgrabe, U. & Diehl, B. (2012)**. ¹H-NMR spectroscopy as a new tool in the assessment of the oxidative state in edible oils. *Journal of the American Oil Chemists' Society*, 1383–1391.

- [163] **D'Amelio, N., Fontanive, L., Uggeri, F., Suggi-Liverani, F. & Navarini, L. (2009).** NMR reinvestigation of the caffeine–chlorogenate complex in aqueous solution and in coffee brews. *Food Biophysics* **4**, 321–330.
- [164] **D'Amelio, N., Papamokos, G., Dreyer, J., Carloni, P. & Navarini, L. (2015).** NMR Studies of Hetero-Association of Caffeine with di-O-Caffeoylquinic Acid Isomers in Aqueous Solution. *Food Biophysics* **10**, 235–243.
- [165] **Vuckovic, D. (2012).** Current trends and challenges in sample preparation for global metabolomics using liquid chromatography-mass spectrometry. *Analytical and Bioanalytical Chemistry* **403**, 1523–1548.
- [166] **Smolinska, A., Blanchet, L., Buydens, L. M. C. & Wijmenga, S. S. (2012).** NMR and pattern recognition methods in metabolomics: from data acquisition to biomarker discovery: a review. *Analytica Chimica Acta* **750**, 82–97.
- [167] **Wishart, D. S. (2008).** Quantitative metabolomics using NMR. *Trends in Analytical Chemistry* **27**, 228–237.
- [168] **Xiao, C., Hao, F., Qin, X., Wang, Y. & Tang, H. (2009).** An optimized buffer system for NMR-based urinary metabolomics with effective pH control, chemical shift consistency and dilution minimization. *The Analyst* **134**, 916–925.
- [169] **De Moura Ribeiro M. V., Boralle, N., Redigolo Pezza, H., Pezza, L. & Toci, A. T. (2017).** Authenticity of roasted coffee using ¹H NMR spectroscopy. *Journal of Food Composition and Analysis* **57**, 24–30.
- [170] **Milani, M. I., Rossini, E. L., Catelani, T. A., Pezza, L., Toci, A. T. & Pezza, H. R. (2020).** Authentication of roasted and ground coffee samples containing multiple adulterants using NMR and a chemometric approach. *Food Control* **112**, 107104.
- [171] **Giraudeau, P., Silvestre, V. & Akoka, S. (2015).** Optimizing water suppression for quantitative NMR-based metabolomics: a tutorial review. *Metabolomics* **11**, 1041–1055.
- [172] **Zheng, G. & Price, W. S. (2010).** Solvent signal suppression in NMR. *Progress in Nuclear Magnetic Resonance Spectroscopy* **56**, 267–288.
- [173] **Price, W. S. (1999).** Water signal suppression in NMR spectroscopy. In *Annual Reports on NMR Spectroscopy Volume 38*. Annual Reports on NMR Spectroscopy. Elsevier, 289–354.
- [174] **Mckay, R. T. (2009).** Recent advances in solvent suppression for solution NMR: a practical reference. *Annual Reports on NMR Spectroscopy* **66**, 33–76.
- [175] **Salomonsen, T; H. M. Jensen; F. H. Larsen; S. B. Engelsen (2009).** The quantitative impact of water suppression on NMR spectra for compositional analysis of alginates. In *Magnetic mesonance in food science*, M. Guðjónsdóttir, P. S. Belton and G. A. Webb, Eds. Special Publications. Royal Society of Chemistry, Cambridge, 12–19.

- [176] **Mckay, R. T. (2011)**. How the 1D-NOESY suppresses solvent signal in metabonomics NMR spectroscopy: an examination of the pulse sequence components and evolution. *Concepts in Magnetic Resonance Part A* **38A**, 197–220.
- [177] **Augustijn, D., Groot, H. J. M. de & Alia, A. (2021)**. HR-MAS NMR applications in plant metabolomics. *Molecules* **26**, 931.
- [178] **Emwas, A.-H.-H., Saccenti, E., Gao, X., Mckay, R. T., Martins dos Santos, V. A. P., Roy, R. & Wishart, D. S. (2018)**. Recommended strategies for spectral processing and post-processing of 1D ¹H-NMR data of biofluids with a particular focus on urine. *Metabolomics* **14**, 31.
- [179] **Karaman, I. (2017)**. Preprocessing and pretreatment of metabolomics data for statistical analysis. *Advances in Experimental Medicine and Biology* **965**, 145–161.
- [180] **Vettukattil, R. (2015)**. Preprocessing of raw metabonomic data. *Methods in Molecular Biology* **1277**, 123–136.
- [181] **Yi, L., Dong, N., Yun, Y., Deng, B., Liu, S., Zhang, Y. & Liang, Y. (2014)**. Recent advances in chemometric methods for plant metabolomics: A review. *Biotechnology advances*.
- [182] **Yi, L., Dong, N., Yun, Y., Deng, B., Ren, D., Liu, S. & Liang, Y. (2016)**. Chemometric methods in data processing of mass spectrometry-based metabolomics: a review. *Analytica Chimica Acta* **914**, 17–34.
- [183] **Vu, T. N. & Laukens, K. (2013)**. Getting your peaks in line: a review of alignment methods for NMR spectral data. *Metabolites* **3**, 259–276.
- [184] **Steliopoulos, P. (2013)**. Validierung PCA-gestützter Analysemethoden zur Authentizitätskontrolle von Lebensmitteln. *Journal für Verbraucherschutz und Lebensmittelsicherheit* **8**, 71–77.
- [185] **Liland, K. H. (2011)**. Multivariate methods in metabolomics – from pre-processing to dimension reduction and statistical analysis. *Trends in Analytical Chemistry* **30**, 827–841.
- [186] **Euceda, L. R., Giskeødegård, G. F. & Bathen, T. F. (2015)**. Preprocessing of NMR metabolomics data. *Scandinavian Journal of Clinical and Laboratory Investigation* **75**, 193–203.
- [187] **Schripsema, J. (2010)**. Application of NMR in plant metabolomics: techniques, problems and prospects. *Phytochemical Analysis* **21**, 14–21.
- [188] **van den Berg, R. A., Hoefsloot, H. C. J., Westerhuis, J. A., Smilde, A. K. & van der Werf, M. J. (2006)**. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC Genomics* **7**, 142.
- [189] **Wegner, A. (2020)**. Computational analysis of metabolic data. In *Metabolomics for Biomedical Research*. Elsevier, 83–103.

- [190] **Saccanti, E., Hoefsloot, H. C. J., Smilde, A. K., Westerhuis, J. A. & Hendriks, M. M. W. B. (2014)**. Reflections on univariate and multivariate analysis of metabolomics data. *Metabolomics* **10**, 361–374.
- [191] **Vinaixa, M., Samino, S., Saez, I., Duran, J., Guinovart, J. J. & Yanes, O. (2012)**. A Guideline to Univariate Statistical Analysis for LC/MS-Based Untargeted Metabolomics-Derived Data. *Metabolites* **2**, 775–795.
- [192] **Feir-Walsh, B. J. (1947)**. An empirical comparison of the ANOVA F-test, normal scores test and Kruskal-Wallis test under violation of assumptions. *Educational and Psychological Measurement* **37**, 789–799.
- [193] **Kruskal, W. H. & Wallis W. A. (1952)**. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association* **47**, 583–621.
- [194] **Nelson, R. (2015)**. Bonferroni and Sidák corrections for multiplicity effects with sub-total scores and comparison question polygraph tests. *Polygraph* **44**, 162–167.
- [195] **Bender, R., Lange, S. & Ziegler, A. (2007)**. Multiples Testen. *Deutsche Medizinische Wochenschrift* **132 Suppl 1**, e26-9.
- [196] **Bowling, F. G. & Thomas, M. (2014)**. Analyzing the metabolome. *Methods in Molecular Biology* **1168**, 31–45.
- [197] **Savorani, F., Rasmussen, M. A., Mikkelsen, M. S. & Engelsen, S. B. (2013)**. A primer to nutritional metabolomics by NMR spectroscopy and chemometrics. *Food Research International* **54**, 1131–1145.
- [198] **Kessler, W. (2006)**. *Multivariate Datenanalyse für die Pharma-, Bio- und Prozessanalytik. Ein Lehrbuch*. WILEY-VCH, Weinheim.
- [199] **Bro, R. & Smilde, A. K. (2014)**. Principal component analysis. *Analytical Methods* **6**, 2812–2831.
- [200] **Daszykowski, M., Kaczmarek, K., Vander Heyden, Y. & Walczak, B. (2007)**. Robust statistics in data analysis - a review. *Chemometrics and Intelligent Laboratory Systems* **85**, 203–219.
- [201] **Daszykowski, M., Walczak, B. & Massart, D. (2003)**. Projection methods in chemistry. *Chemometrics and Intelligent Laboratory Systems* **65**, 97–112.
- [202] **Ebrahimi, P.; N. Viereck; R. Bro; S. B. Engelsen**. Chemometric analysis of NMR Spectra. In *Modern Magnetic Resonance*, 1–20.
- [203] **Samuelsson, L. M. & Larsson, D. G. J. (2008)**. Contributions from metabolomics to fish research. *Molecular Bio Systems* **4**, 974–979.
- [204] **Tharwat, A., Gaber, T., Ibrahim, A. & Hassanien, A. E. (2017)**. Linear discriminant analysis: a detailed tutorial. *AI Communications* **30**, 169–190.

- [205] **Borràs, E., Ferré, J., Boqué, R., Mestres, M., Aceña, L. & Busto, O. (2015)**. Data fusion methodologies for food and beverage authentication and quality assessment - a review. *Analytica Chimica Acta* **891**, 1–14.
- [206] **Siqueira, L. F. S., Araújo J., Raimundo F., Araújo, A. A. de, Morais, C. L. M. & Lima, K. M. G. (2017)**. LDA vs. QDA for FT-MIR prostate cancer tissue classification. *Chemometrics and Intelligent Laboratory Systems* **162**, 123–129.
- [207] **Sharma, A. & Paliwal, K. K. (2015)**. Linear discriminant analysis for the small sample size problem: an overview. *International Journal of Machine Learning and Cybernetics* **6**, 443–454.
- [208] **Yang, J., Zhang, L., Yang, J.-Y. & Zhang, D. (2011)**. From classifiers to discriminators: a nearest neighbor rule induced discriminant analysis. *Pattern Recognition* **44**, 1387–1402.
- [209] **Dunn III, W. J. & Wold, S. (1978)**. A structure-carcinogenicity study of 4-nitroquinoline 1-oxides using the SIMCA method of pattern recognition. *Journal of Medicinal Chemistry* **21**, 1001–1007.
- [210] **Tominaga, Y. (1999)**. Comparative study of class data analysis with PCA-LDA, SIMCA, PLS, ANNs, and k-NN. *Chemometrics and Intelligent Laboratory Systems* **49**, 105–115.
- [211] **Vanden Branden, K. & Hubert, M. (2005)**. Robust classification in high dimensions based on the SIMCA Method. *Chemometrics and Intelligent Laboratory Systems* **79**, 10–21.
- [212] **Webb, A. R., Cawley, G. & Copsey, K. D. (2011)**. *Statistical pattern recognition*. John Wiley & Sons Ltd, Chichester.
- [213] **Brereton, R. G. (2006)**. Consequences of sample size, variable selection, and model validation and optimisation, for predicting classification ability from analytical data. *Trends in Analytical Chemistry* **25**, 1103–1111.
- [214] **Haddad, K., Rahman, A., A Zaman, M. & Shrestha, S. (2013)**. Applicability of Monte Carlo cross validation technique for model development and validation using generalised least squares regression. *Journal of Hydrology* **482**, 119–128.
- [215] **Brereton, R. G., Jansen, J., Lopes, J., Marini, F., Pomerantsev, A., Rodionova, O., R., Roger, J. M., Walczak, B. & Tauler, R. (2018)**. Chemometrics in analytical chemistry-part II: modeling, validation, and applications. *Analytical and Bioanalytical Chemistry* **410**, 6691–6704.
- [216] **Faber, N. M. & Rajkó, R. (2007)**. How to avoid over-fitting in multivariate calibration--the conventional validation approach and an alternative. *Analytica Chimica Acta* **595**, 98–106.
- [217] **Biancolillo, A; R. Boqué; M. Cocchi; F. Marini (2019)**. Data fusion strategies in food analysis. In *Data Fusion Methodology and Applications*. Data Handling in Science and Technology. Elsevier, 271–310.

- [218] **Callao, M. P. & Ruisánchez, I. (2018)**. An overview of multivariate qualitative methods for food fraud detection. *Food Control* **86**, 283–293.
- [219] **Nescatelli, R., Bonanni, R. C., Bucci, R., Magrì, A. L., Magrì, A. D. & Marini, F. (2014)**. Geographical traceability of extra virgin olive oils from Sabina PDO by chromatographic fingerprinting of the phenolic fraction coupled to chemometrics. *Chemometrics and Intelligent Laboratory Systems* **139**, 175–180.
- [220] **Monakhova, Y. B., Godelmann, R., Hermann, A., Kuballa, T., Cannet, C., Schäfer, H., Spraul, M. & Rutledge, D. N. (2014)**. Synergistic effect of the simultaneous chemometric analysis of ^1H NMR spectroscopic and stable isotope (SNIF-NMR, ^{18}O , ^{13}C) data: application to wine analysis. *Analytica Chimica Acta* **833**, 29–39.
- [221] **Jacobsen, N. E. (2007)**. *NMR spectroscopy explained. Simplified theory, applications and examples for organic chemistry and structural biology*. Wiley-Interscience, Hoboken, NJ.
- [222] **Caprioli, G., Cortese, M., Maggi, F., Minnetti, C., Odello, L., Sagratini, G. & Vittori, S. (2014)**. Quantification of caffeine, trigonelline and nicotinic acid in espresso coffee: the influence of espresso machines and coffee cultivars. *International Journal of Food Sciences and Nutrition* **65**, 465–469.
- [223] **Daglia, M., Papetti, A., Grisoli, P., Aceti, C., Spini, V., Dacarro, C. & Gazzani, G. (2007)**. Isolation, identification, and quantification of roasted coffee antibacterial compounds. *Journal of Agricultural and Food Chemistry* **55**, 10208–10213.
- [224] **Fisk, I. D., Kettle, A., Hofmeister, S., Virdie, A. & Kenny, J. S. (2012)**. Discrimination of roast and ground coffee aroma. *Flavour* **1**, 1–8.
- [225] **Gant, A., Leyva, V. E., Gonzalez, A. E. & Maruenda, H. (2015)**. Validated HPLC-diode array detector method for simultaneous evaluation of six quality markers in coffee. *Journal of AOAC International* **98**, 98–102.
- [226] **Luca, S. de, Ciotoli, E., Biancolillo, A., Bucci, R., Magrì, A. D. & Marini, F. (2018)**. Simultaneous quantification of caffeine and chlorogenic acid in coffee green beans and varietal classification of the samples by HPLC-DAD coupled with chemometrics. *Environmental Science and Pollution Research International* **25**, 28748–28759.
- [227] **Del Campo, G., Berregi, I., Caracena, R. & Zuriarrain, J. (2010)**. Quantitative determination of caffeine, formic acid, trigonelline and 5-(hydroxymethyl)furfural in soluble coffees by ^1H NMR spectrometry. *Talanta* **81**, 367–371.
- [228] **Matissek, R., Steiner, G. & Fischer, M. (2014)**. *Lebensmittelanalytik*. Springer Berlin Heidelberg, Berlin, Heidelberg.
- [229] **Ritgen, U. (2019)**. *Analytische Chemie I*. Springer Berlin Heidelberg, Berlin, Heidelberg.
- [230] **Gey, M. H. (2008)**. *Instrumentelle Analytik und Bioanalytik*. Springer Berlin Heidelberg, Berlin, Heidelberg.





- [231] **Monakhova, Y. B., Kohl-Himmelseher, M., Kuballa, T. & Lachenmeier, D. W. (2014)**. Determination of the purity of pharmaceutical reference materials by ¹H NMR using the standardless PULCON methodology. *Journal of Pharmaceutical and Biomedical Analysis* **100**, 381–386.
- [232] **Kromidas, S. (2011)**. *Handbuch Validierung in der Analytik. Ein Lehrbuch*. WILEY-VCH, Weinheim.
- [233] **Gibon, Y. & D. Rolin**. Aspects of experimental design for plant metabolomics experiments and guidelines for growth of plant material. In *Plant metabolomics*, 13–30.
- [234] **Hall, R. D. (2006)**. Plant metabolomics: from holistic hope, to hype, to hot topic. *The New Phytologist* **169**, 453–468.
- [235] **Moco, S., Bino, R. J., Vos, R. C. H. de, Bino, R. & Vervoort, J. (2007)**. Metabolomics technologies and metabolite identification. *Trends in Analytical Chemistry* **26**, 855–866.
- [236] **Wishart, D. S. (2008)**. Metabolomics: applications to food science and nutrition research. *Trends in Food Science & Technology* **19**, 482–493.
- [237] **Defernez, M., Wren, E., Watson, A. D., Gunning, Y., Colquhoun, I. J., Le Gall, G., Williamson, D. & Kemsley, E. K. (2017)**. Low-field (¹H) NMR spectroscopy for distinguishing between arabica and robusta ground roast coffees. *Food Chemistry* **216**, 106–113.
- [238] **Finotello, C., Forzato, C., Gasparini, A., Mammi, S., Navarini, L. & Schievano, E. (2017)**. NMR quantification of 16-O-methylcafestol and kahweol in *Coffea canephora* var. robusta beans from different geographical origins. *Food Control* **75**, 62–69.
- [239] **Baqueta, M. R., Coqueiro, A., Março, P. H., Mandrone, M., Poli, F. & Valderrama, F. (2021)**. Integrated ¹H NMR fingerprint with NIR spectroscopy, sensory properties, and quality parameters in a multi-block data analysis using ComDim to evaluate coffee blends. *Food Chemistry* **355**, 129618.
- [240] **Gniechwitz, D., Reichardt, N., Meiss, E., Ralph, J., Steinhart, H., Blaut, M. & Bunzel, M. (2008)**. Characterization and fermentability of an ethanol soluble high molecular weight coffee fraction. *Journal of Agricultural and Food Chemistry* **56**, 5960–5969.
- [241] **Pinto, L. F. G., Gardner, T., McDermott, C. L. & Ayub, K. O. L. (2014)**. Group certification supports an increase in the diversity of sustainable agriculture network-rainforest alliance certified coffee producers in Brazil. *Ecological Economics* **107**, 59–64.
- [242] **Kidist, T., Zerihun, G. & Biniam, E. (2019)**. Assessment of pre and post-harvest management practices on coffee (*Coffea Arabica* L.) quality determining factors in Gedeo zone, Southern Ethiopia. *African Journal of Agricultural Research* **14**, 1216–1228.
- [243] **Barjolle, D., Quiñones-Ruiz, X. F., Bagal, M. & Comoé, H. (2017)**. The role of the state for geographical indications of coffee: case studies from Colombia and Kenya. *World Development* **98**, 105–119.

- [244] **Ribeiro Costa, B. de (2020)**. Brazilian specialty coffee scenario. In *Coffee Consumption and Industry Strategies in Brazil*. Elsevier, 51–64.
- [245] **Haile, M. & Kang H. (2020)**. *Coffee production and research. Chapter 4: The harvest and post-harvest management practices' impact on coffee quality*. IntechOpen.
- [246] **Belay S., Mideksa D., Gebreziabher S. & Seifu W. (2016)**. Factors affecting coffee (*Coffea Arabica* L.) quality in Ethiopia: a review. *Journal of Multidisciplinary Scientific Research* **4**, 22–28.
- [247] **Melese, Y. Y. & Kolech, S. A. (2021)**. Coffee (*Coffea Arabica* L.): methods, objectives, and future strategies of breeding in Ethiopia-review. *Sustainability* **13**, 10814.
- [248] **Medina, J., Rodríguez, D. C., Arana, V. A., Bernal, A., Esseiva, P. & Wist, J. (2017)**. Comparison of attenuated total reflectance mid-infrared, near infrared, and ¹H-nuclear magnetic resonance spectroscopies for the determination of coffee's geographical origin. *International Journal of Analytical Chemistry* **2017**, 7210463.
- [249] **Núñez, N., Collado, X., Martínez, C., Saurina, J. & Núñez, O. (2020)**. Authentication of the origin, variety and roasting degree of coffee samples by non-targeted HPLC-UV fingerprinting and chemometrics. Application to the detection and quantitation of adulterated coffee samples. *Foods* **9**, 378.
- [250] **Rodrigues, C. I., Maia, R., Miranda, M., Ribeirinho, M., Nogueira, J. M. F. & Máguas, C. (2009)**. Stable isotope analysis for green coffee bean: A possible method for geographic origin discrimination. *Journal of Food Composition and Analysis* **22**, 463–471.
- [251] **Toledo, P. R. A. B. de, Melo, M. M. R. de, Pezza, H. R., Toci, A. T., Pezza, L. & Silva, C. M. (2017)**. Discriminant analysis for unveiling the origin of roasted coffee samples: A tool for quality control of coffee related products. *Food Control* **73**, 164–174.
- [252] **do Carmo Carvalho, D., Ribeiro Pereira Lima Brigagão, M., dos Santos, M. H., Borges Araújo de Paula, F., Giusti-Paiva, A. & Azevedo, L. (2011)**. Organic and conventional *Coffea arabica* L.: a comparative study of the chemical composition and physiological, biochemical and toxicological effects in Wistar rats. *Plant Foods for Human Nutrition* **66**, 114–121.
- [253] **Gordillo-Delgado, F., Marín, E., Mauricio Cortés-Hernández, D., Mejía-Morales, C. & García-Salcedo, A. J. (2012)**. Discrimination of organic coffee via Fourier transform infrared-photoacoustic spectroscopy. *Journal of the Science of Food and Agriculture* **92**, 2316–2319.
- [254] **Ruosi, M. R., Cordero, C., Cagliero, C., Rubiolo, P., Bicchi, C., Sgorbini, B. & Liberto, E. (2012)**. A further tool to monitor the coffee roasting process: aroma composition and chemical indices. *Journal of Agricultural and Food Chemistry* **60**, 11283–11291.
- [255] **Capek, P., Matulová, M., Navarini, L. & Suggi-Liverani, F. (2010)**. Structural features of an arabinogalactan-protein isolated from instant coffee powder of *Coffea arabica* beans. *Carbohydrate Polymers* **80**, 180–185.

- [256] **Nunes, F. M., Reis, A., Silva, A. M. S., Domingues, M. R. M. & Coimbra, M. A. (2008).** Rhamnoarabinosyl and rhamnoarabinoarabinosyl side chains as structural features of coffee arabinogalactans. *Phytochemistry* **69**, 1573–1585.
- [257] **Redgwell, R. J., Curti, D., Wang, J., Dobruchowska, J. M., Gerwig, G. J., Kamerling, J. P. & Bucheli, P. (2011).** Cell wall polysaccharides of Chinese wolfberry (*Lycium barbarum*): part 2. characterisation of arabinogalactan-proteins. *Carbohydrate Polymers* **84**, 1075–1083.
- [258] **Teipel, J. C., Hausler, T., Sommerfeld, K., Scharinger, A., Walch, S. G., Lachenmeier, D. W. & Kuballa, T. (2020).** Application of ¹H nuclear magnetic resonance spectroscopy as spirit drinks screener for quality and authenticity control. *Foods* **9**, 1355.
- [259] **Kolb, M., Bahr, A., Hippich, S. & Schulz, W. (1993).** Ermittlung der Nachweis-, Erfassungs- und Bestimmungsgrenze nach DIN 32645 mit Hilfe eines Programms Calculation of Detection Limit, Identification Limit and Determination Limit according to DIN 32645 with the Aid of a Computer Programs. *Acta Hydrochimica et Hydrobiologica* **21**, 308–311.
- [260] **Galli, V. & Barbas, C. (2004).** Capillary electrophoresis for the analysis of short-chain organic acids in coffee. *Journal of Chromatography A* **1032**, 299–304.
- [261] **Franca, A. S. & Oliveira, L. S. (2008).** *Chemistry of defective coffee beans*. Nova Science Publishers, New York.
- [262] **Marquetti, I., Link, J. V., Lemes, A. L. G., dos Santos Scholz, M. B., Valderrama, P. & Bona, E. (2016).** Partial least square with discriminant analysis and near infrared spectroscopy for evaluation of geographic and genotypic origin of arabica coffee. *Computers and Electronics in Agriculture* **121**, 313–319.
- [263] **Europäische Kommission (2021).** *Durchführungsverordnung (EU) 2021/1165 der Kommission vom 15. Juli 2021 über die Zulassung bestimmter Erzeugnisse und Stoffe zur Verwendung in der ökologischen/biologischen Produktion und zur Erstellung entsprechender Verzeichnisse. VO(EU) 2021/1165.*

7 Chemikalien und Instrumente

7.1 Chemikalien

Äpfelsäure	Merck (99%)
	
<p>H319 Verursacht schwere Augenreizung P280 Schutzhandschuhe/Augenschutz tragen P305+P351+P338 BEI KONTAKT MIT DEN AUGEN: Einige Minuten lang behutsam mit Wasser spülen. Eventuell vorhandene Kontaktlinsen nach Möglichkeit entfernen. Weiter spülen</p>	
γ -Butyrolacton	EDQM (>99%)
	
<p>H302 Gesundheitsschädlich bei Verschlucken H318 Verursacht schwere Augenschäden H336 Kann Schläfrigkeit und Benommenheit verursachen P261 Einatmen von Nebel/Dampf/Aerosol vermeiden. P270 Bei Gebrauch nicht essen, trinken oder rauchen. P280 Schutzhandschuhe/Augenschutz tragen. P305+P351+P338 BEI KONTAKT MIT DEN AUGEN: einige Minuten lang behutsam mit Wasser ausspülen. Eventuell vorhandene Kontaktlinsen nach Möglichkeit entfernen. Weiterspülen. P310 Sofort GIFTINFORMATIONSZENTRUM/Arzt anrufen.</p>	
3-Caffeoylchinasäure	Merck ($\geq 95\%$)
4-Caffeoylchinasäure	Merck ($\geq 98\%$)
5-Caffeoylchinasäure	Carl Roth ($\geq 97\%$)
Cafestol	Phytolab ($\geq 98\%$)
Chinasäure	Sigma Aldrich (98%)
	
<p>H351 Kann vermutlich Krebs erzeugen. P201 Vor Gebrauch besondere Anweisungen einholen. P202 Vor Gebrauch alle Sicherheitshinweise lesen und verstehen. P280 Schutzhandschuhe/ Schutzkleidung/ Augenschutz/ Gesichtsschutz/ Gehörschutz tragen. P308 + P313 BEI Exposition oder falls betroffen: Ärztlichen Rat einholen/ ärztliche Hilfe hinzuziehen. P405 Unter Verschluss aufbewahren. P501 Inhalt/ Behälter einer anerkannten Abfallentsorgungsanlage zuführen.</p>	
Chloroform	Merck ($\geq 99,8\%$)
Chloroform- d_1	Roth (99,8 Atom % D, stabilisiert mit Silber)
	
<p>H302 Gesundheitsschädlich bei Verschlucken.</p>	

H315 Verursacht Hautreizungen.
 H319 Verursacht schwere Augenreizung.
 H331 Giftig bei Einatmen.
 H351 Kann vermutlich Krebs erzeugen.
 H361d Kann vermutlich das Kind im Mutterleib schädigen.
 H372 Schädigt die Organe bei längerer oder wiederholter Exposition.
 P202 Vor Gebrauch alle Sicherheitshinweise lesen und verstehen.
 P260 Nebel/Dampf/Aerosol nicht einatmen.
 P302+P352 Bei Berührung mit der Haut: Mit viel Wasser und Seife waschen.
 P304+P340 Bei Einatmen: Die Person an die frische Luft bringen und für ungehinderte Atmung sorgen.
 P305+P351+P338 Bei Kontakt mit den Augen: einige Minuten lang behutsam mit Wasser ausspülen. Eventuell vorhandene Kontaktlinsen nach Möglichkeit entfernen. Weiterspülen.
 P308+P313 Bei Exposition oder falls betroffen: ärztlichen Rat einholen/ärztliche Hilfe hinzuziehen.

Cholinchlorid	Sigma (98%)
Dimethylsulfon	Sigma (98%)
1,2-Distearoyl-sn-glycero-3-phosphat	Merck (≥98%)
Deuteriumoxid	Deutero (99,9 Atom %)
Ferulasäure	Carl Roth (≥99%)



H315 Verursacht Hautreizungen
 H319 Verursacht schwere Augenreizung
 H335 Kann die Atemwege reizen
 P280 Schutzhandschuhe/Augenschutz tragen
 P302+P352 Bei Berührung mit der Haut: Mit viel Wasser waschen
 P305+P351+P338 Bei Kontakt mit den Augen: Einige Minuten lang behutsam mit Wasser spülen. Eventuell vorhandene Kontaktlinsen nach Möglichkeit entfernen. Weiter spülen

Fumarsäure	Sigma Aldrich (99%)
------------	---------------------



H319 Verursacht schwere Augenreizung
 P264 Nach Gebrauch Haut gründlich waschen.
 P280 Augenschutz/ Gesichtsschutz tragen.
 P305 + P351 + P338 BEI KONTAKT MIT DEN AUGEN: Einige Minuten lang behutsam mit Wasser spülen. Eventuell vorhandene Kontaktlinsen nach Möglichkeit entfernen. Weiter spülen.
 P337 + P313 Bei anhaltender Augenreizung: Ärztlichen Rat einholen/ärztliche Hilfe hinzuziehen

Furfural	Sigma Aldrich (99%)
----------	---------------------



H226 Flüssigkeit und Dampf entzündbar.
 H301 Giftig bei Verschlucken.
 H312 Gesundheitsschädlich bei Hautkontakt.

H315 Verursacht Hautreizungen.
 H319 Verursacht schwere Augenreizung.
 H330 Lebensgefahr bei Einatmen.
 H335 Kann die Atemwege reizen.
 H351 Kann vermutlich Krebs erzeugen.
 H412 Schädlich für Wasserorganismen, mit langfristiger Wirkung.
 P210 Von Hitze, heißen Oberflächen, Funken, offenen Flammen und anderen Zündquellen fernhalten. Nicht rauchen.
 P273 Freisetzung in die Umwelt vermeiden.
 P280 Schutzhandschuhe/ Schutzkleidung/ Augenschutz/ Gesichtsschutz/ Gehörschutz tragen.
 P303 + P361 + P353 Bei Berührung mit der Haut (oder dem Haar): Alle kontaminierten Kleidungsstücke sofort ausziehen. Haut mit Wasser abwaschen.
 P304 + P340 + P310 Bei Einatmen: Die Person an die frische Luft bringen und für ungehinderte Atmung sorgen. Sofort Giftinformationszentrum/ Arzt anrufen.
 P305 + P351 + P338 Bei Kontakt mit den Augen: Einige Minuten lang behutsam mit Wasser spülen. Eventuell vorhandene Kontaktlinsen nach Möglichkeit entfernen. Weiter spülen.

Furfurylalkohol

Sigma Aldrich (98%)



H302 + H312 Gesundheitsschädlich bei Verschlucken oder Hautkontakt.
 H319 Verursacht schwere Augenreizung.
 H331 Giftig bei Einatmen.
 H335 Kann die Atemwege reizen.
 H351 Kann vermutlich Krebs erzeugen.
 H373 Kann die Organe (Nase) schädigen bei längerer oder wiederholter Exposition durch Einatmen
 P201 Vor Gebrauch besondere Anweisungen einholen.
 P280 Schutzhandschuhe/ Schutzkleidung tragen.
 P301 + P312 + P330 Bei Verschlucken: Bei Unwohlsein Giftinformationszentrum/ Arzt anrufen. Mund ausspülen.
 P302 + P352 + P312 Bei Berührung mit der Haut: Mit viel Wasser waschen. Bei Unwohlsein Giftinformationszentrum / Arzt anrufen.
 P304 + P340 + P311 Bei Einatmen: Die Person an die frische Luft bringen und für ungehinderte Atmung sorgen. Giftinformationszentrum / Arzt anrufen.
 P308 + P313 BEI Exposition oder falls betroffen: Ärztlichen Rat einholen/ ärztliche Hilfe hinzuziehen.

5-Hydroxymethylfurfural

Carl Roth (≥97%)



H315 Verursacht Hautreizungen.
 H319 Verursacht schwere Augenreizung
 P280 Schutzhandschuhe/ Schutzkleidung tragen.

P302+P352 Bei Berührung mit der Haut: Mit viel Wasser waschen
 P305+P351+P338 Bei Kontakt mit den Augen: Einige Minuten lang behutsam mit Wasser spülen. Eventuell vorhandene Kontaktlinsen nach Möglichkeit entfernen. Weiter spülen

Kaffeensäure

Merck (98%)



H351 Kann vermutlich Krebs erzeugen
 P201 Vor Gebrauch besondere Anweisungen einholen.
 P202 Vor Gebrauch alle Sicherheitshinweise lesen und verstehen.
 P280 Schutzhandschuhe/ Schutzkleidung/ Augenschutz/ Gesichtsschutz/ Gehörschutz tragen.
 P308 + P313 BEI Exposition oder falls betroffen: Ärztlichen Rat einholen/ ärztliche Hilfe hinzuziehen.
 P405 Unter Verschluss aufbewahren.
 P501 Inhalt/ Behälter einer anerkannten Abfallentsorgungsanlage zuführen.

Kahweol

Phytolab ($\geq 90\%$)

Koffein

Sigma Aldrich (99%)



H302 Gesundheitsschädlich bei Verschlucken.
 P301 + P312 + P330 Bei Verschlucken: Bei Unwohlsein Giftinformationszentrum/ Arzt anrufen. Mund ausspülen.

Linolsäure

Sigma Aldrich (99%)

 α -LinolensäureSigma Aldrich ($\geq 99\%$)

Maleinsäure

Sigma Aldrich ($\geq 99\%$)

H302 + H312 Gesundheitsschädlich bei Verschlucken oder Hautkontakt.
 H315 Verursacht Hautreizungen.
 H317 Kann allergische Hautreaktionen verursachen.
 H318 Verursacht schwere Augenschäden.
 H335 Kann die Atemwege reizen.
 P261 Einatmen von Staub vermeiden.
 P264 Nach Gebrauch Haut gründlich waschen.
 P280 Schutzhandschuhe/ Schutzkleidung/ Augenschutz/ Gesichtsschutz.
 P301 + P312 Bei Verschlucken: Bei Unwohlsein Giftinformationszentrum/ Arzt anrufen.
 P302 + P352 + P312 Bei Berührung mit der Haut: Mit viel Wasser waschen. Bei Unwohlsein Giftinformationszentrum / Arzt anrufen.
 P305 + P351 + P338 Bei Kontakt mit den Augen: Einige Minuten lang behutsam mit Wasser spülen. Eventuell vorhandene Kontaktlinsen nach Möglichkeit entfernen. Weiter spülen.

Methanol MeOH- d_4	Merck ($\geq 99,9\%$) Merck ($\geq 99,8$ Atom%)
-------------------------	--



H225 Flüssigkeit und Dampf leicht entzündbar.
 H301 + H311 + H331 Giftig bei Verschlucken, Hautkontakt oder Einatmen.
 H370 Schädigt die Organe (Augen).
 P210 Von Hitze, heißen Oberflächen, Funken, offenen Flammen und anderen Zündquellen fernhalten. Nicht rauchen.
 P240 Behälter und zu befüllende Anlage erden.
 P280 Schutzhandschuhe/ Schutzkleidung tragen.
 P302 + P352 Bei Kontakt mit den Augen: Mit viel Wasser und Seife waschen.
 P304 + P340 Bei Einatmen: Die Person an die frische Luft bringen und für ungehinderte Atmung sorgen.
 P308 + P310 Bei Exposition oder falls betroffen: Sofort Giftinformationszentrum oder Arzt anrufen.
 P403 + P233 An einem gut belüfteten Ort aufbewahren. Behälter dicht verschlossen halten.

Myo-Inositol	Sigma Aldrich ($\geq 99\%$)
Natriumacetat	Carl Roth ($\geq 98,5\%$)
Natriumdihydrogenphosphat Monoydrat	Merck ($\geq 99\%$)
Natriumformiat	Sigma Aldrich (99%)
Natriumlaktat	Sigma Aldrich (98%)
Natriumtrimethylsilylpropionat	Sigma Aldrich (98 Atom%)
N-Methylpyridiniumiodid	Sigma Aldrich ($\geq 97\%$)



H302 Gesundheitsschädlich bei Verschlucken.
 H318 Verursacht schwere Augenschäden.
 P280 Schutzhandschuhe/ Schutzkleidung tragen.
 P301+P312+P330 Bei Verschlucken: Bei Unwohlsein Giftinformationszentrum/ Arzt anrufen. Mund ausspülen.
 P305+P351+P338+P310 Bei Kontakt mit den Augen: Einige Minuten lang behutsam mit Wasser spülen. Eventuell vorhandene Kontaktlinsen nach Möglichkeit entfernen. Weiter spülen. Sofort Giftinformationszentrum oder Arzt anrufen.

Ölsäure	Sigma Aldrich ($\geq 99\%$)
Ortho-Phosphorsäure	Merck (85%)



H290 Kann gegenüber Metallen korrosiv sein.
 H302 Gesundheitsschädlich bei Verschlucken.
 H314 Verursacht schwere Verätzungen der Haut und schwere Augenschäden.
 P234 Nur in Originalverpackung aufbewahren.
 P270 Bei Gebrauch nicht essen, trinken oder rauchen.

P280 Schutzhandschuhe/ Schutzkleidung/ Augenschutz/ Gesichtsschutz/ Gehörschutz tragen.
 P301 + P312 Bei Verschlucken: Bei Unwohlsein Giftinformationszentrum/ Arzt anrufen.
 P303 + P361 + P353 Bei Berührung mit der Haut (oder dem Haar): Alle kontaminierten Kleidungsstücke sofort ausziehen. Haut mit Wasser abwaschen.
 P305 + P351 + P338 Bei Kontakt mit dem Augen: Einige Minuten lang behutsam mit Wasser spülen. Eventuell vorhandene Kontaktlinsen nach Möglichkeit entfernen. Weiter spülen

p-Cumarsäure

Carl Roth (98%)



H315 Verursacht Hautreizungen.
 H319 Verursacht schwere Augenreizung.
 H335 Kann die Atemwege reizen.
 P261 Einatmen von Staub/ Rauch/ Gas/ Nebel/ Dampf/ Aerosol vermeiden.
 P264 Nach Gebrauch Haut gründlich waschen.
 P271 Nur im Freien oder in gut belüfteten Räumen verwenden.
 P280 Schutzhandschuhe/ Augenschutz/ Gesichtsschutz tragen.
 P302 + P352 Bei Berührung mit der Haut: Mit viel Wasser waschen.
 P305 + P351 + P338 Bei Kontakt mit den Augen: Einige Minuten lang behutsam mit Wasser spülen. Eventuell vorhandene Kontaktlinsen nach Möglichkeit entfernen. Weiter spülen.

Propionsäure

Merck (≥99%)



H226: Flüssigkeit und Dampf entzündbar.
 H314: Verursacht schwere Verätzungen der Haut und schwere Augenschäden.
 H335: Kann die Atemwege reizen.
 P210: Von Hitze, heißen Oberflächen, Funken, offenen Flammen und anderen Zündquellen fernhalten. Nicht rauchen.
 P233: Behälter dicht verschlossen halten.
 P240: Behälter und zu befüllende Anlage erden.
 P280: Schutzhandschuhe/ Schutzkleidung/ Augenschutz/ Gesichtsschutz/ Gehörschutz Tragen.
 P303 + P361 + P353: Bei Berührung mit der Haut(oder dem Haar): Alle kontaminierten Kleidungsstücke sofort ausziehen. Haut mit Wasser abwaschen.
 P305 + P351 + P338: Bei Kontakt mit den Augen: Einige Minuten lang behutsam mit Wasser spülen. Eventuell vorhandene Kontaktlinsen nach Möglichkeit entfernen. Weiter spülen.

Stearinsäure

Sigma Aldrich(≥98,5%)

Trigonellin-Hydrochlorid

Carl Roth (≥97,5%)

Tetramethylsilan

Carl Roth (≥99,9%)



H225 Flüssigkeit und Dampf extrem entzündbar.
 P210: Von Hitze, heißen Oberflächen, Funken, offenen Flammen und anderen Zündquellen fernhalten. Nicht rauchen.
 P240: Behälter und zu befüllende Anlage erden.
 P243: Maßnahmen gegen elektrostatische Aufladung treffen
 P403+P235: An einem gut belüfteten Ort aufbewahren. Kühl halten.

Tripalmitin

VWR ($\geq 98\%$)

Zitronensäure

Merck ($\geq 99,5\%$)

H319: Verursacht schwere Augenreizung.
 H335: Kann die Atemwege reizen.
 P261: Einatmen von Staub/ Rauch/ Gas/ Nebel/ Dampf/ Aerosol vermeiden.
 P264: Nach Gebrauch Haut gründlich waschen.
 P271: Nur im Freien oder in gut belüfteten Räumen verwenden.
 P280: Augenschutz/ Gesichtsschutz tragen.
 P304 + P340 + P312: Bei Einatmen: Die Person an die frische Luft bringen und für ungehinderte Atmung sorgen. Bei Unwohlsein Giftinformationszentrum/Arzt anrufen.
 P305 + P351 + P338: Bei Kontakt mit den Augen: Einige Minuten lang behutsam mit Wasser spülen. Eventuell vorhandene Kontaktlinsen nach Möglichkeit entfernen. Weiter spülen.

7.2 Hergestellte Lösungen und Reagenzien

Lösung	Herstellung
Extraktionslösung für Extraktion der lipophilen Metaboliten	Chloroform : Methanol (1:1, v/v) Dimethylsulfon (200 mg/L)
Messlösung für die 1H-NMR-Messung der lipophilen Metaboliten	CDCl_3 : $\text{MeOH-}d_4$ (1:1, v/v)
CDCl_3 mit 0,1% TMS	500 μL TMS auf 500 mL CDCl_3
Natriumphosphat-Puffer (1M, pH6)	138 g Natriumdihydrogenphosphat Monohydrat auf 1000 mL H_2O ; Einstellung des pH-Wertes mit ortho-Phosphorsäure
TSP-Lösung	50 mg TSP in 50 mL D_2O
QuantRef-Lösung	1,2 g Na-Laktat; 5,0 g Zitronensäure; ca. 10 mg NaN_3 in 1 L destilliertem H_2O gelöst
AfG-Puffer	10 mL destilliertes H_2O + 7,5 g NaH_2PO_4 + 1000 mg H_3PO_4 (85 %)
QuantRef-Messlösung	600 μL der QuantRef-Lösung werden mit 100 μL AfG-Puffer und 70 μL TSP-Lösung versetzt; davon werden 600 μL in ein NMR-Röhrchen überführt

7.3 Instrumente

Probenvorbereitung	
Direktverdränger-Pipette	Volumen 100-1000 µL, BRAND GMBH + CO. KG (Wertheim, Deutschland)
Einkanalpipette	Volumen 100-1000 µL, Eppendorf SE (Hamburg, Deutschland)
Feinwaage	Analysenwaage AX204, Genauigkeit: 0,01 g, Mettler Toledo GmbH (Gießen, Deutschland)
Feinwaage	Elektronische Waage XPE206DR/M, Genauigkeit: 0,001 g, Mettler Toledo GmbH (Gießen, Deutschland)
Mehrfachdispenser	HandyStep, BRAND GMBH + CO. KG (Wertheim, Deutschland)
Moisture Analyzer HC103	Mettler Toledo GmbH (Gießen, Deutschland)
pH-Elektrode	Thermo Scientific Orion Star A211 pH-Tischmessgerät, Thermo Fisher Scientific GmbH (Dreieich, Deutschland)
Probenkonzentrator	Techne Sample Concentrator Dri Block DB-3D, Biostep GmbH, (Burkhardtsdorf, Deutschland)
Reagenzglasschüttler	Multi Reax, Heidolph (Schwabach, Deutschland)
Zentrifuge	Thermo Scientific Heraeus Megafuge 16, Thermo Fisher Scientific GmbH (Dreieich, Deutschland)
HPLC-Messung	
Erfassung der Kaffeemetabolite	HPLC: Agilent 1100/1200 Series Säule: Synergi Polar/RP80 250•2 mm (Phenomenex, Aschaffenburg, Deutschland) Pumpe: Binary Pump G1312A Autosampler: G1329B High Performance Autosampler SL Detektor: Diodanarray-Detektor
¹H-NMR-Messung	
Erfassung der Kaffeemetabolite	Bruker Ascend 400 MHz mit Avance III Konsole (Bruker, Rheinstetten, Germany) Probenkopf: PA BBI 400S1 H-BB-D-05 Probenwechsler: SampleXpress H15000-01
Überprüfung der Robustheit	Bruker UltraShield 400 MHz mit Avance III Konsole (Bruker, Rheinstetten, Germany) Probenkopf: 5mm PASEI 1H/D-13C Z-GRD Probenwechsler: SampleXpress H15040-01
2D NMR Messungen	Bruker Ascend 400 MHz mit Avance III Konsole (Bruker, Rheinstetten, Germany) Probenkopf: PA BBO BB/19F-1H/D Z-GRD Probenwechsler: SampleXpress H15000-01

7.4 Verbrauchsmaterialien

Verbrauchsmaterial	Hersteller
Dispenser-Tips	Für Mehrfachdispenser, Volumen: 1 mL, 5 mL, 50 mL, BRAND GMBH + CO. KG (Wertheim, Deutschland)
Einmalspritzen 2 mL	Braun (Melsungen, Deutschland)
Glaskugeln (Ø6 mm)	
Glasvials, 4 mL mit Schraubverschluss	Klaus Ziemer (Langerwehe, Deutschland)
NMR-Röhrchen O.D. $4,966 \pm 0,004$ mm; I.D. $4,166 \pm 0,004$ mm, Länge 17,78 cm	Deutero (Kastellaun, Deutschland)
Pasteurpipetten	
Pipettenspitzen für Direktverdränger-Pipette	Volumen 100-1000 μ L, Gilson International B.V. (Limburg an der Lahn, Deutschland)
Pipettenspitzen für Einkanalpipette	Volumen 100-1000 μ L, Eppendorf SE (Hamburg, Deutschland)
Spitzenvorsatzfilter (GF-PET, 0,25 μ m)	Macherey-Nagel (Düren, Deutschland)
Spitzenvorsatzfilter (PTFE, 0,45 μ m)	Agilent (Waldbronn, Deutschland)
Zentrifugenfilter (2 kDa, Hydrosart ® Membran)	Sartorius Lab Instruments GmbH and Co. KG (Göttingen, Deutschland)
Zentrifugenfilter (3 kDa, PES-Membran)	Sartorius Lab Instruments GmbH and Co. KG (Göttingen, Deutschland)
Zentrifugenfilter (5 kDa, PES-Membran)	Sartorius Lab Instruments GmbH and Co. KG (Göttingen, Deutschland)
Zentrifugenfilter (1 kDa, PES-Membran)	Pall Deutschland Holding GmbH & Co. KG (Dreieich, Deutschland)
Zentrifugenröhrchen, 15 mL	Sarstedt (Nümbrecht, Deutschland)

8 Arbeitsvorschriften

8.1 Probenpräparation

- Die Kaffeebohnen werden mittels Mahlkönig (Stufe 2,9) zu einer Korngröße von 0,30 mm gemahlen.
- Der Feuchtegehalt der gemahlene Kaffeeproben wird mittels Moisture Analyzer HC103 bestimmt. Nach der Einwaage der Kaffeeprobe von ca. 4,5 g erfolgt die Ermittlung des exakten Gewichtes und die Bestimmung des Feuchtegehaltes automatisch.
- Die Bestimmung des Röstgrades der Kaffeeprobe erfolgt mittels Tonino Farbmesser. Es wird ausreichend Kaffeepulver in einen Siebträgermaschinen-Sieb gegeben, dass der Boden vollständig mit Kaffeepulver bedeckt ist und dieses mittels Tamper festgedrückt. Die Farbwertmessung erfolgt an fünf verschiedenen Messpunkten in den Siebträgermaschinen-Sieb. Als Röstgrad wird der Mittelwert der fünf Messungen angegeben.
- Die Kaffeeproben werden in luft- und lichtundurchlässigen Verpackungen mit Aromaventil bei 5-7°C gelagert.

8.2 Entwicklung der nicht zielgerichteten Methode für die Untersuchung der lipophilen Metaboliten

8.2.1 Optimierung der Lösungsmittelzusammensetzung für die Extraktion

A) Methode nach Folch et al.

- Herstellung einer Mischung bestehend aus Chloroform/Methanol/Wasser (8:4:3, v/v/v), nach Phasentrennung die obere Phase abnehmen und die untere Phase verwerfen. Die obere Phase dient im späteren Schritt als Waschlösung
- Herstellung einer 0,9% NaCl-Lösung (aq)
- 0,5 g Kaffeepulver in einem 50 mL Zentrifugenröhrchen abwiegen.
- Zusatz von 375 µL destilliertes Wasser, 17,5 mL Extraktionsmittel Chloroform/Methanol (2:1, v/v) und fünf Glasperlen (ø6mm)
- Für 20 min bei Raumtemperatur im Reagenzglasschüttler (1600 U/min) schütteln
- Extraktionsgemisch durch einen Faltenfilter in einen Scheidetrichter filtrieren
- Extraktionsgefäß mit 5 mL Extraktionsmittel nachspülen und dieses ebenfalls über den Faltenfilter filtrieren
- Filtrat mit 2,5 mL NaCl-Lösung (0,9%, aq) versetzen und schütteln
- Nach Phasentrennung obere Phase abnehmen und verwerfen
- Untere Phase drei Mal mit je 1 mL hergestellter Waschlösung waschen
- Nach dem Waschen die untere Phase abnehmen und diese im Probenkonzentrator bei 40°C und Stickstoffstrom vom Lösungsmittel befreien

B) Methode nach Ackermann et al.

- Abwiegen von 0,50 g Kaffeepulver in einem 50 mL Zentrifugenröhrchen

- Zugabe von 7,5 mL Extraktionsmittel (Chloroform/Methanol/Wasser (10:5:1, v/v/v) und fünf Glasperlen (\varnothing 6mm)
- Für 20 min bei Raumtemperatur im Reagenzglasschüttler (1600 U/min) schütteln
- Extraktionsgemisch durch einen Faltenfilter in einen Scheidetrichter filtrieren
- Extraktionsgefäß mit 5 mL Extraktionsmittel nachspülen und auch ebenfalls über den Faltenfilter filtrieren
- Filtrat mit 1,0 mL NaCl-Lösung (0,9%, aq) versetzen und schütteln
- Nach Phasentrennung untere Phase abnehmen und diese im Probenkonzentrator bei 40°C und Stickstoffstrom vom Lösungsmittel befreien

C) Methode nach Blight und Dyer

- Abwiegen von 0,50 g Kaffeepulver in einem 50 mL Zentrifugenröhrchen
- Zugabe von 10 mL Extraktionsmittel (Chloroform/Methanol/Wasser (1:2:0,8, v/v/v) und fünf Glasperlen (\varnothing 6mm)
- Für 20 min bei Raumtemperatur im Reagenzglasschüttler (1600 U/min) schütteln
- Extraktionsgemisch durch einen Faltenfilter in einen Scheidetrichter filtrieren
- Extraktionsgefäß mit 5 mL Extraktionsmittel nachspülen und dieses ebenfalls über den Faltenfilter filtrieren
- Zugabe von jeweils 3,75 mL Chloroform und Wasser zum Filtrat
- Gemisch schütteln
- Nach Phasentrennung untere Phase abnehmen und diese im Probenkonzentrator bei 40°C und Stickstoffstrom vom Lösungsmittel befreien

D) Extraktion mit Chloroform

- Abwiegen von 0,50 g Kaffeepulver in einem 50 mL Zentrifugenröhrchen
- Zugabe von 10 mL Chloroform und fünf Glasperlen (\varnothing 6mm)
- Für 20 min bei Raumtemperatur im Reagenzglasschüttler (1600 U/min) schütteln
- Extraktionsgemisch durch einen Faltenfilter in einen Scheidetrichter filtrieren
- Extraktionsgefäß mit 5 mL Chloroform nachspülen und dieses ebenfalls über den Faltenfilter filtrieren
- Filtrat im Probenkonzentrator bei 40°C und Stickstoffstrom vom Lösungsmittel befreien

E) Extraktion mit Chloroform/Methanol (1:1, v/v)

- Abwiegen von 0,50 g Kaffeepulver in einem 50 mL Zentrifugenröhrchen
- Zugabe von 10 mL Extraktionsmittel (Chloroform/Methanol, 1:1, v/v) und fünf Glasperlen (\varnothing 6mm)
- Für 20 min bei Raumtemperatur im Reagenzglasschüttler (1600 U/min) schütteln
- Extraktionsgemisch durch einen Faltenfilter in einen Scheidetrichter filtrieren

- Extraktionsgefäß mit 5 mL Extraktionsmittel nachspülen und dieses ebenfalls über den Faltenfilter filtrieren
- Filtrat im Probenkonzentrator bei 40°C und Stickstoffstrom vom Lösungsmittel befreien

F) Extraktion nach Lin et al.

- Abwiegen von 0,50 g Kaffeepulver in einem 50 mL Zentrifugenröhrchen
- Zugabe von 10 mL Extraktionsmittel (Etanol/Ethylacetat, 1:1, v/v) und fünf Glasperlen (ø6mm)
- Für 20 min bei Raumtemperatur im Reagenzglasschüttler (1600 U/min) schütteln
- Extraktionsgemisch durch einen Faltenfilter in einen Scheidetrichter filtrieren
- Extraktionsgefäß mit 5 mL Extraktionsmittel nachspülen und dieses ebenfalls über den Faltenfilter filtrieren
- Filtrat im Probenkonzentrator bei 40°C und Stickstoffstrom vom Lösungsmittel befreien

G) Extraktion mit Aceton

- Abwiegen von 0,50 g Kaffeepulver in einem 50 mL Zentrifugenröhrchen
- Zugabe von 10 mL Aceton und fünf Glasperlen (ø6mm)
- Für 20 min bei Raumtemperatur im Reagenzglasschüttler (1600 U/min) schütteln
- Extraktionsgemisch durch einen Faltenfilter in einen Scheidetrichter filtrieren
- Extraktionsgefäß mit 5 mL Aceton nachspülen und dieses ebenfalls über den Faltenfilter filtrieren
- Filtrat im Probenkonzentrator bei 40°C und Stickstoffstrom vom Lösungsmittel befreien

Die Extraktionen A-G wurden jeweils in Dreifachbestimmung durchgeführt. Die vom jeweiligen Lösungsmittel befreiten Proben wurden in 700 µL Messlösung (CDCl₃:MeOH-*d*₄, 1:1, v/v) rückgelöst. Davon wurden 600 µL in ein NMR-Röhrchen überführt. Mit der Messlösung erfolgte auch eine Zugabe von 0,1% TMS zum referenzieren der ¹H-NMR-Spektren.

8.2.2 Optimierung der Extraktionsparameter

A) Extraktionszeit

- Einwaage von 0,575 g Kaffee in 15 mL Zentrifugenröhrchen
- Zugabe von 8 mL Extraktionsmittel (Chloroform/Methanol, 1:1, v/v) und fünf Glasperlen (ø6mm)
- Extraktion bei Raumtemperatur für 30, 60, 90, 120 und 180 min im Reagenzglasschüttler (1600 U/min)
- Ein Aliquot des Extraktes von 2 mL abnehmen und membranfiltrieren
- Filtrat im Probenkonzentrator bei 40°C und Stickstoffstrom vom Lösungsmittel befreien
- Rücklösen der vom Lösungsmittel befreiten Proben in 700 µL Messlösung (CDCl₃:MeOH-*d*₄, 1:1, v/v) und Überführen von 600 µL in ein NMR-Röhrchen

Die Aufarbeitung der Kaffeeproben erfolgte für jede Extraktionszeit in Dreifachbestimmung.

B) Verwendung von Glasperlen

- Einwaage von 0,575 g Kaffee in 15 mL Zentrifugenröhrchen
- Zugabe von 8 mL Extraktionsmittel (Chloroform/Methanol, 1:1, v/v)
- Extraktion bei Raumtemperatur für 60 min einmal ohne und einmal mit der Verwendung von fünf Glasperlen (\varnothing 6 mm) im Reagenzglasschüttler (1600 U/min)
- Ein Aliquot des Extraktes von 2 mL abnehmen und membranfiltrieren
- Filtrat im Probenkonzentrator bei 40°C und Stickstoffstrom vom Lösungsmittel befreien
- Rücklösen der vom Lösungsmittel befreiten Proben in 700 μ L Messlösung (CDCl_3 :MeOH- d_4 , 1:1, v/v) und Überführen von 600 μ L in ein NMR-Röhrchen

Die Aufarbeitung der Kaffeeproben erfolgte jeweils in Dreifachbestimmung.

8.2.3 Entwickelte Extraktion für die nicht zielgerichtete Untersuchung der lipophilen Metaboliten

- Einwaage von 0,575 g Kaffee in 15 mL Zentrifugenröhrchen
- Zusatz von fünf Glasperlen (\varnothing 6 mm)
- Zugabe von 8 mL Extraktionsmittel (Chloroform/Methanol, 1:1, v/v) mit DMS als internem Standard (80 mg/400 mL Extraktionsmittel)
- Extraktion bei Raumtemperatur für 60 min im Reagenzglasschüttler (1600 U/min)
- Ein Aliquot des Extraktes von 2 mL abnehmen und membranfiltrieren (PFTE, 0,45 μ m)
- Filtrat im Probenkonzentrator bei 40°C und Stickstoffstrom vom Lösungsmittel befreien
- Rückstand der vom Lösungsmittel befreiten Probe auswiegen

8.2.4 Optimierung der Messlösung und der Messparameter

A) Lösungsmittel für die Messlösung

- Aufarbeitung der Kaffeeproben nach Kapitel 8.2.3
- Die vom Lösungsmittel befreiten Proben jeweils in 700 μ L Messlösung verschiedener Zusammensetzung lösen (CDCl_3 ; CDCl_3 :MeOH- d_4 2:1 (v/v); CDCl_3 :MeOH- d_4 3:2 (v/v); CDCl_3 :MeOH- d_4 1:1 (v/v); CDCl_3 :DMSO- d_6 5:1 (v/v))
- Überführen von 600 μ L in ein NMR-Röhrchen

B) Testen der Anwendung von entschärftem CDCl_3

- Herstellen des entschärftem CDCl_3 : 70 mL CDCl_3 mit 35 mL Natriumcarbonat-Lösung (100 g/L, aq.) in einem Scheidetrichter für 5 min ausschütteln. In einer Braunglasflasche 3,5 g getrocknetes Natriumcarbonat (150°C, 24h) abwiegen und nach der Phasentrennung untere Phase in Braunglasflasche

ablassen. Entschärftes CDCl_3 für zwei h bei geringer Schüttelleistung schüteln und über Nacht stehen lassen. Das entschärftes CDCl_3 ist über Natriumcarbonat gelagert mehrere Wochen stabil.

- Aufarbeitung der Kaffeeproben nach Kapitel 8.2.3
- Die vom Lösungsmittel befreiten Proben jeweils in 700 μL Messlösung lösen ($\text{CDCl}_3:\text{MeOH-}d_4$ 1:1 (v/v); entschärftes $\text{CDCl}_3:\text{MeOH-}d_4$ 1:1 (v/v))
- Die Aufarbeitung der Kaffeeproben und das Rücklösen in der Messlösung erfolgte mehrfach an verschiedenen Tagen

C) Pulsprogramm für die ^1H -NMR-Messung

- Aufarbeitung der Kaffeeproben nach Kapitel 8.2.3
- Rücklösen der vom Lösungsmittel befreiten Probe in 700 μL Messlösung (entschärftes $\text{CDCl}_3:\text{MeOH-}d_4$ 1:1 (v/v)) und Überführen von 600 μL ins NMR-Röhrchen.
- Messung der Probe mittels Experiment in Tabelle 8-1 mit verschiedenen Pulsprogrammen (30° Puls, 90° Puls)

Tabelle 8-1: ^1H -NMR-Messparameter und Prozessierungsparameter für die Optimierung des Pulsprogrammes für die nicht zielgerichtete Methode zur Untersuchung der lipophilen Metaboliten. TD: Time Domain; DS: Anzahl Dummy Scans; NS: Anzahl Scans; AQ: Akquisitionszeit; O1: Sendefrequenz Offset; D1: Relaxationsverzögerung; SI: Size of Real spectrum; RG: Receiver Gain; WDW: Window Fuction; EM: Exponential Function; LB: Line Broadening Factor

^1H-NMR Akquisitionsparameter				
NMR-Spektrometer	Ascend II		O1 [ppm]	6,175
Pulsprogramm	zg	zg30	D1 [s]	15 9
TD	1279	98304	Puls [μs]	8
DS	2	2	Temperatur [K]	290
NS	64	128	Solvent	$\text{CDCl}_3/\text{MeOH-}d_4$
AQ [s]	7,78	5,98	Digmod	baseopt
Spektrale Breite [ppm]	20,5617		RG	40,3
^1H-NMR Prozessierungsparameter				
Software	TopSpin 3.7		LB	0,3
SI	262144		Phasenkorrektur	apk0.noe
WDW	EM		Basislinienkorrektur	aplbc.c

D) Messtemperatur für die ^1H -NMR-Messung

- Aufarbeitung der Kaffeeproben nach Kapitel 8.2.3
- Rücklösen der vom Lösungsmittel befreiten Probe in 700 μL Messlösung (entschärftes $\text{CDCl}_3:\text{MeOH-}d_4$ 1:1 (v/v)) und Überführen von 600 μL ins NMR-Röhrchen.
- Messung der Kaffeeprobe mittels zg-Experiment unter der Verwendung verschiedener Messtemperaturen (280 K, 290 K, 300 K)
- Vor jeder Messung erfolgte eine Kalibrierung der Pulslänge des Anregungspulses

E) Bestimmung der T_1 -Zeit mit dem Inversion Recovery Experiment

- Aufarbeitung der Kaffeeproben nach Kapitel 8.2.3
- Rücklösen der vom Lösungsmittel befreiten Probe in 700 μL Messlösung (entschärftes $\text{CDCl}_3:\text{MeOH-}d_4$ 1:1 (v/v)) und Überführen von 60 μL ins NMR-Röhrchen.

- Messung der Probe mittels Inversion Recovery Experiment der Firma Bruker „t1ir“ bei verschiedenen Relaxationsverzögerungen (0,1 - 40 s)

F) Überprüfung der Stabilität der Messlösung

- Aufarbeitung einer Aufarbeitung der Kaffeeproben nach Kapitel 8.2.3
- Rücklösen der vom Lösungsmittel befreiten Probe in 700 µL Messlösung (entschärftes CDCl₃:MeOH-*d*₄ 1:1 (v/v)) und Überführen von 600 µL ins NMR-Röhrchen
- Messung der Probe mit dem in Tabelle 8-2 aufgeführtem Experiment in einem Zeitfenster von 60 h. Innerhalb der ersten 24 h erfolgt die Messung stündlich anschließend findet eine Messung nach jeweils zwölf h statt

8.2.5 Entwickelte Methode für die nicht zielgerichtete Untersuchung der lipophilen Metaboliten

- Aufarbeitung der Kaffeeproben nach Kapitel 8.2.3
- Vom Lösungsmittel befreite Probe in 700 µL Messlösung (entschärftes CDCl₃:MeOH-*d*₄ 1:1 (v/v)) lösen
- Überführung von 600 µL ins NMR-Röhrchen
- Akquisitions- und Prozessierungsparameter laut Tabelle 8-2.
- Für jede Probe erfolgte eine Kalibrierung der Pulslänge des 90° Pulses
- Überprüfung der Qualität der Spektren durch Ermittlung der Halbwertsbreite des TMS-Signals. Maximal erlaubte Halbwertsbreite des TMS-Signals: 1,2 Hz
- Proben außerhalb dieser Spezifikation wurden wiederholt
- Auf jedes Spektrum wurde eine automatische Phasen- und Basislinienkorrektur angewendet

Tabelle 8-2: ¹H-NMR-Akquisitions- und Prozessierungsparameter für die nicht zielgerichtete Methode zur Untersuchung der lipophilen Metaboliten. TD: Time Domain; DS: Anzahl Dummy Scans; NS: Anzahl Scans; AQ: Akquisitionszeit; O1: Sendefrequenz Offset; D1: Relaxationsverzögerung; SI: Size of Real spectrum; RG: Receiver Gain; WDW: Window Function; EM: Exponential Function; LB: Line Broadening Factor

¹H-NMR Akquisitionsparameter			
NMR-Spektrometer	Ascend 2	O1 [ppm]	6,175
Pulsprogramm	zg	D1 [s]	15
TD	131072	Puls [µs]	8
DS	2	Temperatur [K]	290
NS	64	Solvent	CDCl ₃ /MeOH- <i>d</i> ₄
AQ [s]	7,97	Digmod	baseopt
Spektrale Breite [ppm]	20,5617	RG	40,3
¹H-NMR Prozessierungsparameter			
Software	TopSpin 3.7	LB	0,3
SI	262144	Phasenkorrektur	apk0.noe
WDW	EM	Basislinienkorrektur	aplbc.c

8.3 Entwicklung der nicht zielgerichteten Methode für die Untersuchung der hydrophilen Metaboliten

8.3.1 Optimierung der Extraktionsparameter

A) Verhältnis von Probe zu Extraktionsmittel

- Proben: Eine Kaffeeprobe mit hellen und eine mit dunklem Röstgrad
- Probeneinwaage: 0,10; 0,20; 0,60; 1,20 g in 15 mL Zentrifugenröhrchen
- Zugabe von 8 mL destilliertes Wasser
- Extraktion bei Raumtemperatur für 20 min im Reagenzglasschüttler (1600 U/min)
- Extrakt membranfiltrieren
- Mischen von 700 μ L Filtrat mit 70 μ L TSP in D₂O (1 mg/mL)

B) Extraktionszeit

- Einwaage von 0,2 g Kaffee in 15 mL Zentrifugenröhrchen
- Zugabe von 8 mL destilliertes Wasser
- Extraktion bei Raumtemperatur im Reagenzglasschüttler (1600 U/min) mit unterschiedlichen Extraktionszeiten: 20, 30, 60, 90, 120, 150 und 180 min
- Extrakt membranfiltrieren
- Mischen von 700 μ L Filtrat mit 70 μ L TSP in D₂O (1 mg/mL)
- Für jede Schüttelzeit erfolgte eine Dreifachbestimmung

C) Verwendung von Glasperlen

- Einwaage von 0,2 g Kaffee in 15 mL Zentrifugenröhrchen
- Zugabe von 8 mL destilliertes Wasser
- Extraktion bei Raumtemperatur für 20 Minuten im Reagenzglasschüttler (1600 U/min) mit und ohne Verwendung von fünf Glasperlen (\varnothing 6mm)
- Extrakt membranfiltrieren
- Mischen von 700 μ L Filtrat mit 70 μ L TSP in D₂O (1 mg/mL)
- Die Untersuchung der Verwendung von Glasperlen erfolgte jeweils in Dreifachbestimmung

8.3.2 Optimierung der Kohlenhydratentfernung mittels UF

- Einwaage von 0,2 g in 15 mL Zentrifugenröhrchen
- Zugabe von 8 mL destilliertes Wasser
- Extraktion bei Raumtemperatur für 20 min im Reagenzglasschüttler (1600 U/min)
- Extrakt membranfiltrieren
- Zentrifugenfilter mit verschiedener Porengröße (1; 2; 3 und 5 kDa) zur Entfernung des Glycerins mit Wasser vorspülen. Dazu fünf Mal jeweils 2 mL destilliertes Wasser auf die Filter geben und diese zentrifugieren (15 min; 4000 g; 22°C). Nach dem letzten Waschschrift 30 min zentrifugieren (4000 g; 22°C)
- Nach fünf Waschschriften 1,4 mL des membranfiltrierten Extraktes auf die Filter geben und zentrifugieren (1,5 h; 4000 g; 22°C)

- Mischen von 700 μL Filtrat mit 70 μL TSP in D_2O (1 mg/mL)

8.3.3 Entwickelte Extraktion für die nicht zielgerichtete Untersuchung der hydrophilen Metaboliten

- A) Extraktion der Kaffeeproben ohne die Entfernung der Kohlenhydrate
- Einwaage von 0,20 g in 15 mL Zentrifugenröhrchen
 - Zugabe von 8 mL destilliertes Wasser
 - Extraktion bei Raumtemperatur für 20 min im Reagenzglasschüttler (1600 U/min)
 - Extrakt membranfiltrieren (GF/PET, 0,25 μm)
- B) Extraktion der Kaffeeproben mit der Entfernung der Kohlenhydrate
- Einwaage von 0,20 g in 15 mL Zentrifugenröhrchen
 - Zugabe von 8 mL destilliertes Wasser
 - Extraktion bei Raumtemperatur für 20 min im Reagenzglasschüttler (1600 U/min)
 - Extrakt membranfiltrieren (GF/PET, 0,25 μm)
 - Zentrifugenfilter (1 und 3 kDa) zur Entfernung des Glycerins mit Wasser vorspülen. Dazu fünf Mal jeweils 2 mL destilliertes Wasser auf die Filter geben und diese zentrifugieren (15 min; 4000 g; 22°C). Nach dem letzten Waschschrift 30 min zentrifugieren (4000 g; 22°C)
 - Nach fünf Waschsritten 1,4 mL des membranfiltrierten Extraktes auf die Filter geben und zentrifugieren (1,5 h; 4000 g; 22°C).

8.3.4 Optimierung der Messlösung und der Messparameter

- A) Optimierung der Pufferlösung
- Aufarbeitung der Kaffeeproben nach Kapitel 8.3.3 A)
 - Mischen von 600 μL membranfiltriertem Extrakt mit 100 μL Natriumphosphat Puffer und 70 μL TSP in D_2O (1 mg/mL)
 - Es wurden Natriumphosphatpuffer bestehend aus NaH_2PO_4 und NaOH oder H_3PO_4 mit einem pH-Wert = 6 unterschiedlicher Molarität (0,10; 0,50; 1,00; 2,00; 3,00 mol/L) getestet
 - Es wurden Natriumphosphat-Puffer bestehend aus NaH_2PO_4 und NaOH oder H_3PO_4 mit einer Konzentration von 1 mol/L mit unterschiedlichen pH-Werten (2,00; 3,00; 4,00; 4,50; 5,00; 5,00; 6,00) getestet
- B) Optimierung der Lösungsmittelunterdrückung
- Aufarbeitung einer Kaffeeprobe nach Kapitel 8.3.3 A)
 - Mischen von 600 μL membranfiltriertem Extrakt mit 100 μL Natriumphosphat-Puffer (1M, pH = 6,00) und 70 μL TSP in D_2O (1 mg/mL)
 - Messung der Kaffeeprobe mit den in Tabelle 8-3 beschriebenen Experimenten zur Lösungsmittelunterdrückung der Firma Bruker
 - Vor jeder Messung erfolgte eine Kalibrierung der Pulslänge des Anregungspulses

Tabelle 8-3: ¹H NMR Akquisitions- und Prozessierungsparameter der verwendeten Experimente zur Optimierung der Lösungsmittelunterdrückung. TD: Time Domain; DS: Anzahl Dummy Scans; NS: Anzahl Scans; AQ: Akquisitionszeit; O1: Sendefrequenz Offset; D1: Relaxationsverzögerung; D1-D19: verschiedene Relaxationsverzögerung; P0-P27: Pulse; GPZ: Gradientenparameter in Richtung z; SI: Size of Real spectrum; RG: Receiver Gain; WDW: Window Function; EM: Exponential Function; LB: Line Broadening Factor

Referenzspektrum: zg			
NMR-Spektrometer	Ascend 2	D1 [s]	4
Pulsprogramm	zg	P1 [μs]	9
TD	131072	Temperatur [K]	300
DS	16	Solvent	H ₂ O+D ₂ O
NS	64	Digmod	baseopt
AQ [s]	7,97	RG	3,99
Spektrale Breite	20,5617		
O1 [ppm]	4,702		
Vorsättigung: zgpr			
NMR-Spektrometer	Ascend 2	D1 [s]	4
Pulsprogramm	zgpr	D12 [μs]	20
TD	131072	P1 [μs]	9,31
DS	16	PLdB[1] [dB]	-9,89
NS	64	PLdB[9] [dB]	50,91
AQ [s]	7,97	Temperatur [K]	300
Spektrale Breite	20,5617	Solvent	H ₂ O+D ₂ O
O1 [ppm]	4,702	Digmod	baseopt
		RG	79,08
Vorsättigung: zgcppr			
NMR-Spektrometer	Ascend 2	D1 [s]	4
Pulsprogramm	zgpr	D12 [μs]	20
TD	131072	P1 [μs]	912
DS	16	PLdB[1] [dB]	-9,89
NS	64	PLdB[9] [dB]	50,91
AQ [s]	7,97	Temperatur [K]	300
Spektrale Breite	20,5617	Solvent	H ₂ O+D ₂ O
O1 [ppm]	4,702	Digmod	baseopt
		RG	79,08
Vorsättigung: NOESY			
NMR-Spektrometer	Ascend 2	D1 [s]	4
Pulsprogramm	noesygppr1d	D8 [ms]	10
TD	131072	D12 [μs]	20
DS	16	D16 [μs]	200
NS	64	P0 [μs]	9,12
AQ [s]	7,97	P1 [μs]	9,12
Spektrale Breite	20,5617	P16 [μs]	1000
O1 [ppm]	4,702	PLdB[1] [dB]	-9,89
Temperatur [K]	300	PLdB[9] [dB]	50,91
Solvent	H ₂ O+D ₂ O	RG	79,08
Digmod	baseopt		

Fortsetzung von Tabelle 8-3

Vorsättigung: NOESY mit Nutzung der D7-Zeit			
NMR-Spektrometer	Ascend 2	D1 [s]	4
Pulsprogramm	noesygppr1d_d7.eba	D7 [s]	4
TD	131072	D8 [ms]	10
DS	16	D12 [μs]	20
NS	64	D16 [μs]	200
AQ [s]	7,97	P0 [μs]	9,12
Spektrale Breite	20,5617	P1 [μs]	9,12
O1 [ppm]	4,702	P16 [μs]	1000
Temperatur [K]	300	PLdB[1] [dB]	-9,89
Solvent	H ₂ O+D ₂ O	PLdB[9] [dB]	50,91
Digmod	baseopt	RG	79,08
WET			
NMR-Spektrometer	Ascend 2	D1 [s]	4
Pulsprogramm	wet	D12 [μs]	20
TD	131072	D13 [μs]	4
DS	16	D16 [μs]	20000
NS	64	P1 [μs]	9,12
AQ [s]	7,97	P11 [μs]	20000
Spektrale Breite	20,5617	P16 [μs]	1000
O1 [ppm]	4,702	PLdb [0]	1000
Temperatur [K]	300	PLdB[1] [dB]	-9,89
Solvent	H ₂ O+D ₂ O	PLdB[9] [dB]	50,91
Digmod	baseopt	RG	3,99
WATERGATE: p3919			
NMR-Spektrometer	Ascend 2	D1 [s]	4
Pulsprogramm	p3919	D16 [μs]	200
TD	131072	D19 [μs]	520
DS	16	P1 [μs]	9,12
NS	64	P16 [μs]	1000
AQ [s]	7,97	P27 [μs]	9,12
Spektrale Breite	20,5617	GPZ1 [%]	20
O1 [ppm]	4,702	RG	79,08
Temperatur [K]	300		
Solvent	H ₂ O+D ₂ O		
Digmod	baseopt		
WATERGATE: zggpwwg			
NMR-Spektrometer	Ascend 2	D1 [s]	4
Pulsprogramm	zggpwwg	D12 [μs]	20
TD	131072	D16 [μs]	200
DS	16	P1 [μs]	9,13
NS	64	P2 [μs]	18,26
AQ [s]	7,97	P11 [μs]	1000
Spektrale Breite	20,5617	P16 [μs]	1000
O1 [ppm]	4,702	GPZ1 [%]	20
Temperatur [K]	300	RG	79,08
Solvent	H ₂ O+D ₂ O		
Digmod	baseopt		

Fortsetzung von Tabelle 8-3

Excitation Sculpting: zggpw5			
NMR-Spektrometer	Ascend 2	D1 [s]	4
Pulsprogramm	Zggpw5	D16 [μs]	200
TD	131072	D19 [μs]	520
DS	16	P1 [μs]	9,13
NS	64	P16 [μs]	1000
AQ [s]	7,97	P27 [μs]	9,13
Spektrale Breite	20,5617	GPZ1 [%]	34
O1 [ppm]	4,702	GPZ2 [%]	22
Temperatur [K]	300	RG	79,08
Solvent	H ₂ O+D ₂ O		
Digmod	baseopt		
Excitation Sculpting: zgesgp			
NMR-Spektrometer	Ascend 2	D1 [s]	4
Pulsprogramm	zgesgp	D12 [μs]	20
TD	131072	D16 [μs]	200
DS	16	P1 [μs]	9,12
NS	64	P2 [μs]	18,24
AQ [s]	7,97	P12 [μs]	2000
Spektrale Breite	20,5617	P16 [μs 9]	1000
O1 [ppm]	4,702	GPZ1 [%]	31
Temperatur [K]	300	GPZ2 [%]	11
Solvent	H ₂ O+D ₂ O	RG	79,08
Digmod	baseopt		
Prozessierungsparameter			
Software	TopSpin 3.7	LB	0,3
SI	262144	Phasenkorrektur	manuell
WDW	EM	Basislinienkorrektur	manuell

C) Optimierung der Leistung des Unterdrückungspulses

- Aufarbeitung einer Kaffeeprobe nach Kapitel 8.3.3 A)
- Mischen von 600 μL membranfiltriertem Extrakt mit 100 μL Natriumphosphat-Puffer (1M, pH = 6,00) und 70 μL TSP in D₂O (1 mg/mL)
- Messung der Kaffeeprobe mittels NOESY-Experiment unter der Verwendung der *D7*-Zeit (Tabelle 8-4) und verschiedenen Leistungen des Unterdrückungspulses (10-50 Hz)
- Vor jeder Messung erfolgte eine Kalibrierung der Pulslänge des Anregungspulses

D) Optimierung der Relaxationsverzögerung

- Aufarbeitung einer Kaffeeprobe nach Kapitel 8.3.3 A)
- Mischen von 600 μL membranfiltriertem Extrakt mit 100 μL Natriumphosphat-Puffer (1M, pH = 6,00) und 70 μL TSP in D₂O (1 mg/mL)
- Messung der Kaffeeprobe mittels NOESY-Experiment ohne der Verwendung der *D7*-Zeit (Tabelle 8-4) und verschiedenen *D1*-Zeiten (0,1-40 s)
- Vor jeder Messung erfolgte eine Kalibrierung der Pulslänge des Anregungspulses
- Ermittlung der *D1*-Zeit durch Auftragen des Integrals gegen die *D1*-Zeit

E) Optimierung der Messtemperatur

- Aufarbeitung einer Kaffeeprobe nach Kapitel 8.3.3 A)
- Mischen von 600 μL membranfiltriertem Extrakt mit 100 μL Natriumphosphat-Puffer (1M, pH = 6,00) und 70 μL TSP in D_2O (1 mg/mL)
- Messung der Kaffeeprobe mittels NOESY-Experiment unter der Verwendung der $D7$ -Zeit (Tabelle 8-4) und verschiedenen Messtemperaturen (280 K, 290 K, 300 K)
- Vor jeder Messung erfolgte eine Kalibrierung der Pulslänge des Anregungspulses

F) Überprüfung der Stabilität der Messlösung

- Aufarbeitung einer Kaffeeprobe nach Kapitel 8.3.3 A)
- Mischen von 600 μL membranfiltriertem Extrakt mit 100 μL Natriumphosphat-Puffer (1M, pH = 6) und 70 μL TSP in D_2O (1 mg/mL).
- Messung der Kaffeeprobe mit dem in Tabelle 8-4 dargestellten Experiment in einem Zeitfenster von 2 Wochen. Innerhalb der ersten 24 h erfolgt die Messung der Kaffeeprobe stündlich, anschließend jeweils nach zwölf h.

8.3.5 Entwickelte Methode für die nicht zielgerichtete Untersuchung der hydrophilen Metaboliten

- Aufarbeitung der Kaffeeproben nach Kapitel 8.3.3
- Mischen von 600 μL membranfiltriertem Extrakt mit 100 μL Natriumphosphat-Puffer (1M, pH = 6,00) und 70 μL TSP in D_2O (1 mg/mL)
- Akquisitions- und Prozessierungsparameter laut Tabelle 8-4
- Überprüfung der Qualität der Spektren durch Ermittlung der Halbwertsbreite des TSP-Signals. Maximal erlaubte Halbwertsbreite des TSP-Signals: 1,2 Hz
- Proben außerhalb dieser Spezifikation wurden wiederholt
- Auf jedes Spektrum wurde eine automatische Phasen- und Basislinienkorrektur angewendet

Tabelle 8-4: ^1H -NMR-Akquisitions- und Prozessierungsparameter für die nicht zielgerichtete Methode zur Untersuchung der hydrophilen Metaboliten. SI: Size of Real spectrum; RG: Receiver Gain; WDW: Window Function; EM: Exponential Function; LB: Line Broadening Factor

^1H-NMR Akquisitionsparameter			
NMR-Spektrometer	Ascend 2	D1 [s]	4
Pulsprogramm	noesygppr1d_d7.eba	D7 [s]	4
TD	131072	D8 [ms]	10
DS	16	D12 [μs]	20
NS	64	D16 [μs]	200
AQ [s]	7,97	P0 [μs]	9,12
Spektrale Breite [ppm]	20,5617	P1 [μs]	9,12
O1 [ppm]	4,702	P16 [μs]	1000
Temperatur [K]	300	PLdB[1] [dB]	-9,89
Solvent	$\text{H}_2\text{O}+\text{D}_2\text{O}$	PLdB[9] [dB]	50,91
Digmod	baseopt		
^1H-NMR Prozessierungsparameter			
Software	TopSpin 3.7	LB	0,3
SI	262144	Phasenkorrektur	apk0.noe
WDW	EM	Basislinienkorrektur	aplbc.c

8.4 Methoden zur multivariaten Datenanalyse

8.4.1 Datenvorbehandlung für die Durchführung der PCA-LDA

Die Datenvorbehandlung erfolgte mittels eines vorgefertigten Skriptes in der MATLAB® Version R2019b (MathWorks, Natick, MA, USA), welches von der Firma Bruker Biospin (Rheinstetten, Deutschland) zur Verfügung gestellt wurde. Das vollständige Skript ist der digitalen Version dieser Arbeit beigelegt. Es wurde ein Alignment, eine Normalisierung, eine Bucketierung und eine log-Transformation durchgeführt, wobei die jeweiligen Buckets mit den Lösungsmittelsignalen ausgeschlossen wurden (Tabelle 8-5)

Tabelle 8-5: Parameter für die Datenvorbehandlung der $^1\text{H-NMR}$ -Spektren für die PCA-LDA der lipophilen und hydrophilen Metaboliten.

	Methode für die Analyse der lipophilen Metaboliten	Methode für die Analyse der hydrophilen Metaboliten
Bereich für die Bucketierung [ppm]	0,50 – 9,80	0,50 – 9,80
Anzahl der Buckets	2000	1000
Normalisierung	3,00 – 3,05	-0,10 – 0,10
Ausschlüsse [ppm]	8,20 – 8,53; 7,58 – 7,64; 4,60 – 4,95; 3,60 – 3,67; 3,30 – 3,39; 3,00 – 3,10; 1,18 – 1,23	4,72 – 4,88
Transformation	log-Transformation mit log-Faktor = 1	log-Transformation mit log-Faktor = 1

- A) Überprüfung der Minimierung der biologischen Variation durch Normalisierung der $^1\text{H-NMR}$ -Spektren der lipophilen Metaboliten auf das jeweilige Gewicht der von dem Lösungsmittel befreiten Probe
- Jeweiliges $^1\text{H-NMR}$ -Spektrum der lipophilen Metaboliten nach der Normalisierung auf den Bereich 3,00 – 3,05 ppm durch die im Kapitel 8.2.5 notierte Auswaage der entsprechenden Probe teilen
- B) Überprüfung der Minimierung der biologischen Variation durch Normalisierung der $^1\text{H-NMR}$ -Spektren der lipophilen und hydrophilen Metaboliten auf den jeweiligen Feuchtegehalt der Probe
- Jeweiliges $^1\text{H-NMR}$ -Spektrum der lipophilen Metaboliten nach der Normalisierung auf den Bereich 3,00 – 3,05 ppm durch den in Kapitel 8.1 bestimmten Feuchtegehalt der entsprechenden Probe teilen
 - Jeweiliges $^1\text{H-NMR}$ -Spektrum der hydrophilen Metaboliten nach der Normalisierung auf den Bereich 3,00 – 3,05 ppm durch den in Kapitel 8.1 bestimmten Feuchtegehalt der entsprechenden Probe teilen

8.4.2 Univariate Datenanalyse mittels Kruskal-Wallis Test

Die Durchführung des Kruskal-Wallis Tests (Tabelle 8-6) erfolgte mittels eines vorgefertigten Skriptes in der Matlab Version R2019b (MathWorks, Natick, MA, USA), welches von der Firma Bruker Biospin (Rheinstetten, Deutschland) zur Verfügung gestellt wurde. Das vollständige Skript ist der digitalen Version dieser Arbeit beigelegt.

Tabelle 8-6: Parameter für die Durchführung des Kruskal-Wallis Tests für die lipophilen und hydrophilen Metaboliten.

	Analyse der lipophilen Metaboliten	Analyse der hydrophilen Metaboliten
Bereich für die Bucketierung [ppm]	0,50 – 9,80	0,50 – 9,80
Anzahl der Buckets	1775	983
Normalisierung	3,00 – 3,05	-0,10 – 0,10
Korrigiertes Signifikanzniveau	$\alpha_{SID} = 2,89 \cdot 10^{-5}$	$\alpha_{SID} = 5,22 \cdot 10^{-5}$

8.4.3 Multivariate Datenanalyse mittels PCA-LDA

- PCA zur Dimensionsreduktion
- LDA zur Erstellung eines Klassifizierungsmodells
- Interne Validierung mittels zehnfacher MCKV und einem Prognoseintervall von 95%
- Externe Validierung mit 20% des Probandensatzes jeder Klasse. Die Proben wurden nicht zum Modelaufbau und zur MCKV verwendet. Datenvorbehandlung nach Kapitel 8.4.1.
- Es wurden folgende zwei-Klassen-Modelle zur Untersuchung des geographischen Ursprungs nach Kontinent gebildet: Afrika-Asien; Afrika-Südamerika; Afrika-Nordamerika; Asien-Südamerika; Asien-Nordamerika; Nordamerika-Südamerika.
- Es wurden folgende zwei-Klassen-Modelle zur Untersuchung des geographischen Ursprungs nach Länder gebildet: Brasilien-Afrika; Brasilien-Asien; Brasilien-Nordamerika; Äthiopien-Asien; Äthiopien-Südamerika; Äthiopien-Nordamerika; Kolumbien-Afrika; Kolumbien-Asien; Kolumbien-Nordamerika
- Es wurden folgende zwei-Klassen-Modelle zur Untersuchung des Anbaus gebildet: bio-konventionell
- Es wurden folgende zwei-Klassen-Modelle zur Untersuchung des Röstgrades gebildet: hell (Tonino-Wert >90,0) - dunkel (Tonino-Wert <80,0)
- Es wurden folgende zwei-Klassen-Modelle zur Untersuchung Röstverfahrens gebildet: Trommelröstung – Heißlufttröstung (Ikawa)

8.4.4 Datenfusionierung

A) Low-Level Datenfusionierung

- Datenvorbehandlung der ¹H-NMR-Daten nach Tabelle 8-5
- Fusionierung der Buckets der ¹H-NMR-Spektren der lipophilen und hydrophilen Metaboliten
- PCA zur Dimensionsreduktion
- LDA mit den Scores der ersten 29 bzw. 19 PC der PCA
- Interne Validierung mittels zehnfacher MCKV und einem Prognoseintervall von 95%

B) Mid-Level Datenfusionierung

- Datenvorbehandlung der ¹H-NMR-Daten nach Tabelle 8-5

- PCA jeweils der Buckets der ^1H -NMR-Spektren der hydrophilen und lipophilen ^1H -NMR-Spektren
- Fusionierung jeweils der ersten 29 bzw. 19 PC der PCA der Daten der lipophilen und hydrophilen Metaboliten
- LDA mit den fusionierten Scores der beiden PCA
- Interne Validierung mittels zehnfacher MCKV und einem Prognoseintervall von 95%

8.5 Identifizierung der Kaffeemetaboliten

A) Aufstockungsversuche

- Herstellung der Referenzlösungen mit einer Konzentration von 10 mg/mL
- Für die Untersuchung lipophiler Metaboliten Herstellung der Referenzlösungen in (Chloroform/Methanol, 1:1, v/v). Aufarbeitung und Messung einer Kaffeeprobe nach Kapitel 8.2.5. Zusatz von 0,1 mL Referenzlösung von dem Befreien des Lösungsmittels
- Für die Untersuchung der hydrophilen Metaboliten Herstellung der Referenzlösungen in H_2O . Aufarbeitung und Messung einer Kaffeeprobe nach Kapitel 8.3.5. Mischen von 600 μL membranfiltrierten Extrakt mit 100 μL Natriumphosphat-Puffer (1 mol/L, pH = 6,00), 80 μL TSP in D_2O (1 mg/mL) und 100 μL Referenzlösung

B) Strukturaufklärung mittels 2D-Experimente

- Es erfolgte eine manuelle Phasen- und eine automatische Basislinienkorrektur
- Messung mittels der in Tabelle 8-7 aufgeführten Experimente

Tabelle 8-7: Akquisitionsparameter der durchgeführten heteronuklearen 2D-NMR-Experimente. HSQC Hetero Nuklear Single Quantum Coherence; HSQC-TOCSY Hetero Nuklear Single Quantum Coherence with Total Correlation Spectroscopy; NUS Non Uniform Sampling

Akquisitionsparameter		
	F_2 -Dimension	F_1 -Dimension
Pulsprogramme	hsqcetpg (^1H - ^{13}C -HSQC)	
	hsqcgpmlph (^1H - ^{13}C -HSQC-TOCSY)	
Aufgenommene Datenpunkte	1024	64
Anzahl Dummy Scans	2048	
Anzahl Scans	512	
Temperatur [K]	300	
Spektrale Breite [ppm]	15,9793	200,3651
Akquisitionszeit [s]	0,080	0,002
D1-Zeit [s]	2	
NUS-Anteil [%]	50	
Senderfrequenz-Offset [ppm]	7,500	99,391
Prozessierungsparameter		
Software	TopSpin 3.7	
Size of Real Spectrum	2048	1024
Window Function	QSINE	QSINE
Algorithmus zur Wiederherstellung der Datenpunkte	Iterative Soft Thresholding (IST)	
Phasenkorrektur	manuell	
Basislinienkorrektur	automatisch	

8.6 Zielgerichtete Untersuchung zur Quantifizierung ausgewählter hydrophiler Kaffeemetaboliten

8.6.1 Referenzmethode: HPLC-DAD

Die Entwicklung der HPLC-DAD-Referenzmethode sowie die Vermessung der Proben mittels der entwickelten Methode erfolgte durch Frau Anabel Frölich im Rahmen eines Forschungspraktikums.

A) Probenvorbereitung

- 50 mg gemahlene Kaffeeprobe in 15 mL Zentrifugenröhrchen einwiegen, genaues Gewicht notieren
- Zugabe von 8 mL destilliertes Wasser
- Extraktion bei Raumtemperatur für 60 min im Reagenzglasschüttler (1600 U/min)
- Extrakt membranfiltrieren und in HPLC-Vials abfüllen

B) Messung mittels HPLC-DAD

- Messung von Trigonellin, Koffein, HMF, 3-CQA, 4-CQA und 5-CQA
- Säule: Synergi Polar/RP80 250•2 mm (Phenomenex, Aachen, Deutschland)
- Fluss: 0,3 mL/min (Gradient laut Tabelle 8-8)
- Injektionsvolumen: 5 µL
- Säulentemperatur: 35°C
- Detektor: Diodenarray (Messwellenlängen: 260 nm für NMP, 270 nm für Trigonellin, HMF, Koffein und 330 nm für 3-CQA, 4-CQA und 5-CQA)
- Kalibrierung: Wässrige Lösung der Analyten Trigonellin, Koffein, 3-CQA, 4-CQA und 5-CQA im Konzentrationsbereich von 0,50 – 205 mg/L
- Auswertung über eine lineare Anpassung der Kalibriergeraden

C) Validierung

- Richtigkeit: Anhand der Wiederfindung durch Aufstockung einer Stammlösung von Koffein, Trigonellin, HMF, 3-CQA, 4-CQA und 5-CQA in acht verschiedenen Konzentrationen.
- Linearität: Bestimmung durch lineare Regression und Residuenanalyse der acht Aufstockungen.
- NG, BG: Bestimmung nach DIN 32645 anhand der acht Aufstockungen
- Präzision: Fünffache Aufarbeitung einer Kaffeeprobe von zwei verschiedenen Personen an unterschiedlichen Tagen

Tabelle 8-8: Gradientenprogramm der Hochleistungsflüssigkeitschromatographie mit Diodenarray-Detektor (HPLC-DAD) zur Analyse von Trigonellin, Koffein, 5-Hydroxymethylfurfural, 3-Caffeoylchinasäure, 4-Caffeoylchinasäure und 5-Caffeoylchinasäure. Eluent A 0,3% Natriumformiatpuffer pH = 2,4; Eluent B Methanol

Zeitfenster [min]	Anteil Eluent A [%]	Anteil Eluent B [%]
0-5	100	0
5-20	1	99
20-25	1	99
25-25,5	100	0
25,5-35	100	0

8.6.2 Amtliche Untersuchungsmethode L 46.00-3 für die Bestimmung von Koffein mittels HPLC-DAD

Zur Quantifizierung von Koffein und der Ermittlung eventueller Korrekturfaktoren wurden die amtliche Methode „Untersuchung von Kaffee und Kaffee-Erzeugnissen. Bestimmung des Koffeingehaltes mittels HPLC-Referenzverfahren“ auf Kaffeeproben angewandt. Die Untersuchung der Kaffeeproben mittels amtlicher Methode wurde von Frau Anabel Frölich im Rahmen eines Forschungspraktikums untersucht.

A) Probenvorbereitung

- Einwaage von 1 g gemahlene Kaffee in 250 mL Erlenmeyerkolben und Notieren des genauen Gewichtes
- Zugabe von 5 g Magnesiumoxid und 150 mL destilliertes Wasser
- Für 20 min im Schüttelbad bei 90°C extrahieren
- Quantitative Überführung des Extraktes in 250 mL Messkolben
- Nach Temperieren auf 20°C Auffüllen des Messkolbens bis zur Marke
- Aliquot des Extraktes abfüllen und in ein HPLC-Vial abfüllen.

B) Messung mittel HPLC-DAD

- Messung von Koffein
- Säule: Synergi Polar/RP80 250•2 mm (Phenomenex, Aachen, Deutschland)
- Fluss: 1,0 mL/min (Eluent: 24% Methanol in Wasser (v/v))
- Injektionsvolumen: 10 µL
- Säulentemperatur: 35°C
- Detektor: Diodenarray (Messwellenlänge: 272 nm)
- Kalibrierung: Wässrige Lösung des Analyten Koffein im Konzentrationsbereich von 0,0005 – 0,205 mg/mL
- Auswertung über eine lineare Anpassung der Kalibriergeraden

8.6.3 ¹H-qNMR-spektroskopische Methode

A) Probenvorbereitung und Erfassung der Metaboliten nach Kapitel 8.3.5

- Referenzierung der ¹H-NMR-Spektren auf TSP ($\delta_H = 0,00$ ppm)

B) Quantifizierung

- Quantifizierung nach PULCON
- Konzentrationen der Standardsubstanzen in der QuantRef: Zitronensäure (4,24 g/L); Milchsäure (0,92 g/L)
- Messung der QuantRef mit den gleichen Akquisitions- und Prozessierungsparametern zu Beginn jeder Messserie
- Berechnung des ERETIC-Faktors und der Analyt-Konzentration nach Formel 13 und 14 in Kapitel 3.7
- Berechnung der Konzentration durch Integration der in Tabelle 8-9 aufgeführten Signale und Parameter
- Anwendung eines Line-Shape-Fitting Verfahrens für CQA, Koffein, Essigsäure und Milchsäure
- Die Anwendung der Line-Shape-Fitting Verfahren und die automatische Auswertung erfolgte mittels eines vorgefertigten Skriptes in der MATLAB®

Version R2019b (MathWorks, Natick, MA, USA), welches von der Firma Bruker Biospin (Rheinstetten, Deutschland) zur Verfügung gestellt wurde. Das vollständige Skript ist der digitalen Version dieser Arbeit beigelegt.

- Berechnung des VF aufgrund des Mischens von 600 µL membranfiltriertem Extrakt mit 100 µL Natriumphosphat-Puffer (1M, pH = 6,00) und 70 µL TSP in D₂O (1 mg/mL): $VF = 600 \mu\text{L} / 770 \mu\text{L} = 0,78$.

Tabelle 8-9: Parameter für die Quantifizierung der einzelnen Kaffeemetaboliten mittels PULCON. HMF 5-Hydroxymethylfurfural, NMP N-Methylpyridinium, CQA Isomere der Caffeoylchinasäure

Analyt	Molare Masse[g/mol]	Anzahl Protonen	Integrations-Bereich [ppm]	Korrekturfaktor*
HMF	126,11	1	9,440-9,480	1,0
Trigonellin	137,14	1	9,080-9,170	1,1*
NMP	94,13	1	8,490-8,600	1,0
Ameisensäure	46,03	1	8,420-8,490	1,2*
CQA***	354,31	1	6,890-7,010	1,0
Koffein	194,19	3	3,315-3,362	1,2**
Essigsäure	60,05	3	1,915-1,940	1,1*
Milchsäure	90,08	3	1,320-1,490	1,0
Referenzsubstanzen in der QuantRef				
Zitronensäure	192,12	2	2,920-3,026	-
Milchsäure	90,08	3	1,315-1,360	-

* Korrekturfaktor ermittelt aus T1-Zeit Bestimmung und Überprüfung des NOE (Kapitel 8.3.4)

** Korrekturfaktor von Koffein ermittelt zusätzlich durch Vergleichsuntersuchungen mittels Hochleistungsflüssigkeitschromatographie (Kapitel 8.6.1 A)

*** Bestimmung von 3-, 4- und 5-Caffeoylchinasäure als Summenparameter CQA

C) Validierungsparameter: Richtigkeit

- Herstellung von Stammlösungen für Ameisensäure, Essigsäure, Milchsäure, Koffein, NMP, HMF, CQA
- Zusatz der Stammlösungen in acht verschiedenen Konzentrationen zu jeweils einer Kaffeeprobe
- Aufarbeitung der Kaffeeprobe nach Kapitel 8.3.5
- Aufarbeitung einer Probe ohne Zusatz der Stammlösungen nach Kapitel 8.3.5 (Leerprobe)
- Konzentration der Analyten in den Aufstockungen nach Tabelle 8-9

Tabelle 8-10: Konzentrationen der Aufstockungen in g/kg der zu validierenden Kaffeemetaboliten. Dot Aufstockung; HMF 5-Hydroxymethylfurfural; NMP N-Methylpyridinium; CQA Isomere der Caffeoylchinasäure

Analyt	Dot1	Dot2	Dot3	Dot4	Dot5	Dot6	Dot7	Dot8
HMF	0,20	0,41	0,61	0,82	1,02	1,22	1,43	1,63
Trigonellin	1,87	3,75	7,50	11,2	15,0	18,7	22,5	26,2
NMP	0,21	0,42	0,84	1,68	3,37	5,05	6,74	8,42
Ameisensäure	0,87	1,75	2,62	3,50	4,37	5,24	6,12	6,99
CQA*	4,04	8,08	12,1	16,2	24,2	32,3	40,4	48,5
Koffein	2,35	4,70	9,40	14,1	18,8	23,5	28,2	32,9
Essigsäure	0,91	1,83	3,66	5,49	7,32	9,15	11,0	12,8
Milchsäure	0,23	0,46	0,91	1,82	2,74	3,65	4,56	9,12

* Zusatz von CQA in Form von 5-Caffeoylchinasäure (5-CQA)

D) Validierungsparameter: Präzision

- Messpräzision: Eine Kaffeeprobe nach Kapitel 8.3.5 aufarbeiten und fünfmal vermessen. Dabei wurden die Temperatureinstellung, das Locken, das Shimming und das Matching und Tuning von jeder Messung durchgeführt
- Wiederholpräzision: Aufarbeitung einer Kaffeeprobe nach Kapitel 8.3.5 fünfmal von einer Person (Person A)
- Vergleichspräzision: Aufarbeitung einer Kaffeeprobe nach Kapitel 8.3.5 fünfmal jeweils von zwei verschiedenen Personen (Person A und Person B)
- Laborpräzision: Aufarbeitung einer Kaffeeprobe nach Kapitel 8.3.5 fünfmal jeweils von zwei verschiedenen Personen (Person A und Person B) an verschiedenen Tagen

E) Linearität

- Zur Bestimmung der Linearität wurden die ¹H-NMR-Spektren der in Kapitel 8.6.3 C) aufgearbeiteten Aufstockungen verwendet.

F) Wiederfindungsrate

- Zur Bestimmung der Wiederfindungsrate wurden die ¹H-NMR-Spektren der in Kapitel 8.6.3 C) aufgearbeiteten Aufstockungen verwendet.

G) Robustheit

- Aufarbeitung einer Kaffeeprobe nach Kapitel 8.3.5 fünfmal und Messung der Proben an jeweils zwei verschiedenen NMR-Geräten

H) NG und BG

- Zur Bestimmung der Nachweis- und Bestimmungsgrenze wurden die ¹H-NMR-Spektren der in Kapitel 8.6.3 C) aufgearbeiteten Aufstockungen verwendet

9 Anhang

9.1 Probenmaterial

Tabelle 9-1: Im Rahmen dieser Arbeit verwendeten Kaffeeproben mit den Metadaten über geographischen Ursprung, Anbaumethode, Röstverfahren, Wassergehalt, Fettauswaage und Röstgrad. Pr.-Nr.: Proben-Nummer; AF: Afrika; AS: Asien; NA: Nordamerika; SA: Südamerika; Ind: Indonesien; El Sal.: El Salvador; Guat.: Guatemala; Costa R.: Costa Rica; Nicarag.: Nicaragua; Galapag.: Galapagos; Kolum.: Kolumbien; Kon: Konventionell; HWB: Halbwertsbreite; TMS: Tetramethylsilan; TSP: Trimethylsilylpropionat

Pr.-Nr.	Kontinent	Land	Anbaumethode	Röstgrad	Fettauswaage [mg]	Feuchtegehalt [%]	Röstverfahren	HWB TMS [Hz]	HWB TTSP [Hz]
AF100	AF	Kenia	Kon.	94,5	22,3	4,6	Trom-	0,9	0,8
AF101	AF	Kenia	Kon.	101,3	23,3	3,1	Trom-	0,8	0,9
AF102	AF	Kenia	Kon.	95,7	21,1	3,5	Trom-	0,8	0,8
AF104	AF	Kenia	Kon.	107,9	20,8	4,2	Trom-	0,8	0,8
AF105	AF	Kenia	Kon.	106,8	23,8	3,6	Trom-	0,8	0,8
AF106	AF	Kenia	Bio	80,7	22,7	1,8	Trom-	0,7	0,9
AF107	AF	Kenia	Kon.	74,1	24,2	1,8	Trom-	0,8	0,8
AF10	AF	Äthiopien	Kon.	112,8	23,8	2,5	Trom-	0,8	0,9
AF11	AF	Äthiopien	Kon.	70,6	26,1	3,0	Trom-	0,8	0,8
AF12	AF	Äthiopien	Kon.	110,9	23,9	5,2	Trom-	0,8	0,8
AF13	AF	Äthiopien	Kon.	101,7	21,6	3,3	Trom-	0,8	0,8
AF14	AF	Äthiopien	Kon.	100,8	21,4	3,3	Trom-	0,8	0,8
AF150	AF	Tansania	Kon.	77,8	20,6	4,2	Trom-	0,8	0,8
AF153	AF	Tansania	Kon.	112,8	19,4	3,4	Trom-	0,8	0,8
AF155	AF	Tansania	Kon.	83,3	23,6	2,0	Trom-	0,8	0,9
AF15	AF	Äthiopien	Bio	88,8	23,3	4,5	Trom-	0,8	0,8
AF15a	AF	Äthiopien	Bio	76,1	22,9	4,8	Trom-	0,8	0,8
AF15b	AF	Äthiopien	Bio	86,9	23,4	4,3	Trom-	0,8	0,8
AF15f	AF	Äthiopien	Bio	101,8	23,4	4,9	Trom-	0,8	0,8
AF15g	AF	Äthiopien	Bio	99,5	23,3	4,1	Trom-	0,8	0,8
AF16	AF	Äthiopien	Bio	84,3	21,2	4,8	Trom-	0,8	0,8
AF16a	AF	Äthiopien	Bio	86,5	21,1	4,5	Trom-	0,8	0,8
AF17	AF	Äthiopien	Kon.	102,2	20,8	3,1	Trom-	0,9	0,8
AF18	AF	Äthiopien	Kon.	95,4	22,9	3,9	Trom-	0,8	0,8
AF19	AF	Äthiopien	Kon.	84,8	22,9	4,2	Trom-	0,8	0,8
AF200	AF	Uganda	Kon.	82,5	20,5	4,0	Trom-	0,9	0,8
AF201	AF	Uganda	Kon.	101,7	19,9	4,6	Trom-	0,9	0,8
AF203	AF	Uganda	Kon.	97,9	22,1	3,7	Trom-	0,8	0,8
AF204	AF	Uganda	Kon.	73,4	23,9	1,8	Trom-	0,9	0,8
AF207	AF	Uganda	Kon.	82,7	20,2	1,2	Trom-	0,8	0,8
AF208	AF	Uganda	Bio	87,3	22,9	2,7	Trom-	0,9	0,8
AF20	AF	Äthiopien	Kon.	109,8	21,9	3,1	Trom-	0,9	0,8
AF21	AF	Äthiopien	Kon.	100,1	22,9	3,6	Trom-	0,8	0,9
AF22	AF	Äthiopien	Kon.	83,6	21,6	4,4	Trom-	0,9	0,8
AF22a	AF	Äthiopien	Kon.	85,2	18,4	4,6	Trom-	0,7	1,0
AF23	AF	Äthiopien	Kon.	72,9	22,6	3,2	Trom-	0,8	0,9
AF24	AF	Äthiopien	Kon.	89,0	24,4	3,3	Trom-	0,8	0,9
AF24a	AF	Äthiopien	Kon.	88,8	20,9	2,5	Trom-	0,8	0,9
AF250	AF	Malawi	Kon.	105,0	19,6	3,4	Trom-	0,8	0,8
AF25	AF	Äthiopien	Kon.	94,7	21,9	2,7	Trom-	0,8	0,9
AF26	AF	Äthiopien	Kon.	103,6	19,6	1,3	Trom-	0,7	0,8
AF26a	AF	Äthiopien	Kon.	94,8	20,3	4,2	Trom-	0,8	0,8
AF27	AF	Äthiopien	Bio	89,0	15,5	1,0	Trom-	0,8	0,8
AF28	AF	Äthiopien	Bio	110,2	22,2	3,3	Trom-	0,7	0,8
AF29	AF	Äthiopien	Kon.	105,3	21,7	5,3	Trom-	38,9	0,9
AF30	AF	Äthiopien	Bio	96,6	20,4	2,1	Trom-	0,8	0,9
AF31	AF	Äthiopien	Kon.	86,2	20,3	2,3	Trom-	0,8	0,9

Anhang

AF31a	AF	Äthiopien	Kon.	80,0	20,0	2,4	Trom-	0,7	0,9
AF32	AF	Äthiopien	Kon.	93,7	21,9	2,8	Trom-	0,9	0,9
AF33	AF	Äthiopien	Kon.	111,8	20,9	2,5	Trom-	0,8	0,8
AF33a	AF	Äthiopien	Kon.	111,3	21,4	1,8	Trom-	0,9	0,9
AF34	AF	Äthiopien	Kon.	97,0	23,6	4,3	Trom-	0,8	0,9
AF350	AF	Ruanda	Kon.	79,4	22,5	2,5	Trom-	0,8	0,9
AF350a	AF	Ruanda	Kon.	114,3	21,2	2,6	Trom-	0,9	0,9
AF351	AF	Ruanda	Kon.	80,4	22,9	2,8	Trom-	0,8	0,9
AF352	AF	Ruanda	Kon.	86,2	24,9	3,8	Trom-	0,8	0,8
AF35	AF	Äthiopien	Kon.	100,3	21,9	1,7	Trom-	0,8	0,9
AF36	AF	Äthiopien	Kon.	106,4	20,9	2,1	Trom-	0,8	0,9
AF37	AF	Äthiopien	Kon.	90,6	22,6	3,8	Trom-	0,8	0,8
AF38	AF	Äthiopien	Bio	88,3	22,9	1,3	Trom-	0,9	0,8
AF39	AF	Äthiopien	Kon.	90,4	22,4	1,1	Trom-	0,8	0,8
AF3	AF	Äthiopien	Bio	89,4	19,5	2,8	Trom-	0,8	0,8
AF400	AF	Kongo	Bio	85,9	18,3	2,9	Trom-	38,8	0,9
AF400a	AF	Kongo	Bio	79,7	18,7	2,6	Trom-	0,8	0,9
AF40	AF	Äthiopien	Kon.	74,4	24,1	4,3	Trom-	0,9	0,8
AF41	AF	Äthiopien	Kon.	92,5	21,5	1,1	Trom-	0,8	0,8
AF42	AF	Äthiopien	Bio	84,0	23,3	1,4	Trom-	1,1	0,8
AF43	AF	Äthiopien	Bio	108,6	16,4	3,8	Trom-	0,9	0,8
AF44	AF	Äthiopien	Bio	79,9	22,6	2,7	Trom-	0,8	0,8
AF450	AF	Kamerun	Kon.	75,3	21,4	2,8	Trom-	0,7	0,9
AF45	AF	Äthiopien	Bio	76,2	24,8	3,2	Trom-	1,3	0,8
AF46	AF	Äthiopien	Bio	72,7	24,3	3,4	Trom-	0,8	0,8
AF47	AF	Äthiopien	Kon.	88,6	23,1	2,0	Trom-	0,9	0,8
AF48	AF	Äthiopien	Kon.	86,6	24,7	1,5	Trom-	0,8	0,8
AF4	AF	Äthiopien	Bio	86,6	21,4	4,1	Trom-	0,8	0,8
AF50	AF	Burundi	Kon.	55,8	24,9	3,7	Trom-	0,8	0,8
AF51	AF	Burundi	Kon.	79,0	23,7	2,5	Trom-	0,8	0,8
AF52	AF	Burundi	Kon.	95,2	24,3	3,6	Trom-	0,8	0,8
AF53	AF	Burundi	Kon.	97,9	20,9	3,3	Trom-	0,8	0,9
AF54	AF	Burundi	Kon.	118,3	19,4	2,3	Trom-	0,7	0,8
AF55	AF	Burundi	Kon.	92,4	21,6	3,5	Trom-	0,8	0,9
AF56	AF	Burundi	Kon.	93,9	20,4	2,3	Trom-	0,9	1,0
AF5	AF	Äthiopien	Bio	88,2	23,3	4,3	Trom-	0,8	0,8
AF6	AF	Äthiopien	Kon.	105,6	20,9	4,1	Trom-	0,8	0,9
AF8	AF	Äthiopien	Kon.	113,7	21,5	4,3	Trom-	0,8	0,8
AF900	AF		Kon.	94,3	24,5	6,1	Trom-	0,8	0,8
AF900a	AF		Kon.	101,3	23,9	4,4	Trom-	0,7	0,8
AF901	AF		Kon.	90,8	23,8	4,2	Trom-	0,9	0,8
AF9	AF	Äthiopien	Kon.	68,9	22,1	2,2	Trom-	0,8	0,8
20BA039	AS	Indien		106,0	17,8	5,1	Trom-	0,6	0,8
20BA085	AS	Indien		106,0	17,4	4,6	Trom-	0,6	0,9
21BA001	AS	Indien		106,0	18,4	5,0	Trom-	0,7	0,8
21BA002	AS	Indien		106,0	19,6	5,0	Trom-	0,6	0,8
21BA003	AS	Indien		106,0	18,8	5,1	Trom-	0,7	0,8
21BA004	AS	Indien		106,0	19,5	5,3	Trom-	0,8	0,9
21BA005	AS	Indien		106,0	17,6	5,2	Trom-	0,7	0,9
21BA006	AS	Indien		106,0	17,4	5,2	Trom-	0,8	0,8
21BA007	AS	Indien		106,0	17,4	5,0	Trom-	0,6	0,8
21BA009	AS	Indien		106,0	16,3	5,4	Trom-	0,7	0,9
21BA00A	AS	Indien		106,0	16,5	5,1	Trom-	0,7	0,9
21BA00B	AS	Indien		106,0	16,8	4,9	Trom-	0,6	0,8
21BA011	AS	Indien		106,0	27,5	5,0	Trom-	0,6	0,9
21BA012	AS	Indien		106,0	18,6	5,0	Trom-	0,6	0,9
21BA013	AS	Indien		106,0	18,5	5,3	Trom-	0,8	0,8
21BA014	AS	Indien		106,0	18,7	5,5	Trom-	0,7	0,9
21BA015	AS	Indien		106,0	16,5	4,9	Trom-	0,7	0,9
21BA016	AS	Indien		106,0	17,6	4,9	Trom-	0,7	0,9
21BA017	AS	Indien		106,0	16,2	5,0	Trom-	0,6	0,9
21BA027	AS	Indien		106,0	17,9	4,9	Trom-	0,6	0,9

Anhang

AS101	AS	Sumatra	Kon.	104,1	21,2	2,0	Trom-	0,8	0,8
AS11	AS	Indien	Kon.	99,9	17,2	2,5	Trom-	0,9	0,8
AS12	AS	Indien	Kon.	106,6	17,4	1,6	Trom-	0,8	0,8
AS13	AS	Indien	Kon.	107,9	17,0	3,3	Trom-	0,8	0,8
AS14	AS	Indien	Kon.	111,8	17,9	2,9	Trom-	0,8	0,8
AS155	AS	Thailand	Kon.	84,6	19,5	2,7	Trom-	0,7	0,9
AS156	AS	Thailand	Kon.	91,5	19,6	3,1	Trom-	0,7	0,9
AS15	AS	Indien	Kon.	114,9	17,1	3,3	Trom-	0,9	0,8
AS163	AS	Thailand	Kon.	81,4	19,7	2,5	Trom-	0,8	0,8
AS164	AS	Thailand	Kon.	106,8	18,9	3,9	Trom-	0,8	0,8
AS165	AS	Thailand	Kon.	91,7	19,1	2,9	Trom-	0,7	0,8
AS166	AS	Thailand	Kon.	110,0	18,7	3,9	Trom-	0,7	0,8
AS16	AS	Indien	Kon.	101,6	17,2	4,9	Trom-	0,8	0,8
AS16a	AS	Indien	Kon.	103,7	16,6	4,6	Trom-	0,6	0,8
AS19	AS	Indien	Kon.	108,0	16,6	2,3	Trom-	0,8	0,8
AS200	AS	Jemen	Kon.	90,6	23,8	3,6	Trom-	0,8	0,8
AS20	AS	Indien	Kon.	105,4	17,2	2,5	Trom-	0,8	0,8
AS21	AS	Indien	Bio	104,6	18,3	1,4	Trom-	0,8	0,8
AS22	AS	Indien	Kon.	110,9	16,7	1,5	Trom-	0,8	0,9
AS23	AS	Indien	Kon.	108,9	17,8	1,8	Trom-	0,8	0,9
AS23a	AS	Indien	Kon.	111,8	17,4	2,2	Trom-	0,8	0,9
AS250	AS	Nepal	Bio	99,1	20,8	3,5	Trom-	0,8	0,8
AS251	AS	Nepal	Bio	78,9	20,6	2,4	Trom-	0,8	0,9
AS252	AS	Nepal	Bio	102,5	19,1	1,3	Trom-	0,7	0,8
AS25	AS	Indien	Kon.	96,3	19,0	1,9	Trom-	0,9	0,9
AS26	AS	Indien	Kon.	89,4	17,4	3,7	Trom-	0,7	0,9
AS27	AS	Indien	Kon.	101,1	18,8	1,6	Trom-	0,8	0,8
AS28	AS	Indien	Kon.	112,3	17,3	3,4	Trom-	0,8	0,8
AS29	AS	Indien	Kon.	105,6	18,5	1,4	Trom-	0,8	0,8
AS300	AS	Myanmar	Kon.	92,5	18,8	3,6	Trom-	0,8	0,8
AS400	AS	Vietnam	Kon.	84,0	19,8	1,6	Trom-	0,7	0,9
AS401	AS	Vietnam	Kon.	108,1	20,3	2,6	Trom-	0,8	0,8
AS450	AS	Laos	Bio	81,1	19,9	2,8	Trom-	0,7	0,8
AS50	AS	Ind.	Kon.	86,2	20,6	1,9	Trom-	0,8	0,8
AS51	AS	Ind.	Kon.	67,0	22,7	2,6	Trom-	0,8	0,8
AS52	AS	Ind.	Kon.	103,7	21,9	2,7	Trom-	0,8	0,8
AS53	AS	Ind.	Kon.	102,0	19,9	3,1	Trom-	0,8	0,9
AS54	AS	Ind.	Kon.	108,8	20,3	3,2	Trom-	0,8	0,8
AS55	AS	Ind.	Kon.	101,4	21,6	3,2	Trom-	0,8	0,8
AS56	AS	Ind.	Kon.	85,2	21,5	2,5	Trom-	0,7	0,8
AS57	AS	Ind.	Kon.	91,5	19,3	1,7	Trom-	0,8	0,8
AS58	AS	Ind.	Kon.	99,4	20,2	2,2	Trom-	0,8	0,8
AS59	AS	Ind.	Kon.	72,2	20,1	3,4	Trom-	0,8	0,8
AS60	AS	Ind.	Kon.	99,4	20,4	1,6	Trom-	0,8	0,9
AS61	AS	Ind.	Bio	82,7	19,0	1,7	Trom-	0,7	0,8
AS62	AS	Ind.	Kon.	111,5	20,1	2,2	Trom-	0,8	0,8
19BE007	NA	El Sal.		110,0	22,5	2,8	Trom-	0,8	1,0
19BE011	NA	El Sal.		110,0	21,8	2,6	Trom-	0,6	0,9
19BE013	NA	El Sal.		110,0	20,6	2,3	Trom-	0,8	0,8
19BE021	NA	El Sal.		110,0	22,9	2,4	Trom-	0,6	0,8
19BE022	NA	El Sal.		110,0	22,4	2,0	Trom-	0,7	0,8
19BE025	NA	El Sal.		110,0	18,3	2,6	Trom-	0,8	0,9
19BE027	NA	El Sal.		110,0	21,4	2,9	Trom-	0,6	0,9
19BE031	NA	El Sal.		110,0	22,5	2,0	Trom-	0,7	0,8
19BE032	NA	El Sal.		110,0	22,0	2,7	Trom-	0,8	0,9
19BE033	NA	El Sal.		110,0	22,4	1,8	Trom-	0,6	0,9
19BE034	NA	El Sal.		110,0	19,0	1,9	Trom-	0,7	0,8
19BE045	NA	El Sal.		110,0	22,1	2,4	Trom-	0,6	0,9
19BE048	NA	El Sal.		110,0	23,1	2,7	Trom-	0,8	0,9
19BE059	NA	El Sal.		110,0	21,7	2,4	Trom-	0,8	0,9
19BE064	NA	El Sal.		110,0	21,9	2,7	Trom-	0,8	0,9
19BE073	NA	El Sal.		110,0	23,4	1,9	Trom-	0,7	0,9

Anhang

19BE080	NA	El Sal.		110,0	21,6	3,0	Trom-	0,6	0,9
19BE090	NA	El Sal.		110,0	22,6	2,0	Trom-	0,8	0,9
19BE091	NA	El Sal.		110,0	22,3	2,2	Trom-	0,8	0,8
19BE093	NA	El Sal.		110,0	22,5	2,0	Trom-	0,8	0,9
19BE098	NA	El Sal.		110,0	21,7	2,3	Trom-	0,7	0,8
19BE100	NA	El Sal.		110,0	21,5	2,4	Trom-	0,7	0,8
19BE103	NA	El Sal.		110,0	21,3	1,8	Trom-	0,6	0,8
19BE105	NA	El Sal.		110,0	22,1	2,0	Trom-	0,8	0,9
19BE107	NA	El Sal.		110,0	20,7	1,6	Trom-	0,8	0,9
19BE114	NA	El Sal.		110,0	21,1	1,7	Trom-	0,7	0,9
19BE115	NA	El Sal.		110,0	21,4	2,1	Trom-	0,7	0,9
19BE122	NA	El Sal.		110,0	21,3	2,5	Trom-	0,8	0,9
19BE124	NA	El Sal.		110,0	21,0	2,4	Trom-	0,6	0,8
19BE126	NA	El Sal.		110,0	22,0	2,5	Trom-	0,8	0,9
19BE127	NA	El Sal.		110,0	20,8	1,6	Trom-	0,6	0,8
19BE130	NA	El Sal.		110,0	20,3	2,3	Trom-	0,7	0,9
19BE131	NA	El Sal.		110,0	16,2	3,2	Trom-	0,8	0,8
19BE134	NA	El Sal.		110,0	20,2	3,2	Trom-	0,7	0,9
19BE135	NA	El Sal.		110,0	20,7	2,0	Trom-	0,6	0,9
19BE136	NA	El Sal.		110,0	22,4	1,9	Trom-	0,6	0,9
19BE137	NA	El Sal.		110,0	19,5	1,6	Trom-	0,8	0,9
19BE139	NA	El Sal.		110,0	21,3	2,0	Trom-	0,7	0,8
19BE147	NA	El Sal.		110,0	19,9	2,3	Trom-	0,6	0,8
19BE148	NA	El Sal.		110,0	20,9	2,5	Trom-	0,7	0,9
19BE150	NA	El Sal.		110,0	22,8	3,0	Trom-	0,8	1,0
19BE152	NA	El Sal.		110,0	21,4	2,2	Trom-	0,7	0,8
19BE165	NA	El Sal.		110,0	19,4	2,7	Trom-	0,6	0,9
19BE168	NA	El Sal.		110,0	20,8	2,3	Trom-	0,8	0,9
19BE300	NA	El Sal.		110,0	22,9	1,9	Trom-	0,6	0,8
19HA000	NA	Mexiko		110,0	19,2	3,3	Trom-	0,6	0,8
19HA000	NA	Mexiko		110,0	19,6	3,0	Trom-	0,6	0,8
19HA000	NA	Mexiko		110,0	19,6	3,2	Trom-	0,8	0,8
19HA004	NA	Mexiko		110,0	17,0	3,1	Trom-	0,7	0,8
19HA005	NA	Mexiko		110,0	17,4	3,1	Trom-	0,6	0,8
20HA001	NA	Mexiko		110,0	19,5	3,5	Trom-	0,7	0,9
20HA002	NA	Mexiko		110,0	19,8	3,3	Trom-	0,7	0,8
20HA003	NA	Mexiko		110,0	19,5	3,1	Trom-	0,8	0,9
20HA004	NA	Mexiko		110,0	20,7	3,2	Trom-	0,6	0,8
20HA005	NA	Mexiko		110,0	18,8	3,2	Trom-	0,8	0,8
20HA006	NA	Mexiko		110,0	17,7	3,3	Trom-	0,7	0,8
20HA007	NA	Mexiko		110,0	19,7	2,8	Trom-	0,6	0,8
20HA008	NA	Mexiko		110,0	18,2	4,0	Trom-	0,7	0,9
20HA009	NA	Mexiko		110,0	20,0	3,0	Trom-	0,8	0,9
20HA010	NA	Mexiko		110,0	18,9	3,2	Trom-	0,8	0,8
20HA011	NA	Mexiko		106,0	18,3	3,4	Trom-	0,6	0,9
20HA012	NA	Mexiko		110,0	19,3	3,4	Trom-	0,7	0,8
20HA013	NA	Mexiko		110,0	19,6	3,0	Trom-	0,7	1,0
20HA014	NA	Mexiko		106,0	17,8	3,0	Trom-	0,7	0,8
20HA015	NA	Mexiko		110,0	20,1	3,1	Trom-	0,8	1,0
20HA016	NA	Mexiko		110,0	19,1	3,1	Trom-	0,8	0,8
20HA018	NA	Mexiko		110,0	16,5	3,4	Trom-	0,8	0,9
21HA001	NA	Mexiko		110,0	18,5	3,4	Trom-	0,6	0,8
N101	NA	Guat.	Bio	107,7	17,0	4,6	Trom-	0,8	0,9
N103	NA	Guat.	Kon.	86,2	20,8	1,9	Trom-	0,8	0,9
N104	NA	Guat.	Kon.	82,2	23,7	4,6	Trom-	0,8	0,8
N105	NA	Guat.	Kon.	93,3	22,3	3,2	Trom-	0,8	0,8
N106	NA	Guat.	Kon.	92,5	22,8	3,8	Trom-	0,8	0,8
N107	NA	Guat.	Kon.	107,8	21,4	2,8	Trom-	0,9	0,9
N108	NA	Guat.	Kon.	110,9	21,3	3,1	Trom-	0,9	0,8
N109	NA	Guat.	Kon.	108,0	20,8	3,2	Trom-	0,8	0,8
N10	NA	Costa R.	Kon.	79,8	22,7	1,7	Trom-	0,8	0,8
N110	NA	Guat.	Kon.	77,4	23,2	3,0	Trom-	0,8	0,9
N111	NA	Guat.	Kon.	73,9	22,1	1,5	Trom-	1,0	0,8

Anhang

N112	NA	Guat.	Kon.	45,3	27,4	3,2	Trom-	0,9	0,9
N113	NA	Guat.	Kon.	97,2	18,9	3,2	Trom-	0,8	0,9
N114	NA	Guat.	Kon.	99,6	20,4	1,5	Trom-	0,8	0,8
N115	NA	Guat.	Bio	87,0	24,6	2,9	Trom-	0,7	0,9
N116	NA	Guat.	Kon.	69,9	24,1	1,5	Trom-	0,8	0,9
N117	NA	Guat.	Kon.	92,3	22,2	4,6	Trom-	0,8	0,8
N118	NA	Guat.	Kon.	95,3	23,1	2,0	Trom-	0,8	0,9
N119	NA	Guat.	Kon.	90,7	20,7	5,3	Trom-	0,7	0,9
N11	NA	Costa R.	Kon.	81,3	21,2	4,0	Trom-	0,8	0,9
N13	NA	Costa R.	Kon.	85,8	22,7	2,2	Trom-	0,7	0,9
N150	NA	Honduras	Bio	84,7	21,7	2,9	Trom-	0,8	0,8
N151	NA	Honduras	Bio	102,5	19,7	4,6	Trom-	0,8	0,8
N159	NA	Honduras	Kon.	62,9	22,4	3,1	Trom-	0,8	0,8
N15	NA	Costa R.	Kon.	80,8	23,7	5,6	Trom-	0,9	0,8
N162	NA	Honduras	Bio	91,1	19,1	4,4	Trom-	0,9	0,9
N163	NA	Honduras	Kon.	86,7	20,3	1,7	Trom-	0,8	0,8
N164	NA	Honduras	Kon.	104,9	21,5	3,1	Trom-	0,8	0,8
N165	NA	Honduras	Kon.	103,7	19,7	4,4	Trom-	0,8	0,9
N166	NA	Honduras	Kon.	82,0	23,5	1,4	Trom-	0,8	0,8
N167	NA	Honduras	Kon.	101,8	20,4	3,8	Trom-	1,0	0,9
N168	NA	Honduras	Bio	87,2	22,1	3,2	Trom-	0,8	0,8
N169	NA	Honduras	Bio	80,6	21,2	4,4	Trom-	0,9	0,8
N16	NA	Costa R.	Kon.	98,9	19,7	1,3	Trom-	0,9	0,8
N170	NA	Honduras	Bio	95,9	21,6	3,7	Trom-	0,8	0,9
N171	NA	Honduras	Kon.	82,4	21,5	1,7	Trom-	0,9	0,9
N172	NA	Honduras	Bio	80,1	20,8	2,2	Trom-	0,9	0,9
N173	NA	Honduras	Kon.	99,2	22,2	2,3	Trom-	0,8	0,8
N174	NA	Honduras	Kon.	67,8	22,1	1,1	Trom-	0,8	0,9
N175	NA	Honduras	Bio	95,2	20,1	5,1	Trom-	0,9	0,8
N176	NA	Honduras	Kon.	80,2	22,3	3,9	Trom-	0,7	0,8
N177	NA	Honduras	Kon.	89,5	22,1	2,2	Trom-	0,8	0,9
N178	NA	Honduras	Bio	89,9	21,8	2,5	Trom-	0,8	0,9
N179	NA	Honduras	Kon.	120,2	19,1	2,5	Trom-	0,8	0,8
N17	NA	Costa R.	Kon.	81,5	21,3	3,3	Trom-	0,8	0,9
N180	NA	Honduras	Kon.	99,7	21,3	1,4	Trom-	0,8	0,8
N181	NA	Honduras	Kon.	81,7	19,9	4,4	Trom-	0,8	0,8
N1	NA	Costa R.	Kon.	76,4	23,5	4,6	Trom-	0,8	0,8
N1a	NA	Costa R.	Kon.	85,6	22,8	2,4	Trom-	0,8	0,9
N1b	NA	Costa R.	Kon.	85,1	23,7	1,4	Trom-	0,8	0,9
N200	NA	Kuba	Kon.	68,8	19,4	4,1	Trom-	0,8	0,8
N202	NA	Kuba	Kon.	96,8	20,1	3,9	Trom-	0,8	0,8
N203	NA	Kuba	Kon.	118,0	17,8	3,5	Trom-	0,8	0,8
N204	NA	Kuba	Kon.	104,7	20,2	3,3	Trom-	0,8	0,9
N205	NA	Kuba	Kon.	50,8	19,0	2,8	Trom-	0,6	0,9
N259	NA	Mexiko	Kon.	79,0	19,6	2,7	Trom-	0,8	0,8
N268	NA	Mexiko	Bio	89,4	21,5	3,1	Trom-	0,9	0,8
N269	NA	Mexiko	Kon.	83,8	21,5	3,6	Trom-	0,8	0,8
N270	NA	Mexiko	Kon.	97,2	20,6	3,9	Trom-	0,9	0,8
N271	NA	Mexiko	Bio	94,6	22,5	2,8	Trom-	0,8	0,8
N272	NA	Mexiko	Bio	92,7	21,2	2,2	Trom-	0,8	0,9
N273	NA	Mexiko	Kon.	111,0	20,1	3,4	Trom-	0,8	0,9
N274	NA	Mexiko	Kon.	104,9	19,9	3,1	Trom-	0,8	0,9
N275	NA	Mexiko	Bio	77,9	21,9	2,4	Trom-	0,8	0,8
N276	NA	Mexiko	Bio	96,1	22,0	1,9	Trom-	0,8	0,9
N277	NA	Mexiko	Bio	85,6	20,5	2,8	Trom-	0,9	0,9
N278	NA	Mexiko	Bio	96,9	21,2	3,0	Trom-	0,8	0,8
N279	NA	Mexiko	Bio	87,8	22,2	2,3	Trom-	0,8	0,8
N280	NA	Mexiko	Kon.	89,2	22,9	4,0	Trom-	0,8	0,8
N281	NA	Mexiko	Bio	83,1	21,3	3,6	Trom-	0,9	0,9
N282	NA	Mexiko	Kon.	107,8	15,6	4,4	Trom-	0,6	0,8
N283	NA	Mexiko	Bio	132,4	18,0	3,2	Trom-	0,8	0,8
N284	NA	Mexiko	Kon.	112,2	17,1	2,7	Trom-	0,8	0,8
N285	NA	Mexiko	Kon.	105,7	18,3	3,5	Trom-	0,7	0,8

Anhang

N296	NA	Mexiko	Kon.	115,1	17,6	3,7	Trom-	0,8	0,9
N300	NA	Nicarag.	Kon.	96,6	19,0	3,6	Trom-	0,8	0,8
N301	NA	Nicarag.	Kon.	102,4	17,7	4,6	Trom-	0,8	0,8
N302	NA	Nicarag.	Kon.	84,5	19,8	4,2	Trom-	0,8	0,9
N303	NA	Nicarag.	Bio	72,9	17,3	3,7	Trom-	0,8	0,9
N303b	NA	Nicarag.	Bio	74,0	18,9	3,1	Trom-	0,9	0,8
N304	NA	Nicarag.	Kon.	106,2	19,7	2,9	Trom-	0,8	0,8
N305	NA	Nicarag.	Kon.	110,2	17,8	3,1	Trom-	0,8	0,8
N306	NA	Nicarag.	Kon.	100,7	17,0	5,1	Trom-	0,7	0,8
N307	NA	Nicarag.	Bio	74,4	20,3	2,9	Trom-	0,8	0,8
N350	NA	Panama	Kon.	104,3	20,1	4,9	Trom-	0,8	0,8
N351	NA	Panama	Kon.	88,4	22,3	1,6	Trom-	0,8	0,8
N352	NA	Panama	Kon.	90,6	23,7	2,0	Trom-	0,8	0,8
N3	NA	Costa R.	Bio	79,7	21,2	3,3	Trom-	0,8	0,9
N400	NA	DomRep	Kon.	100,5	20,8	3,3	Trom-	0,8	0,8
N401	NA	DomRep	Kon.	114,6	21,2	4,6	Trom-	0,8	0,8
N450	NA	Grenada	Kon.	60,8	17,8	4,3	Trom-	0,8	0,9
N4	NA	Costa R.	Kon.	96,9	19,2	4,3	Trom-	0,9	0,9
N51	NA	El Sal.	Kon.	100,5	21,8	4,4	Trom-	0,8	0,8
N52	NA	El Sal.	Kon.	108,1	21,0	3,4	Trom-	0,8	0,8
N53	NA	El Sal.	Kon.	108,6	18,8	4,7	Trom-	0,8	0,8
N54	NA	El Sal.	Kon.	112,2	20,6	4,5	Trom-	0,8	0,8
N55	NA	El Sal.	Kon.	107,5	19,8	2,5	Trom-	0,8	0,8
N56	NA	El Sal.	Kon.	105,5	19,6	4,4	Trom-	0,6	0,9
N57	NA	El Sal.	Kon.	111,6	21,8	3,8	Trom-	0,8	0,8
N58	NA	El Sal.	Kon.	104,8	19,9	2,5	Trom-	1,0	0,9
N59	NA	El Sal.	Bio	89,6	20,3	4,0	Trom-	0,7	0,9
N5	NA	Costa R.	Kon.	64,6	23,2	4,3	Trom-	0,8	0,8
N60	NA	El Sal.	Kon.	87,4	21,8	4,5	Trom-	0,8	0,9
N6	NA	Costa R.	Kon.	79,9	23,6	2,9	Trom-	0,8	0,8
N7	NA	Costa R.	Kon.	96,8	20,3	3,9	Trom-	0,8	0,8
N8	NA	Costa R.	Kon.	110,2	21,1	3,3	Trom-	0,8	0,9
N900	NA	Karibik	Kon.	97,2	20,2	4,0	Trom-	0,7	0,9
N9	NA	Costa R.	Kon.	101,4	21,0	3,4	Trom-	0,8	0,8
16DU23	SA	Brasilien		134,2	20,6	3,9	Trom-	0,8	0,8
20DU001	SA	Brasilien		106,0	20,3	5,7	Trom-	0,8	0,8
20DU002	SA	Brasilien		106,0	20,3	5,3	Trom-	0,6	0,8
20DU003	SA	Brasilien		110,0	20,1	5,6	Trom-	0,8	0,8
20DU004	SA	Brasilien		106,0	18,9	5,5	Trom-	0,8	0,8
20DU006	SA	Brasilien		106,0	20,0	5,3	Trom-	0,7	0,9
20DU007	SA	Brasilien		106,0	21,1	6,0	Trom-	0,8	1,0
20DU008	SA	Brasilien		106,0	20,6	5,2	Trom-	0,6	0,8
20DU010	SA	Brasilien		106,0	19,8	5,7	Trom-	0,9	0,9
20DU013	SA	Brasilien		106,0	21,2	5,6	Trom-	0,7	0,9
20DU014	SA	Brasilien		106,0	20,7	5,7	Trom-	0,7	0,9
20DU015	SA	Brasilien		106,0	20,1	5,7	Trom-	0,8	0,8
20DU016	SA	Brasilien		106,0	21,2	5,6	Trom-	0,8	0,8
20DU017	SA	Brasilien		106,0	19,6	5,8	Trom-	0,7	0,9
20DU018	SA	Brasilien		106,0	20,5	5,1	Trom-	0,7	0,8
20DU019	SA	Brasilien		106,0	19,3	5,6	Trom-	0,7	0,8
20DU020	SA	Brasilien		106,0	19,8	5,4	Trom-	0,7	0,9
20DU021	SA	Brasilien		106,0	19,9	5,3	Trom-	0,7	0,8
20DU023	SA	Brasilien		106,0	20,9	5,7	Trom-	0,7	0,9
20DU024	SA	Brasilien		106,0	23,4	5,7	Trom-	0,8	0,8
20DU025	SA	Brasilien		106,0	19,5	5,6	Trom-	0,8	1,0
20DU027	SA	Brasilien		106,0	19,7	5,0	Trom-	0,6	0,9
20DU028	SA	Brasilien		106,0	18,9	5,9	Trom-	0,8	0,8
20DU030	SA	Brasilien		106,0	21,2	5,3	Trom-	0,8	0,8
20DU031	SA	Brasilien		106,0	19,2	5,8	Trom-	0,8	0,9
20DU032	SA	Brasilien		106,0	22,1	5,8	Trom-	0,8	0,8
20DU033	SA	Brasilien		106,0	19,3	5,1	Trom-	0,6	0,8
20DU034	SA	Brasilien		106,0	18,8	5,4	Trom-	0,7	0,8

Anhang

20DU035	SA	Brasilien		106,0	20,2	5,7	Trom-	0,8	0,9
20DU036	SA	Brasilien		106,0	20,1	5,3	Trom-	0,6	1,1
20DU037	SA	Brasilien		106,0	20,4	5,7	Trom-	0,7	0,8
20DU038	SA	Brasilien		106,0	20,5	5,7	Trom-	0,8	0,9
20DU039	SA	Brasilien		106,0	20,8	5,5	Trom-	0,7	0,8
20DU040	SA	Brasilien		106,0	19,9	5,3	Trom-	0,7	0,9
20DU041	SA	Brasilien		106,0	19,2	5,8	Trom-	0,8	0,9
20DU042	SA	Brasilien		106,0	18,5	5,4	Trom-	0,7	0,8
20DU043	SA	Brasilien		106,0	19,7	5,3	Trom-	0,6	0,8
20DU044	SA	Brasilien		106,0	20,5	5,4	Trom-	0,8	0,9
20DU045	SA	Brasilien		106,0	21,4	5,2	Trom-	0,7	0,8
20DU046	SA	Brasilien		106,0	20,8	5,1	Trom-	0,6	0,9
20DU047	SA	Brasilien		106,0	21,7	5,5	Trom-	0,8	0,8
20DU048	SA	Brasilien		110,0	20,0	5,1	Trom-	0,6	0,9
20DU049	SA	Brasilien		106,0	19,9	5,9	Trom-	0,7	0,9
20DU050	SA	Brasilien		106,0	19,8	5,4	Trom-	0,7	0,9
20DU051	SA	Brasilien		106,0	20,0	5,3	Trom-	0,8	0,9
20DU054	SA	Brasilien		106,0	19,9	5,3	Trom-	0,6	0,9
20DU055	SA	Brasilien		106,0	20,3	5,3	Trom-	0,7	0,9
20DU056	SA	Brasilien		106,0	20,2	5,3	Trom-	0,7	0,8
20DU057	SA	Brasilien		106,0	19,0	5,8	Trom-	0,8	0,8
20DU058	SA	Brasilien		106,0	20,3	5,7	Trom-	0,8	0,9
20DU059	SA	Brasilien		106,0	23,0	6,0	Trom-	0,9	0,8
20DU060	SA	Brasilien		106,0	21,7	5,4	Trom-	0,7	0,9
20DU061	SA	Brasilien		106,0	20,6	5,4	Trom-	0,7	0,9
20DU062	SA	Brasilien		106,0	21,7	5,3	Trom-	0,8	0,9
20DU063	SA	Brasilien		110,0	19,5	4,9	Trom-	0,6	0,8
20DU065	SA	Brasilien		106,0	19,4	5,2	Trom-	0,6	0,8
20DU065	SA	Brasilien		106,0	19,6	5,1	Trom-	0,7	0,9
20DU066	SA	Brasilien		106,0	23,7	5,8	Trom-	0,9	0,8
20DU067	SA	Brasilien		106,0	19,9	5,7	Trom-	0,7	0,9
20DU068	SA	Brasilien		106,0	18,9	5,5	Trom-	0,9	0,9
20DU069	SA	Brasilien		106,0	17,9	5,8	Trom-	0,7	0,8
20DU070	SA	Brasilien		106,0	17,5	5,6	Trom-	0,7	0,8
20DU071	SA	Brasilien		106,0	19,0	5,0	Trom-	0,7	0,8
20DU072	SA	Brasilien		106,0	19,4	5,0	Trom-	0,6	0,8
20DU073	SA	Brasilien		106,0	19,6	5,3	Trom-	0,6	0,8
S100	SA	Peru	Bio	81,4	17,7	2,9	Trom-	0,8	0,9
S101	SA	Peru	Bio	96,6	21,6	1,7	Trom-	0,8	0,8
S104	SA	Peru	Bio	101,6	20,5	2,3	Trom-	0,8	0,9
S105	SA	Peru	Bio	112,3	19,1	3,5	Trom-	0,8	0,8
S106	SA	Peru	Bio	94,9	21,3	3,2	Trom-	0,8	0,8
S107	SA	Peru	Bio	79,8	22,3	3,0	Trom-	0,8	0,9
S108	SA	Peru	Bio	90,4	22,0	2,3	Trom-	0,8	0,8
S108a	SA	Peru	Bio	97,1	18,1	1,7	Trom-	0,7	0,9
S109	SA	Peru	Bio	73,6	21,5	3,5	Trom-	0,8	0,8
S10	SA	Brasilien	Kon.	94,6	24,7	3,0	Trom-	0,8	0,8
S110	SA	Preu	Bio	71,0	23,5	3,8	Trom-	0,8	0,9
S111	SA	Peru	Kon.	92,3	20,8	2,5	Trom-	0,7	0,9
S112	SA	Peru	Kon.	64,6	22,7	2,8	Trom-	0,8	0,8
S113	SA	Peru	Bio	95,4	20,6	3,6	Trom-	0,7	0,8
S113a	SA	Peru	Bio	87,6	21,7	3,8	Trom-	0,7	0,9
S113b	SA	Peru	Bio	72,0	24,0	3,4	Trom-	0,8	0,8
S114	SA	Peru	Bio	107,2	20,7	2,5	Trom-	0,8	0,9
S115	SA	Peru	Bio	99,2	23,1	2,7	Trom-	0,8	0,8
S116	SA	Peru	Bio	81,9	21,6	4,2	Trom-	0,8	0,9
S117	SA	Peru	Bio	83,6	20,3	2,7	Trom-	1,0	0,9
S117a	SA	Peru	Bio	69,5	21,9	2,6	Trom-	0,8	0,8
S118	SA	Peru	Bio	88,7	20,9	3,3	Trom-	0,7	0,9
S119	SA	Peru	Bio	77,7	22,3	2,5	Trom-	0,8	0,9
S119a	SA	Peru	Bio	80,6	23,0	1,8	Trom-	2,1	0,8
S11	SA	Brasilien	Kon.	77,6	23,4	3,2	Trom-	0,9	0,9
S120	SA	Peru	Bio	93,0	19,6	4,2	Trom-	0,8	0,9

Anhang

S120a	SA	Peru	Bio	104,1	21,0	3,9	Trom-	0,7	0,9
S121	SA	Peru	Bio	88,2	17,9	4,5	Trom-	0,8	0,8
S122	SA	Peru	Bio	97,2	20,8	3,1	Trom-	0,8	0,9
S123	SA	Peru	Bio	101,5	22,7	2,1	Trom-	1,0	0,9
S124	SA	Peru	Kon.	103,1	21,7	2,4	Trom-	0,8	0,8
S125	SA	Peru	Kon.	106,1	20,7	3,0	Trom-	0,8	0,8
S126	SA	Peru	Bio	76,7	22,4	3,9	Trom-	0,7	0,8
S127	SA	Peru	Bio	102,6	17,7	2,1	Trom-	0,8	0,9
S127a	SA	Peru	Bio	103,5	20,1	1,8	Trom-	0,8	0,8
S128	SA	Peru	Bio	61,6	21,8	4,5	Trom-	0,8	0,9
S12	SA	Brasilien	Kon.	95,1	20,8	3,5	Trom-	0,8	0,8
S13	SA	Brasilien	Kon.	107,5	21,2	3,4	Trom-	0,8	0,8
S13a	SA	Brasilien	Kon.	108,1	20,6	2,4	Trom-	0,8	0,9
S14	SA	Brasilien	Kon.	96,7	23,5	4,0	Trom-	0,8	0,8
S150	SA	Galap.	Kon.	106,0	18,1	4,1	Trom-	0,8	0,8
S15	SA	Brasilien	Kon.	100,6	22,6	3,6	Trom-	0,8	0,8
S16	SA	Brasilien	Kon.	84,4	22,1	3,4	Trom-	0,8	0,8
S17	SA	Brasilien	Kon.	114,4	21,8	3,0	Trom-	0,8	0,8
S18	SA	Brasilien	Kon.	103,3	22,6	2,8	Trom-	0,8	0,8
S19	SA	Brasilien	Kon.	81,1	22,0	3,6	Trom-	0,8	0,8
S19a	SA	Brasilien	Kon.	82,5	21,8	4,8	Trom-	0,9	0,8
S200	SA	Bolivien	Kon.	114,0	19,5	3,5	Trom-	0,8	0,8
S20	SA	Brasilien	Kon.	89,7	22,6	5,6	Trom-	0,8	0,8
S21	SA	Brasilien	Kon.	85,0	22,7	3,3	Trom-	0,8	1,0
S21a	SA	Brasilien	Kon.	93,4	23,1	2,1	Trom-	0,7	0,8
S22	SA	Brasilien	Kon.	101,0	24,4	3,1	Trom-	0,7	0,8
S23	SA	Brasilien	Bio	107,8	21,9	2,9	Trom-	0,8	0,8
S23a	SA	Brasilien	Bio	102,3	21,9	2,6	Trom-	0,8	0,9
S24	SA	Brasilien	Kon.	72,2	23,4	4,1	Trom-	0,8	0,8
S251	SA	Ecuador	Kon.	94,6	21,3	5,9	Trom-	0,7	0,8
S25	SA	Brasilien	Kon.	89,3	17,6	1,2	Trom-	0,7	0,8
S26	SA	Brasilien	Kon.	103,3	21,3	2,6	Trom-	0,8	0,8
S27	SA	Brasilien	Kon.	93,1	22,3	2,2	Trom-	0,8	0,8
S28	SA	Brasilien	Kon.	82,4	17,8	3,5	Trom-	0,8	0,8
S29	SA	Brasilien	Kon.	90,9	23,1	3,4	Trom-	0,9	0,8
S30	SA	Brasilien	Kon.	87,0	23,3	4,6	Trom-	0,8	0,8
S31	SA	Brasilien	Kon.	102,4	23,8	2,6	Trom-	0,9	0,9
S31a	SA	Brasilien	Kon.	97,2	21,6	1,2	Trom-	0,7	0,9
S32	SA	Brasilien	Kon.	91,5	21,2	1,2	Trom-	0,7	0,9
S33	SA	Brasilien	Kon.	97,7	21,0	4,0	Trom-	0,8	0,9
S34	SA	Brasilien	Kon.	97,9	21,8	2,9	Trom-	0,8	0,8
S35	SA	Brasilien	Kon.	105,5	20,0	2,4	Trom-	0,8	0,8
S36	SA	Brasilien	Kon.	103,9	20,3	2,1	Trom-	0,8	0,8
S37	SA	Brasilien	Kon.	109,4	18,6	2,4	Trom-	0,8	0,9
S39	SA	Brasilien	Kon.	112,9	22,0	1,8	Trom-	0,8	0,9
S40	SA	Brasilien	Kon.	80,3	20,9	2,1	Trom-	0,8	0,9
S40a	SA	Brasilien	Kon.	75,6	21,9	1,6	Trom-	0,8	0,9
S41	SA	Brasilien	Kon.	98,5	20,6	1,4	Trom-	0,6	0,8
S42	SA	Brasilien	Kon.	74,1	22,6	1,1	Trom-	0,6	0,8
S43	SA	Brasilien	Kon.	83,9	23,0	4,6	Trom-	0,7	0,9
S44	SA	Brasilien	Bio	92,4	22,2	1,8	Trom-	0,8	0,9
S45	SA	Brasilien	Kon.	100,5	21,8	1,4	Trom-	0,8	0,8
S46	SA	Brasilien	Kon.	107,5	22,6	1,6	Trom-	0,9	1,0
S47	SA	Brasilien	Kon.	87,6	22,9	0,9	Trom-	0,8	0,8
S48	SA	Brasilien	Kon.	81,5	24,5	1,5	Trom-	0,7	0,8
S49	SA	Brasilien	Kon.	87,4	22,7	1,4	Trom-	0,7	0,8
S49_1	SA	Brasilien	Kon.	106,6	20,4	1,6	Trom-	0,7	0,8
S4	SA	Brasilien	Kon.	99,5	22,0	2,5	Trom-	0,8	0,8
S50	SA	Kolum.	Kon.	79,6	20,0	2,7	Trom-	0,9	0,8
S52	SA	Kolum.	Kon.	96,6	20,4	4,3	Trom-	0,8	0,8
S54	SA	Kolum.	Bio	96,1	19,6	4,4	Trom-	0,8	0,8
S55	SA	Kolum.	Kon.	82,4	20,9	4,2	Trom-	0,8	0,9
S56	SA	Kolum.	Kon.	81,4	20,7	2,4	Trom-	0,9	0,9

Anhang

S57	SA	Kolum.	Kon.	75,8	20,8	3,0	Trom-	0,9	0,8
S58	SA	Kolum.	Kon.	89,6	24,1	3,2	Trom-	0,8	0,8
S59	SA	Kolum.	Kon.	95,8	21,5	3,5	Trom-	0,8	0,8
S5	SA	Brasilien	Kon.	99,7	21,3	2,3	Trom-	0,9	0,8
S60	SA	Kolum.	Kon.	99,6	20,7	4,5	Trom-	0,8	0,8
S61	SA	Kolum.	Kon.	80,4	21,2	3,4	Trom-	0,8	0,8
S62	SA	Kolum.	Kon.	105,0	21,0	3,1	Trom-	0,9	0,9
S63	SA	Kolum.	Kon.	94,2	22,6	2,6	Trom-	0,8	0,8
S64	SA	Kolum.	Kon.	79,9	21,4	2,3	Trom-	0,8	0,8
S65	SA	Kolum.	Bio	100,2	24,8	2,1	Trom-	0,9	0,8
S66	SA	Kolum.	Kon.	112,0	24,6	4,8	Trom-	0,9	0,8
S67	SA	Kolum.	Kon.	81,0	22,6	3,6	Trom-	0,8	0,8
S68	SA	Kolum.	Kon.	83,1	21,1	4,6	Trom-	0,8	0,8
S69	SA	Kolum.	Kon.	81,5	23,0	3,2	Trom-	0,9	0,8
S70	SA	Kolum.	Bio	72,1	22,7	3,7	Trom-	0,9	0,9
S70a	SA	Kolum.	Bio	71,2	21,9	3,1	Trom-	0,7	0,9
S71	SA	Kolum.	Kon.	95,9	16,9	2,9	Trom-	0,9	0,9
S72	SA	Kolum.	Kon.	101,9	18,2	4,2	Trom-	0,8	0,8
S73	SA	Kolum.	Bio	108,9	21,6	2,1	Trom-	0,8	0,8
S74	SA	Kolum.	Bio	57,3	21,3	2,9	Trom-	0,7	0,9
S74a	SA	Kolum.	Bio	74,0	22,1	2,5	Trom-	0,8	0,9
S75	SA	Kolum.	Bio	71,7	21,6	3,3	Trom-	0,7	0,9
S76	SA	Kolum.	Bio	68,3	22,9	2,1	Trom-	0,6	0,9
S77	SA	Kolum.	Kon.	99,6	23,5	2,8	Trom-	0,9	0,9
S78	SA	Kolum.	Kon.	96,3	23,1	4,6	Trom-	0,8	0,8
S79	SA	Kolum.	Kon.	91,8	21,7	1,6	Trom-	0,8	0,8
S7	SA	Brasilien	Bio	104,2	21,3	4,7	Trom-	0,8	0,8
S80	SA	Kolum.	Bio	70,7	21,3	1,2	Trom-	0,7	0,8
S81	SA	Kolum.	Kon.	114,7	22,0	2,0	Trom-	0,7	0,8
S82	SA	Kolum.	Kon.	102,2	19,7	1,6	Trom-	0,9	0,8
S83	SA	Kolum.	Kon.	107,7	20,0	2,1	Trom-	0,8	0,8
S84	SA	Kolum.	Kon.	106,6	20,2	2,0	Trom-	0,8	0,8
S8	SA	Brasilien	Kon.	80,4	25,3	2,0	Trom-	0,8	0,8
S900	SA		Bio	80,0	22,7	3,1	Trom-	0,9	0,8
S901	SA		Kon.	107,4	20,8	1,9	Trom-	0,8	0,8
S9	SA	Brasilien	Kon.	105,6	20,8	3,1	Trom-	0,9	0,8
B10			Bio	84,3	20,0	4,8	Trom-	0,9	0,9
B11			Bio	105,6	23,1	3,4	Trom-	0,8	0,9
B12			Bio	69,4	21,8	2,4	Trom-	2,7	0,8
B13			Bio	78,8	21,8	3,4	Trom-	1,0	0,9
B14			Bio	93,6	22,8	4,1	Trom-	0,8	0,8
B15			Bio	88,8	26,2	3,0	Trom-	0,9	0,9
B16			Bio	110,1	21,5	2,1	Trom-	0,9	0,9
B17			Bio	82,5	21,0	4,1	Trom-	0,9	0,8
B18			Bio	89,0	20,4	4,9	Trom-	0,8	0,8
B19			Bio	107,0	20,3	2,7	Trom-	0,8	0,8
B1			Bio	81,5	17,6	4,3	Trom-	0,9	0,9
B22			Bio	84,6	20,6	3,3	Trom-	0,8	0,8
B23			Bio	90,8	21,2	2,6	Trom-	0,8	0,8
B24			Bio	75,5	22,5	4,8	Trom-	0,9	0,9
B25			Bio	88,3	21,0	4,5	Trom-	0,7	0,8
B26			Bio	76,1	23,8	4,4	Trom-	0,9	0,8
B27			Bio	68,3	23,2	3,7	Trom-	0,9	0,9
B28			Bio	77,6	23,1	2,5	Trom-	0,8	0,8
B29			Bio	73,6	19,6	4,7	Trom-	0,9	0,9
B29a			Bio	84,7	21,6	4,6	Trom-	0,8	0,8
B29b			Bio	84,4	20,4	4,5	Trom-	0,8	0,8
B2			Bio	99,6	22,9	4,6	Trom-	0,9	0,9
B30			Bio	83,4	19,8	4,6	Trom-	0,9	0,9
B30a			Bio	90,4	19,6	4,8	Trom-	0,7	0,8
B31			Bio	92,4	21,7	4,8	Trom-	0,9	0,8
B32			Bio	94,0	27,4	2,2	Trom-	0,9	0,9

Anhang

B33		Bio		73,4	22,0	4,3	Trom-	0,8	0,9
B34		Bio		82,7	18,6	4,8	Trom-	0,9	0,9
B35		Bio		84,9	21,2	4,0	Trom-	0,7	0,9
B36		Bio		88,9	21,3	3,7	Trom-	0,8	0,9
B37		Bio		75,5	18,9	3,9	Trom-	0,8	0,8
B38		Bio		86,3	23,6	3,2	Trom-	0,8	0,9
B3		Bio		84,3	20,3	3,9	Trom-	0,9	0,8
B40		Bio		89,2	23,4	3,1	Trom-	0,9	0,9
B42		Bio		80,5	21,4	4,5	Trom-	0,8	0,9
B43		Bio		75,9	22,2	3,3	Trom-	0,8	0,8
B43a		Bio		83,6	22,7	3,3	Trom-	0,8	0,9
B44		Bio		67,9	22,4	4,1	Trom-	0,9	0,8
B45		Bio		90,6	20,8	4,2	Trom-	0,8	0,9
B46		Bio		96,4	22,7	4,0	Trom-	0,8	0,9
B48		Bio		71,1	24,1	4,5	Trom-	0,9	0,9
B4		Bio		84,9	23,2	1,8	Trom-	0,7	0,8
B50		Bio		89,0	22,8	3,0	Trom-	0,8	0,8
B52		Bio		78,1	20,5	4,3	Trom-	0,9	0,9
B53		Bio		75,7	21,4	4,2	Trom-	0,9	0,9
B55		Bio		87,1	22,0	4,3	Trom-	0,9	0,9
B56		Bio		85,9	20,9	2,8	Trom-	0,9	0,9
B57		Bio		79,8	22,1	3,8	Trom-	1,0	0,9
B58		Bio		77,5	23,1	3,8	Trom-	0,8	0,8
B59		Bio		84,2	22,8	5,6	Trom-	0,8	0,9
B5		Bio		80,0	22,6	3,6	Trom-	0,7	0,8
B60		Bio		81,8	22,6	3,5	Trom-	0,8	0,9
B61		Bio		69,5	22,1	3,1	Trom-	2,1	0,9
B62		Bio		82,4	20,7	3,8	Trom-	0,9	0,8
B63		Bio		99,2	21,1	4,2	Trom-	0,7	0,8
B64		Bio		75,7	22,9	3,5	Trom-	0,7	0,8
B65		Bio		69,9	22,3	2,9	Trom-	0,8	0,8
B66		Bio		81,2	19,1	4,1	Trom-	1,0	0,8
B67		Bio		74,1	23,7	2,7	Trom-	0,8	0,8
B68		Bio		90,1	23,4	3,1	Trom-	0,7	0,8
B69		Bio		65,0	21,4	2,8	Trom-	0,8	0,8
B6		Bio		98,7	21,1	5,3	Trom-	1,7	0,8
B70		Bio		74,6	18,4	3,6	Trom-	0,8	0,8
B71		Bio		77,7	21,2	4,4	Trom-	0,7	0,8
B72		Bio		67,8	23,3	2,5	Trom-	0,7	0,8
B73		Bio		94,9	20,2	4,6	Trom-	0,7	0,8
B74		Bio		82,0	20,9	4,6	Trom-	0,7	0,8
B75		Bio		88,3	20,1	1,7	Trom-	0,8	0,8
B76		Bio		93,4	20,3	1,2	Trom-	0,7	0,8
B77		Bio		72,2	19,2	4,7	Trom-	0,7	0,8
B78		Bio		85,5	19,9	4,8	Trom-	1,0	0,8
B79		Bio		96,4	22,7	4,0	Trom-	0,8	0,9
B7		Bio		70,1	21,5	4,4	unk-	0,9	0,9
B80		Bio		79,7	18,9	4,4	Trom-	1,0	0,8
B81		Bio		84,0	20,9	4,3	Trom-	0,8	0,8
B82		Bio		87,1	22,0	4,6	Trom-	0,8	0,8
B83		Bio		88,8	21,9	2,8	Trom-	0,8	0,8
B83a		Bio		75,3	21,9	3,1	Trom-	0,8	0,9
B84		Bio		81,6	22,1	4,8	Trom-	1,1	0,8
B85		Bio		86,7	21,4	3,0	Trom-	0,8	0,8
B8		Bio		78,1	21,5	4,7	Trom-	0,9	0,9
B9		Bio		85,8	18,4	4,6	Trom-	0,9	0,9
AF151	AF	Tansania	Kon.	66,4	21,4	1,4	Ikawa	0,8	0,9
AF152	AF	Tansania	Kon.	99,6	17,6	1,3	Ikawa	0,9	0,8
AF154	AF	Tansania	Kon.	110,6	16,5	1,2	Ikawa	1,0	0,8
AF1	AF	Äthiopien	Kon.	116,1	22,4	2,8	Ikawa	0,8	0,8
AF205	AF	Uganda	Kon.	72,5	24,3	1,3	Ikawa	0,7	0,9
AF206	AF	Uganda	Kon.	76,8	23,6	1,1	Ikawa	0,7	0,9

Anhang

AF2	AF	Äthiopien	Kon.	81,1	21,8	1,0	Ikawa	0,8	0,9
AF7	AF	Äthiopien	Bio	112,9	25,0	2,3	Ikawa	0,8	0,8
AS100	AS	Sumatra	Bio	105,1	22,5	2,4	Ikawa	0,8	0,8
AS150	AS	Thailand	Kon.	81,2	18,7	1,0	Ikawa	0,7	0,8
AS151	AS	Thailand	Kon.	74,0	19,9	1,0	Ikawa	0,8	0,9
AS152	AS	Thailand	Kon.	85,0	21,3	1,5	Ikawa	0,8	0,8
AS153	AS	Thailand	Kon.	80,7	21,2	1,1	Ikawa	0,7	0,8
AS154	AS	Thailand	Kon.	86,4	19,7	2,7	Ikawa	0,8	0,8
AS157	AS	Thailand	Kon.	85,6	19,8	1,4	Ikawa	0,8	0,8
AS158	AS	Thailand	Kon.	83,2	22,4	1,2	Ikawa	0,8	0,9
AS159	AS	Thailand	Kon.	81,3	19,9	1,0	Ikawa	0,7	0,9
AS160	AS	Thailand	Kon.	93,4	22,4	1,4	Ikawa	0,7	0,9
AS161	AS	Myanmar	Kon.	85,0	22,3	1,8	Ikawa	0,8	0,8
AS162	AS	Myanmar	Kon.	69,5	20,8	1,1	Ikawa	0,7	0,9
AS167	AS	Thailand	Kon.	72,2	19,0	1,4	Ikawa	0,8	0,8
AS17	AS	Indien	Kon.	98,7	17,2	1,4	Ikawa	0,7	0,8
AS1	AS	Indien	Kon.	81,4	21,9	1,0	Ikawa	0,8	0,9
AS24	AS	Indien	Kon.	96,3	16,0	1,3	Ikawa	0,8	0,9
AS2	AS	Indien	Kon.	103,3	20,0	1,2	Ikawa	0,8	0,9
AS350	AS	Malaysia	Kon.	81,6	20,2	1,4	Ikawa	0,8	0,9
AS351	AS	Malaysia	Kon.	94,6	17,7	1,6	Ikawa	0,9	0,8
AS3	AS	Indien	Kon.	85,6	20,1	1,3	Ikawa	0,7	0,9
AS4	AS	Indien	Kon.	89,1	19,3	1,3	Ikawa	0,8	0,9
AS5	AS	Indien	Kon.	103,3	19,9	1,1	Ikawa	0,8	0,9
AS6	AS	Indien	Kon.	81,0	20,6	1,2	Ikawa	0,7	0,9
AS7	AS	Indien	Kon.	85,9	20,2	1,1	Ikawa	0,8	0,9
AS8	AS	Indien	Kon.	87,3	19,5	1,1	Ikawa	0,7	0,9
AS9	AS	Indien	Kon.	84,1	21,5	3,1	Ikawa	0,8	0,9
N100	NA	Guat.	Kon.	90,1	20,5	1,2	Ikawa	0,8	0,9
N102	NA	Guat.	Bio	105,2	22,8	2,7	Ikawa	0,9	0,8
N152	NA	Honduras	Kon.	76,8	21,5	1,3	Ikawa	0,8	1,0
N153	NA	Honduras	Kon.	77,9	21,9	1,3	Ikawa	0,7	0,9
N154	NA	Honduras	Kon.	81,1	22,9	1,4	Ikawa	0,7	0,9
N155	NA	Honduras	Kon.	95,6	24,1	1,2	Ikawa	0,8	0,9
N156	NA	Honduras	Kon.	79,6	23,6	1,4	Ikawa	0,8	1,0
N157	NA	Honduras	Kon.	85,5	17,8	1,4	Ikawa	0,8	0,9
N158	NA	Honduras	Kon.	67,8	23,3	1,1	Ikawa	0,7	0,9
N160	NA	Honduras	Kon.	77,8	21,4	1,2	Ikawa	0,8	0,9
N161	NA	Honduras	Bio	118,1	21,5	2,2	Ikawa	0,8	0,8
N1	NA	Costa R.	Kon.	85,1	22,5	1,4	Ikawa	0,9	0,8
N201	NA	Kuba	Kon.	119,8	20,0	2,3	Ikawa	0,8	0,9
N250	NA	Mexiko	Kon.	87,7	21,2	1,2	Ikawa	0,7	0,9
N251	NA	Mexiko	Kon.	93,6	21,2	1,1	Ikawa	0,8	0,9
N252	NA	Mexiko	Kon.	87,2	20,8	1,1	Ikawa	0,8	0,9
N253	NA	Mexiko	Kon.	92,9	20,1	1,2	Ikawa	0,8	0,9
N254	NA	Mexiko	Kon.	85,2	20,8	1,0	Ikawa	0,8	0,9
N255	NA	Mexiko	Kon.	92,2	20,9	1,2	Ikawa	0,8	0,9
N256	NA	Mexiko	Kon.	96,4	18,9	1,1	Ikawa	0,8	0,9
N257	NA	Mexiko	Kon.	105,0	19,8	2,0	Ikawa	0,8	0,8
N261	NA	Mexiko	Kon.	80,2	20,5	0,9	Ikawa	0,8	0,9
N262	NA	Mexiko	Kon.	92,4	20,7	1,3	Ikawa	0,7	0,9
N263	NA	Mexiko	Kon.	85,2	18,9	1,0	Ikawa	0,8	0,9
N264	NA	Mexiko	Bio	86,4	18,7	0,9	Ikawa	0,8	0,9
N265	NA	Mexiko	Bio	88,2	19,1	1,0	Ikawa	0,8	0,9
N266	NA	Mexiko	Bio	90,1	19,7	1,0	Ikawa	0,7	0,9
N267	NA	Mexiko	Bio	78,6	21,6	1,1	Ikawa	0,8	0,9
N286	NA	Mexiko	Kon.	86,7	21,4	1,4	Ikawa	0,8	0,9
N287	NA	Mexiko	Kon.	86,5	18,0	1,4	Ikawa	0,9	0,9
N288	NA	Mexiko	Kon.	99,5	19,8	1,3	Ikawa	0,8	0,8
N289	NA	Mexiko	Kon.	88,2	19,0	1,3	Ikawa	0,8	0,9
N290	NA	Mexiko	Kon.	88,2	20,7	1,2	Ikawa	0,8	0,8
N291	NA	Mexiko	Kon.	91,9	19,8	1,4	Ikawa	0,9	0,8
N292	NA	Mexiko	Kon.	87,3	19,3	1,2	Ikawa	0,7	1,0

Anhang

N293	NA	Mexiko	Kon.	93,6	18,4	1,5	Ikawa	0,7	0,9
N294	NA	Mexiko	Kon.	85,6	19,9	1,4	Ikawa	0,8	0,9
N295	NA	Mexiko	Kon.	85,6	18,3	1,3	Ikawa	0,8	1,0
N2	NA	Costa R.	Kon.	83,8	19,0	1,3	Ikawa	0,7	0,9
N500	NA	Jamaika	Kon.	55,3	23,3	1,3	Ikawa	0,7	0,9
N50	NA	El Sal.	Kon.	86,2	22,1	1,2	Ikawa	0,8	0,9
S102	SA	Peru	Bio	104,2	22,8	2,4	Ikawa	0,8	0,9
S103	SA	Peru	Bio	84,0	21,3	1,2	Ikawa	0,8	0,9
S1	SA	Brasilien	Kon.	84,1	23,8	0,9	Ikawa	0,7	0,9
S250	SA	Ecuador	Kon.	83,3	24,4	1,3	Ikawa	0,8	0,8
S2	SA	Brasilien	Kon.	87,2	24,3	1,0	Ikawa	0,7	0,9
S38	SA	Brasilien	Kon.	91,7	22,3	0,9	Ikawa	0,8	0,9
S3	SA	Brasilien	Kon.	92,3	23,3	1,1	Ikawa	0,8	0,9
S51_1	SA	Kolum.	Bio	79,3	20,0	1,4	Ikawa	0,7	0,9
S53_1	SA	Kolum.	Bio	94,5	20,6	2,2	Ikawa	0,7	0,8
S6	SA	Brasilien	Kon.	98,0	21,1	1,3	Ikawa	0,8	0,9

Tabella 9-2: Übersicht über die im Rahmen der zielgerichteten Analyse gemessenen Kaffeeproben mit Defekten. NY: New Yorker Klassifizierung

Nummer	Defekt	Klassifizierung	Nummer	Defekt	Klassifizierung
1	Kein		27	Dark Green Beans	NY4
2	Unripe Beans	NY2	28	Dark Green Beans	NY6
3	Unripe Beans	NY4	29	Bleached Beans	NY2
4	Unripe Beans	NY6	30	Bleached Beans	NY4
5	Eaten Beans	NY2	31	Bleached Beans	NY6
6	Eaten Beans	NY4	32	Pulper Cuts	NY2
7	Eaten Beans	NY6	33	Pulper Cuts	NY4
8	Well Eaten Beans	NY2	34	Pulper Cuts	NY6
9	Well Eaten Beans	NY4	35	Broken Beans	NY2
10	Well Eaten Beans	NY6	36	Broken Beans	NY4
11	Faded Beans	NY2	37	Broken Beans	NY6
12	Faded Beans	NY4	38	Husks	NY2
13	Faded Beans	NY6	39	Husks	NY4
14	Quaker	NY2	40	Husks	NY6
15	Quaker	NY4	41	Shells	NY2
16	Quaker	NY6	42	Shells	NY4
17	Malformed	NY2	43	Shells	NY6
18	Malformed	NY4	44	Parchment	NY2
19	Malformed	NY6	45	Parchment	NY4
20	Black Beans	NY2	46	Parchment	NY6
21	Black Beans	NY4	47	Pods	NY2
22	Black Beans	NY6	48	Pods	NY4
23	Sour/Brown Beans	NY2	49	Pods	NY6
24	Sour/Brown Beans	NY4	50	Twigs	NY2
25	Sour/Brown Beans	NY6	51	Twigs	NY4
26	Dark Green Beans	NY2	52	Twigs	NY6

9.2 Entwicklung einer Methode für die nicht zielgerichtete Untersuchung der lipophilen Metaboliten

Tabelle 9-3: Mittelwerte der Integrale der Signale 1 – 12 bestimmt aus einer Dreifachbestimmung für sieben getestete Extraktionsmethoden, welche zur Ermittlung der Extraktionseffizienz herangezogen wurden. Extraktionsmethoden: Folch et al.: Chloroform: Methanol: 2:1 (v/v); Ackermann et al.: Chloroform: Methanol: Wasser 10:5:1 (v/v/v); Bligh und Dyer: Chloroform: Methanol 1:2 (v/v); Lin et al.: Ethylacetat: Ethanol 1:1 (v/v), Aceton, Chloroform, Chloroform: Methanol 1:1 (v/v). Die Werte der Integrale wurden durch den Faktor 1000000 geteilt.

Signal	δ [ppm]	Folch et al.	Ackermann et al.	Bligh und Dyer	Lin et al.	Aceton	Chloroform	Chloroform : Methanol (1:1, v/v)
Signal 1	0,76-0,87	5465,5	5253,3	890,2	5710,4	5196,9	4702,4	4889,8
Signal 2	1,22-1,34	29068,6	26544,5	3716,2	34087,7	31097,6	28919,5	27690,1
Signal 3	1,46-1,66	5209,0	4871,7	1307,4	5807,1	5273,2	5019,5	4901,1
Signal 4	2,68-2,78	1400,8	1337,3	192,0	1623,0	1474,6	1319,0	1361,6
Signal 5	3,53-3,56	412,4	740,5	674,0	204,0	167,8	120,2	323,7
Signal 6	3,96-4,00	418,4	739,5	668,9	216,8	173,4	127,2	356,1
Signal 7	4,08-4,15	948,3	908,9	128,7	1103,3	996,8	895,5	949,8
Signal 8	4,21-4,26	269,0	281,8	122,0	290,2	257,5	289,9	278,7
Signal 9	5,21-5,34	3265,7	2938,4	274,2	3922,3	3575,4	3247,2	3162,8
Signal 10	5,84-5,90	74,2	75,3	36,0	82,0	73,9	101,1	71,9
Signal 11	7,21-7,23	84,2	107,9	45,1	85,9	78,0	55,1	89,8
Signal 12	7,25-7,29	92,8	99,3	54,4	80,1	71,8	98,7	80,4

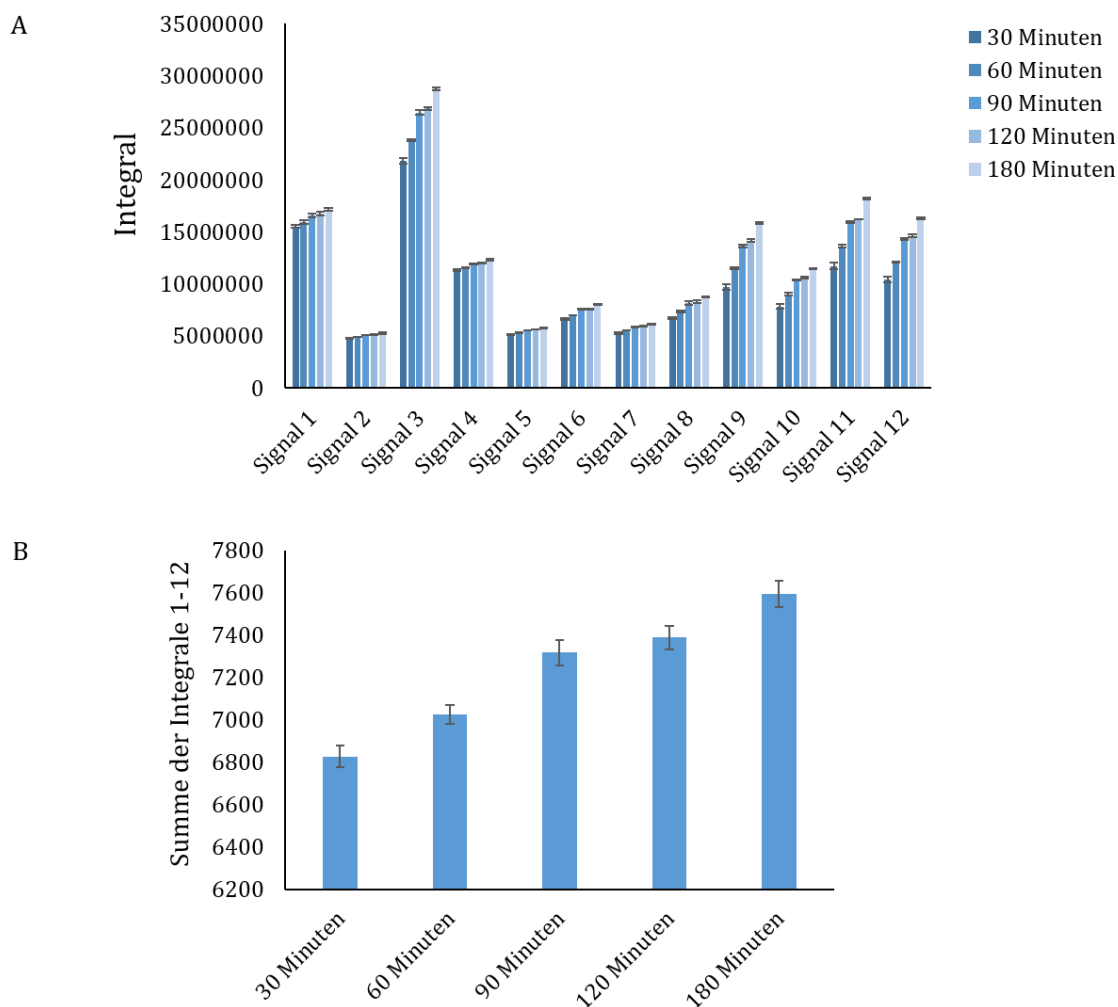


Abbildung 9-1: (A) Mittlere Integrale der Signale 1 – 12 für verschiedene Extraktionszeiten gebildet aus einer Dreifachbestimmung. Extraktionszeiten: 30, 60, 90, 120 und 180 min. Extraktionsmittel: Chloroform: Methanol (1:1, v/v), Messlösung (CDCl₃:MeOH-d₄ 1:1 (v/v)). Signal 1: $\delta_H = 0,85 - 0,91$ ppm; Signal 2: $\delta_H = 2,69 - 2,77$ ppm; Signal 3: $\delta_H = 3,55 - 3,58$ ppm; Signal 4: $\delta_H = 5,22 - 5,36$ ppm; Signal 5: $\delta_H = 5,85 - 5,97$ ppm; Signal 6: $\delta_H = 7,22 - 7,24$ ppm; Signal 7: $\delta_H = 7,27 - 7,29$ ppm; Signal 8: $\delta_H = 7,74 - 7,77$ ppm; Signal 9: $\delta_H = 8,27 - 8,32$ ppm; Signal 10: $\delta_H = 8,51 - 8,56$ ppm; Signal 11: $\delta_H = 8,77 - 8,81$ ppm; Signal 12: $\delta_H = 9,16 - 9,19$ ppm; Zur Vergleichbarkeit wurden die Integrale der Signale 1,2 und 4 um den Wert 200 und die Signale 3 und 5-8 um den Wert 10 geteilt (B) Gesamtintegral gebildet aus der Summe der Integrale 1-12 für verschiedene Extraktionszeiten. Extraktionszeiten: 30, 60, 90, 120 und 180 min. Die Werte des Gesamtintegrals wurden um den Wert 1000000 geteilt.

Tabelle 9-4: Mittlere Verfahrensvariationskoeffizienten gebildet aus der Dreifachbestimmung der Signalintegrale 1 – 12 für verschiedene Extraktionszeiten. Signal 1: $\delta_H = 0,85 - 0,91$ ppm; Signal 2: $\delta_H = 2,69 - 2,77$ ppm; Signal 3: $\delta_H = 3,55 - 3,58$ ppm; Signal 4: $\delta_H = 5,22 - 5,36$ ppm; Signal 5: $\delta_H = 5,85 - 5,97$ ppm; Signal 6: $\delta_H = 7,22 - 7,24$ ppm; Signal 7: $\delta_H = 7,27 - 7,29$ ppm; Signal 8: $\delta_H = 7,74 - 7,77$ ppm; Signal 9: $\delta_H = 8,27 - 8,32$ ppm; Signal 10: $\delta_H = 8,51 - 8,56$ ppm; Signal 11: $\delta_H = 8,77 - 8,84$ ppm; Signal 12: $\delta_H = 9,16 - 9,19$ ppm

Extraktionszeit	30 Minuten	60 Minuten	90 Minuten	120 Minuten	180 Minuten
ØVK [%]	1,5	0,7	0,7	0,7	0,7

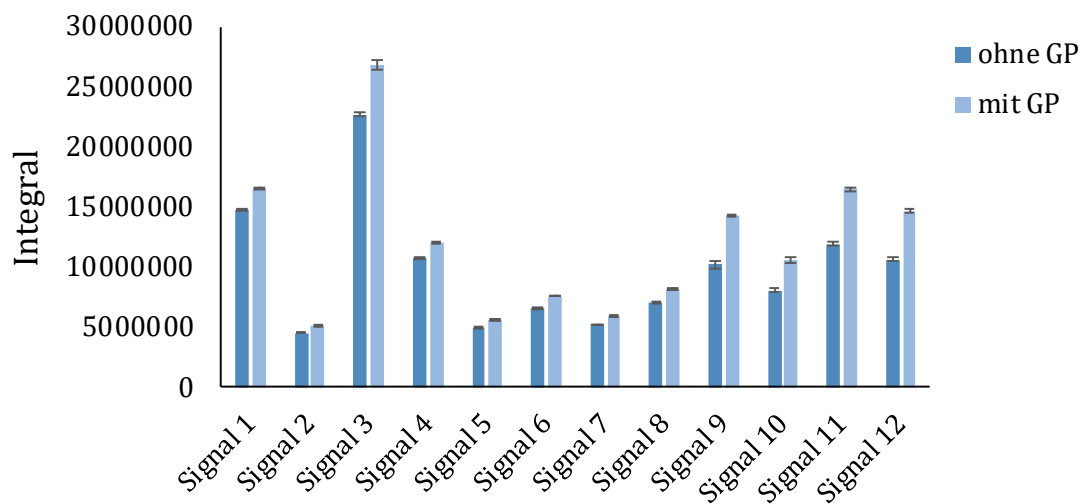


Abbildung 9-2: Mittlere Integrale der Signale 1 – 12 für die Extraktion mit und ohne Glasperlen (GP) gebildet aus einer Dreifachbestimmung. Extraktionszeit: 60 min, Extraktionsmittel: Chloroform: Methanol (1:1, v/v), Messlösung (CDCl₃:MeOH-d₄ 1:1 (v/v)). Signal 1: $\delta_H = 0,85 - 0,91$ ppm; Signal 2: $\delta_H = 2,69 - 2,77$ ppm; Signal 3: $\delta_H = 3,55 - 3,58$ ppm; Signal 4: $\delta_H = 5,22 - 5,36$ ppm; Signal 5: $\delta_H = 5,85 - 5,97$ ppm; Signal 6: $\delta_H = 7,22 - 7,24$ ppm; Signal 7: $\delta_H = 7,27 - 7,29$ ppm; Signal 8: $\delta_H = 7,74 - 7,77$ ppm; Signal 9: $\delta_H = 8,27 - 8,32$ ppm; Signal 10: $\delta_H = 8,51 - 8,56$ ppm; Signal 11: $\delta_H = 8,77 - 8,84$ ppm; Signal 12: $\delta_H = 9,16 - 9,19$ ppm

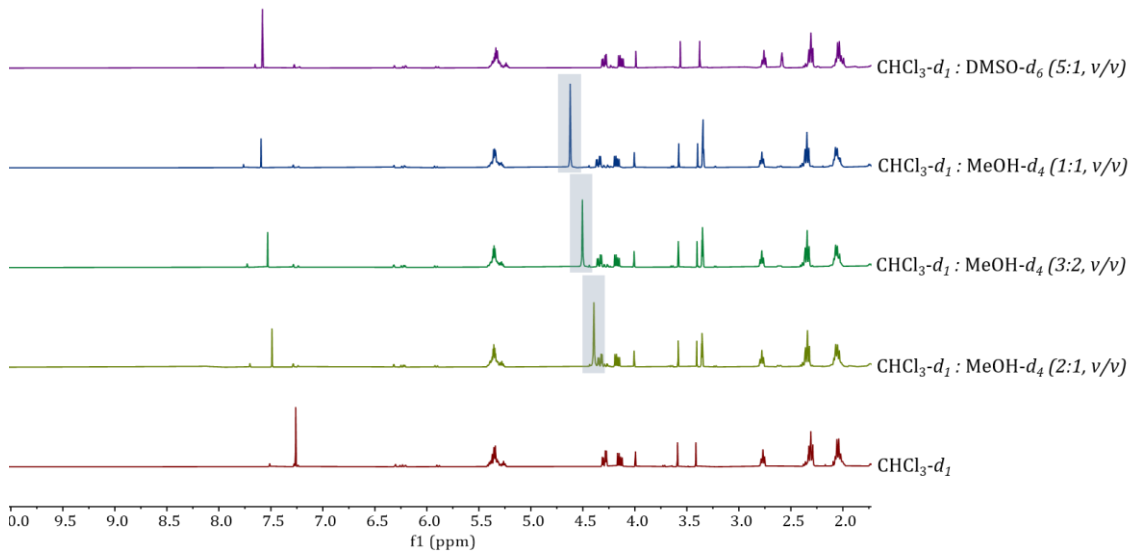


Abbildung 9-3: Ausschnitte verschiedener 400 MHz ¹H-NMR-Spektren im Bereich $\delta_H = 2,25 - 10,0$ ppm der lipophilen Kaffeemetaboliten, gemessen in verschiedenen Messlösungen. Die chemische Verschiebung wurde auf $\delta_{TMS} = 0,00$ ppm referenziert. Der grau hinterlegte Bereich stellt das Signal des enthaltenen Restwassers dar.

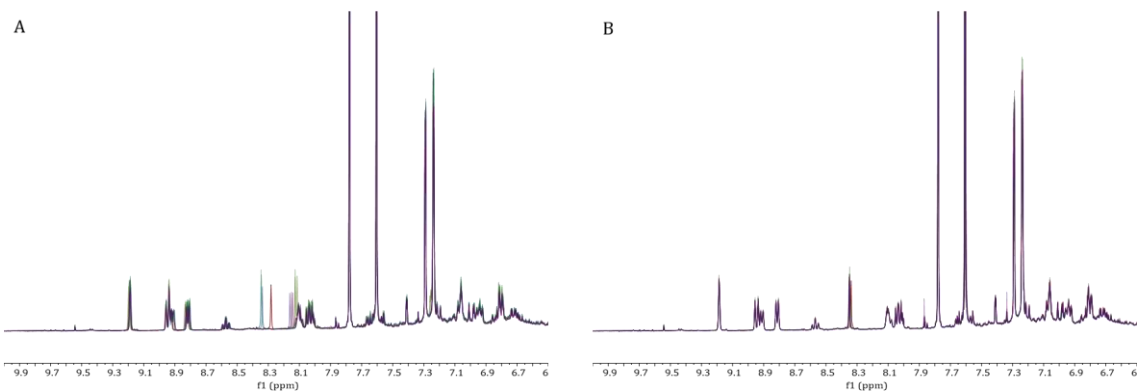


Abbildung 9-4: Ausschnitte verschiedener 400 MHz ¹H-NMR-Spektren im Bereich $\delta_H = 6,50 - 10,0$ ppm der lipophilen Kaffeemetaboliten, aufgearbeitet und gemessen an unterschiedlichen Tagen (Arbeitsvorschrift 8.2.4 B). Die chemische Verschiebung wurde auf $\delta_{TMS} = 0,00$ ppm referenziert. (A) Verwendung von unbehandeltem CDCl₃ in der Messlösung. (B) Verwendung von entschärftem CDCl₃ in der Messlösung. Vorgang der Entschärfung von CDCl₃ zur Entfernung reaktiver Abbauprodukte: 70 mL CDCl₃ mit 35 mL Natriumcarbonat-Lösung (100 g/L, aq.) in einem Scheidetrichter für 5 min ausschütteln. In einer Braunglasflasche 3,5 g getrocknetes Natriumcarbonat (150°C, 24h) abwiegen und nach der Phasentrennung untere Phase in Braunglasflasche ablassen. Entschärftes CDCl₃ über Nacht stehen lassen

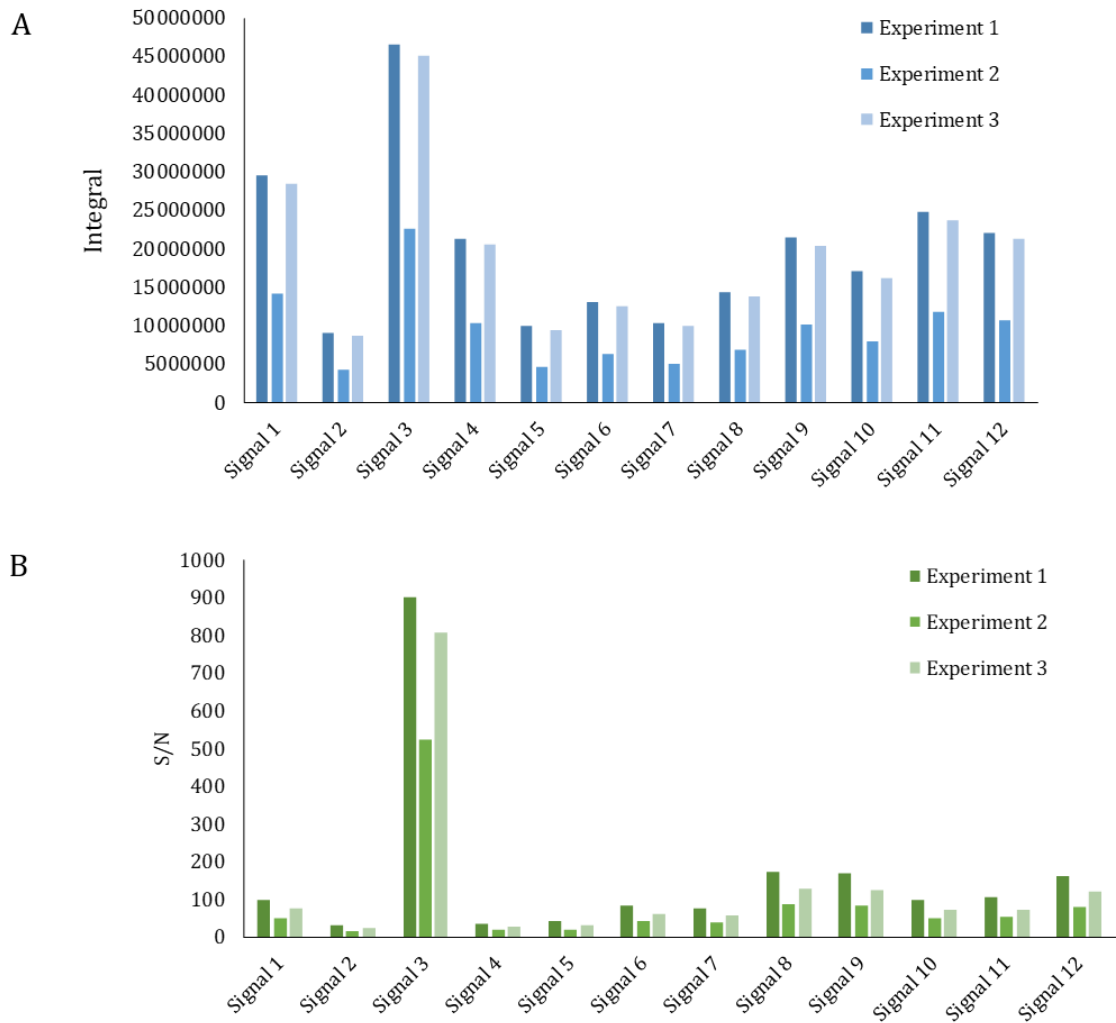


Abbildung 9-5: (A) Integrale und (B) Signal-zu-Rauschen-Verhältnisse (S/N) der Signale 1 – 12 für die Anwendung verschiedener Pulswinkel. Messlösung $CDCl_3:MeOH-d_4$ 1:1 (v/v); Signal 1: $\delta_H = 0,85 - 0,91$ ppm; Signal 2: $\delta_H = 2,69 - 2,77$ ppm; Signal 3: $\delta_H = 3,55 - 3,58$ ppm; Signal 4: $\delta_H = 5,22 - 5,36$ ppm; Signal 5: $\delta_H = 5,85 - 5,97$ ppm; Signal 6: $\delta_H = 7,22 - 7,24$ ppm; Signal 7: $\delta_H = 7,27 - 7,2$ ppm; Signal 8: $\delta_H = 7,74 - 7,77$ ppm; Signal 9: $\delta_H = 8,27 - 8,32$ ppm; Signal 10: $\delta_H = 8,51 - 8,56$ ppm; Signal 11: $\delta_H = 8,77 - 8,84$ ppm; Signal 12: $\delta_H = 9,16 - 9,19$ ppm; Zur Vergleichbarkeit wurden die Integrale der Signale 1,2 und 4 um den Wert 200 und die Signale 3 und 5-8 um den Wert 10 geteilt. Experiment 1: Pulswinkel: 90° ; Acquisitionszeit: 7,78 s; Anzahl Scans: 64; Relaxationsverzögerung: 15 s; Gesamtdauer: 25,8 min; Experiment 2: Pulswinkel: 30° ; Acquisitionszeit: 7,78 s; Anzahl Scans: 64; Relaxationsverzögerung: 15 s; Gesamtdauer: 25,8 min; Experiment 3: Pulswinkel: 30° ; Acquisitionszeit: 5,98 s; Anzahl Scans: 128; Relaxationsverzögerung: 9 s; Gesamtdauer: 31,5 min

Tabelle 9-5: T_1 -Zeiten verschiedener Signale im 400 MHz ^1NMR -Spektrum der lipophilen Kaffeemetaboliten gemessen in $\text{CDCl}_3:\text{MeOH-d}_4$ 1:1 (v/v) und bestimmt mittels Inversion Recovery Experiment.

Proton/Verbindung	δ_{H} [ppm]	T_1 [s]
Trigonellin	9,71-9,19	1,62
Trigonellin	8,70-8,81	1,61
N-Methylpyridinium	8,46-8,58	1,64
Ameisensäure	8,22-8,29	1,68
Koffein	7,67-7,77	1,71
Furfurylalkohol	7,34-7,38	1,50
Kahweol	7,22-7,27	1,57
Cafestol	7,15-7,22	1,25
Cafestol	6,12-6,18	1,53
Kahweol	5,82-5,92	1,81
Trigonellin	4,48-4,50	0,87
Glycerin	4,38-4,42	0,34
Glycerin	4,17-4,27	0,65
Koffein	3,92-3,98	0,25
Koffein	3,50-3,55	1,08
Koffein	3,34-3,38	2,73
Dimethylsulfon	2,93-3,03	0,82
Bis-allylische Protonen	2,66-2,80	0,76
CH_2 in α -Position der Carbonylklasse	2,26-2,34	2,14
CH_2 in α -Position einer Allylklasse	1,91-2,10	0,94
CH_2 in β -Position der Carbonylklasse	1,50-1,68	1,42
$(\text{CH}_2)_n$	1,18-1,39	0,24
Endständige CH_3	0,81-0,92	1,41

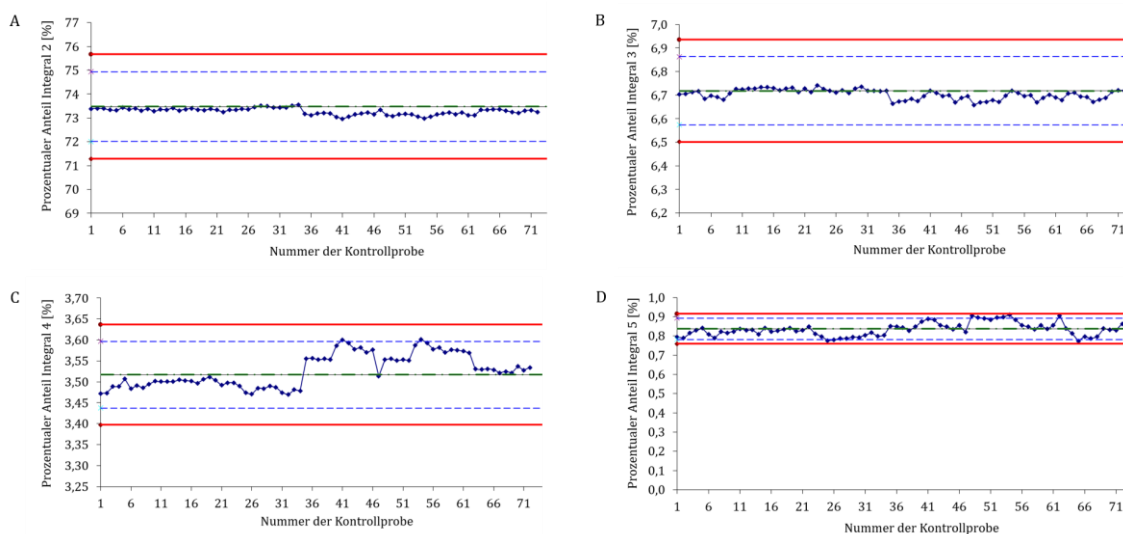


Abbildung 9-6: Kontrollkarten von (A) Integral 2 ($\delta_{\text{H}} = 1,23 - 1,33$ ppm); (B) Integral 3 ($\delta_{\text{H}} = 2,31 - 2,34$ ppm); (C) Integral 4 ($\delta_{\text{H}} = 2,72 - 2,78$ ppm); (D) Integral 5 ($\delta_{\text{H}} = 3,55 - 3,60$ ppm) ausgewählter Signale lipophiler Kaffeemetaboliten für insgesamt 74 Kontrollproben. Aus den Integralen der zehn Signale wurde das Gesamtintegral berechnet und anschließend der prozentuale Anteil jeden Integrals bezogen auf das Gesamtintegral bestimmt. Die grüne Linie zeigt den Mittelwert. Die rote, durchgezogene Linie stellt die Kontrollgrenze dar (Mittelwert ± 3 -Standardabweichung). Die blaue, gestrichelte Linie beschreibt die Warngrenze (Mittelwert ± 2 -Standardabweichung).

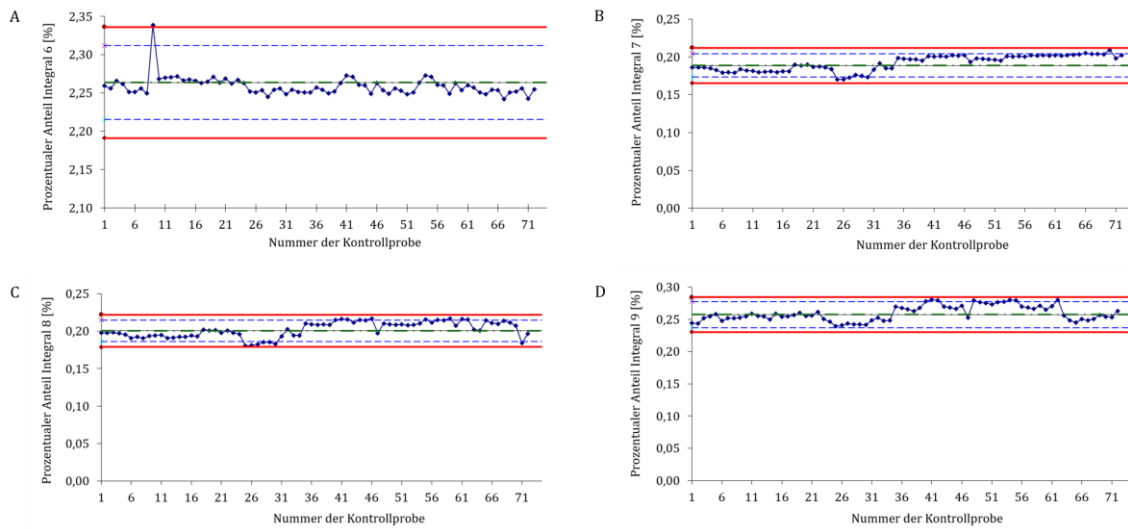


Abbildung 9-7: Kontrollkarten von (A) Integral 6 ($\delta_H = 4,31 - 4,39$ ppm); (B) Integral 7 ($\delta_H = 5,87 - 5,96$ ppm); (C) Integral 8 ($\delta_H = 7,26 - 7,31$ ppm); (D) Integral 9 ($\delta_H = 7,74 - 7,81$ ppm) ausgewählter Signale lipophiler Kaffeemataboliten für insgesamt 74 Kontrollproben. Aus den Integralen der zehn Signale wurde das Gesamtintegral berechnet und anschließend der prozentuale Anteil jeden Integrals bezogen auf das Gesamtintegral bestimmt. Die grüne Linie zeigt den Mittelwert. Die rote, durchgezogene Linie stellt die Kontrollgrenze dar (Mittelwert ± 3 ·Standardabweichung). Die blaue, gestrichelte Linie beschreibt die Warngrenze (Mittelwert ± 2 ·Standardabweichung).

9.3 Entwicklung einer Methode für die nicht zielgerichtete Untersuchung der hydrophilen Metaboliten

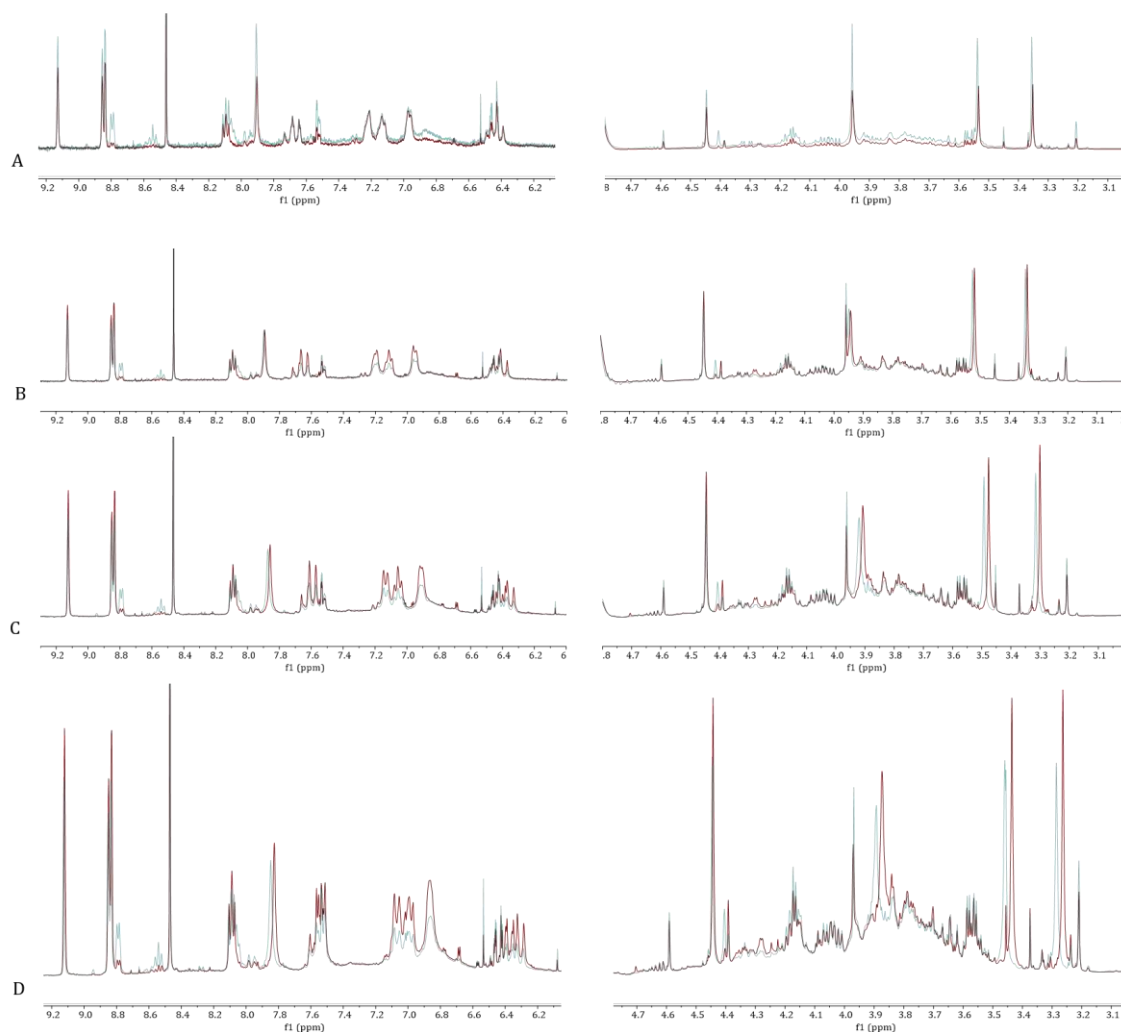


Abbildung 9-8: Ausschnitte verschiedener 400 MHz ^1H -NMR-Spektren im Bereich $\delta_{\text{H}} = 6,00 - 9,30$ ppm und $\delta_{\text{H}} = 3,00 - 4,80$ ppm einer Kaffeeprobe mit hellem Röstgrad (blau) und dunklem Röstgrad (rot), gemessen in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die Lösungsmittelunterdrückung erfolgte mit dem NOESY-Pulsprogramm. Die chemische Verschiebung wurde auf $\delta_{\text{TSP}} = 0,00$ ppm referenziert. Getestet wurden die Probeneinwaagen (A) 0,10g; (B) 0,20 g; (C) 0,60 g und (D) 1,20 g. Die Extraktion erfolgte in jeweils 8 mL H_2O

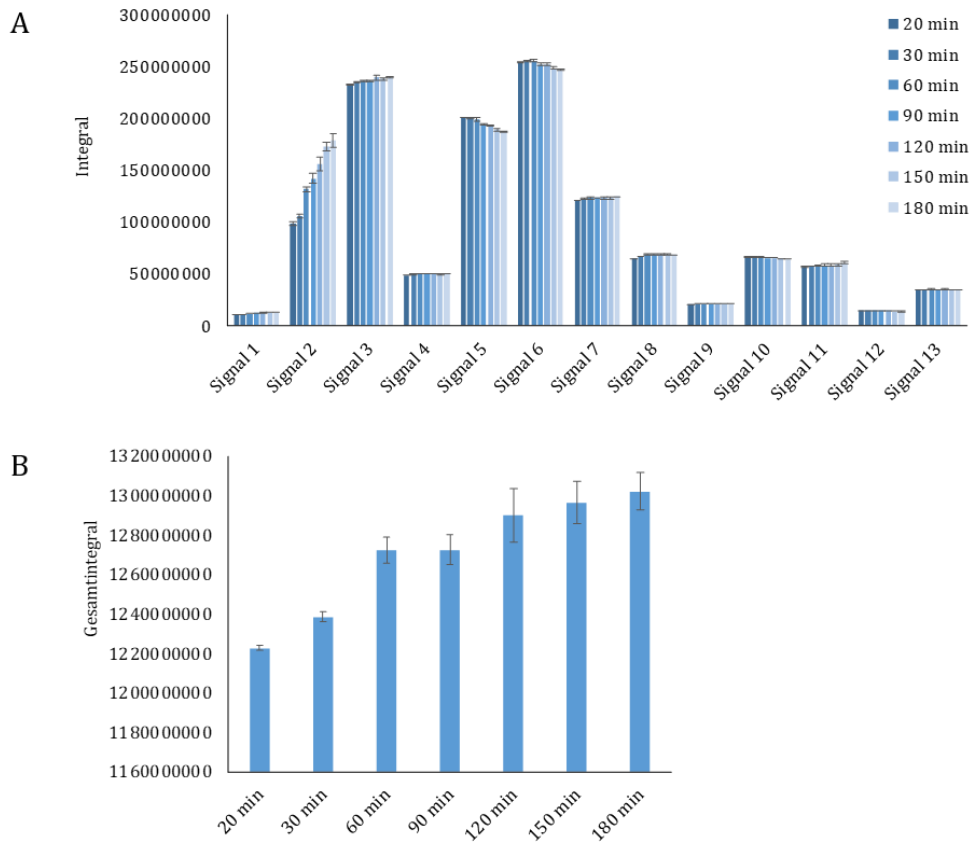


Abbildung 9-9: (A) Mittlere Integrale der Signale 1-13 für verschiedene Extraktionszeiten gebildet aus einer Dreifachbestimmung. Extraktionszeiten: 20, 30, 60, 90, 120, 150 und 180 min. Extraktionsmittel: H₂O; Messlösung H₂O:D₂O 9:1 (v/v). Signal 1: $\delta_H = 1,05 - 1,06$ ppm; Signal 2: $\delta_H = 1,32 - 1,33$ ppm; Signal 3: $\delta_H = 1,92 - 1,93$ ppm; Signal 4: $\delta_H = 3,19 - 3,20$ ppm; Signal 5: $\delta_H = 3,33 - 3,34$ ppm; Signal 6: $\delta_H = 3,49 - 3,51$ ppm; Signal 7: $\delta_H = 4,42 - 4,44$ ppm; Signal 8: $\delta_H = 7,05 - 7,11$ ppm; Signal 9: $\delta_H = 7,51 - 7,53$ ppm; Signal 10: $\delta_H = 7,86 - 7,89$ ppm; Signal 11: $\delta_H = 8,42 - 8,45$ ppm; Signal 12: $\delta_H = 8,49 - 8,54$ ppm; Signal 13: $\delta_H = 9,09 - 9,12$ ppm; Extraktionszeiten: 20, 30, 60, 90, 120, 150 und 180 min

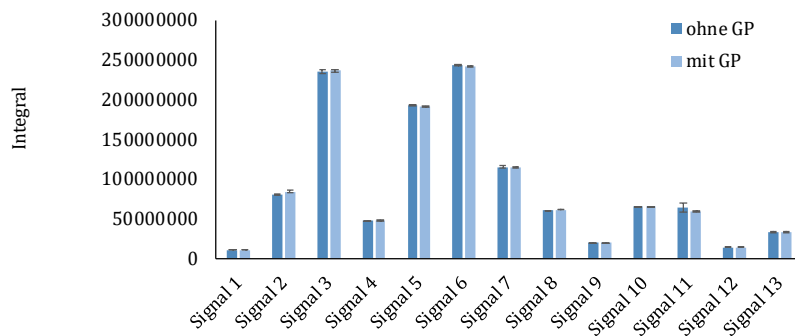


Abbildung 9-10: Mittlere Integrale der Signale 1 - 13 für die Extraktion mit und ohne Glasperlen (GP) gebildet aus einer Dreifachbestimmung. Extraktionszeit: 2 min, Extraktionsmittel: H₂O; Messlösung H₂O:D₂O 9:1 (v/v). Signal 1: $\delta_H = 1,05 - 1,06$ ppm; Signal 2: $\delta_H = 1,32 - 1,33$ ppm; Signal 3: $\delta_H = 1,92 - 1,93$ ppm; Signal 4: $\delta_H = 3,19 - 3,20$ ppm; Signal 5: $\delta_H = 3,33 - 3,34$ ppm; Signal 6: $\delta_H = 3,49 - 3,51$ ppm; Signal 7: $\delta_H = 4,42 - 4,44$ ppm; Signal 8: $\delta_H = 7,05 - 7,11$ ppm; Signal 9: $\delta_H = 7,51 - 7,53$ ppm; Signal 10: $\delta_H = 7,86 - 7,89$ ppm; Signal 11: $\delta_H = 8,42 - 8,45$ ppm; Signal 12: $\delta_H = 8,49 - 8,54$ ppm; Signal 13: $\delta_H = 9,09 - 9,12$ ppm

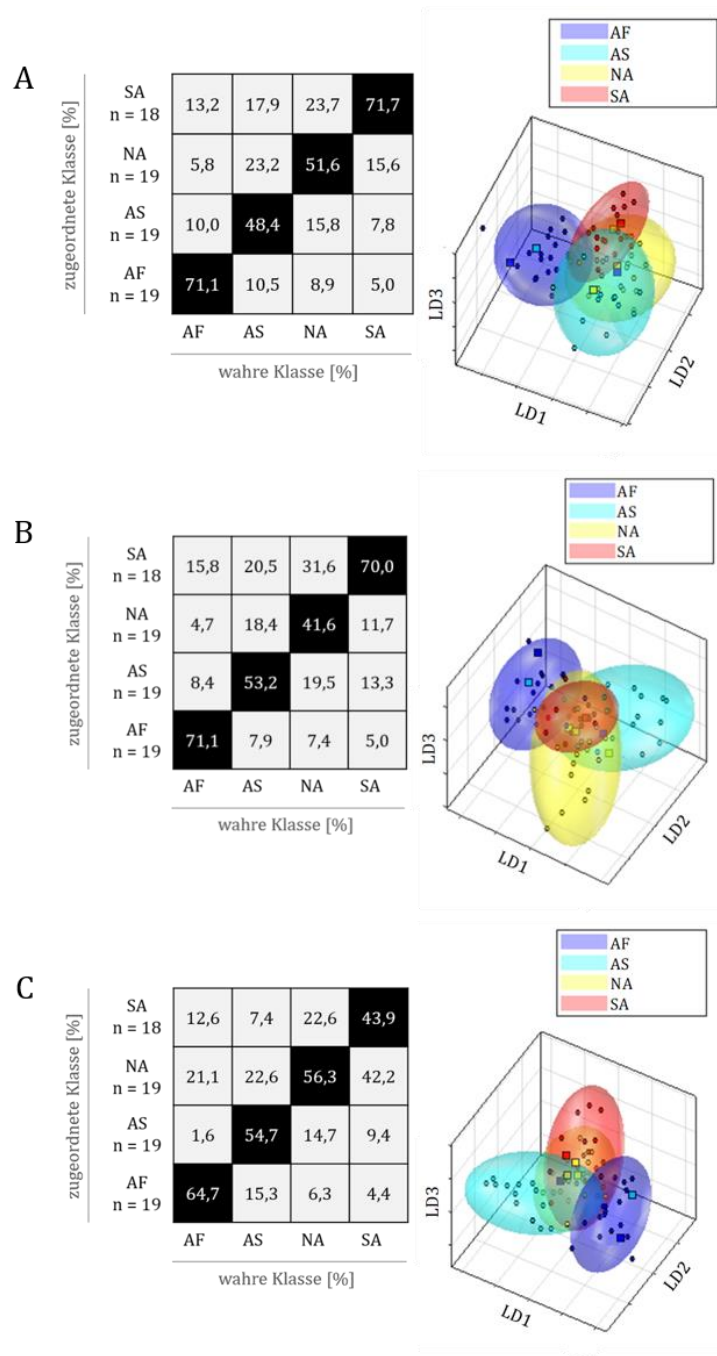


Abbildung 9-11: Ergebnisse der Erstellung der Klassifizierungsmodelle mittels PCA-LDA zur Vorhersage des geographischen Ursprungs von Kaffee bezüglich des Kontinents aus dem ¹H-NMR-Spektrum der hydrophilen Kaffeemetaboliten unter Anwendung verschiedener Filter. (A) Membranfiltration während der Probenvorbereitung. (B) Anwendung der Ultrafiltration mit einem 2 kDa-Zentrifugenfilter. (C) Anwendung der Ultrafiltration mit einem 1 kDa-Zentrifugenfilter. Als interne Validierung wurde eine zehnfache MCKV durchgeführt und der Datensatz wurde zu 90% Trainingsdatensatz und 10% Testdatensatz eingeteilt. Jeweils links in der Abbildung (A)-(C) ist der LDA-Raum eines Kreuzvalidierungsschrittes dargestellt. Das 95%ige Prognoseellipsoid des Trainingsdatensatzes ist als Wolke dargestellt, die Testdatensatz-Proben werden durch Punkte symbolisiert. Rechts in der Abbildung (A)-(C) ist die Konfusionsmatrix der MCKV dargestellt. Die x-Achse und die y-Achse zeigen die zugeordnete Klasse bzw. die wahre Klasse. In der Konfusionsmatrix ist die Wahrscheinlichkeit des Vorhersageergebnisses in Prozent angegeben. Verwendetes Experiment NOESY-Pulssequenz: $\delta_H=0,5 - 9,8$ ppm; 1000 Bucket; Ausschluss $\delta_H = 4,72 - 4,88$ ppm; Normalisierung: $\delta_H = -0,10 - 0,10$ ppm; log-Transformation; PCA-Dimensionen: 15; n: Anzahl der Proben; AF: Afrika; AS: Asien; NA: Nordamerika; SA: Südamerika; LD: lineare Diskriminanzfunktion; PCA: Principle Component Analysis; MCKV: Monte-Carlo-Kreuzvalidierung; LDA: Lineare Diskriminanzanalyse

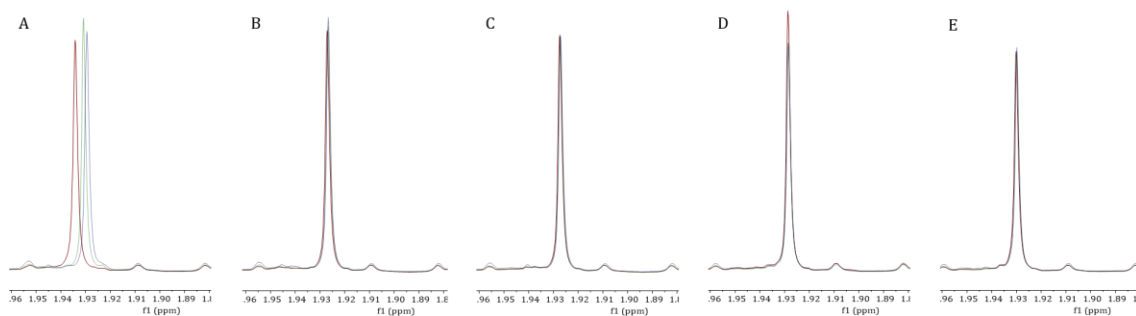


Abbildung 9-12: Ausschnitte verschiedener 400 MHz ^1H -NMR-Spektren im Bereich $\delta_{\text{H}} = 1,80 - 1,96$ ppm einer wässrigen Kaffeeprobe mit hellem Röstgrad (rot), mittlerem Röstgrad (blau) und dunklem Röstgrad (grün), gemessen in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die Lösungsmittelunterdrückung erfolgte mit dem NOESY-Pulsprogramm. Die chemische Verschiebung wurde auf $\delta_{\text{TSP}} = 0,00$ ppm referenziert. Getestet wurde die Verwendung eines Natriumphosphatpuffers mit dem pH-Wert = 6 und verschiedenen Konzentrationen. (A) $c = 0,10$ mol/L; (B) $c = 0,50$ mol/L; (C) $c = 1,00$ mol/L; (D) $c = 2,00$ mol/L; (E) $c = 3,00$ mol/L

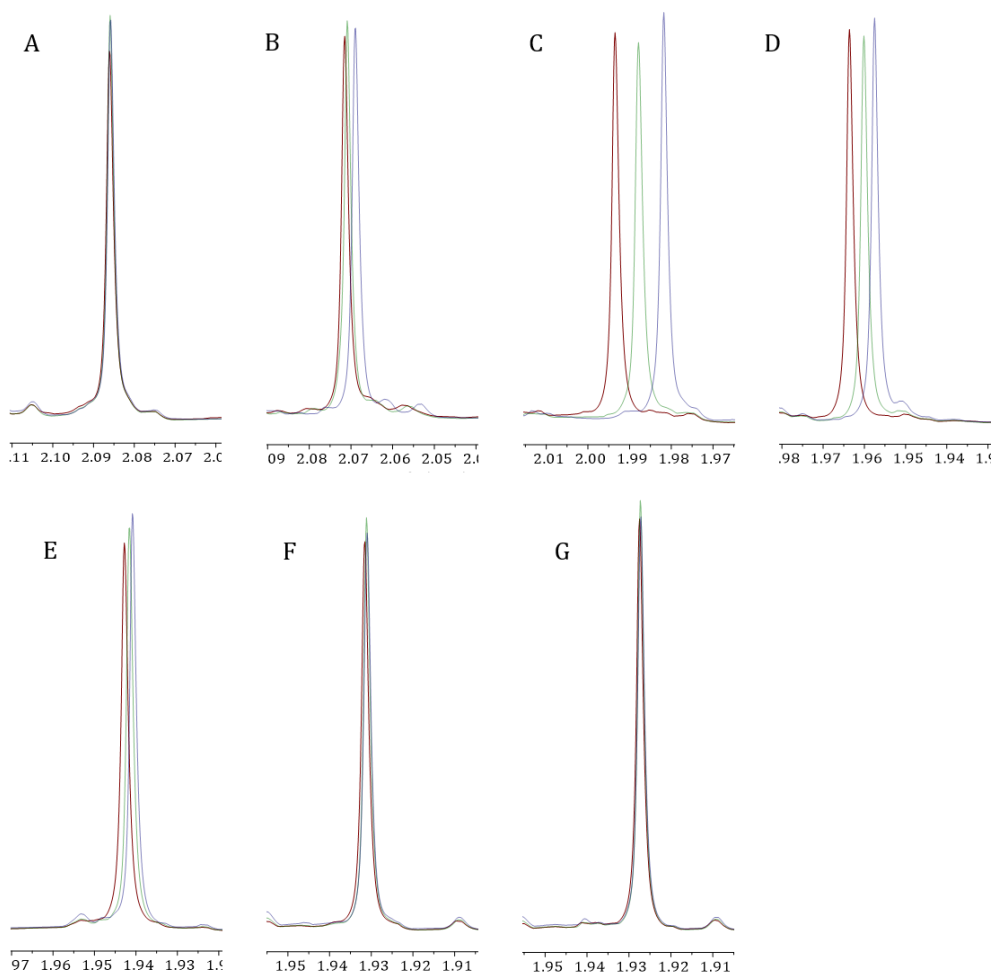


Abbildung 9-13: Ausschnitte verschiedener 400 MHz ^1H -NMR-Spektren im Bereich $\delta_{\text{H}} = 1,90 - 2,10$ ppm einer wässrigen Kaffeeprobe mit hellem Röstgrad (rot), mittlerem Röstgrad (blau) und dunklem Röstgrad (grün), gemessen in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die Lösungsmittelunterdrückung erfolgte mit dem NOESY-Pulsprogramm. Die chemische Verschiebung wurde auf $\delta_{\text{TSP}} = 0,00$ ppm referenziert. Getestet wurde die Verwendung eines Natriumphosphatpuffers mit der Konzentration von 1 mol/L und verschiedenen pH-Werten. (A) pH = 2,00; (B) pH = 3,00; (C) pH = 4,00; (D) pH = 4,50; (E) pH = 5,00; (F) pH = 5,50; (G) pH = 6,00

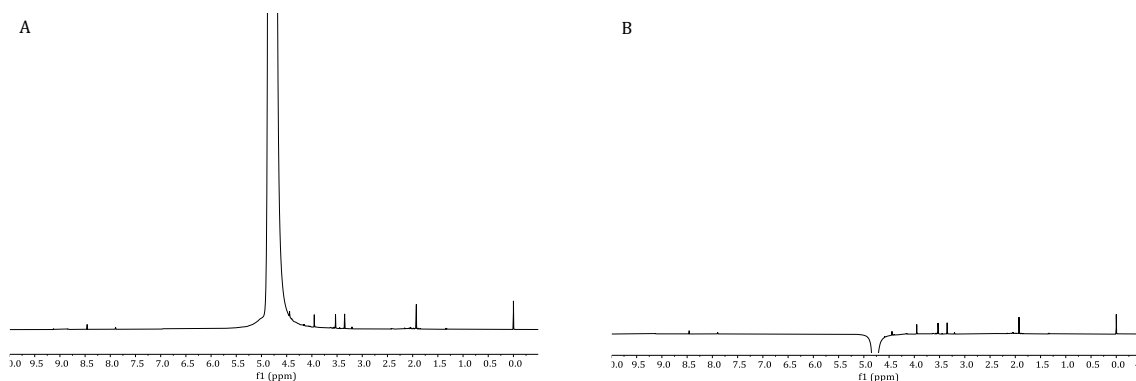


Abbildung 9-14: Ausschnitte verschiedener 400 MHz ^1H -NMR-Spektren im Bereich $\delta_{\text{H}} = -0,50 - 10,0$ ppm der hydrophilen Kaffeemetaboliten, gemessen in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die chemische Verschiebung wurde auf $\delta_{\text{TSP}} = 0,00$ ppm referenziert. (A) Ohne Lösungsmittelunterdrückung (Pulsprogramm: zg); (B) Lösungsmittelunterdrückung mittels Water Supression Enhanced Through T_1 Effects (WET, Pulsprogramm: wet)

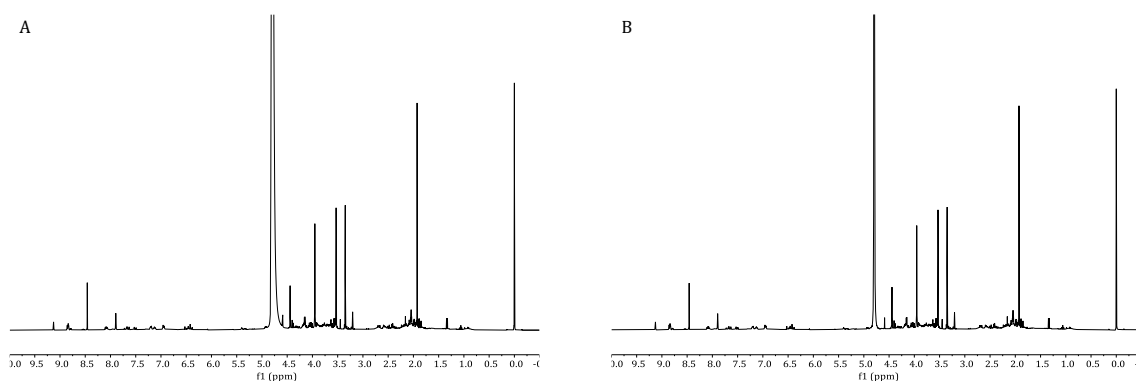


Abbildung 9-15: Ausschnitte verschiedener 400 MHz ^1H -NMR-Spektren im Bereich $\delta_{\text{H}} = -0,50 - 10,0$ der hydrophilen Kaffeemetaboliten, gemessen in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die chemische Verschiebung wurde auf $\delta_{\text{TSP}} = 0,00$ ppm referenziert. (A) Lösungsmittelunterdrückung mit Vorsättigungspuls (Pulsprogramm: zgpr); (B) Lösungsmittelunterdrückung mit Vorsättigungspuls und nachfolgendem Composite Puls (Pulsprogramm: zgcppr)

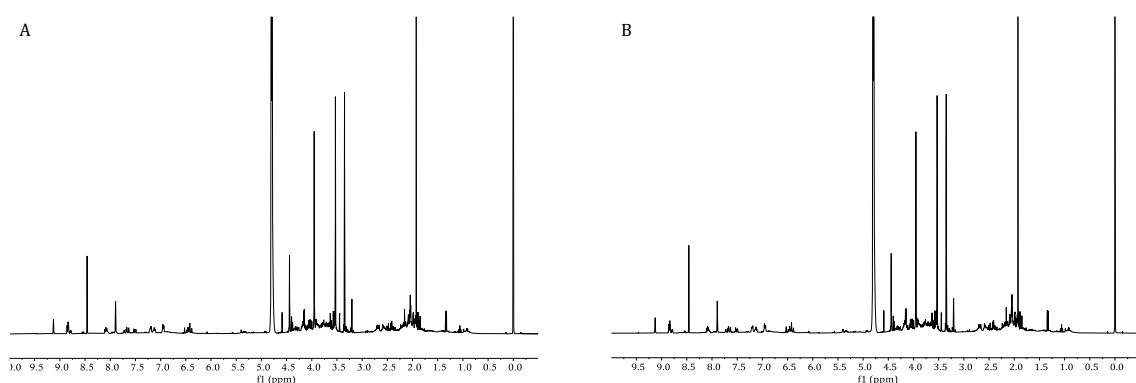


Abbildung 9-16: Ausschnitte verschiedener 400 MHz ^1H -NMR-Spektren im Bereich $\delta_{\text{H}} = -0,50 - 10,0$ ppm der hydrophilen Kaffeemetaboliten, gemessen in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die chemische Verschiebung wurde auf $\delta_{\text{TSP}} = 0,00$ ppm referenziert. (A) Lösungsmittelunterdrückung mittels 1D- Nuclear Overhauser Effect Spectroscopy (NOESY) (Pulsprogramm: noesyppr1d); (B) Lösungsmittelunterdrückung mittels 1D-NOESY unter Verwendung der zusätzlichen Wartezeit D7 (Pulsprogramm: noesyppr1d_d7)

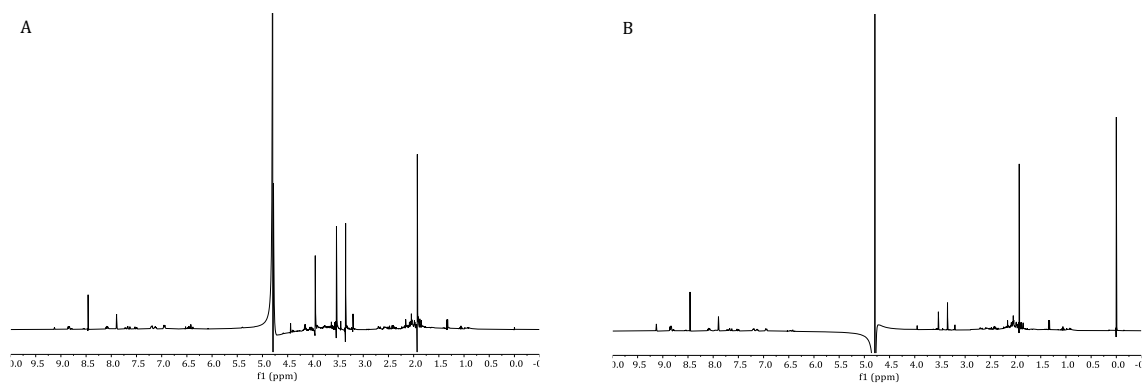


Abbildung 9-17: Ausschnitte verschiedener 400 MHz ^1H -NMR-Spektren im Bereich $\delta_{\text{H}} = -0,50 - 10,0$ ppm der hydrophilen Kaffeemetaboliten, gemessen in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die chemische Verschiebung wurde auf $\delta_{\text{TSP}} = 0,00$ ppm referenziert. (A) Lösungsmittelunterdrückung mittels Water Supression by Gradient Tailored Excitation (WATERGATE) (Pulsprogramm: p3919); (B) Lösungsmittelunterdrückung mittels WATERGATE (Pulsprogramm: zggpwg)

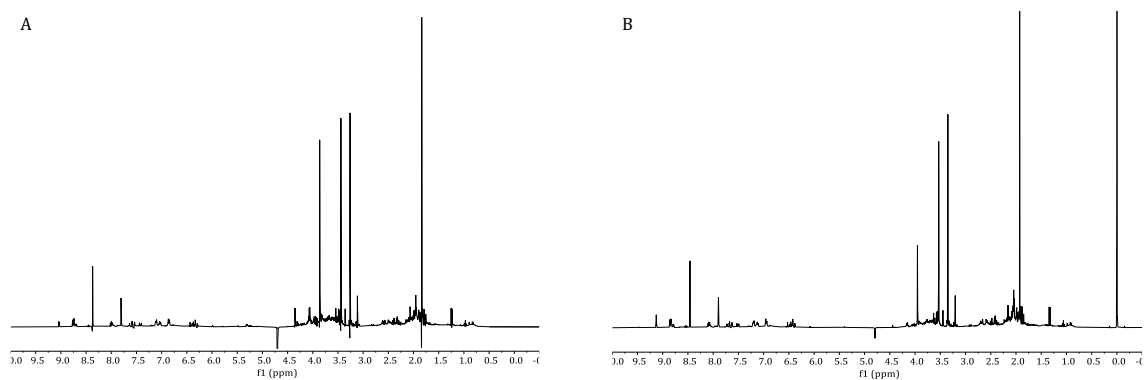


Abbildung 9-18: Ausschnitte verschiedener 400 MHz ^1H -NMR-Spektren im Bereich $\delta_{\text{H}} = -0,50 - 10,0$ ppm der hydrophilen Kaffeemetaboliten, gemessen in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die chemische Verschiebung wurde auf $\delta_{\text{TSP}} = 0,00$ ppm referenziert. (A) Lösungsmittelunterdrückung mittels Excitation-Sculpting (Pulsprogramm: zggpw5); (B) Lösungsmittelunterdrückung mittels Excitation-Sculpting (Pulsprogramm: zgesgp)

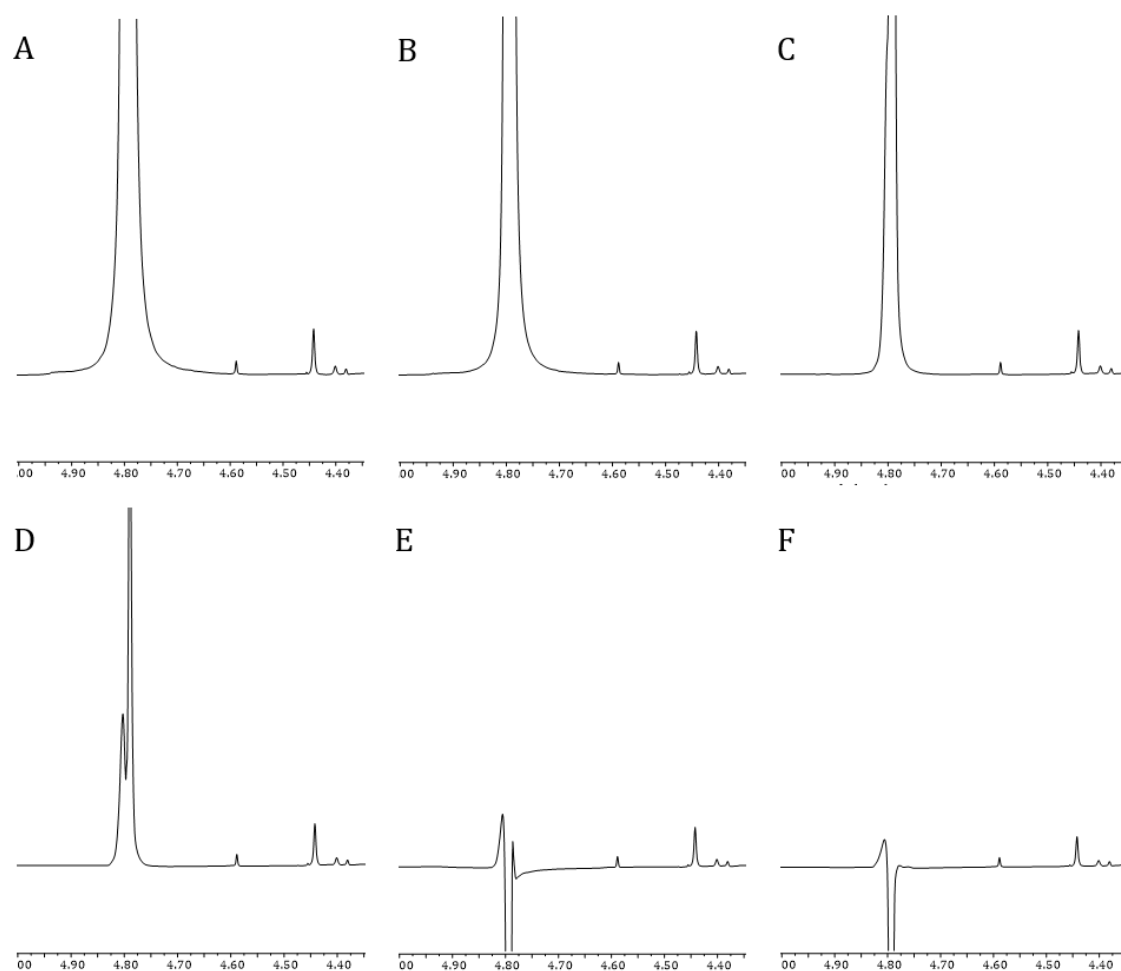


Abbildung 9-19: Ausschnitte verschiedener 400 MHz ^1H -NMR-Spektren im Bereich $\delta_{\text{H}} = 4,35 - 5,00$ ppm der hydrophilen Kaffeemetaboliten, gemessen in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die chemische Verschiebung wurde auf $\delta_{\text{TSP}} = 0,00$ ppm referenziert. Wasserunterdrückung mittels 1D-NOESY (Pulsprogramm: noesygppr1d_d7) unter Verwendung verschiedener Leistungen des Sättigungspulses. (A) Leistung = 10 Hz, (B) Leistung = 15 Hz; (C) Leistung = 20 Hz; (D) Leistung = 25 Hz; (E) Leistung = 30 Hz; (F) Leistung = 50 Hz

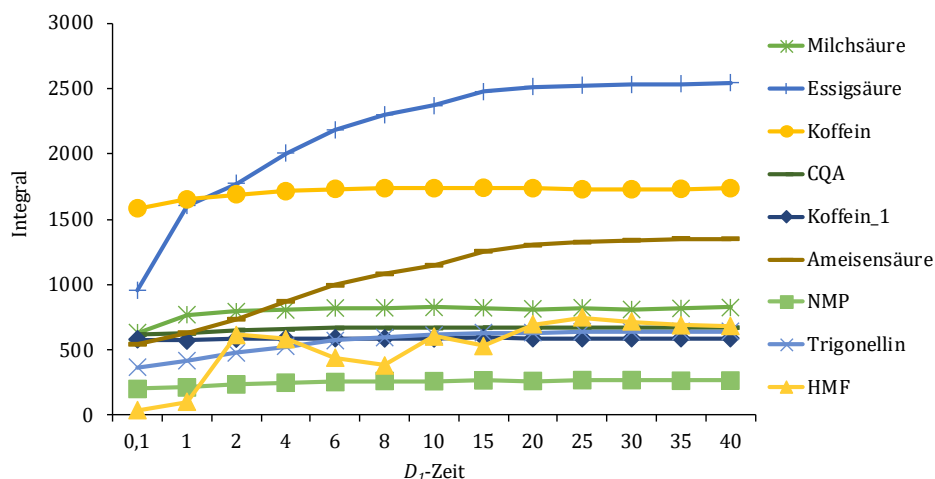


Abbildung 9-20: Absolute Integrale der Signale von Milchsäure ($\delta_H = 1,32 - 1,36$ ppm), Essigsäure ($\delta_H = 1,92 - 1,94$ ppm), Koffein ($\delta_H = 3,32 - 3,36$ ppm), CQA ($\delta_H = 6,90 - 6,99$ ppm), Koffein_1 ($\delta_H = 7,87 - 7,92$ ppm), Ameisensäure ($\delta_H = 8,42 - 8,49$ ppm), NMP ($\delta_H = 8,49 - 8,60$ ppm), Trigonellin ($\delta_H = 9,08 - 9,17$ ppm) und HMF ($\delta_H = 9,40 - 9,48$ ppm) in Abhängigkeit unterschiedlich gewählter D_1 -Zeiten des Experimentes „noesygppr1d“. Ausgewählte D_1 -Zeiten: 0,1 s; 1,0 s; 2,0 s; 4,0 s; 6,0 s; 8,0 s; 10,0 s; 15,0 s; 20,0 s; 25,0 s; 30,0 s; 40,0 s. Zur besseren Darstellung wurden die Werte der Integrale der Milchsäure, Ameisensäure, NMP und Trigonellin um den Faktor 50000, die Werte der Integrale von Essigsäure, Koffein, CQA und Koffein_1 um den Faktor 100000 und der Wert der Integrale von HMF um den Faktor 1000 geteilt. CQA: Caffeoylchinasäure-Isomere; NMP: N-Methylpyridinium; HMF: 5-Hydroxymethylfurfural

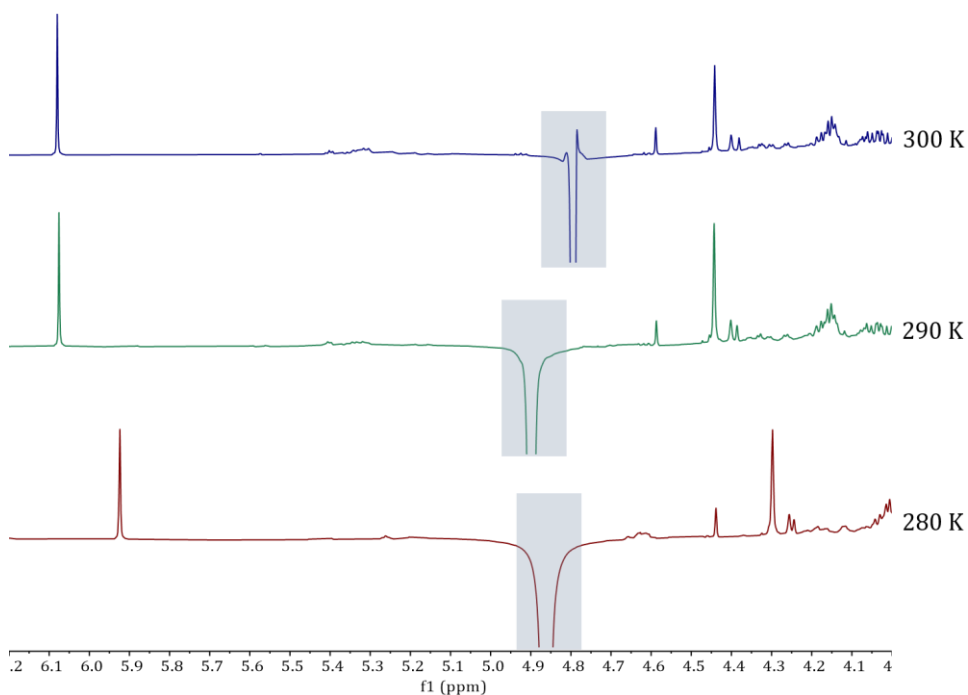


Abbildung 9-21: Ausschnitte verschiedener 400 MHz ^1H -NMR-Spektren im Bereich $\delta_H = 4,00 - 6,20$ ppm der lipophilen Kaffeemetaboliten, gemessen in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die chemische Verschiebung wurde auf $\delta_{\text{TSP}} = 0,00$ ppm referenziert. Wasserunterdrückung mittels 1D-NOESY (Pulsprogramm: noesygppr1d_d7) unter Verwendung verschiedener Messtemperaturen. Der grau hinterlegte Bereich stellt das Signal des Restwassers dar.

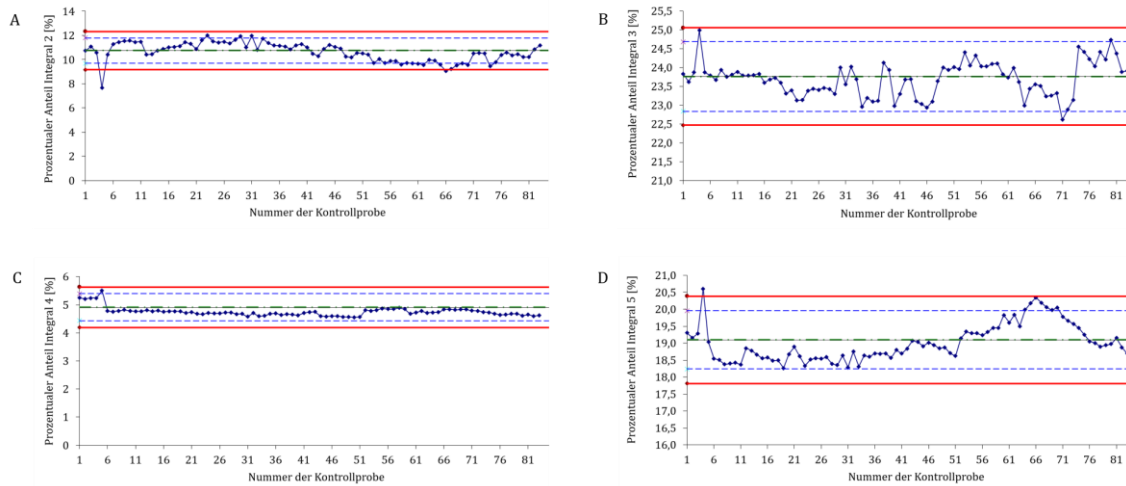


Abbildung 9-22: Kontrollkarten von (A) Integral 2 ($\delta_H = 1,3 - 1,37$ ppm); (B) Integral 3 ($\delta_H = 1,92 - 1,94$ ppm); (C) Integral 4 ($\delta_H = 3,18 - 3,22$ ppm); (D) Integral 5 ($\delta_H = 3,33 - 3,36$ ppm) der Signale ausgewählter hydrophiler Kaffeemetaboliten für insgesamt 84 Kontrollproben. Aus den Integralen der zehn Signalen wurde das Gesamtintegral berechnet und anschließend der prozentuale Anteil jeden Integrals bezogen auf das Gesamtintegral bestimmt. Die grüne Linie zeigt den Mittelwert. Die rote, durchgezogene Linie stellt die Kontrollgrenze dar (Mittelwert ± 3 -Standardabweichung). Die blaue, gestrichelte Linie beschreibt die Warngrenze (Mittelwert ± 2 -Standardabweichung).

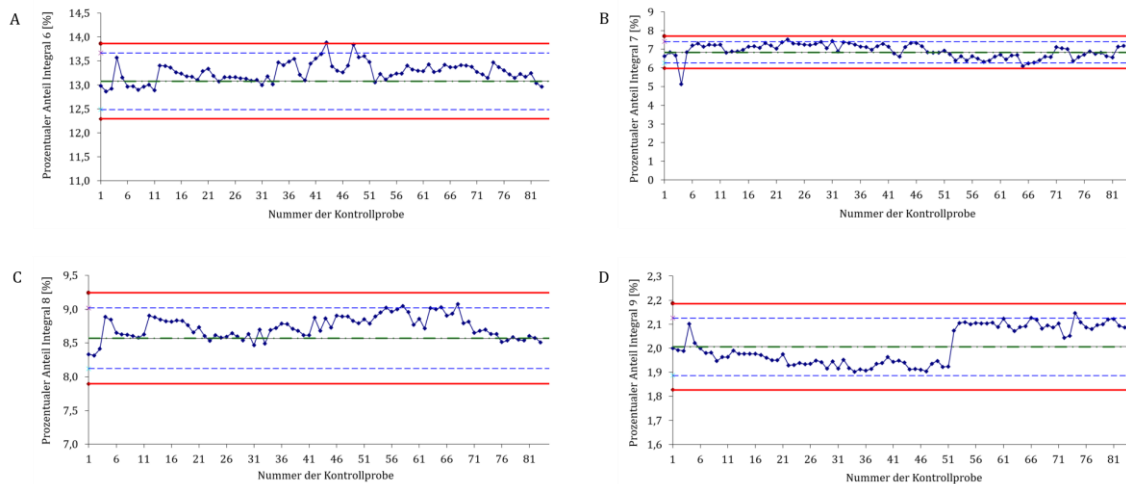


Abbildung 9-23: Kontrollkarten von (A) Integral 6 ($\delta_H = 4,42 - 4,47$ ppm); (B) Integral 7 ($\delta_H = 5,27 - 5,38$ ppm); (C) Integral 8 ($\delta_H = 6,89 - 7,00$ ppm); (D) Integral 9 ($\delta_H = 7,51 - 7,55$ ppm) der Signale ausgewählter hydrophiler Kaffeemetaboliten für insgesamt 84 Kontrollproben. Aus den Integralen der zehn Signalen wurde das Gesamtintegral berechnet und anschließend der prozentuale Anteil jeden Integrals bezogen auf das Gesamtintegral bestimmt. Die grüne Linie zeigt den Mittelwert. Die rote, durchgezogene Linie stellt die Kontrollgrenze dar (Mittelwert ± 3 -Standardabweichung). Die blaue, gestrichelte Linie beschreibt die Warngrenze (Mittelwert ± 2 -Standardabweichung).

9.4 Nicht zielgerichtete Untersuchung mittels Verfahren der uni- und multivariaten Statistik

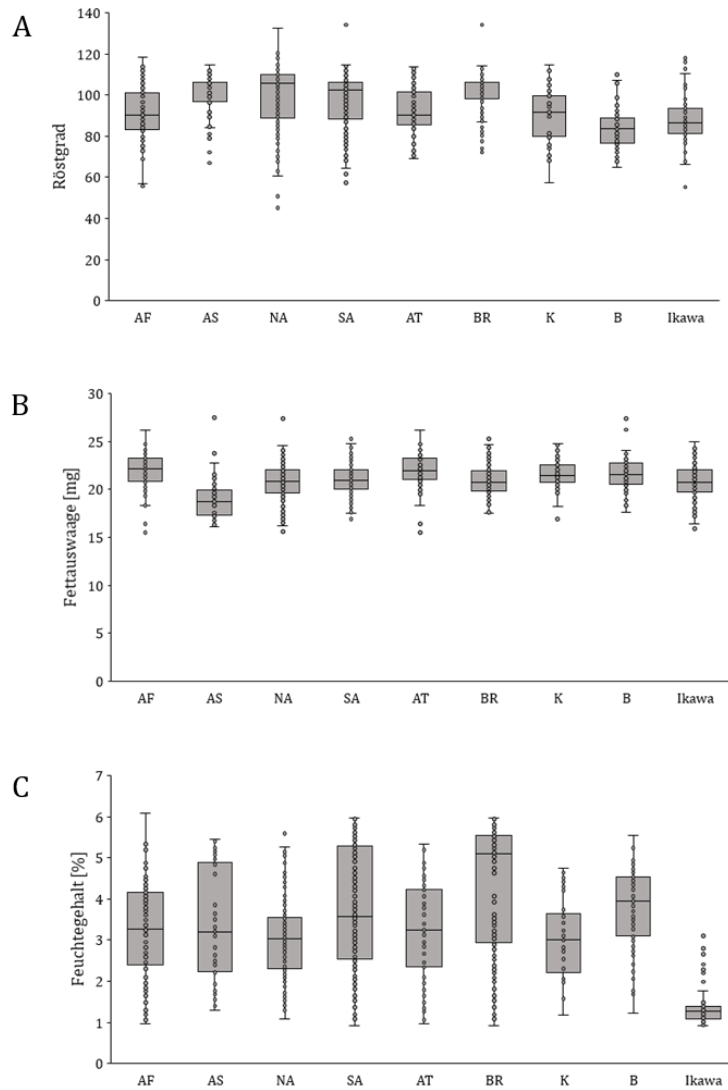


Abbildung 9-24: Darstellung der Metadaten (A) Röstgrad, (B) Fettsäure nach der Extraktion der lipophilen Kaffeemetaboliten und (C) Feuchtegehalt der untersuchten Kaffeeproben in Box-Whisker-Plots. AF: Afrika; AS: Asien; NA: Nordamerika; SA: Südamerika; AT: Äthiopien; BR: Brasilien; K: Kolumbien; B: Biologischer Anbau; Ikawa: Röstung mittel Heißluft

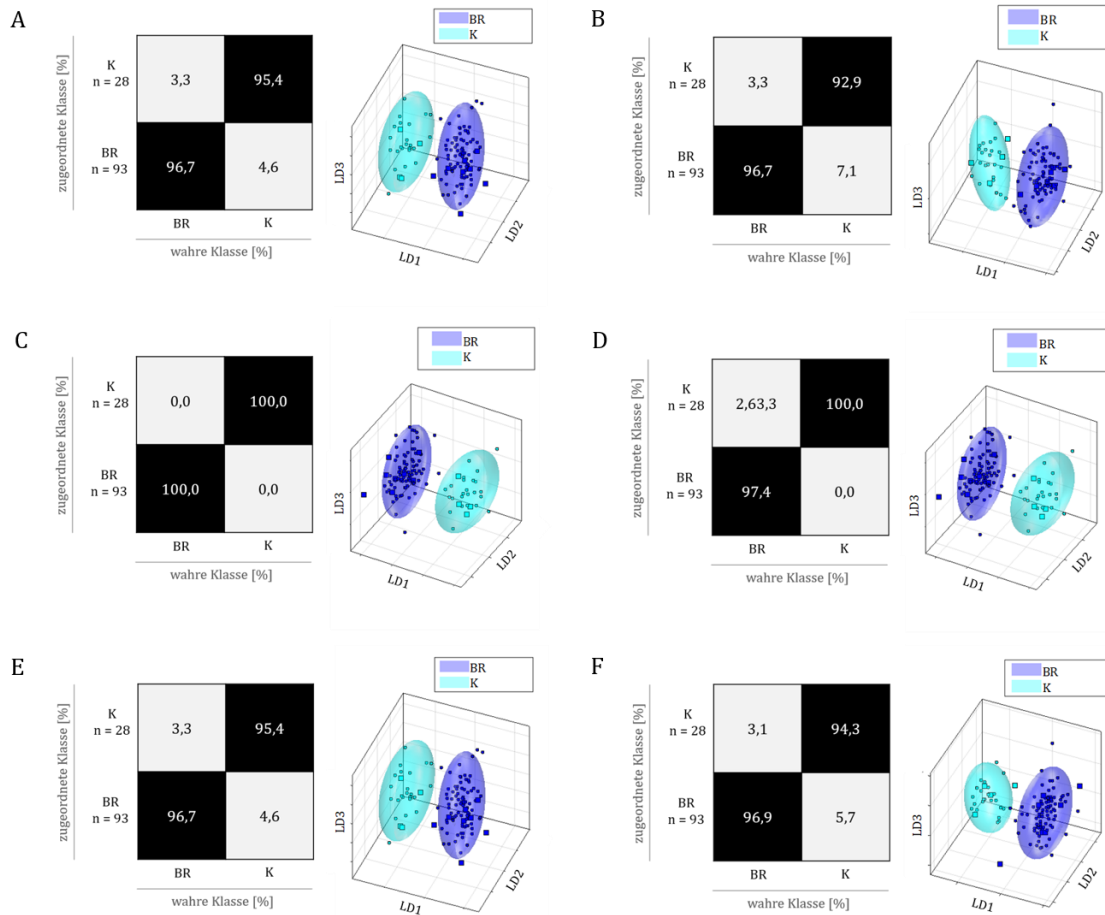


Abbildung 9-25: Ergebnisse der MCKV der PCA-LDA zur Vorhersage des geographischen Ursprungs von Kaffee bezüglich der Länder Brasilien und Kolumbien aus den ^1H -NMR-Spektren der lipophilen und hydrophilen Kaffeemetaboliten. Jeweils rechts in der Abbildung (A)-(E) ist der LDA-Raum eines Kreuzvalidierungsschrittes dargestellt. Das 95%ige Prognoseellipsoid des Trainingsdatensatzes ist als Wolke dargestellt, die Testdatensatz-Proben werden durch Rechtecke symbolisiert. Links in der Abbildung (A)-(E) ist die Konfusionsmatrix der MCKV dargestellt. Die x-Achse und die y-Achse zeigen die zugeordnete Klasse bzw. die wahre Klasse. In der Konfusionsmatrix ist die Wahrscheinlichkeit des Vorhersageergebnisses in Prozent angegeben. (A) Modell der lipophilen Kaffeemetaboliten mit der Datenvorbehandlung wie in Kapitel 8.4.1 beschrieben. (B) Modell der lipophilen Kaffeemetaboliten mit der Datenvorbehandlung wie in Kapitel 8.4.1 beschrieben mit zusätzlicher Normalisierung auf einen einheitlichen Feuchtegehalt von 3%. (C) Modell der hydrophilen Kaffeemetaboliten mit der Datenvorbehandlung wie in Kapitel 8.4.1 beschrieben. (D) Modell der hydrophilen Kaffeemetaboliten mit der Datenvorbehandlung wie in Kapitel 8.4.1 beschrieben mit zusätzlicher Normalisierung auf einen einheitlichen Feuchtegehalt von 3%. (E) Modell der lipophilen Kaffeemetaboliten mit der Datenvorbehandlung wie in Kapitel 8.4.1 beschrieben. (F) Modell der lipophilen Kaffeemetaboliten mit der Datenvorbehandlung wie in Kapitel 8.4.1 beschrieben mit zusätzlicher Normalisierung auf eine einheitliche Fettauswaage von 20 mg

Tabelle 9-6: Ergebnisse der externen Validierung des Klassifizierungsmodells Afrika-Nordamerika basierend auf der Analyse der lipophilen und hydrophilen Metaboliten. Der *conf*-Wert dient als Beurteilungskriterium für die richtige Zuordnung. Bei einem *conf*-Wert von 0 liegt die Probe im Klassenmittelpunkt, bei einem *conf*-Wert <1 innerhalb des 95%igen Prognoseellipsoids der Klasse. AF: Afrika; NA: Nordamerika

Testprobe	Modell basierend auf der Analyse der lipophilen Metaboliten		Modell basierend auf der Analyse der hydrophilen Metaboliten	
	<i>conf</i> (AF)	<i>conf</i> (NA)	<i>conf</i> (AF)	<i>conf</i> (NA)
19BE007	1,85	0,08	2,35	0,07
19BE032	1,87	0,07	2,71	0,32
19BE064	1,78	0,15	2,48	0,09
19BE098	2,04	0,16	2,32	0,10
19BE115	2,27	0,43	2,66	0,27
19BE134	1,86	0,09	2,46	0,09
19BE137	2,01	0,17	1,76	0,67
19BE150	1,87	0,06	2,23	0,18
19HA005	2,38	0,56	3,15	0,78
20HA004	2,46	0,65	2,40	0,05
20HA010	2,19	0,33	2,39	0,06
21HA0001	2,85	1,10	2,93	0,56
AF102	0,86	1,20	0,47	2,02
AF107	0,57	1,54	0,89	1,58
AF10	0,24	1,96	0,43	2,05
AF155	0,55	1,56	0,96	1,51
AF15f	0,15	2,36	0,10	2,41
AF17	0,34	2,58	0,64	3,17
AF208	0,39	1,75	0,52	1,97
AF22a	0,44	1,69	1,84	0,59
AF29	0,08	2,15	0,05	2,49
AF31a	0,43	1,69	0,13	2,37
AF33a	0,14	2,30	0,13	2,64
AF350a	0,45	1,68	1,00	1,47
AF3	1,77	0,21	1,89	0,53
AF400a	0,23	1,93	0,32	2,84
AF41	0,19	1,99	0,21	2,29
AF48	0,16	2,09	0,22	2,28
AF53	0,23	1,93	0,16	2,68
AF901	0,36	1,78	0,72	1,76
N101	1,24	0,76	1,69	0,74
N105	1,68	0,25	2,08	0,34
N109	1,19	0,81	1,66	0,77
N115	1,60	0,34	2,14	0,28
N118	1,91	0,09	2,73	0,34
N15	1,25	0,75	2,05	0,37
N162	1,92	0,11	2,44	0,07
N168	1,96	0,09	2,16	0,25
N174	2,13	0,28	2,58	0,19
N178	1,86	0,08	2,06	0,36
N181	1,39	0,59	1,98	0,44
N259	1,85	0,16	2,78	0,40
N270	1,84	0,07	2,60	0,21
N273	1,38	0,60	1,72	0,71
N280	1,78	0,17	2,46	0,07
N296	1,43	0,54	1,83	0,60
N303b	2,32	0,49	2,61	0,22
N350	1,58	0,37	2,06	0,36
N450	1,16	0,85	0,66	1,81
N51	2,07	0,21	2,09	0,33
N5	0,08	2,14	0,35	2,14
N8	1,78	0,14	1,93	0,49
N900	2,03	0,16	1,96	0,46

Tabelle 9-7: Ergebnisse der externen Validierung des Klassifizierungsmodells Brasilien-Afrika basierend auf der Analyse der lipophilen und hydrophilen Metaboliten. Der *conf*-Wert dient als Beurteilungskriterium für die richtige Zuordnung. Bei einem *conf*-Wert von 0 liegt die Probe im Klassenmittelpunkt, bei einem *conf*-Wert <1 innerhalb des 95%igen Prognoseellipsoids der Klasse. BR: Brasilien; AF: Afrika

Testprobe	Modell basierend auf der Analyse der lipophilen Metaboliten		Modell basierend auf der Analyse der hydrophilen Metaboliten	
	<i>conf</i> (BR)	<i>conf</i> (AF)	<i>conf</i> (BR)	<i>conf</i> (AF)
20DU002	0,64	2,71	0,39	2,76
20DU016	0,10	2,13	0,15	2,41
20DU025	0,09	2,14	0,41	2,77
20DU030	0,15	2,31	0,09	2,53
20DU037	0,86	2,88	0,15	2,60
20DU045	0,89	1,48	0,80	1,98
20DU050	0,15	2,32	0,11	2,55
20DU054	0,70	2,76	0,61	2,90
20DU058	0,96	1,43	0,74	2,02
20DU059	0,29	1,97	0,51	2,84
20DU062	0,07	2,24	0,54	2,14
20DU068	0,74	1,61	0,68	2,05
20DU071	0,18	2,32	0,40	2,77
AF102	2,20	0,44	3,67	0,11
AF107	1,93	0,66	3,04	0,50
AF10	1,48	1,01	2,29	1,00
AF155	2,06	0,56	2,78	0,67
AF15f	3,00	0,20	4,18	0,25
AF17	3,41	0,52	4,88	0,71
AF208	1,94	0,65	2,97	0,55
AF22a	2,71	0,12	2,49	0,86
AF29	3,41	0,53	3,70	0,08
AF31a	2,49	0,21	3,55	0,17
AF33a	2,48	0,22	4,07	0,19
AF350a	2,33	0,34	2,95	0,56
AF3	1,43	1,05	1,82	1,31
AF400a	2,36	0,31	4,18	0,25
AF41	2,32	0,34	3,43	0,25
AF48	2,47	0,23	3,29	0,33
AF53	2,21	0,43	3,75	0,06
AF901	1,84	0,73	2,97	0,54
S11	0,15	2,29	0,08	2,49
S16	0,48	1,81	1,32	1,63
S20	0,33	2,43	0,43	2,22
S23	0,08	2,20	0,47	2,19
S27	0,28	1,98	0,09	2,50
S31	0,37	1,91	0,52	2,16
S36	0,19	2,33	0,19	2,39
S42	0,62	1,71	0,41	2,24
S45	0,09	2,25	0,67	2,06
S49	0,13	2,17	0,93	3,12

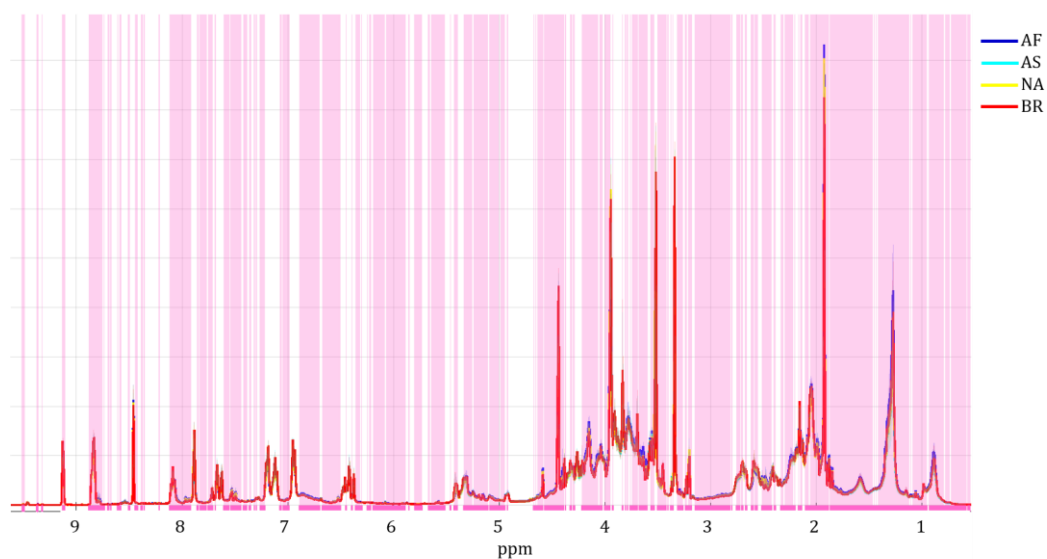


Abbildung 9-26: Univariate Datenanalyse zur Untersuchung spektraler Unterschiede zwischen den hydrophilen Kaffeemetaboliten der Klassen Afrika, Asien, Nordamerika und Brasilien mittels Kruskal-Wallis Test. Pink hinterlegt sind die Buckets mit einem kleineren p-Wert als das Signifikanzniveau. $\alpha = 0,05$, Anzahl Buckets: 983, nach Šidák korrigiertes Signifikanzniveau $\alpha_{\text{korrt}} = 5,22 \cdot 10^{-5}$. AF: Afrika; AS: Asien; NA: Nordamerika; BR: Brasilien

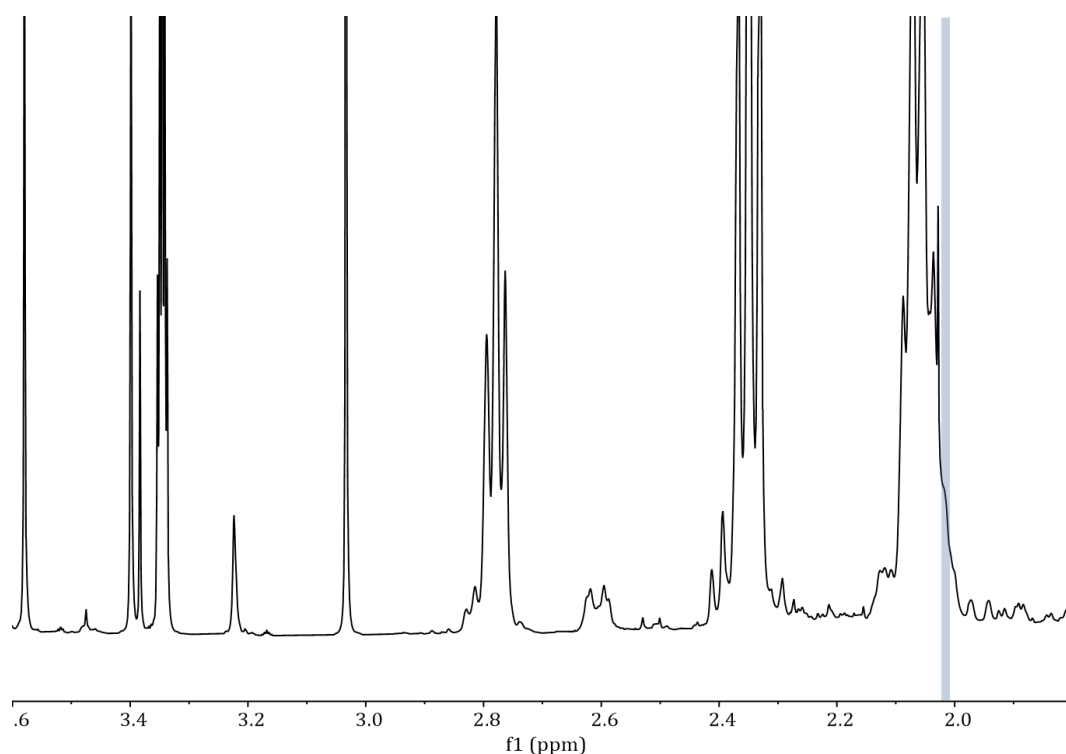


Abbildung 9-27: Ausschnitt des 400 MHz ^1H -NMR-Spektrums im Bereich $\delta_{\text{H}} = 1,80 - 3,60$ ppm der lipophilen Metaboliten in der Probe AF15 gemessen in $\text{CDCl}_3:\text{MeOH-d}_4$ 1:1 (v/v). Blau hinterlegt sind die spektralen Bereiche der Variablen $\delta_{\text{H}} = 2,02$ ppm und 3,47 ppm, welche die Abgrenzung der Klasse Brasilien im Klassifizierungsmodell Brasilien-Nordamerika beeinflussten.

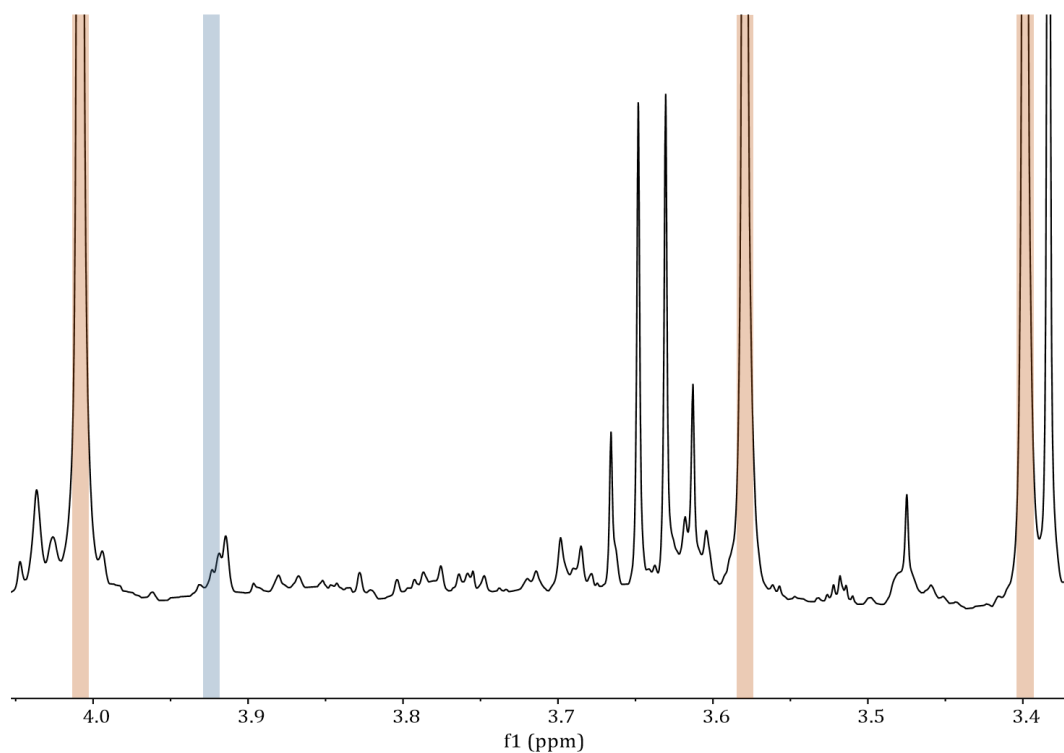


Abbildung 9-28: Ausschnitt des 400 MHz ^1H -NMR-Spektrums im Bereich $\delta_{\text{H}} = 1,80 - 3,60$ ppm der lipophilen Metaboliten in der Probe AF15 gemessen in $\text{CDCl}_3:\text{MeOH-d}_4$ 1:1 (v/v). Blau hinterlegt ist der spektrale Bereich der Variablen $\delta_{\text{H}} = 3,92$ ppm, welche die Abgrenzung der Klasse Brasilien im Klassifizierungsmodell Äthiopien-Nordamerika beeinflusste. Orange hinterlegt sind die spektralen Bereiche der Variablen $\delta_{\text{H}} = 3,40$; 3,58 und 4,01 ppm, welche die Trennung der Klasse Nordamerika im Klassifizierungsmodell Äthiopien-Nordamerika beeinflussten.

9.5 Datenfusionierung der lipophilen und hydrophilen ^1H -NMR-spektroskopischen Daten

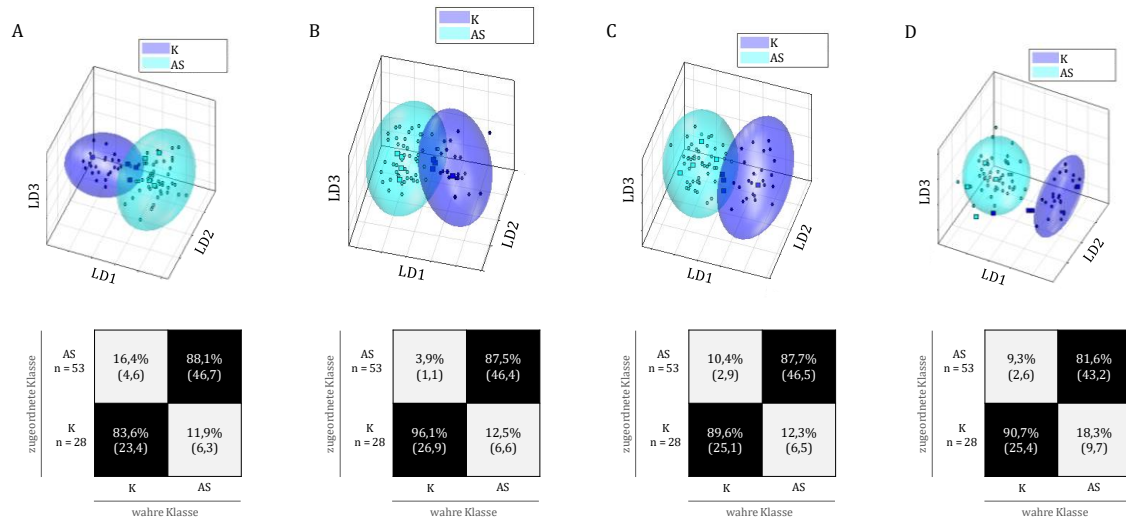


Abbildung 9-29: Ergebnisse der MCKV des Klassifizierungsmodells zur Vorhersage des geographischen Ursprungs von Kaffee bezüglich des Landes Kolumbien und den Kontinent Asien. (A) Klassifizierungsmodell durch die Analyse der ^1H -NMR-Spektren der lipophilen Metaboliten; (B) Klassifizierungsmodell durch die Analyse der ^1H -NMR-Spektren der hydrophilen Kaffeemetaboliten; (C) Klassifizierungsmodell durch die Low-Level Datenfusionierung der ^1H -NMR-Spektren der lipophilen und hydrophilen Metaboliten; (D) Klassifizierungsmodell durch die Mid-Level Datenfusionierung der ^1H -NMR-Spektren der lipophilen und hydrophilen Metaboliten. Oben in (A)-(D) ist der LDA-Raum eines Kreuzvalidierungsschrittes dargestellt. Die Wolken kennzeichnen das 95%ige Prognoseellipsoid des Trainingsdatensatzes. Unten in (A)-(D) ist die Konfusionsmatrix der MCKV dargestellt. Die x-Achse und die y-Achse zeigen die zugeordnete Klasse bzw. die wahre Klasse. In der Konfusionsmatrix ist die Wahrscheinlichkeit des Vorhersageergebnisses in Prozent angegeben. Anzahl der für die LDA verwendeten PC: (A)-(C) 19; PC: (D) 38. MCKV: Monte-Carlo Kreuzvalidierung; PCA-LDA: Hauptkomponentenanalyse mit linearer Diskriminanzanalyse; LD: Lineare Diskriminanzfunktion; K: Kolumbien; AS: Asien

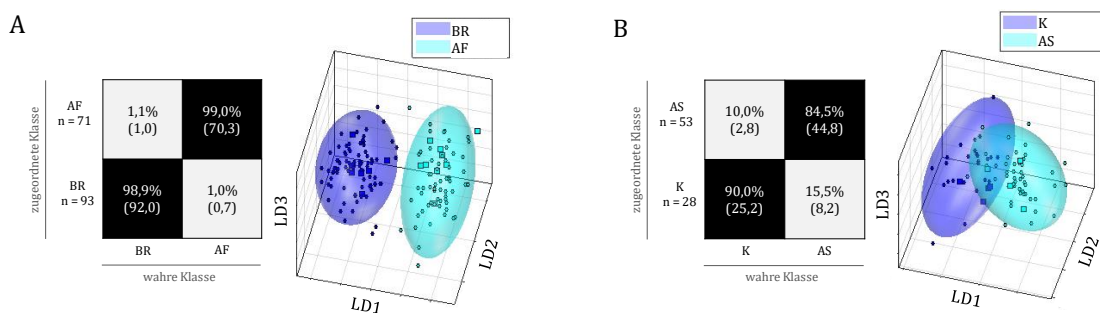


Abbildung 9-30: Ergebnisse der MCKV des Klassifizierungsmodells der Mid-Level Datenfusionierung der ^1H -NMR-Spektren der lipophilen und hydrophilen Metaboliten. (A) Klassifizierungsmodell Brasilien-Afrika. Für die Datenfusionierung verwendete PC: 29. (B) Klassifizierungsmodell Kolumbien-Asien. Für die Datenfusionierung verwendete PC: 19. Rechts in (A) und (B) ist der LDA-Raum eines Kreuzvalidierungsschrittes dargestellt. Die Wolken kennzeichnen das 95%ige Prognoseellipsoid des Trainingsdatensatzes. Links in (A) und (B) ist die Konfusionsmatrix der MCKV dargestellt. Die x-Achse und die y-Achse zeigen die zugeordnete Klasse bzw. die wahre Klasse. In der Konfusionsmatrix ist die Wahrscheinlichkeit des Vorhersageergebnisses in Prozent angegeben. MCKV: Monte-Carlo Kreuzvalidierung; PCA-LDA: Hauptkomponentenanalyse mit linearer Diskriminanzanalyse; LD: Lineare Diskriminanzfunktion; BR: Brasilien; AF: Afrika; K: Kolumbien; AS: Asien

9.6 Identifizierung lipophiler Kaffeemetaboliten mittels NMR-Spektroskopie

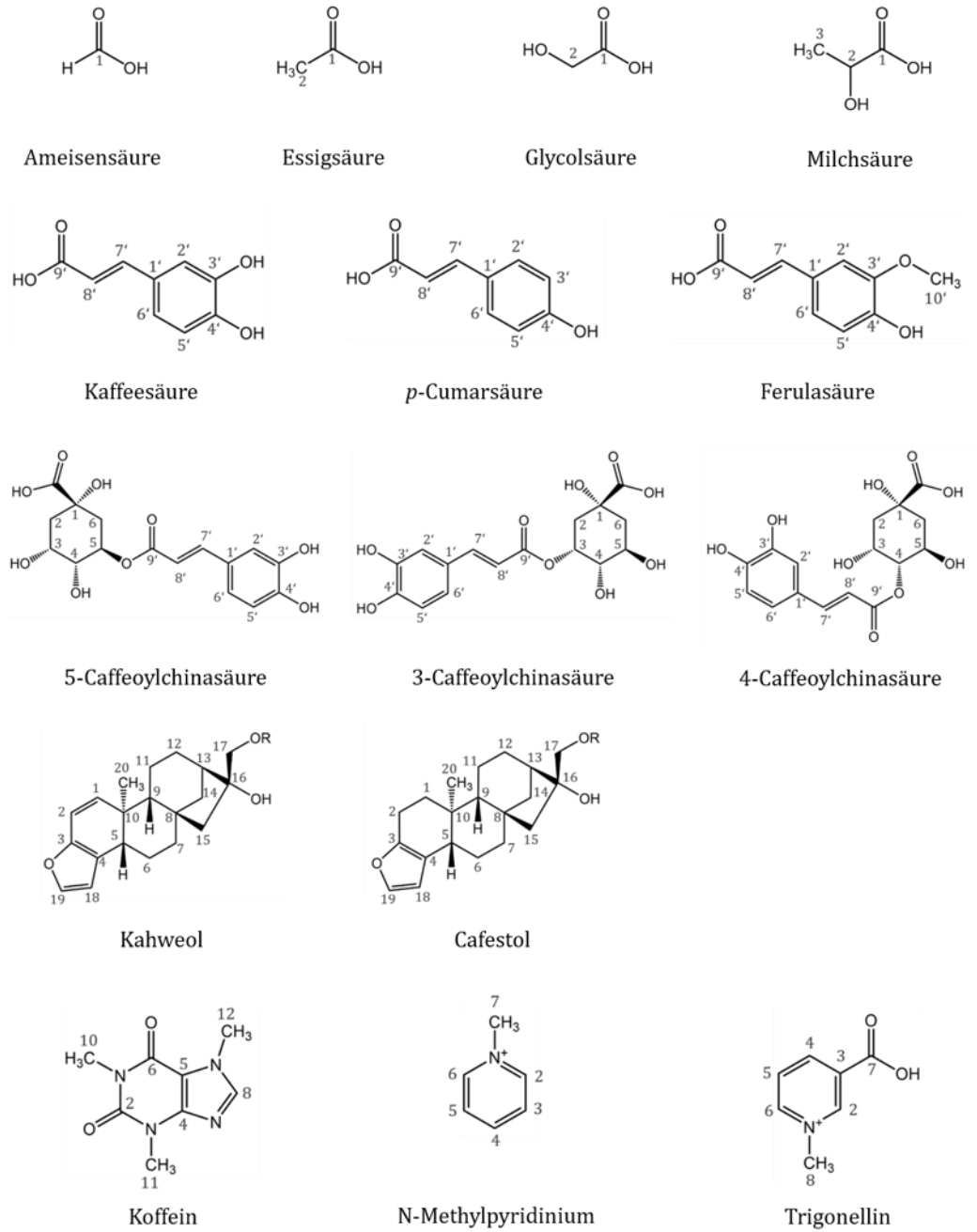


Abbildung 9-31: Strukturformeln ausgewählter in Kaffee vorkommenden Primär- und Sekundärmetabolite. In den Strukturen sind die ^1H -Kerne durchnummeriert.

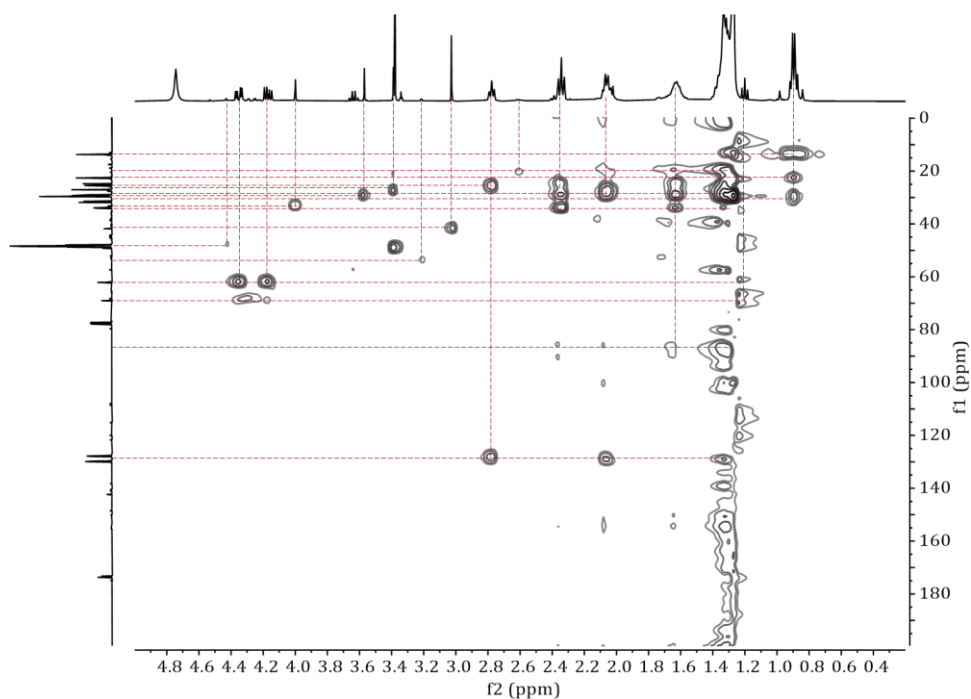


Abbildung 9-32: Ausschnitt eines 400 MHz ^1H - ^{13}C -korrelierten HSQC-TOCSY-NMR-Spektrums der lipophilen Kaffeemetaboliten im Bereich $\delta_{\text{H}} = 0,20 - 5,00$ ppm und $\delta_{\text{C}} = 0,00 - 190$ ppm, gemessen in $\text{CDCl}_3:\text{MeOH-d}_4$ 1:1 (v/v). Die Spinsysteme sind in rot gestrichelten Linien dargestellt.

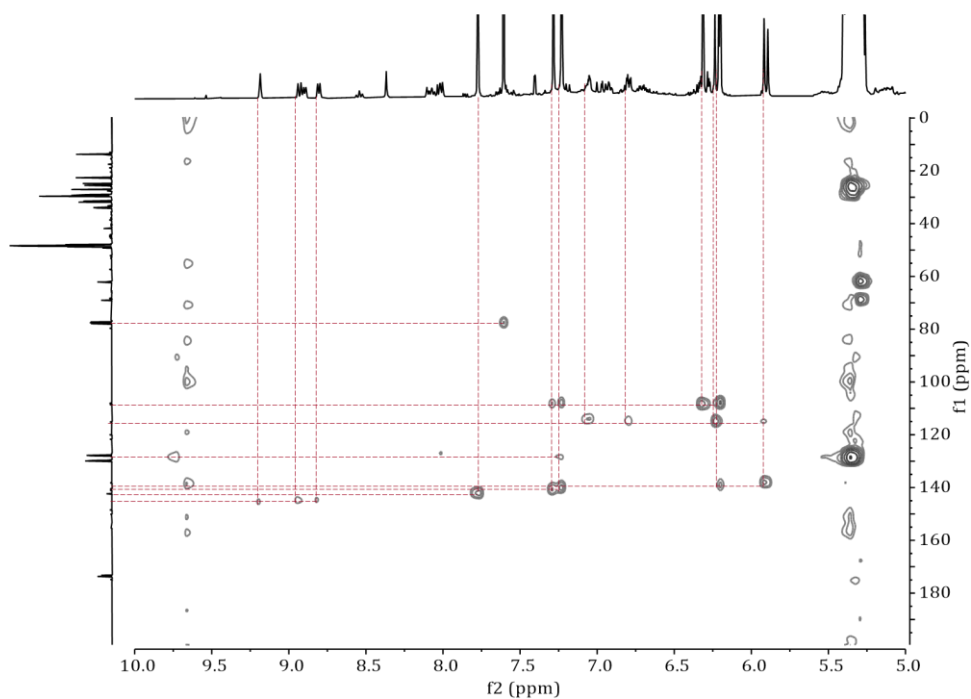


Abbildung 9-33: Ausschnitt eines 400 MHz ^1H - ^{13}C -korrelierten HSQC-TOCSY-NMR-Spektrums der lipophilen Kaffeemetaboliten im Bereich $\delta_{\text{H}} = 5,00 - 10,0$ ppm und $\delta_{\text{C}} = 0,00 - 190$ ppm, gemessen in $\text{CDCl}_3:\text{MeOH-d}_4$ 1:1 (v/v). Die Spinsysteme sind in rot gestrichelten Linien dargestellt.

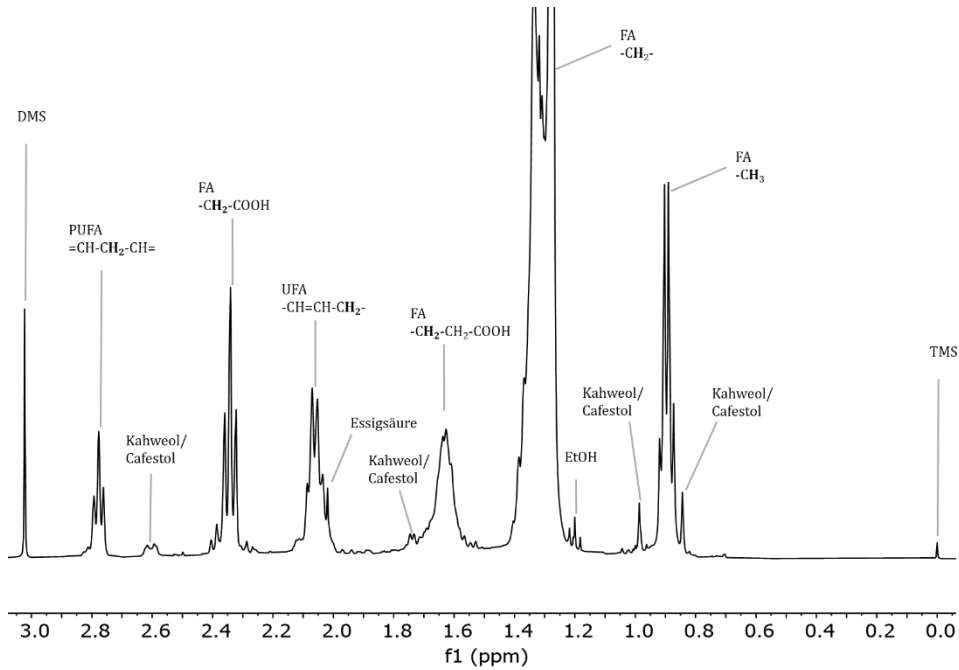


Abbildung 9-34: Ausschnitt eines 400 MHz ^1H -NMR-Spektrums im Bereich $\delta_{\text{H}} = -0,20 - 3,20$ ppm der lipophilen Kaffee-metaboliten, gemessen in $\text{CDCl}_3:\text{MeOH-d}_4$ 1:1 (v/v). Die chemische Verschiebung wurde auf $\delta_{\text{TMS}} = 0,00$ ppm referenziert. TMS: Tetramethylsilan; FA: Fettsäuren; EtOH: Ethanol; UFA: ungesättigte Fettsäuren; PUFA: mehrfach ungesättigte Fettsäuren; DMS: Dimethylsulfon

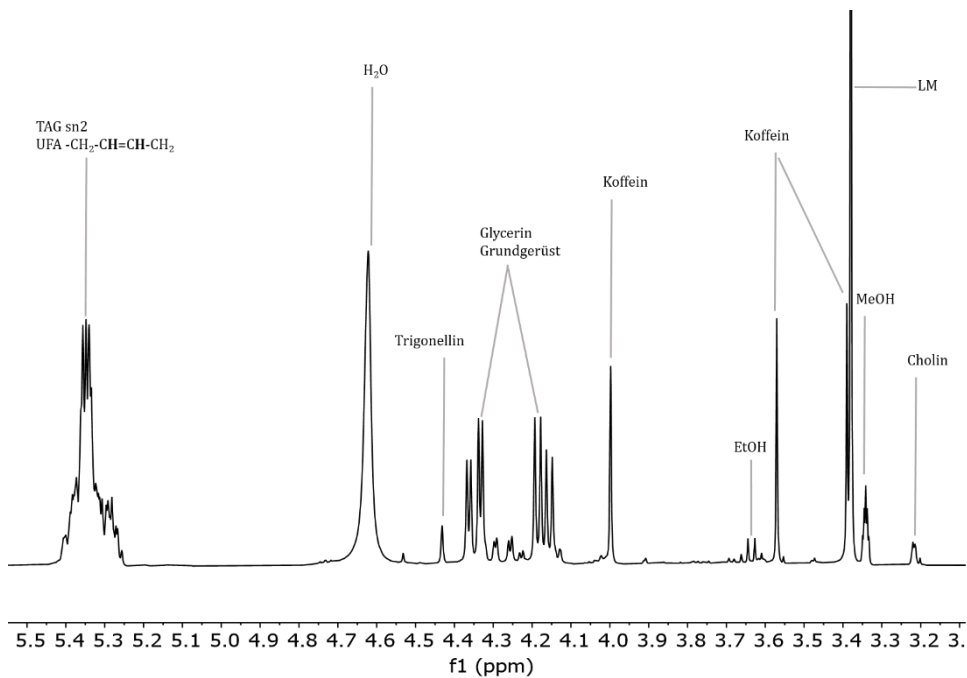


Abbildung 9-35: Ausschnitt eines 400 MHz ^1H -NMR-Spektrums im Bereich $\delta_{\text{H}} = 3,00 - 5,50$ ppm der lipophilen Kaffee-metaboliten, gemessen in $\text{CDCl}_3:\text{MeOH-d}_4$ 1:1 (v/v). Die chemische Verschiebung wurde auf $\delta_{\text{TMS}} = 0,00$ ppm referenziert. LM: Signal erzeugt durch Lösungsmittel; MeOH: Methanol; EtOH: Ethanol; TAG: Triacylglyceride; UFA: ungesättigte Fettsäuren

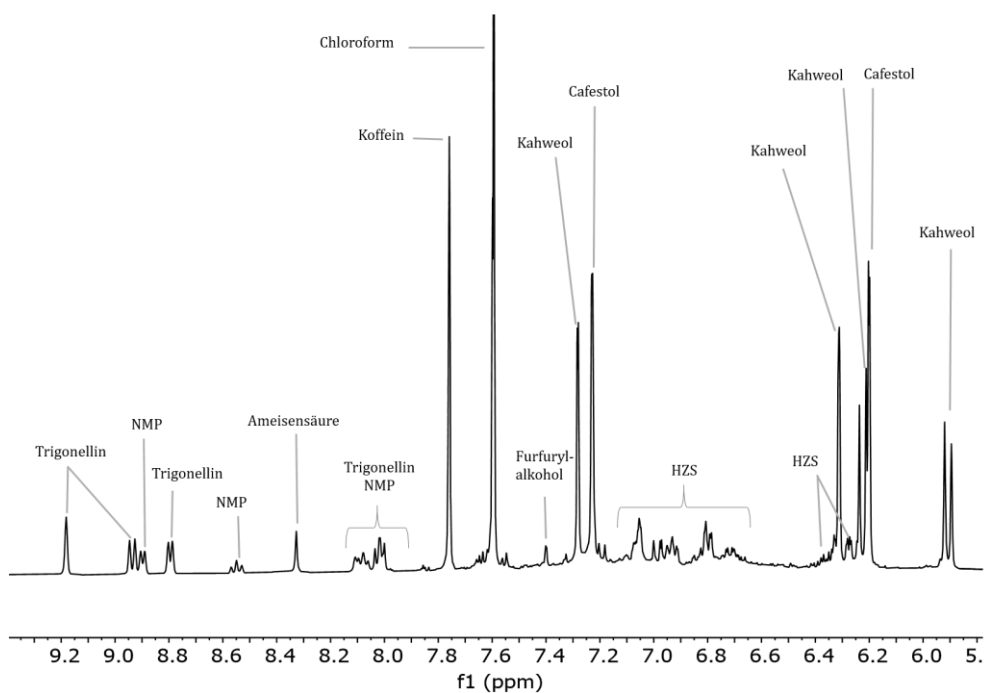


Abbildung 9-36: Ausschnitt eines 400 MHz ^1H -NMR-Spektrums im Bereich $\delta_{\text{H}} = 5,50 - 9,40$ ppm der lipophilen Kaffeemetaboliten, gemessen in $\text{CDCl}_3:\text{MeOH-d}_4$ 1:1 (v/v). Die chemische Verschiebung wurde auf $\delta_{\text{TMS}} = 0,00$ ppm referenziert HZS: Hydroxyzimtsäureisomere (Kaffeensäure, Ferulasäure, p-Cumarsäure); NMP: N-Methylpyridinium

9.7 Identifizierung lipophiler Kaffeemetaboliten mittels NMR-Spektroskopie

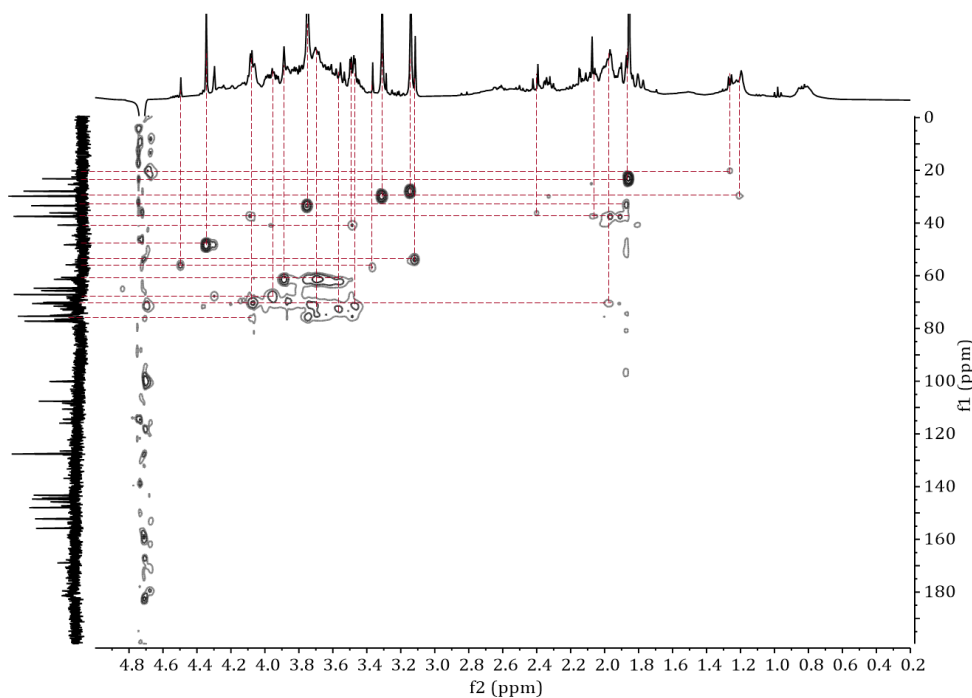


Abbildung 9-37: Ausschnitt eines 400 MHz ^1H - ^{13}C -korrelierten HSQC-TOCSY-NMR-Spektrums der hydrophilen Kaffeemetaboliten im Bereich $\delta_{\text{H}} = 0,20 - 5,0$ ppm und $\delta_{\text{C}} = 0,00 - 190$ ppm, gemessen in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die Spinsysteme sind in rot gestrichelten Linien dargestellt.

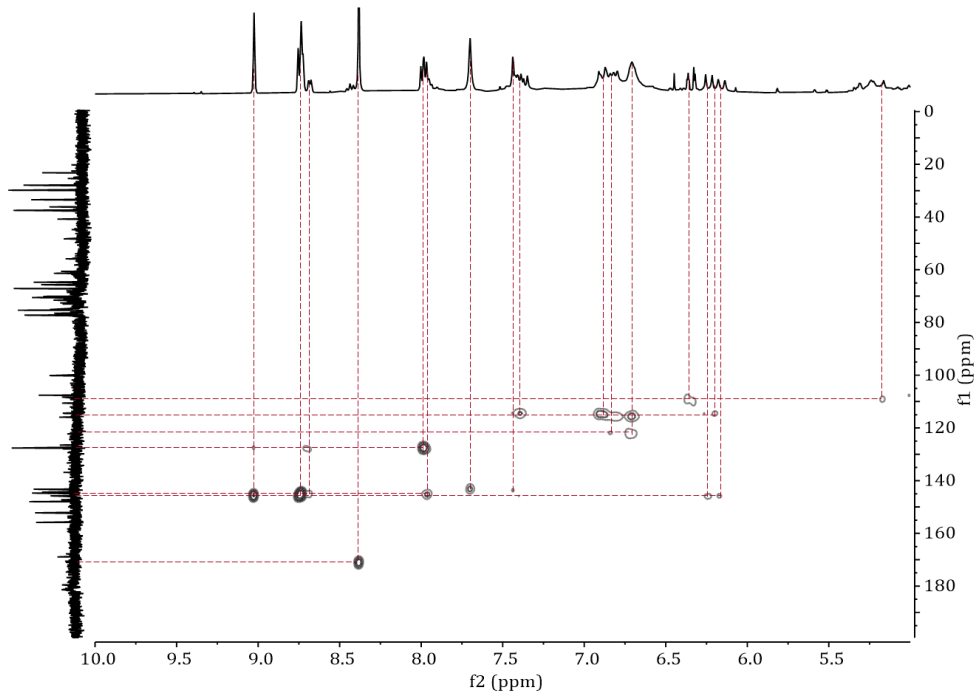


Abbildung 9-38: Ausschnitt eines 400 MHz ^1H - ^{13}C -korrelierten HSQC-NMR-Spektrums der hydrophilen Kaffeemetaboliten im Bereich $\delta_{\text{H}} = 5,00 - 10,0$ ppm und $\delta_{\text{C}} = 0,00 - 190$ ppm, gemessen in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die Spinsysteme sind in rot gestrichelten Linien dargestellt.

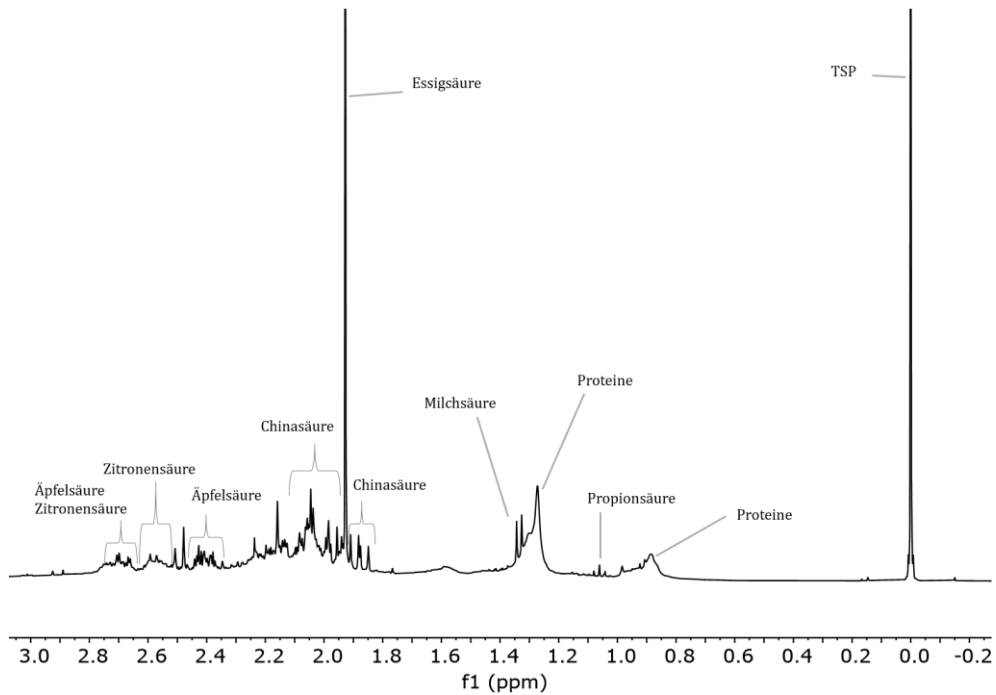


Abbildung 9-39: Ausschnitt eines 400 MHz ^1H -NMR-Spektrums im Bereich $\delta_{\text{H}} = -0,20 - 3,20$ ppm der hydrophilen Kaffeemetaboliten in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die Lösungsmittelunterdrückung erfolgte mit dem NOESY-Pulsprogramm. Die chemische Verschiebung wurde auf $\delta_{\text{TSP}} = 0,00$ ppm referenziert. TSP: Trimethylsilylpropionat

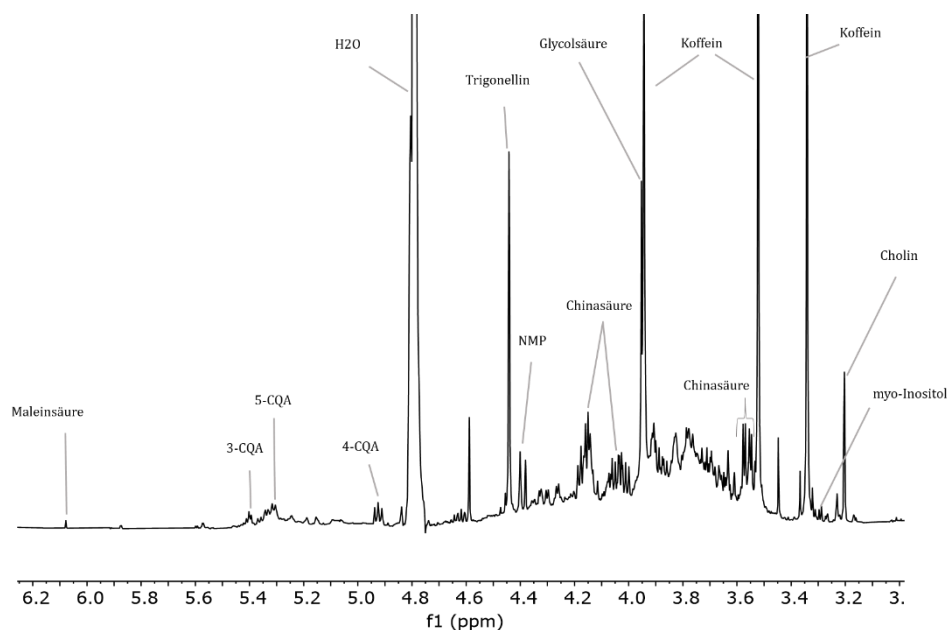


Abbildung 9-40: Ausschnitt eines 400 MHz ^1H -NMR-Spektrums im Bereich $\delta_{\text{H}} = -3,20 - 6,20$ ppm der hydrophilen Kaffeemetaboliten in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die Lösungsmittelunterdrückung erfolgte mit dem NOESY-Pulsprogramm. Die chemische Verschiebung wurde auf $\delta_{\text{TSP}} = 0,00$ ppm referenziert 4-CQA: 4-Caffeoylchinasäure; 5-CQA: 5-Caffeoylchinasäure; 3-CQA: 3-Caffeoylchinasäure

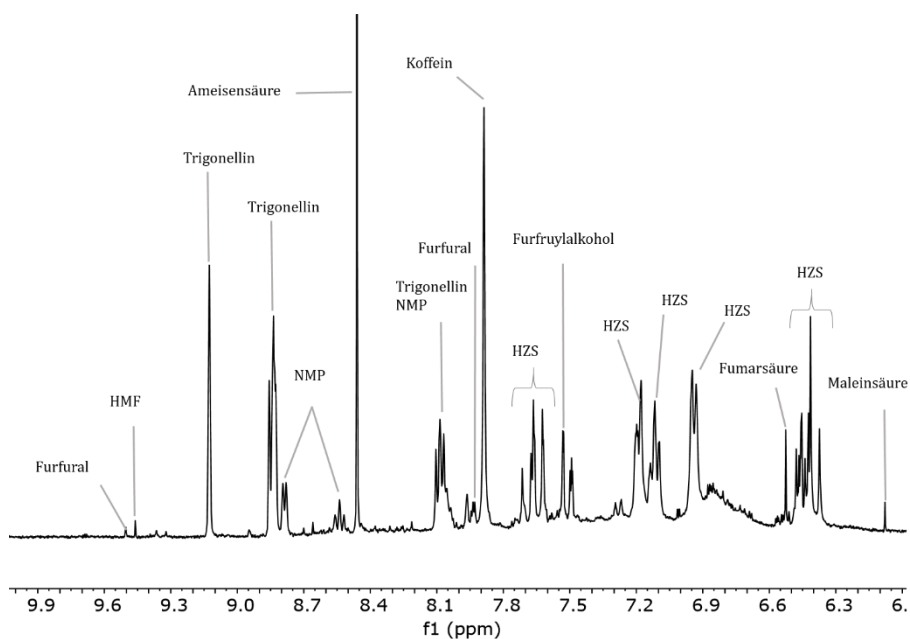


Abbildung 9-41: Ausschnitt eines 400 MHz ^1H -NMR-Spektrums im Bereich $\delta_{\text{H}} = 6,20 - 10,0$ ppm der hydrophilen Kaffeemetaboliten in $\text{H}_2\text{O}:\text{D}_2\text{O}$ 9:1 (v/v). Die Lösungsmittelunterdrückung erfolgte mit dem NOESY-Pulsprogramm. Die chemische Verschiebung wurde auf $\delta_{\text{TSP}} = 0,00$ ppm referenziert HZS: Hydroxyzimtsäureisomere (Kaffeensäure, Ferulasäure, p-Cumarsäure); NMP: N-Methylpyridinium; HMF: 5-Hydroxymethylfurfural

9.8 Zielgerichtete Analyse zur Quantifizierung ausgewählter Kaffeemetaboliten

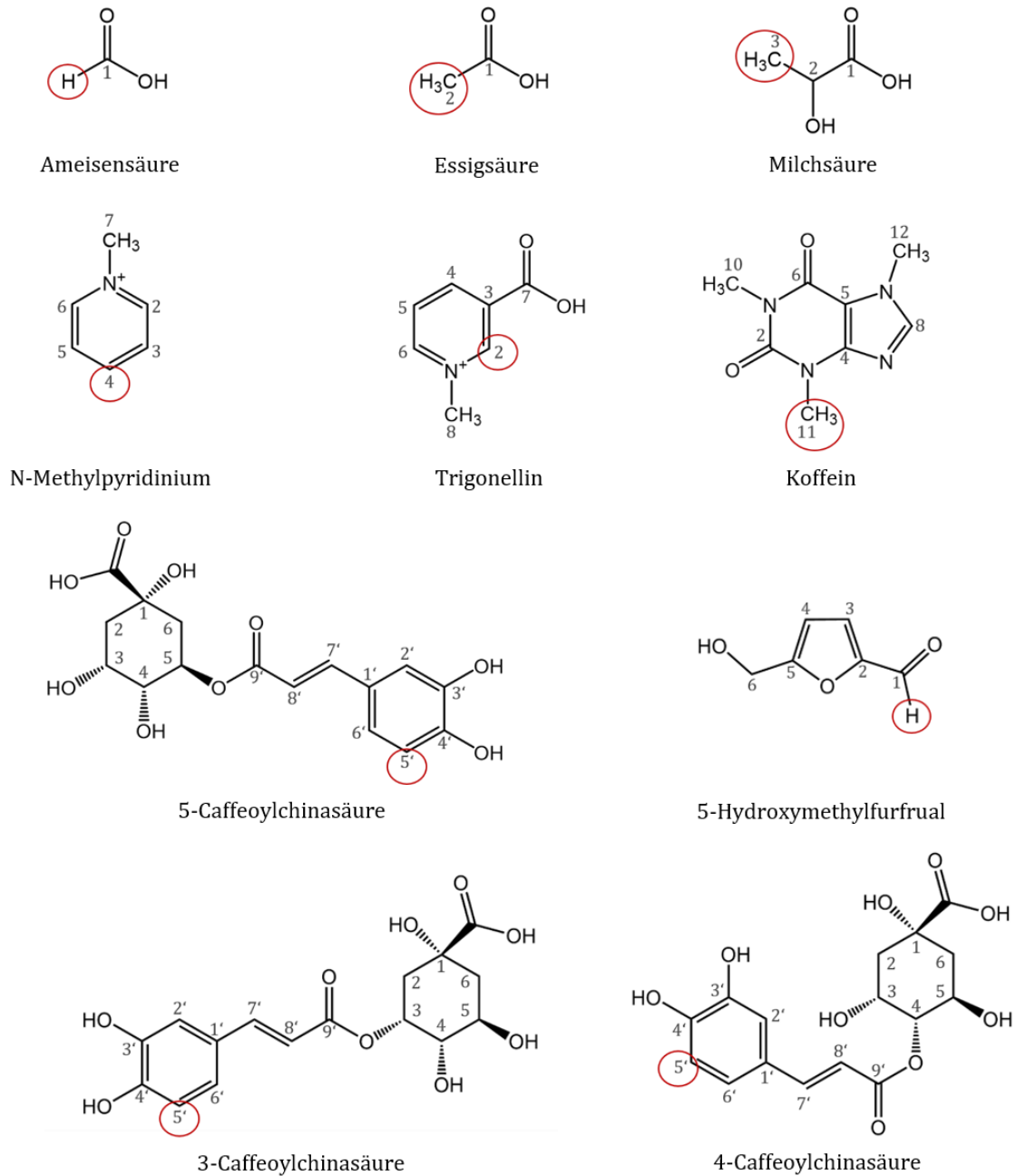


Abbildung 9-42: Strukturen der in den Kaffeeproben quantifizierten Metaboliten. Rot eingekreist sind die ¹H-Kerne, welche die Signale erzeugen, mit denen die Verbindungen quantifiziert wurden.

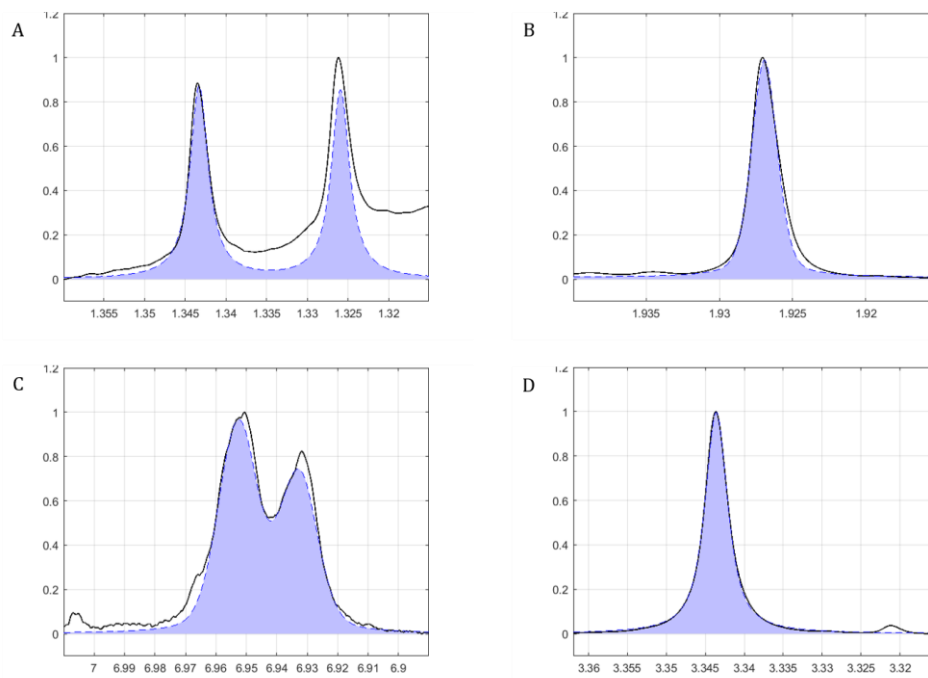


Abbildung 9-43: Kurvenanpassungsmethoden durch die Voigt-Funktion an die Signale von (A) Milchsäure, (B) Essigsäure, (C) Caffeoylchinasäure-Isomere (D) Koffein, über welche die Quantifizierung erfolgte. Die schwarze Linie stellt das Signal dar und der blau hinterlegte Bereich kennzeichnet das Integral, welches nach der Anpassung der Voigt-Funktion an das entsprechende Signal berechnet wurde.

Tabelle 9-8: Wiederfindungen bestimmt als Quotient des bestimmten durch den tatsächlich vorhandenen Gehalt und angegeben in Prozent der Analyten Milchsäure, Essigsäure, Koffein, Caffeoylchinasäure-Isomere (CQA), Ameisensäure, N-Methylpyridinium (NMP), Trigonellin und 5-Hydroxymethylfurfural (HMF) für die ^1H -qNMR-Methode. Die Reinsubstanzen der Analyten wurden in acht unterschiedlichen Konzentrationen auf eine Kaffeematrix aufgestockt und sind als Aufstockung1-Aufstockung8 (A 1 – A 8) gekennzeichnet.

	Milch- säure [%]	Essig- säure [%]	Koffein [%]	CQA [%]	Ameisen- säure [%]	NMP [%]	Tri- gonellin [%]	HMF [%]
A 1	100	105	86,1	80,6	104	110	105	116
A 2	97,5	96,2	85,7	78,4	101	110	103	106
A 3	103	109	84,0	76,2	97,4	107	102	108
A 4	96,7	104	84,5	77,8	97,9	106	103	112
A 5	96,3	104	85,6	78,7	101	105	104	
A 6	99,8	97,2	84,6	78,7	101	103	102	109
A 7	98,7	104	85,0	77,8	100	105	103	108
A 8	98,7	104	85,4	78,6	101	108	104	111

Tabelle 9-9: Variationskoeffizienten (VKs) zur Überprüfung der Messpräzision (MP), Vergleichspräzision (VP) und Robustheit ® für die Analyten Milchsäure, Essigsäure, Koffein, Caffeoylchinasäure-Isomere (CQA), Ameisensäure, N-Methylpyridinium (NMP), Trigonellin und 5-Hydroxymethylfurfural (HMF) für die ^1H -qNMR-Methode.

	Milch- säure	Essig- säure	Koffein	CQA	Ameisen- säure	NMP	Tri- gonellin	HMF
VK(MP)[%]	1,90	5,20	1,40	2,20	2,20	2,40	1,20	10,1
VK(VP)[%]	2,30	5,40	1,70	2,90	2,50	2,50	1,60	10,3
VK(R) [%]	2,90	5,10	1,30	1,90	1,00	4,00	0,40	7,50

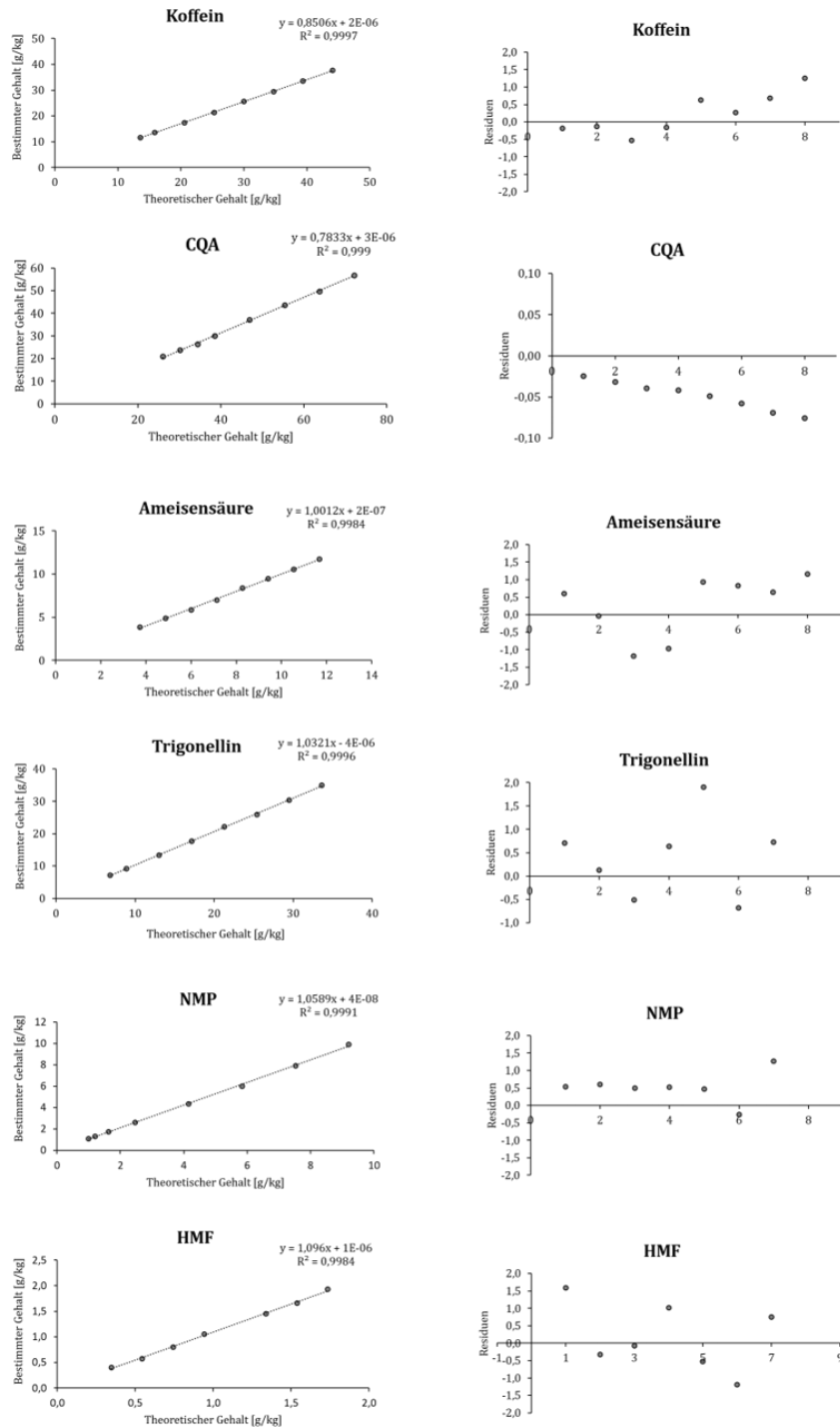


Abbildung 9-44: Ergebnisse der linearen Regression und Residuenanalyse zur Überprüfung der Linearität der ^1H -qNMR-Methode für die Analyten Koffein, Caffeoylchinasäure-Isomere (CQA), Ameisensäure, Trigonellin, N-Methylpyridinium (NMP) und 5-Hydroxymethylfurfural (HMF). Links dargestellt ist die Auftragung der bestimmten Konzentration gegenüber der tatsächlich vorhandenen Konzentration für die acht Aufstockungen der jeweiligen Reinsubstanz auf die Kaffeematrix unter Angabe der Geradengleichung und des Bestimmtheitsmaß R^2 . Rechts dargestellt ist die Auftragung der Residuen gegen die acht Aufstockungen

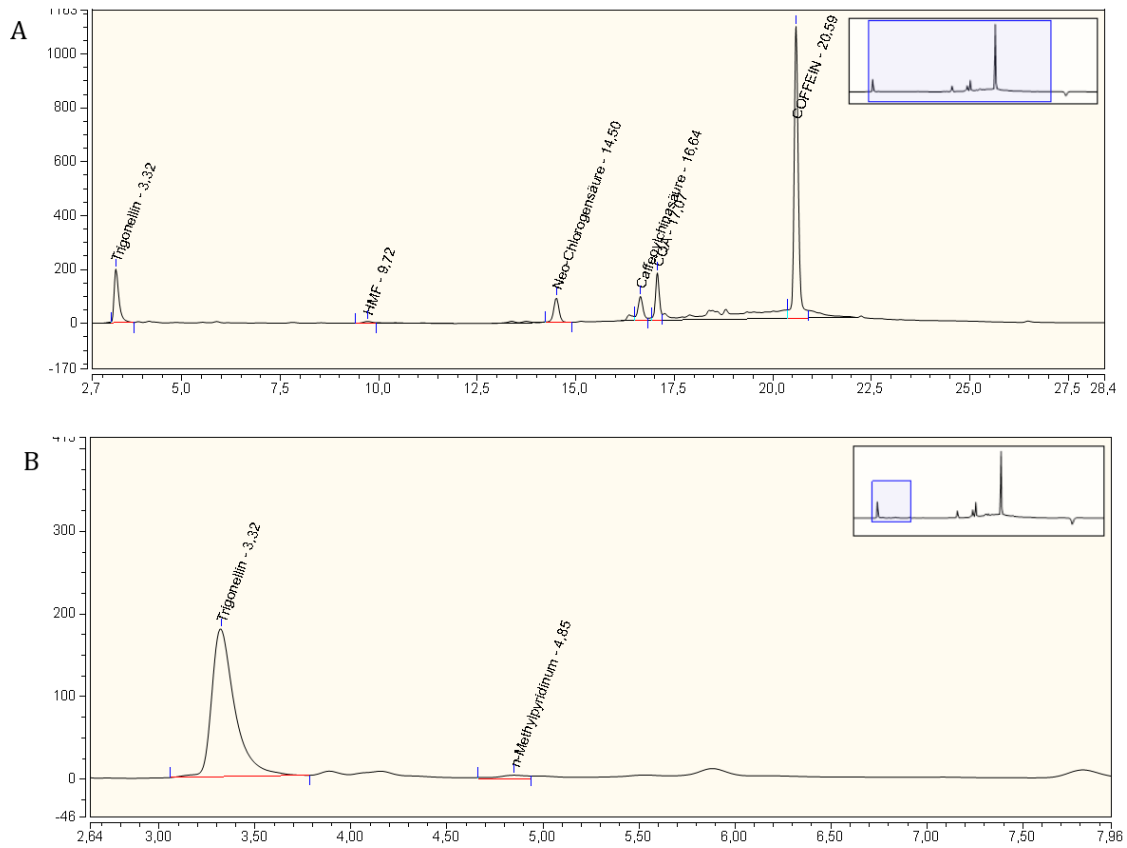


Abbildung 9-45: HPLC-DAD-Chromatogramme im Bereich (A) 2,70 – 29,4 min und (B) 2,64 – 7,96 min. Gekennzeichnet sind die Peaks der Analyten Trigonellin, N-Methylpyridinium, 5-Hydroxymethylfurfural (HMF), Neochlorogensäure (3-Caffeoylchinasäure), Caffeoylchinasäure (Caffeoylchinasäure), CQA (5-Caffeoylchinasäure) und Koffein.

Tabelle 9-10: Wiederfindungen bestimmt als Quotient des bestimmten durch den tatsächlich vorhandenen Gehalt und angegeben in Prozent der Analyten Trigonellin, N-Methylpyridinium (NMP), 3-Caffeoylchinasäure (3-CQA), 4-Caffeoylchinasäure (4-CQA), 5-Caffeoylchinasäure (5-CQA) und Koffein für die HPLC-DAD-Quantifizierungsmethode. Die Reinsubstanzen der Analyten wurden in acht unterschiedlichen Konzentrationen auf eine Kaffeematrix aufgestockt und sind als Aufstockung1-Aufstockung8 gekennzeichnet.

	Trigonellin [%]	NMP [%]	HMF [%]	3-CQA [%]	4-CQA [%]	5-CQA [%]	Koffein [%]
A 1	100	-0,90	82,5	99,2	90,5	98,4	90,8
A 2	100	26,2	91,5	97,1	94,3	101	93,8
A 3	101	67,4	95,5	96,0	97,0	97,0	93,5
A 4	100	83,0	97,0	95,7	107	98,1	93,3
A 5	101	98,1	98,0	97,9	104	95,8	93,7
A 6	101	104	97,9	96,7	112	94,6	93,0
A 7	100	120	97,2	96,1	105	94,0	92,8
A 8	101	170	97,0	95,2	109	90,3	92,5

Tabelle 9-11: Nachweisgrenzen (NG) und Bestimmungsgrenzen (BG) (angegeben in g/kg ermittelt nach DIN 32645) für die Analyten Trigonellin, 5-Hydroxymethylfurfural (HMF), 3-Caffeoylchinasäure (3-CQA), 4-Caffeoylchinasäure (4-CQA), 5-Caffeoylchinasäure (5-CQA) und Koffein für die HPLC-DAD-Quantifizierungsmethode

	Trigonellin	NMP	HMF	3-CQA	4-CQA	5-CQA	Koffein
NG [g/kg]	0,27	3,09	0,38	0,39	1,26	1,33	0,82
BG [g/kg]	1,01	8,44	1,41	1,42	3,86	4,47	2,89

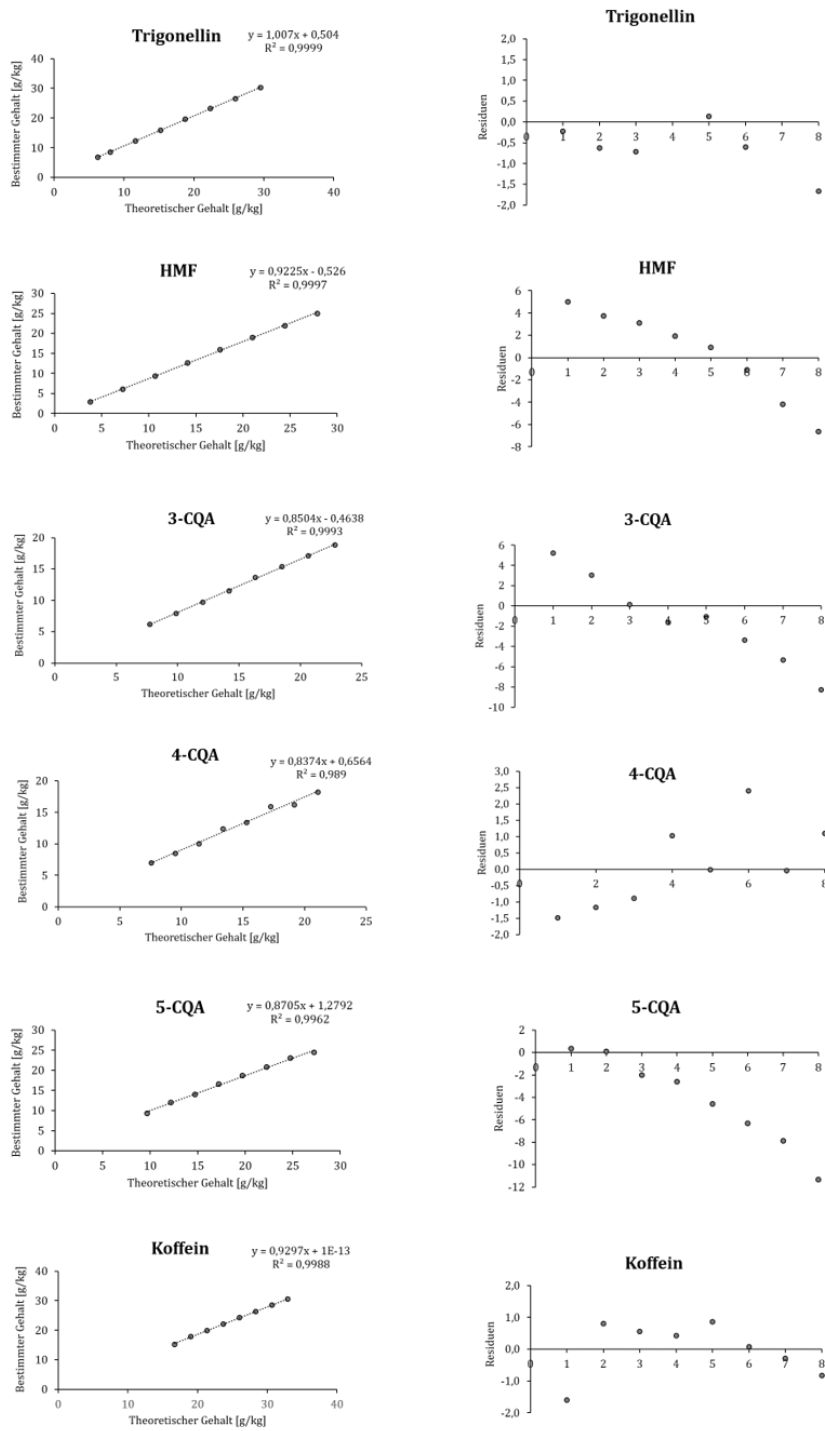


Abbildung 9-46: Ergebnisse der linearen Regression und Residuenanalyse zur Überprüfung der Linearität der HPLC-DAD-Quantifizierungsmethode für die Analyten Trigonellin, N-Methylpyridinium (NMP), 3-Caffeoylchinasäure (3-CQA), 4-Caffeoylchinasäure (4-CQA), 5-Caffeoylchinasäure (5-CQA) und Koffein. Links dargestellt ist die Auftragung der bestimmten Konzentration gegenüber der tatsächlich vorhandenen Konzentration für die acht Aufstockungen der jeweiligen Reinsubstanz auf die Kaffeematrix unter Angabe der Geradengleichung und des Bestimmtheitsmaß R^2 . Rechts dargestellt ist die Auftragung der Residuen gegen die acht Aufstockungen.

Tabelle 9-12: Bezeichnung der Kaffeeproben, welche neben der ^1H -qNMR-Methode zusätzlich mittels HPLC-DAD-Referenzmethode zur Quantifizierung von Trigonellin, HMF, 3-CQA, 4-CQA, 5-CQA und Koffein und mittels amtlicher Untersuchungsmethode L 46.00-3 zur Bestimmung von Koffein vermessen wurden.

Proben mittels HPLC-DAD-Referenzverfahren vermessen		Proben mittels amtlicher Untersuchungsmethode L 46.00-3 vermessen	
AF5	N169	AF15b	
AF15b	N170	AF20	
AF17	N171	AF52	
AF20	N172	AF204	
AF22	N204	AS1	
AF150	N258	AS7	
AF204	N260	AS8	
AS15	N270	AS16	
AS16	N274	AS56	
AS155	N275	AS155	
AS156	N277	AS156	
AS164	N281	AS162	
B1	N304	B30a	
B9	N305	B40	
B10	N400	B45	
B30a	S5	B53	
B36	S7	B59	
B38	S9	K	
B40	S11	N1a	
B42	S13	N8	
B45	S14	N112	
B48	S23	N150	
B52	S23a	N151	
B53	S27	N158	
B55	S29	N172	
B59	S30	N252	
B60	S31	N253	
N1a	S35	N256	
N3	S50	N270	
N6	S62	N275	
N8	S68	S1	
N101	S69	S23a	
N105	S70	S31	
N107	S71	S32	
N108	S101	S50	
N110	S107	S62	
N112	S109	S63	
N164	S110	S69	
N167	K	S111	

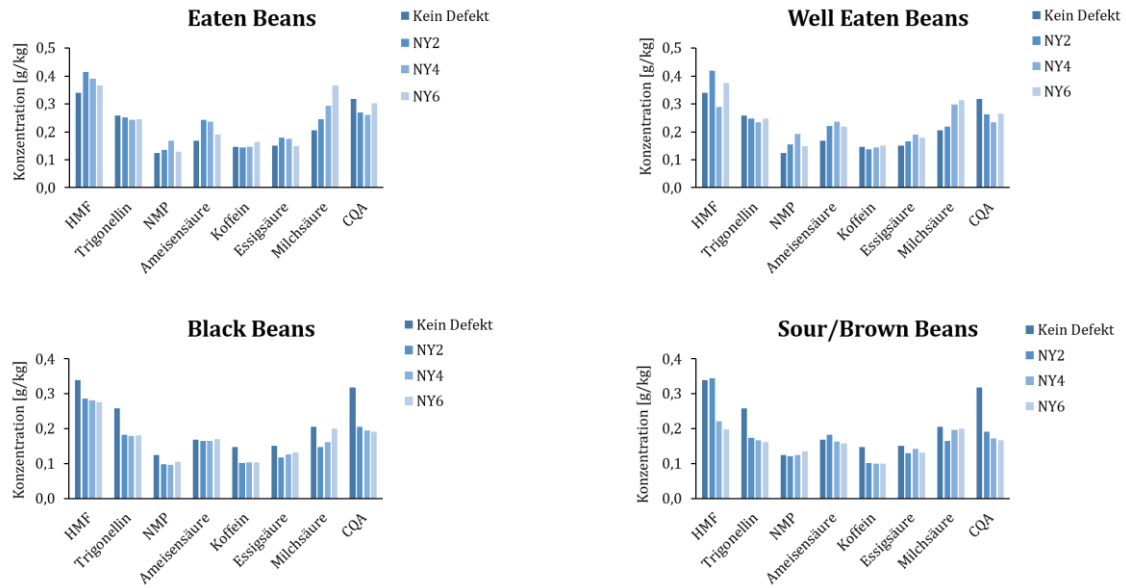


Abbildung 9-47: Die mittels ^1H -qNMR-Methode bestimmten Gehalte von 5-Hydroxymethylfurfural (HMF), Trigonellin, N-Methylpyridinium (NMP), Ameisensäure, Koffein, Essigsäure, Milchsäure und Caffeoylchinasäure-Isomere (CQA) der Kaffeedefekte „Eaten Beans“, „Well Eaten Beans“, „Black Beans“, „Sour/Brown Beans“. Die Defekte sind nach New Yorker Methode klassifiziert. Es sind die Klassifizierungsstufen NY2, NY4, NY6 und die Kaffeeprob ohne Defekte aufgetragen. Zur Vergleichbarkeit wurden die ermittelten Konzentrationen der einzelnen Analyten durch unterschiedliche Faktoren dividiert: HMF (1), Trigonellin (30), NMP (3), Ameisensäure (10), Koffein (100), Essigsäure (20), Milchsäure (3), CQA (100)

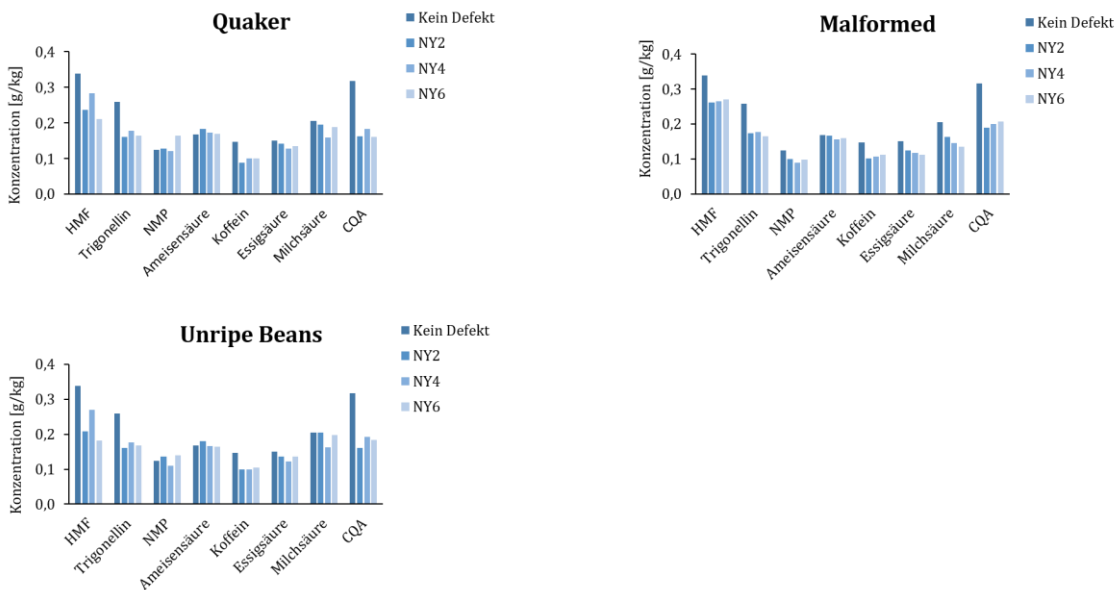


Abbildung 9-48: Die mittels ^1H -qNMR-Methode bestimmten Gehalte von 5-Hydroxymethylfurfural (HMF), Trigonellin, N-Methylpyridinium (NMP), Ameisensäure, Koffein, Essigsäure, Milchsäure und Caffeoylchinasäure-Isomere (CQA) der Kaffeedefekte „Quaker“, „Malformed“, „Unripe Beans“. Die Defekte sind nach New Yorker Methode klassifiziert. Es sind die Klassifizierungsstufen NY2, NY4, NY6 und die Kaffeeprob ohne Defekte aufgetragen. Zur Vergleichbarkeit wurden die ermittelten Konzentrationen der einzelnen Analyten durch unterschiedliche Faktoren dividiert: HMF (1), Trigonellin (30), NMP (3), Ameisensäure (10), Koffein (100), Essigsäure (20), Milchsäure (3), CQA (100)

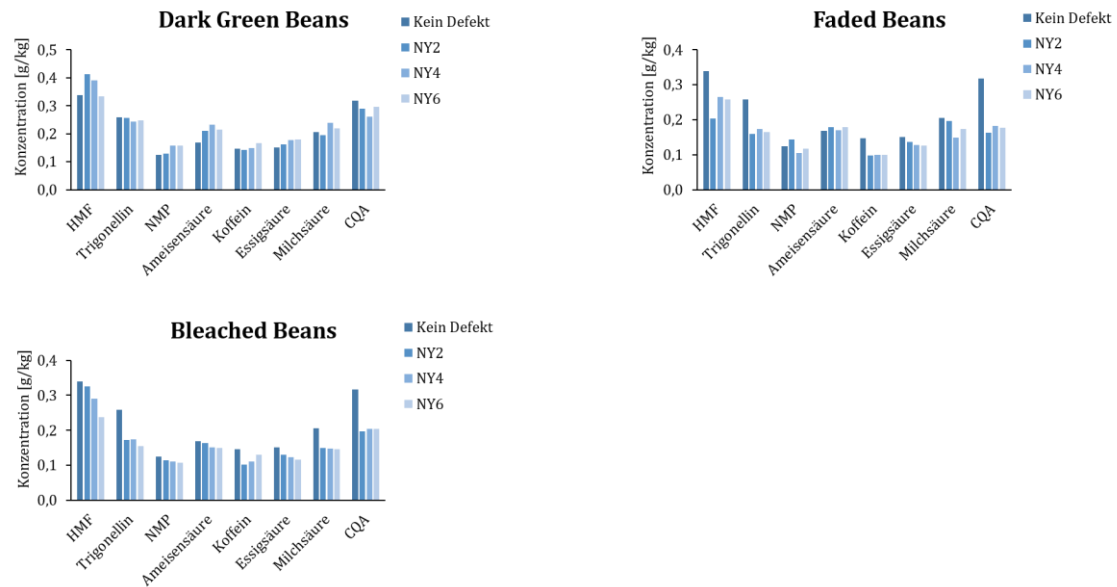


Abbildung 9-49: Die mittels ^1H -qNMR-Methode bestimmten Gehalte von 5-Hydroxymethylfurfural (HMF), Trigonellin, N-Methylpyridinium (NMP), Ameisensäure, Koffein, Essigsäure, Milchsäure und Caffeoylchinasäure-Isomere (CQA) der Kaffeedefekte „Dark Green Beans“, „Faded Beans“, „Bleached Beans“. Die Defekte sind nach New Yorker Methode klassifiziert. Es sind die Klassifizierungsstufen NY2, NY4, NY6 und die Kaffeeprob ohne Defekte aufgetragen. Zur Vergleichbarkeit wurden die ermittelten Konzentrationen der einzelnen Analyten durch unterschiedliche Faktoren dividiert: HMF (1), Trigonellin (30), NMP (3), Ameisensäure (10), Koffein (100), Essigsäure (20), Milchsäure (3), CQA (100)

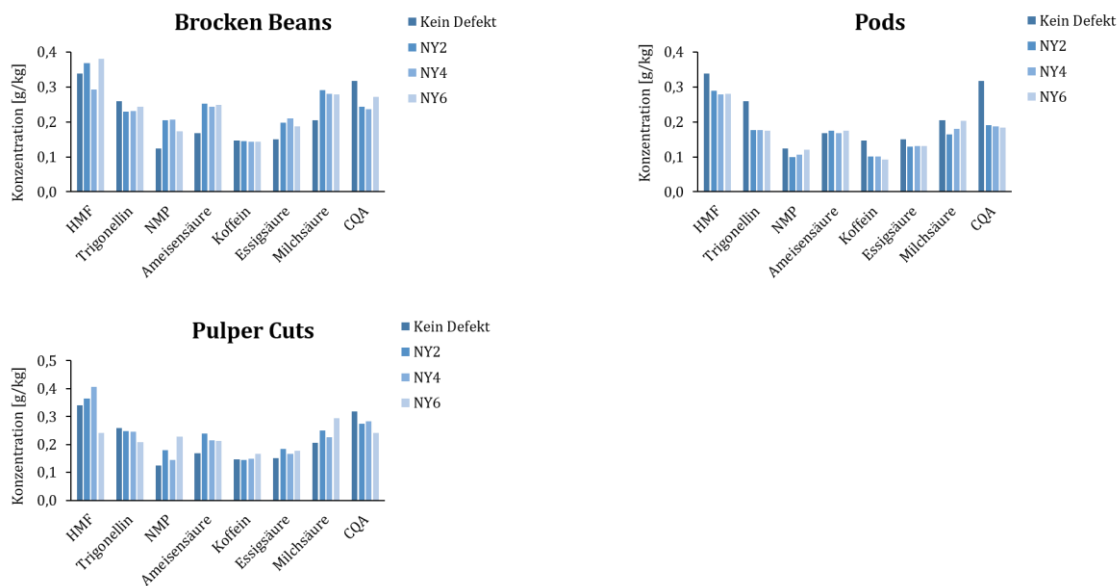


Abbildung 9-50: Die mittels ^1H -qNMR-Methode bestimmten Gehalte von 5-Hydroxymethylfurfural (HMF), Trigonellin, N-Methylpyridinium (NMP), Ameisensäure, Koffein, Essigsäure, Milchsäure und Caffeoylchinasäure-Isomere (CQA) der Kaffeedefekte „Broken Beans“, „Pods“, „Pulper Cuts“. Die Defekte sind nach New Yorker Methode klassifiziert. Es sind die Klassifizierungsstufen NY2, NY4, NY6 und die Kaffeeprob ohne Defekte aufgetragen. Zur Vergleichbarkeit wurden die ermittelten Konzentrationen der einzelnen Analyten durch unterschiedliche Faktoren dividiert: HMF (1), Trigonellin (30), NMP (3), Ameisensäure (10), Koffein (100), Essigsäure (20), Milchsäure (3), CQA (100)

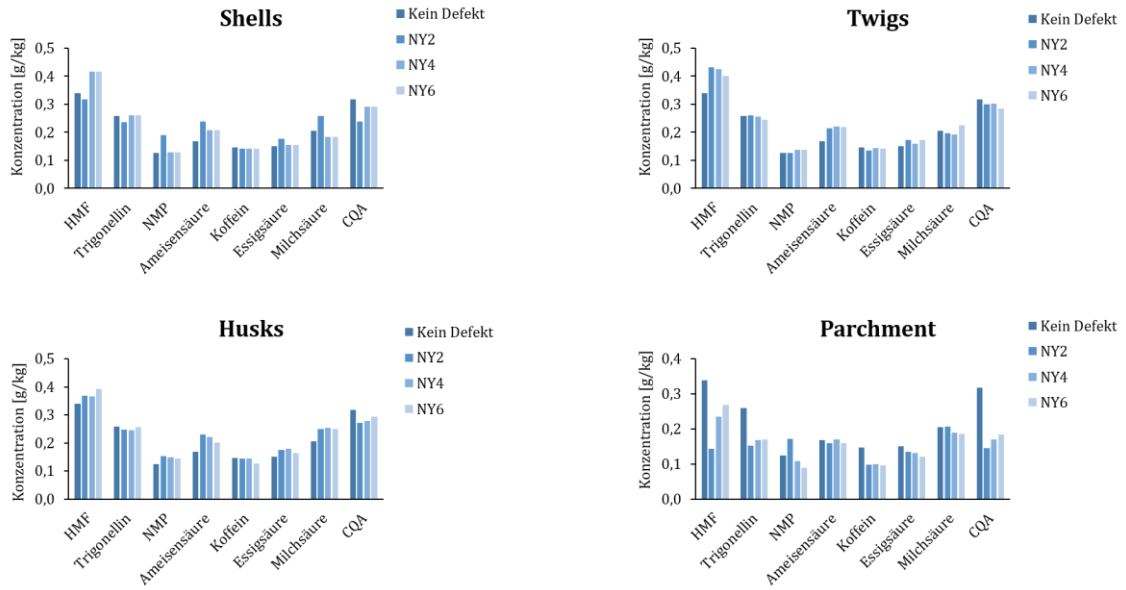


Abbildung 9-51: Die mittels ^1H -qNMR-Methode bestimmten Gehalte von 5-Hydroxymethylfurfural (HMF), Trigonellin, N-Methylpyridinium (NMP), Ameisensäure, Koffein, Essigsäure, Milchsäure und Caffeoylechinolchinasäure-Isomere (CQA) der Kaffeedefekte „Shells“, „Twigs“, „Husks“, „Parchment“. Die Defekte sind nach New Yorker Methode klassifiziert. Es sind die Klassifizierungsstufen NY2, NY4, NY6 und die Kaffeeprob ohne Defekte aufgetragen. Zur Vergleichbarkeit wurden die ermittelten Konzentrationen der einzelnen Analyten durch unterschiedliche Faktoren dividiert: HMF (1), Trigonellin (30), NMP (3), Ameisensäure (10), Koffein (100), Essigsäure (20), Milchsäure (3), CQA (100)

9.9 MATLAB®-Skript zur Erstellung von Klassifizierungsmodellen basierend auf der Analyse der lipophilen bzw. hydrophilen Metaboliten zur Abgrenzung von Kaffee nach Kontinenten bzw. nach ausgewählten Ländern bzw. nach Anbaumethode und nach Röstverfahren und zur Identifizierung möglicher lipophiler bzw. hydrophiler Metaboliten als Ursache für die Abgrenzung der entsprechenden Klassen

```
%% Erstellung der PCA-LDA mit interner und externer Validierung
%Erstellt von Vera Gottstein, Thomas Kuballa und Birk Schütz
%Vera.Gottstein@cvuaka.bwl.de; Thomas.Kuballa@cvuaka.bwl.de

%% Einlesen der Daten
% Definition der Datenpfade
% Defitnition der Datenpfade beispielhaft für die Analyse der lipophilen
% Kaffeemetaboliten

Path1 = 'N:\Praktikanten\Vera Gottstein\Kaffee\Multivariate Datenana-
lyse\fett\Spektren\Afrika';
Path2 = 'N:\Praktikanten\Vera Gottstein\Kaffee\Multivariate Datenana-
lyse\fett\Spektren\Asien';
Path3 = 'N:\Praktikanten\Vera Gottstein\Kaffee\Multivariate Datenana-
lyse\fett\Spektren\Nordamerika';
Path4 = 'N:\Praktikanten\Vera Gottstein\Kaffee\Multivariate Datenana-
lyse\fett\Spektren\Suedamerika';
Path5 = 'N:\Praktikanten\Vera Gottstein\Kaffee\Multivariate Datenana-
lyse\fett\Spektren\Bio';
Path6 = 'N:\Praktikanten\Vera Gottstein\Kaffee\Multivariate Datenanalyse\fett\Mit
Ikawa geröstet';

files1 = bbio_all_1r_files(Path1,10,1); %Einlesen der spektralen Daten
files2 = bbio_all_1r_files(Path2,10,1);
files3 = bbio_all_1r_files(Path3,10,1);
files4 = bbio_all_1r_files(Path4,10,1);
files5 = bbio_all_1r_files(Path5,10,1);
files6 = bbio_all_1r_files(Path6,10,1);
```

```

files = [files1(:); files2(:); files3(:); files4(:); files5(:), files6(:)];      % create file
list for data import

```

```

SPEC = bbio_spec_read1d(files);      % Einlesen aller Spektren

```

```

V_coffee = cell(length(SPEC),15);      % Erstellung einer Matrix für das Einlesen der
xml Datei

```

```

% Import der Metadaten aus der xml-Datei in die Matrix V_coffee

```

```

for i=1:length(SPEC)

```

```

    V_coffee{i,1} = SPEC(i).file;

```

```

    xmlfile = [SPEC(i).file(1:end-10) 'sample.xml'];

```

```

    if exist(xmlfile,'file')

```

```

        tmp = XMLFile2Struct(xmlfile);

```

```

        V_coffee{i,2} = tmp.Sample.SampleID;

```

```

        V_coffee{i,3} = tmp.Sample.ProjectID;

```

```

        V_coffee{i,4} = tmp.Sample.Matrix;

```

```

        V_coffee{i,5} = tmp.Sample.Kaffeeart;

```

```

        V_coffee{i,6} = tmp.Sample.KaffeeKontinent;

```

```

        V_coffee{i,7} = tmp.Sample.KaffeeLand;

```

```

        V_coffee{i,8} = tmp.Sample.KaffeeRegion;

```

```

        V_coffee{i,9} = tmp.Sample.KaffeeAnbau;

```

```

        V_coffee{i,10} = tmp.Sample.KaffeeFW;

```

```

        V_coffee{i,11} = str2double(tmp.Sample.KaffeeFWTonino);

```

```

        V_coffee{i,12} = str2double(tmp.Sample.Kaffeereinwaage);

```

```

        V_coffee{i,13} = str2double(tmp.Sample.Wassergehalt);

```

```

    a = isfield(tmp.Sample, 'Roestart');

```

```

    if a>0;

```

```

        V_coffee{i,14} = tmp.Sample.Roestart;

```

```

    else

```

```

        V_coffee{i,14} = 'Trommel';

```

```
end

else
    V_coffee{i,2} = 'unknown'; %Set value " if not exist
    V_coffee{i,3} = 'unknown'; %Set value " if not exist
    V_coffee{i,4} = 'unknown'; %Set value " if not exist
    V_coffee{i,5} = 'unknown'; %Set value " if not exist
    V_coffee{i,6} = 'unknown'; %Set value " if not exist
    V_coffee{i,7} = 'unknown'; %Set value " if not exist
    V_coffee{i,8} = 'unknown'; %Set value " if not exist
    V_coffee{i,9} = 'unknown'; %Set value " if not exist
    V_coffee{i,10} = 'unknown'; %Set value " if not exist
    V_coffee{i,11} = 'unknown'; %Set value " if not exist
    V_coffee{i,12} = 'unknown'; %Set value " if not exist
    V_coffee{i,13} = 'unknown'; %Set value " if not exist
    V_coffee{i,14} = 'unknown'; %Set value " if not exist
end

end

% Vervollständigen der Matrix V_coffee durch Einfügen von "unknown" in
% leere Felder
complete_V_coffee;

% Berechnung der Halbwertsbreite des TMS- oder TSP-Signals

reg = [-0.2 0.2];

for i = 1:length(SPEC)
    V_coffee{i,15} = bbio_sig_GetWidthSinglett(SPEC(i), reg(1), reg(2));
end
```



```
% Erstellung von Vektoren für die Metadaten in V_Coffee
```

```
SampleID      = V_coffee(:,2);
ProjectID     = V_coffee(:,3);
Matrix        = V_coffee(:,4);
Kaffeeart     = V_coffee(:,5);
KaffeeKontinent = V_coffee(:,6);
KaffeeLand    = V_coffee(:,7);
KaffeeRegion  = V_coffee(:,8);
KaffeeAnbau   = V_coffee(:,9);
KaffeeFW      = V_coffee(:,10);
KaffeeFWTonino = V_coffee(:,11);
KaffeeFWTonino = cell2mat(KaffeeFWTonino);
Kaffeeeinwaage = V_coffee(:,12);
Kaffeeeinwaage = cell2mat(Kaffeeeinwaage);
Wassergehalt  = V_coffee(:,13);
Wassergehalt  = cell2mat(Wassergehalt);
Roestart      = V_coffee(:,14);
TMS           = V_coffee(:,15); % TSP bei Analyse hydrophiler Metaboliten
TMS           = cell2mat(TSP); % TSP bei Analyse hydrophiler Metaboliten
TMS_control   = TMS > 1.2; % TSP bei Analyse hydrophiler Metaboliten
```

```
%% Datenvorbehandlung
```

```
% Normalisierung bezüglich ein Referenz-Signal
```

```
scaling_region = [3.0 3.05]; % Analyse lipophile Metaboliten
%scaling_region = [-0.1 0.1]; % Analyse hydrophile Metaboliten
SPEC_SCALE = bbio_spec_scale(SPEC,'ppm',scaling_region);
```

```
% Normalisierung auf einen einheitlichen Feuchtegehalt von 3%
% Die Normalisierung auf einen einheitlichen Feuchtegehalt wurde nur
% gesetzt und bei der Erstellung der Klassifizierungsmodelle nicht
% durchgeführt
%
```

```

% SPEC_SCALE_W = SPEC_SCALE;
% for i=1:length(SPEC_SCALE_W)
%     SPEC_SCALE_W(i).Data = (SPEC_SCALE(i).Data/Wassergehalt(i))*3; % norma-
lize to 3% water
%     % SPEC_SCALE_W(i).Data = (SPEC_SCALE_W(i).Data/Wassergehalt{i,1})*3; %
normalize to 3% water
% end;
%
% Normalisierung auf eine einheitliche Auswaage des evaporiterten
% Rückstands von 20 mg. Die Normalisierung auf eine einheitliche Auswaage des eva-
poriterten Rückstands wurde nur
% gesetzt und bei der Erstellung der Klassifizierungsmodelle nicht durchgeführt
%
% SPEC_SCALE_N = SPEC_SCALE;
% for i=1:length(SPEC_SCALE_N)
%     SPEC_SCALE_N(i).Data = (SPEC_SCALE(i).Data/Kaffeeinwaage(i))*20; % nor-
malize to 20 mg fat
% end;

% Alignment auf TMS- oder TSP-Signal

reg = [-0.2 0.2];
fpos = 0.0;

SPEC_SCALE_AL = SPEC_SCALE;

for i=1:length(SPEC_SCALE_AL)
    bucketing_region = bbio_spec_ppm(SPEC_SCALE_AL(i));
    idx = find(bucketing_region>reg(1) & bucketing_region<reg(2));
    D = SPEC_SCALE_AL(i).Data(idx);
    [~,P]=max(D);
    maxpos = bucketing_region(idx(P));
    SPEC_SCALE_AL(i) = bbio_spec_shift1d(SPEC_SCALE_AL(i),+maxpos-fpos);
end;

```

```

%% Datenvorbehandlung
% Bucketierung

bucketing_region = [0.5 9.8];
buckets = 2000; % Analyse lipophile Metaboliten: 2000,
% buckets = 1000; % Analyse hydrophile Metaboliten: 1000
exclusions = [8.2 8.53; 7.58 7.64; 4.6 4.95; 3.6 3.67; 3.3 3.39; 3.0 3.1; 1.18 1.23];
%Ausgeschlossener Bereich für Analyse lipophiler Metaboliten
%exclusions = [4.72 4.88]; % Ausgeschlossener Bereich für Analyse hydrophiler Me-
taboliten

A = bbio_spec_bucket1d(SPEC_SCALE_AL,'ppm',bucketing_region,'buckets',bu-
ckets,'null',exclusions,'deletenulls',true);

%% Erstellung der Gruppen für die zwei-Klassen Klassifizierungsmodelle
% Beispielhaft dargestellt für das zwei-Klassen Modell Afrika-Nordamerika
% Je nach gewünschter Klasse müssen die Metadaten in index1 und index2
% abgeändert werden

index1 = find(ismember(KaffeeKontinent,'Afrika'));
index2 = find(ismember(KaffeeKontinent,'Nordamerika'));
index = [index1(:);index2(:)];

groups_o_Afrika = KaffeeKontinent(index,:);

% Ändern der Felder in Afrika und Nordamerika

for i=1:length(groups_o_Afrika)
    z = strcmp(groups_o_Afrika{i,1},'Afrika');
    if z > 0;
        groups_o_Afrika{i,1} = 'Afrika';
    else groups_o_Afrika{i,1} = 'Nordamerika';
    end
end

%% Datenvorbehandlung

```

```
% log-Transformation

LOG_PARAMETER = 1;
Konfidenz = 0.99;
DATA_a = A.DATA(index,:);
DATA_a = bbio_LogTransform(DATA_a, LOG_PARAMETER);
PPM = A.PPM;

%% Kruskal-Wallis Test

DATA = DATA_a; % bucketierte und log-transformierte Daten
GROUPS = groups_o_Afrika; % Gruppeninformation

clf;
bbio_data_unidifn(DATA,GROUPS,'ppm',PPM,'ALPHA',0.0000289); %Kruskal Wallis
Test für lipophile Metaboliten
%bbio_data_unidifn(DATA,GROUPS,'ppm',PPM,'ALPHA',0.0000522); %Kruskal Wal-
lis Test für hydrophile Metaboliten

%% Erstellung des Klassifizierungsmodells mittels PCA-LDA und interne Validierung

LDA = lda_create(DATA, GROUPS, 10, 10, 29, false, 1); % Erstellung das Klassifizie-
rungsmodells mit interner Validierung
clf;

%Erstellung der Abbildungen
subplot(1,2,1);
lda_visu_confusion(LDA)

subplot(1,2,2);
MODEL_FOR_VISU = 1;
lda_visu_clouds(LDA,MODEL_FOR_VISU, 0.95);
```

```
%% Externe Validierung
%Einlesen der Daten und Datenvorbehandlung

% Definition des Prognoseintervalls
    Confidence = 0.95;

%Datenpfad des externen Testdatensatzes angeben
    filestest = bbio_all_1r_files('N:\Praktikanten\Vera Gottstein\Kaffee\Multivariate
Datenanalyse\fett\Spektren\Testsatz fuer GU',10,1);
    SPECTEST = bbio_spec_read1d(filestest);

% Datenvorbehandlung

    STEST = bbio_spec_scale(SPECTEST,'ppm',scaling_region);
    BTEST = bbio_spec_bucket1d(STEST,'ppm',bucketing_region,'buckets',bu-
ckets,'null',exclusions,'deletenulls',true);
    TDATA = bbio_LogTransform(BTEST.DATA, 1);
%% Externe Validierung
%Anwendung des Klassifizierungsmodells auf den externen Probandensatz

    LDAAPPLY = lda_apply(TDATA,LDA, Confidence);

subplot(1,2,2);
cla;

MODEL_FOR_VISU = 1;

lda_visu_clouds(LDA,MODEL_FOR_VISU, Confidence, false);
SC = LDAAPPLY.SCORES{MODEL_FOR_VISU};
%plot3(SC(:,1),SC(:,2),SC(:,3),'pr','MarkerFaceColor','r','MarkerSize',13);
for i=1:size(SC,1)
    text(SC(i,1),SC(i,2),SC(i,3),sprintf('%d',i),'FontSize',14,'FontWeight','Bold');
end;

n = 1:length(SPECTEST);
```

```

[bbio_mat2cell(n)' bbio_spec_info(SPECTEST,'SAMPLEID')]
%% Exportieren der Ergebnisse der externen Validierung in eine Excel-Tabelle
%Berechnung der conf-Werte
g1 = LDA.GROUP_LEGEND';
g2 = LDA.GROUP_LEGEND';
for i=1:length(g1)
    g1{i} = ['pval (' g1{i} ')'];
    g2{i} = sprintf('conf(%s, %.2f)',g2{i}, LDAAPPLY.Confidence);
end;

n = 1:length(SPECTEST);
H = ['Nr' 'SampleID' g1 g2 'DateiPfad']; % 'Herkunft laut Etikett'
T = [bbio_mat2cell(n)' bbio_spec_info(SPECTEST,'SAMPLEID')' bbio_mat2cell(LDA-
APPLY.PVAL) bbio_mat2cell(LDAAPPLY.CONF) {SPECTEST.file}'];
C = [H;T];

fn = sprintf('ergebnisse-%s.xlsx',strrep(datestr(now),':','')); %bei der Uhrzeit im
Dateinamen muss der Doppelpunkt ersetzt/entfernt werden
cd 'H:\Kaffee\Schriftliche Arbeit\Bilder Ergebnisse Diskussion\Klassifizierungsmo-
delle'
xlswrite(fn,C);
winopen(fn);
%% Erstellung der PCA-LDA zur Analyse des Loadings-Plots

%{
Finale Workflow

[MONTE-CARLO /Kreuzvalidierung über folgende Schritte]
1) (Dimensionsreduktion via Kruskal-Wallis)
2) Unit-Variance-Scaling
3) Dimensionsreduktion via PCA
4) Trennung via LDA

%}

```

```
% Definitionen
DEF_CV = 10; % 10 bedeutet in 10 Teile zerlegen
MONTE_CARLO_RUNS = 1;
PCADIM = 29; % Anzahl Dimensionen für PCA
ConfEllipsoidAlpha = 0.95; % Ellipsoide zeigen x% der Modelldaten

UNIT_SCALE = false; % true=ja, false=nein

KRUSKAL_REDUCTION = 0; % 0 = keine Reduktion, Anzahl der Buckets

TAB = tabulate(GROUPS);
m = size(TAB,1); % Anzahl der Gruppen
n = size(DATA,1); % Größe der Stichprobe, der Daten
COLORMAP = lines(m);

LDADIM = m-1;
clc;

GROUPS_predicted = cell(n,MONTE_CARLO_RUNS);

for MC = 1:MONTE_CARLO_RUNS

    figure(MC);
    set(gcf,'Renderer','opengl');

% Unterteilung der Kreuzvalidierung
    idx = randperm(n);
    CV_SETS = cell(DEF_CV,1);
    d = floor(n/DEF_CV);
    for i=1:DEF_CV
        CV_SETS{i} = idx((1:d)+(i-1)*d);
        if i==DEF_CV
            CV_SETS{i} = idx(((1)+(i-1)*d):n);
        end;
    end;
```

```
end;

clf;

for CV_STEP = 1:DEF_CV

% Modelldaten und TestDaten auftrennen
    testset = CV_SETS{CV_STEP};
    modelset = setdiff(1:n,testset);

    DATA_m = DATA(modelset,:);
    DATA_t = DATA(testset,:);
    GROUPS_m = GROUPS(modelset);
    GROUPS_t = GROUPS(testset);

% Modell bauen
% -----

% Schritt 2 / unit variance scaling
    if UNIT_SCALE
        Mstd = std(DATA_m);
        DATA_m = DATA_m./repmat(Mstd, size(DATA_m,1),1);
    end;

    if KRUSKAL_REDUCTION>0
        kruskalp = bbio_anova(DATA_m,GROUPS_m,'kruskal');
        [~,KRUKALSORT] = sort(kruskalp);
        DATA_m = DATA_m(:,KRUKALSORT(1:KRUSKAL_REDUCTION));
    end;

% Schritt 3 - PCA (zur Dimensionsreduktion)
% Mittelwert der Modelldaten berechnen
    Mmean = mean(DATA_m);
```

```

DATA_m = DATA_m-repmat(Mmean, size(DATA_m,1),1);
[coeff,~,~,~,EXPLAINED] = pca(DATA_m);

DATA_m = DATA_m*coeff; % Scores der PCA
% Dimensionsreduktion
DATA_m = DATA_m(:,1:PCADIM);
fprintf(' Explained Variance (%d dimensions): %.2f%%\n', PCADIM,sum(EX-
PLAINED(1:PCADIM)));

% Schritt 4 - LDA berechnen
% Berechne LDA anhand der Scores der PCA
[~,~,LDA] = manova1(DATA_m, GROUPS_m);
% Berechne die Scores der LDA
DATA_m = DATA_m * LDA.eigenvec;
% Scores = DATA_m * (coeff * LDA.eigenvec)
% Loadings sind PCACoeff*LDACoeff

% Gruppenzentren bestimmen
GROUP_MEANS = NaN(m,LDADIM);
for i=1:m
    tmpidx = find(ismember(GROUPS_m,TAB{i,1}));
    GROUP_MEANS(i,:) = mean(DATA_m(tmpidx,1:LDADIM));
end;

% Plot Loadings
figure(MC+100);
subplot(4,3,CV_STEP);
LOADINGS = coeff(:,1:PCADIM)*LDA.eigenvec;
plot(LOADINGS(:,1),LOADINGS(:,2),'ko','MarkerSize',3);
ltext = cell(size(LOADINGS,1),1);
for l=1:length(ltext)
    ltext{l} = sprintf('%d',l);
    ltext{l} = sprintf('%.3f',PPM(l));
    if KRUSKAL_REDUCTION>0

```

```

        ltext{1} = sprintf('%.3f',PPM(KRUKALSORT(1)));
    end;
end;
text(LOADINGS(:,1),LOADINGS(:,2),ltext);

% Anwendung des Modells auf interne Testdaten
% -----
    if UNIT_SCALE
        DATA_t = DATA_t./repmat(Mstd, size(DATA_t,1),1);
    end;

    if KRUSKAL_REDUCTION>0
        DATA_t = DATA_t(:,KRUKALSORT(1:KRUSKAL_REDUCTION));
    end;

    DATA_t = DATA_t-repmat(Mmean, size(DATA_t,1),1);
    DATA_t = DATA_t*coeff; % Scores der PCA
    DATA_t = DATA_t(:,1:PCADIM);
    DATA_t = DATA_t * LDA.eigenvec;

% Distanzen berechnen
    DIST = pdist2(DATA_t(:,1:LDADIM),GROUP_MEANS,'euclidean');
    [~,b]=min(DIST,[],2);
    GROUPS_predicted(testset, MC)=TAB(b,1);

    figure(MC);
    subplot(4,3,CV_STEP);
    hold on;
    L = perms([-1 0 1]);
    for i=1:size(L,1)
        light('Position',L(i,:), 'Style','infinite');
    end;
    axis vis3d;

```

```

    for i=1:m
        tmpidx = find(ismember(GROUPS_m,TAB{i,1}));
        plot3(DATA_m(tmpidx,1),DATA_m(tmpidx,2),DATA_m(tmpidx,3),'ko','Marker-
FaceColor',COLORMAP(i,:),'MarkerSize',3);
        % Darstellung Ellipsoide
        tmpEData = DATA_m(tmpidx,1:3);
        mypatch = bbio_conf3dellipsoid(tmpEData, ConfEllipsoidAlpha, 30);
        patch(mypatch,'FaceColor',COLORMAP(i,:),'FaceAlpha',0.3,'EdgeColor','no-
ne','FaceLighting','phong');

        tmpidx = find(ismember(GROUPS_t,TAB{i,1}));
        plot3(DATA_t(tmpidx,1),DATA_t(tmpidx,2),DATA_t(tmpidx,3),'ks','Marker-
FaceColor',COLORMAP(i,:));

    end;
    box on;
end;

% Konfusionsmatrix
CONFUSION = NaN(m,m);
for i=1:m
    for j=1:m
        tmpidx = find(ismember(GROUPS,TAB{i,1}));
        tmpidxp = find(ismember(GROUPS_predicted(:, MC),TAB{j,1}));
        cl = intersect(tmpidx,tmpidxp);
        fprintf('Von %d Proben der Gruppe %s wurden %d als %s bestimmt.\n',
length(tmpidx),TAB{i,1},length(cl),TAB{j,1})
        CONFUSION(i,j) = length(cl);
    end;
end;
CONFUSION
CONFUSION_PERCENTAGE = CONFUSION./repmat(sum(CONFU-
SION,2),1,size(CONFUSION,2));
CONFUSION_PERCENTAGE*100
end;

```

9.10 MATLAB®-Skript zur Datenfusionierung der ¹H-NMR spektroskopischen Daten der lipophilen und hydrophilen Metaboliten und zur Erstellung von Klassifizierungsmodellen für die Abgrenzung von Kaffee nach ausgewählten Ländern

```
%% Datenfusionierung der NMR-Daten der lipophilen und hydrophilen NMR spektroskopischen Daten
```

```
%Erstellt von Vera Gottstein, Thomas Kuballa und Birk Schütz
```

```
%Vera.Gottstein@cvuaka.bwl.de; Thomas.Kuballa@cvuaka.bwl.de
```

```
%% Einlesen der Daten
```

```
% Definition der Datenpfade
```

```
Path_f = 'N:\Praktikanten\Vera Gottstein\Kaffee\Multivariate Datenanalyse\Datafusion\Fett';
```

```
Path_w = 'N:\Praktikanten\Vera Gottstein\Kaffee\Multivariate Datenanalyse\Datafusion\Wässrig';
```

```
files_f = bbio_all_1r_files(Path_f,10,1); %Einlesen der spektralen Daten
```

```
files_w = bbio_all_1r_files(Path_w,10,1);
```

```
SPEC_f = bbio_spec_read1d(files_f); % Einlesen aller Spektren
```

```
SPEC_w = bbio_spec_read1d(files_w);
```

```
V_coffee = cell(length(SPEC_f),15); % Erstellung einer Matrix für das Einlesen der xml Datei
```

```
% Import der Metadaten aus der xml-Datei in die Matrix V_coffee
```

```
for i=1:length(SPEC_f)
```

```
    V_coffee{i,1} = SPEC_f(i).file;
```

```
    xmlfile = [SPEC_f(i).file(1:end-10) 'sample.xml'];
```

```
    if exist(xmlfile,'file')
```

```
        tmp = XMLFile2Struct(xmlfile);
```

```
        V_coffee{i,2} = tmp.Sample.SampleID;
```

```
V_coffee{i,3} = tmp.Sample.ProjectID;
V_coffee{i,4} = tmp.Sample.Matrix;
V_coffee{i,5} = tmp.Sample.Kaffeeart;
V_coffee{i,6} = tmp.Sample.KaffeeKontinent;
V_coffee{i,7} = tmp.Sample.KaffeeLand;
V_coffee{i,8} = tmp.Sample.KaffeeRegion;
V_coffee{i,9} = tmp.Sample.KaffeeAnbau;
V_coffee{i,10} = tmp.Sample.KaffeeFW;
V_coffee{i,11} = str2double(tmp.Sample.KaffeeFWTonino);
V_coffee{i,12} = str2double(tmp.Sample.Kaffeeeinwaage);
V_coffee{i,13} = str2double(tmp.Sample.Wassergehalt);

a = isfield(tmp.Sample, 'Roestart');
if a>0;
    V_coffee{i,14} = tmp.Sample.Roestart;
else
    V_coffee{i,14} = 'Trommel';
end

else
    V_coffee{i,2} = 'unknown'; %Set value " if not exist
    V_coffee{i,3} = 'unknown'; %Set value " if not exist
    V_coffee{i,4} = 'unknown'; %Set value " if not exist
    V_coffee{i,5} = 'unknown'; %Set value " if not exist
    V_coffee{i,6} = 'unknown'; %Set value " if not exist
    V_coffee{i,7} = 'unknown'; %Set value " if not exist
    V_coffee{i,8} = 'unknown'; %Set value " if not exist
    V_coffee{i,9} = 'unknown'; %Set value " if not exist
    V_coffee{i,10} = 'unknown'; %Set value " if not exist
    V_coffee{i,11} = 'unknown'; %Set value " if not exist
    V_coffee{i,12} = 'unknown'; %Set value " if not exist
    V_coffee{i,13} = 'unknown'; %Set value " if not exist
    V_coffee{i,14} = 'unknown'; %Set value " if not exist
end
```

end

% Vervollständigen der Matrix V_coffee durch Einfügen von "unknown" in

% leere Felder

complete_V_coffee;

% Berechnung der Halbwertsbreite des TSP-Signal

reg = [-0.2 0.2];

for i = 1:length(SPEC_f)

V_coffee{i,15} = bbio_sig_GetWidthSinglett(SPEC_f(i), reg(1), reg(2));

end

% Erstellung von Vektoren für die Metadaten in V_Coffee

SampleID = V_coffee(:,2);

ProjectID = V_coffee(:,3);

Matrix = V_coffee(:,4);

Kaffeeart = V_coffee(:,5);

KaffeeKontinent = V_coffee(:,6);

KaffeeLand = V_coffee(:,7);

KaffeeRegion = V_coffee(:,8);

KaffeeAnbau = V_coffee(:,9);

KaffeeFW = V_coffee(:,10);

KaffeeFWTonino = V_coffee(:,11);

KaffeeFWTonino = cell2mat(KaffeeFWTonino);

Kaffeeeinwaage = V_coffee(:,12);

Kaffeeeinwaage = cell2mat(Kaffeeeinwaage);

Wassergehalt = V_coffee(:,13);

Wassergehalt = cell2mat(Wassergehalt);

Roestart = V_coffee(:,14);

```
TSP      = V_coffee(:,15);
TSP      = cell2mat(TSP);
TSP_control  = TSP > 1.2;
```

```
%% Datenvorbehandlung
```

```
% Normalisierung bezüglich ein Referenz-Signal und Alignment
```

```
% Parameter
```

```
scaling_region_f = [3.0 3.06]; % Referenz-Signal lipophile Metaboliten
```

```
scaling_region_w = [-0.1 0.1]; % Referenz-Signal hydrophile Metaboliten
```

```
reg = [-0.2 0.2]; % Bereich für TMS- und TSP-Signal
```

```
fpos = 0.0; % Zielbereich für TMS- und TSP-Signal
```

```
% Normalisierung
```

```
SPEC_SCALE_f = bbio_spec_scale(SPEC_f,'ppm',scaling_region_f); % Normalisierung
hydrophile Metaboliten
```

```
SPEC_SCALE_w = bbio_spec_scale(SPEC_f,'ppm',scaling_region_w); % Normalisierung
lipophile Metaboliten
```

```
% Alignment
```

```
SPEC_SCALE_AL_f = SPEC_SCALE_f; % Alignment lipophile Metaboliten
```

```
for i=1:length(SPEC_SCALE_AL_f)
```

```
    bucketing_region = bbio_spec_ppm(SPEC_SCALE_AL_f(i));
```

```
    idx = find(bucketing_region>reg(1) & bucketing_region<reg(2));
```

```
    D = SPEC_SCALE_AL_f(i).Data(idx);
```

```
    [~,P]=max(D);
```

```
    maxpos = bucketing_region(idx(P));
```

```
    SPEC_SCALE_AL_f(i) = bbio_spec_shift1d(SPEC_SCALE_AL_f(i),+maxpos-fpos);
```

```
end;
```

```
SPEC_SCALE_AL_w = SPEC_SCALE_w; % Alignment hydrophile Metaboliten
```

```
for i=1:length(SPEC_SCALE_AL_w)
```

```

bucketing_region = bbio_spec_ppm(SPEC_SCALE_AL_w(i));
idx = find(bucketing_region>reg(1) & bucketing_region<reg(2));
D = SPEC_SCALE_AL_w(i).Data(idx);
[~,P]=max(D);
maxpos = bucketing_region(idx(P));
SPEC_SCALE_AL_w(i) = bbio_spec_shift1d(SPEC_SCALE_AL_w(i),+maxpos-fpos);
end;

%% Datenvorbehandlung
% Bucketierung und log-Transformation

% Defentition der Paramerer
bucketing_region = [0.5 9.8];
buckets_f = 2000;
buckets_w = 1000;
exclusions_f = [8.2 8.53; 7.58 7.64; 4.6 4.95; 3.6 3.67; 3.3 3.39; 3.0 3.1; 1.18 1.23];
exclusions_w = [4.72 4.88];

scf = 1.0; % Skalieurngsfaktor für Daten der lipophilen Metaboliten (low level Daten-
fusionieurng)
LOG_PARAMETER = 1; % Log Parameter
Konfidenz = 0.99; % Confidence intervall
PCAdim = 29; % PCA dimensions
MONTE_CARLO = 10; % Monte Carlo cross validations
CROSS_SETS = 10; % Cross sets for Monte Carlo validations
AutoScaling = 'false'; % use autoscaling with 'true' or not with 'false'

% Bucketierung der Daten
O_f = bbio_spec_bucket1d(SPEC_SCALE_AL_f,'ppm',bucketing_region,'buckets',bu-
ckets_f,'null',exclusions_f,'deletenulls',true); % buckets from fat spectra
O_w = bbio_spec_bucket1d(SPEC_SCALE_AL_w,'ppm',bucketing_region,'buckets',bu-
ckets_w,'null',exclusions_f,'deletenulls',true); % buckets from water spectra

%% Erstellung der Gruppen für die zwei-Klassen Klassifizierungsmodelle
% Beispielhaft dargestellt für das zwei-Klassen Modell Afrika-Nordamerika

```



```
% Je nach gewünschter Gruppe müssen die Metadaten in index1 und index2
% abgeändert werden

index1 = find(ismember(KaffeeLand,'Kolumbien'));
index2 = find(ismember(KaffeeKontinent,'Asien'));
index = [index1(:);index2(:)];

groups_o_Afrika = KaffeeLand(index,:);

% Ändern der Felder in Afrika und Nordamerika

for i=1:length(groups_o_Afrika)
    z = strcmp(groups_o_Afrika{i,1},'Kolumbien');
    if z > 0;
        groups_o_Afrika{i,1} = 'Kolumbien';
    else groups_o_Afrika{i,1} = 'Asien';
    end
end

%% Fusionierung der Daten für low-level-Datenfusionierung

O = O_f;
O.DATA = horzcat(O_f.DATA*scf,O_w.DATA); % low level Datenfusionierung

%% Durchführung der low-level Datenfusionierung

DATA = O.DATA(index,:); % Daten der low level Datenfusionierung
DATA = bbio_LogTransform(DATA, LOG_PARAMETER); % Log-Transformation data
fusion
DATA_f = bbio_LogTransform(O_f.DATA, LOG_PARAMETER); % Log-Transformation
lipophile Metaboliten
DATA_w = bbio_LogTransform(O_w.DATA, LOG_PARAMETER); % Log-Transforma-
tion hydrophile Metaboliten
LDA_ll = lda_create(DATA, groups_o_Afrika, MONTE_CARLO, CROSS_SETS, PCAdim,
AutoScaling, 1); % LDA data fusion
```

```
% print figures

figure(30)
clf; % print
subplot(1,2,1);
lda_visu_confusion(LDA_ll);
subplot(1,2,2);
MODEL_FOR_VISU = 1;
lda_visu_clouds(LDA_ll,MODEL_FOR_VISU, Konfidenz);
title('PCA-LDA low level Datenfusionierung');

%% Durchführung der mid level Datenfusionierung

% Berechnung der PCA-Scores für die lipophilen Metaboliten

PCADIM = 29; %
ConfEllipsoidAlpha = 0.95;
UNIT_SCALE = false;
KRUSKAL_REDUCTION = 0;
GROUPS = groups_o_Afrika;

clc;

% PCA
Mmean = mean(DATA_f);
DATA_f = DATA_f-repmat(Mmean, size(DATA_f,1),1);
[coeff,~,~,EXPLAINED] = pca(DATA_f);

DATA_f = DATA_f*coeff; % Scores der PCA
DATA_f = DATA_f(:,1:PCADIM);
fprintf(' Explained Variance (%d dimensions): %.2f%%\n', PCADIM,sum(EXPLAINED(1:PCADIM)));

% Berechnung der PCA-Scores für die hydrophilen Metaboliten
```

```

Mmean = mean(DATA_w);
DATA_w = DATA_w-repmat(Mmean, size(DATA_w,1),1);
[coeff,~,~,~,EXPLAINED] = pca(DATA_w);

DATA_w = DATA_w*coeff; % Scores der PCA
DATA_w = DATA_w (:,1:PCADIM);
fprintf(' Explained Variance (%d dimensions): %.2f%%\n', PCADIM,sum(EXPLAI-
NED(1:PCADIM)));

% Fusionierung der PCA-Scores der lipophilen und hydrophilen Daten

DATA_ml = horzcat(DATA_f,DATA_w);

% PCA-LDA der mid level fusionierten lipophilen und hydrophilen Daten

LDA_ml = lda_create( DATA_ml, GROUPS, 10, 10, 38 , false , 1); %führt eine PCA-LDA
durch mit der optimierten PCA-Dimension

% Dimensionen(bestd) %damit kann der optimale Wert der PCA angezeigt werden

clf;
figure(31);
subplot(1,2,1);
lda_visu_confusion(LDA_ml)

subplot(1,2,2);
MODEL_FOR_VISU = 1;
lda_visu_clouds(LDA_ml,MODEL_FOR_VISU, Konfidenz);
title('PCA-LDA mid level Datenfusionierung');

```

9.11 Verwendete MATLAB®-Unterskripte für die in Kapitel 9.9 und Kapitel 9.10 verwendeten Hauptskripte

```
function FILELIST = bbio_all_1r_files( START_DIR, ExpNo, ProcNo)
```

```
if nargin<2
    ExpNo = [];
    ProcNo = [];
end;

FILELIST = recursive_reading(START_DIR, ExpNo, ProcNo);

function FILELIST = recursive_reading(BASEDIR, ExpNo, ProcNo)

    FILELIST = {};

    D = dir(BASEDIR);

    for i=1:length(D)
        s = D(i).name;
        c = [BASEDIR '\' s];
        if strcmp(s, '.') || strcmp(s, '..')
            continue;
        end;
        if isdir(c)
            tmpL = recursive_reading(c, ExpNo, ProcNo);
            FILELIST = union(FILELIST, tmpL);
        end;
        if strcmp(s, '1r') && (isempty(findstr(c, '\99999\pdata\')))

            c = strrep(c, '/', '\');
            %c = strrep(c, '\\', '\');

            % check the filter
            try
                tokens = regexp(c, '(\\d+)\\pdata\\(\\d+)\\1r', 'tokens');
                tokens = tokens{1};
                E = str2double(tokens{1});
            catch
            end
        end
    end
end
```

```
    P = str2double(tokens{2});
catch %#ok<CTCH>
    continue;
end;

if ~isempty(ExpNo)
    if (E~=ExpNo) || (P~=ProcNo)
        continue;
    end;
end;

FILELIST = union(FILELIST, c);
end;
end;
```

function P = bbio_anova(DATA, GROUPS, Kruskal)

```
if nargin<3
    Kruskal = "";
end;

if strcmp(Kruskal,'kruskal')
    UseKruskalWallis = true;
else
    UseKruskalWallis = false;
end;

P = zeros(size(DATA,2),1);
for i=1:size(DATA,2)
    if UseKruskalWallis
        P(i) = kruskalwallis(DATA(:,i),GROUPS,'off');
    else
        P(i) = anova1(DATA(:,i),GROUPS,'off');
    end;
end;
```

```
end;
```

```
function [mypatch, s,v,m] = bbio_conf3dellipsoid(X, alpha, points)
```

```
if nargin<3
```

```
    points = 30;
```

```
end;
```

```
if nargin<2
```

```
    alpha = 0.95;
```

```
end;
```

```
m    = mean(X);
```

```
c    = cov(X);
```

```
[u,s,v] = svd(c); %#ok<ASGLU>
```

```
s    = sqrt(diag(s));
```

```
limit = sqrt(chi2inv(alpha,2));
```

```
mypatch = bbio_conf3dellipsoid_internal(limit, s,v,m, points);
```

```
function RESULT = bbio_data_unidiffn(DATA, GROUPS, varargin)
```

```
% Dealing with the input
```

```
p = inputParser;
```

```
p.addParamValue('ppm',[]);
```

```
p.addParamValue('plotQlow',0.125);
```

```
p.addParamValue('plotQup',0.875);
```

```
p.addParamValue('colors',[]);
```

```
p.addParamValue('FaceAlpha',0.2);
```

```
p.addParamValue('DoKruskal',true);
```

```
p.addParamValue('ALPHA',0.01);
```

```
p.parse(varargin{:});
```

```
PARS = p.Results;

if nargin==0
    RESULT = PARS;
    return;
end;

RESULT = [];

[G, GINV] = grp2idx(GROUPS);

NrOfGroups = max(G);

if isempty(PARS.colors)
    colors = jet(NrOfGroups);
    if NrOfGroups == 2
        colors = [1 0 0;0 0 1];
    end;
    if NrOfGroups == 3
        colors = [1 0 0;0 0 1; 0 1 0];
    end;
else
    colors = PARS.colors;
end;

%% Plotting
% -----

clf;
set(gca,'xdir','reverse');
hold on;

ppm = PARS.ppm;
```

```
for i=1:NrOfGroups

    myDATA = DATA(G == i,:);
    mycol = colors(i,:);
    % Plotting the MEDIAN
    plot(ppm,nanmedian(myDATA), 'color',mycol,'LineWidth',2);

end;

for i=1:NrOfGroups

    myDATA = DATA(G == i,:);
    mycol = colors(i,:);

    % Plotting the Patches
    C1 = quantile(myDATA, PARS.plotQlow);
    C2 = quantile(myDATA, PARS.plotQup);
    patch([ppm fliplr(ppm)], [C1 fliplr(C2)],...
        'k','EdgeColor',mycol,'LineStyle','none','FaceAlpha',PARS.FaceAlpha,...
        'FaceColor', mycol);
end;

xlabel('[ppm]');
grid on;
set(gca,'xlim',[min(ppm) max(ppm)]);
legend(GINV);

%% Kruskal-Wallis
% -----
```



```
if PARS.DoKruskal

    kruskal = bbio_anova(DATA, GROUPS,'kruskal');
    RESULT.kruskal = kruskal;
    m = length(kruskal);

    ultimax = max(max(DATA));
    ultimmin = min(min(DATA));

    d = (ultimax - ultimmin)/20;

    for i=1:m
        delta = (ppm(2)-ppm(1))/1;
        ppmpos = ppm(i);
        if kruskal(i) < PARS.ALPHA
            patch([ppmpos-delta/2 ppmpos+delta/2 ppmpos+delta/2 ppmpos-delta/2],...
                [ultimmin ultimmin ultimmax ultimmax],'w','LineWidth',...
                0.1, 'FaceColor',[1,0.4,0.8], 'EdgeColor','none',...
                'FaceAlpha',0.3);

            patch([ppmpos-delta/2 ppmpos+delta/2 ppmpos+delta/2 ppmpos-delta/2],...
                [ultimmin ultimmin ultimmin-d ultimmin-d],'w','LineWidth',...
                0.1, 'FaceColor',[1,0.4,0.8], 'EdgeColor',[1,0.4,0.8]',...
                'FaceAlpha',1);

            patch([ppmpos-delta/2 ppmpos+delta/2 ppmpos+delta/2 ppmpos-delta/2],...
                [ultimax ultimax ultimax+d ultimax+d],'w','LineWidth',...
                0.1, 'FaceColor',[1,0.4,0.8], 'EdgeColor',[1,0.4,0.8]',...
                'FaceAlpha',1);
        end;
    end;

end;
```

function Dout = bbio_LogTransform(D, minlevel)

```
Dout = D;

for i=1:size(D,2)
    % values less minlevel
    ind = D(:,i)<= minlevel;
    ind2 = D(:,i)>minlevel;

    % value less noiselevel stay
    Dout(ind,i) = D(ind,i);

    % Log-Transformation
    % stetig/diffbar fortsetzung von linear
    temp = D(ind2,i);
    temp = (log(temp) - log(minlevel))/(1/minlevel) + minlevel;
    Dout(ind2,i) = temp;
end;
```

function CELL = bbio_mat2cell(MAT)

```
CELL = cell(size(MAT));

for i=1:size(CELL,1)
    for j=1:size(CELL,2)
        CELL{i,j} = MAT(i,j);
    end;
end;
```

function W = bbio_sig_GetWidthSinglett(RAW, minppm, maxppm)

```
if nargin<2
    % Width of TSP
    minppm = -0.05;
    maxppm = 0.05;
end;
```

```
W = zeros(length(RAW),1).*NaN;

for j = 1:length(W)
    SPEC = RAW(j);
    try
        ppm = bbio_spec_ppm(SPEC);

        idx = ((ppm>minppm) & (ppm<maxppm));
        D = SPEC.Data(idx);
        D = D - min(D);

        ppm2 = ppm(idx);
        tmpD = diff(sign(D-max(D)/2));

        i1 = find(tmpD == 2,1,'first');
        i2 = find(tmpD == -2,1,'last');

        DT = D - max(D)/2;
        a1 = -DT(i1)/(DT(i1+1)-DT(i1));
        a2 = -DT(i2)/(DT(i2+1)-DT(i2));

        delta = abs(ppm2(2)-ppm2(1));

        p1 = ppm2(i1) - a1*delta;
        p2 = ppm2(i2) - a2*delta;

        W(j) = abs(p2-p1) * SPEC.SF;
    catch e %#ok<NASGU>
    end;
end;
```

function MDATA = bbio_spec_bucket1d(RAW, varargin)

```
% Dealing with the input
p = inputParser;
```

```
p.addParamValue('ppm',[0.5 9.5]);
p.addParamValue('buckets',[]);
p.addParamValue('delta',[]);
p.addParamValue('null',[]);
p.addParamValue('deletenulls',false);
p.parse(varargin{:});
PARS = p.Results;

if nargin==0
    MDATA = PARS;
    return;
end;

n = length(RAW);

% Build the output struct
MDATA = [];

if isempty(PARS.buckets) && isempty(PARS.delta)
    error('Either variable "buckets" or "delta" must be defined!');
end;

ppm = PARS.ppm;
maxppm = max(ppm);
minppm = min(ppm);

if ~isempty(PARS.buckets)
    % number of buckets is given
    buckets = PARS.buckets;
    bsize = (maxppm-minppm)/(buckets-1);
else
    % size of each bucket is given
    bsize = PARS.delta;
    buckets = floor((maxppm-minppm)/bsize)+1;
end;
```

```
% Generate PPM-Axis

PPM = ones(1, buckets);

BucketList = zeros(buckets, 2);
for i=1:buckets
    BucketList(i,1) = (bsize*(i-1)) + minppm - bsize/2;
    BucketList(i,2) = (bsize*(i)) + minppm - bsize/2;
    PPM(i) = (BucketList(i,1) + BucketList(i,2))/2;
end;

MADATA.PPM = PPM;
MADATA.BLIST = BucketList;
MADATA.DATA = zeros(n, buckets);

if isempty(RAW)
    return;
end;

% test, if PPM is finer or as fine as RAW(1)
INTERPOL = false;
p = bbio_spec_ppm(RAW(1));
deltaPPM = abs(PPM(2)-PPM(1));
deltaRAW = abs(p(2)-p(1));
if deltaPPM<deltaRAW*2
    % turn on interpolation-Mode
    INTERPOL = true;
end;

TMPDATA = MADATA.DATA;
for k=1:n

    % simple but slow method
    %{
```

```

for i=1:buckets
    p = bbio_spec_ppm(RAW(k));
    idx = p>BucketList(i,1) & p<=BucketList(i,2);
    MDATA.DATA(k,i) = mean(RAW(k).Data(idx));
end;
%}

% fast method
if INTERPOL == false
    SPEC = RAW(k);

    si = SPEC.size;
    ofs = SPEC.maxppm;
    sw = SPEC.sw;
    S = SPEC.Data;

    tmpV = zeros(1,buckets);
    for i=1:buckets
        % fast method
        [tmp, minind] = internal_ppmindex(ofs, sw, si, BucketList(i,1)); %#ok<ASGLU>
        [tmp, maxind] = internal_ppmindex(ofs, sw, si, BucketList(i,2)); %#ok<ASGLU>

        if round(minind) ~= minind
            minind = ceil(minind);
        end;
        if round(maxind) ~= maxind
            maxind = ceil(maxind)+1;
        end;
        c = (minind - maxind)+1;
        tmpV(i)=sum(S(maxind:minind))./c;
    end;
    TMPDATA(k,:) = tmpV;
else
    % Bucketing via Interpolation
    p = bbio_spec_ppm(RAW(k));

```

```

    TMPDATA(k,:) = interp1(p,RAW(k).Data,PPM);
end;
end;
MDATA.DATA = TMPDATA;

% applying the "null-space", exclusions
NULLS = [];
if ~isempty(PARS.null)
    for i=1:size(PARS.null,1)
        NULLS = union( NULLS, find(PPM>PARS.null(i,1) & PPM<PARS.null(i,2)));
    end;
end;

MDATA.DATA(:,NULLS) = 0;

```

```

if PARS.deletenulls
    ii = setdiff(1:size(MDATA.DATA,2), NULLS);
    MDATA.DATA = MDATA.DATA(:, ii);
    MDATA.PPM = MDATA.PPM(ii);
end;

```

```
function [index, exactindex] = internal_ppmindex(ofs, sw, size, ppm)
```

```

    exactindex = (ofs-ppm)/sw*(size-1);
    index = ceil(exactindex);

```

```
function INFO = bbio_spec_info(RAWData, Parameter)
```

```

INFO = [];

DirectNumbers = {'LB','SI','Date','O1','PHC0','PHC1','SF','RG','NS','TD','BF1','SW'};
Arrays        = {'PULSES','DELAYS','POWERLEVELS', 'POWERLEVELSW', 'POWERLE-
VELSDB'};
Singles       = {'P1','D1','SR','EXPNO','PROCNO', 'SAMPLEID','PROBHD'};
DirectStrings = {'PULPROG','SPECT','TITLE','file'};

```

```
if nargin==0
    INFO = [];
    INFO = union(INFO, Arrays);
    INFO = union(INFO, DirectNumbers);
    INFO = union(INFO, Singles);
    INFO = union(INFO, DirectStrings);
    return;
end;

if ~isempty(find(ismember(DirectStrings, Parameter), 1))
    INFO = ({RAWData.(Parameter)});
    return;
end;

if ~isempty(find(ismember(DirectNumbers, Parameter), 1))
    if strcmp(Parameter,'SI')
        Parameter = 'size';
    end;
    if strcmp(Parameter,'SW')
        Parameter = 'sw';
    end;
    INFO = cell2mat({RAWData.(Parameter)});
    return;
end;

PULSES    = cell2mat({RAWData.PULSES});
DELAYS    = cell2mat({RAWData.D});
POWERLEVEL = cell2mat({RAWData.POWERLEVEL});
POWERLEVELW = cell2mat({RAWData.POWERLEVELW});
POWERLEVELDB = cell2mat({RAWData.POWERLEVELDB});

if strcmp(Parameter,'PL1');
    INFO = POWERLEVEL(:,2);
```



```
    if RAWData(1).TOPSPIN>=3
        INFO = POWERLEVELDB(:,2);
    end;
    return;
end;

if strcmp(Parameter,'PL9');
    INFO = POWERLEVEL(:,10);
    if RAWData(1).TOPSPIN>=3
        INFO = POWERLEVELDB(:,10);
    end;
    return;
end;

if strcmp(Parameter,'PULSES');
    INFO = PULSES;
    return;
end;

if strcmp(Parameter,'DELAYS');
    INFO = DELAYS;
    return;
end;

if strcmp(Parameter,'POWERLEVELS');
    INFO = POWERLEVEL;
    return;
end;

if strcmp(Parameter,'POWERLEVELSW');
    INFO = POWERLEVELW;
    return;
end;

if strcmp(Parameter,'POWERLEVELSDB');
```

```
    INFO = POWERLEVELDB;
    return;
end;

if strcmp(Parameter,'P1');
    INFO = PULSES(:,2);
    return;
end;

if strcmp(Parameter,'D1');
    INFO = DELAYS(:,2);
    return;
end;

files = {RAWData.file};
files = strrep(files,'/','\');

if strcmp(Parameter,'EXPNO');
    INFO = cellfun(@(x) str2double(x{1}),regexp(files,'(\d+)\pdata\\(\d+)\\d+r','tokens'));
    return;
end;

if strcmp(Parameter,'PROBHD');
    INFO = cell(length(RAWData),1);
    for i=1:length(RAWData)
        s = strtrim(RAWData(i).ACQUS.PROBHD(2:end-1));
        INFO{i} = s;
    end;
    return;
end;

if strcmp(Parameter,'PROCNO');
    INFO = cellfun(@(x) str2double(x{1}),regexp(files,'(\d+)\pdata\\(\d+)\\d+r','tokens'));
end;
```

```

    return;
end;

if strcmp(Parameter,'SAMPLEID');
    INFO = cellfun(@(x) (x{1}),regexp(files,'\\([^\s]*)\\d+\\pdata\\d+\\','tokens'));
    return;
end;

if strcmp(Parameter,'SR');
    SF = bbio_spec_info(RAWData, 'SF');
    BF = bbio_spec_info(RAWData, 'BF1');
    INFO = (SF-BF)*1E6;
    return;
end;

if strcmp(Parameter,'SWH');
    SW = bbio_spec_info(RAWData, 'SW');
    BF = bbio_spec_info(RAWData, 'BF1');
    O1 = bbio_spec_info(RAWData, 'O1');
    INFO = SW .* (BF + 1E-6*O1);
    return;
end;

```

function [ppm] = bbio_spec_ppm(Spec)

```

% results ppm axis for spectrum
step = Spec.sw / (Spec.size-1);
ppm = Spec.maxppm:-step:Spec.minppm;

```

function RAW = bbio_spec_read1d(FileList, varargin)

```

%
% -----
% ----- Bruker Biospin Toolbox -----
% -----

```

```
% bbio_spec_read1d = Read 1R-NMR files
%
% This function reads the spectra of all given filenames
% and returns a struct RAW with the Data and almost all
% available parameters
%
% INPUT:
%   FileList:
%       is a cell-array containing the filenames
%       of the files to be read (with 1r)
% Parameters
%   ppmRange: [min max]
%       is a 1x2-array defining the ppm-Range for
%       reading (optional)
%   imaginary: (true/false)
%       boolean, "true" for reading imaginary-part
%
% -----

% Dealing with the input
p = inputParser;
p.addParamValue('imaginary',false);
p.addParamValue('ppmRange',[-1E99 1E99]);
p.addParamValue('readShim',true);
p.addParamValue('readSCON',true);
p.parse(varargin{:});
PARS = p.Results;

if ischar(FileList)
    FileList = {FileList};
end;

n = length(FileList);
RAW = struct('Data',cell(n,1),'Imag',cell(n,1));
```

```
if n>=10
    fprintf('1D-Reading (%d spectra)...\n', n);
end;
for k=1:n

    cFileName = FileList{k};
    cDir      = cFileName(1:(end-2));

    if mod(k,100)==0 && k>2
        fprintf(' %d (%.02f%%)\n', k, 100*k/n);
    end;

    if mod(k,10)==1 && n>=10
        fprintf('.');
    end;

    if exist(cFileName, 'file')~=2
        fprintf('\nFile "%s" does not exist!\n', cFileName);
        continue;
    end;

    % loading jcamp-info files acqu, procs

    [ACQUS, ACQUSTXT] = bbio_internal_JCAMP_read( [cDir '..\..\acqus']);
    [PROCS, PROCSTXT] = bbio_internal_JCAMP_read( [cDir 'procs']);

    AUDITA = bbio_internal_loadtext([cDir '..\..\audita.txt']);
    AUDITP = bbio_internal_loadtext([cDir 'auditp.txt']);

    titlefile = [cDir 'title'];
    if exist(titlefile,'file')
        myTitle = strtrim(bbio_internal_loadtext(titlefile));
    else
        myTitle = '';
```

```

end;

RAW(k).TITLE = myTitle;
RAW(k).ACQUS = ACQUS;
RAW(k).PROCS = PROCS;
RAW(k).ACQUESTXT = ACQUESTXT;
RAW(k).PROCSTXT = PROCSTXT;
RAW(k).AUDITA = AUDITA;
RAW(k).AUDITP = AUDITP;

% save important parameters at top-level
if ~isempty(ACQUS)
    RAW(k).O1 = ACQUS.O1;
    RAW(k).RG = ACQUS.RG;
    RAW(k).PULPROG = ACQUS.PULPROG;
    RAW(k).BF1 = ACQUS.BF1;
    RAW(k).TD = ACQUS.TD;
    RAW(k).NS = ACQUS.NS;
    RAW(k).Date = bbio_internal_UnixToMatLabDate(ACQUS.DATE);

    % Real Aquisition Date with correct TimeZone
    v = regexp(RAW(k).AUDITA,'1,<[^\>]*>,<[^\>]*>,<[^\>]*>,<go>','tokens');
    if ~isempty(v)
        ds = v{1}{1};
        RAW(k).DateReal = datenum(ds(1:19),'yyyy-mm-dd HH:MM:SS');
        RAW(k).TimeZone = str2double(ds(25:27));
    else
        RAW(k).DateReal = NaN;
        RAW(k).TimeZone = NaN;
    end;

    % extract lists like pulses, powerlevels, delays
    [PULSES, PL, D, SPECT] = internal_add_customfields_acqus(ACQUS,...
        ACQUESTXT); %#ok<NASGU>

```

```
% getting the spectrometer is complicated (WRPA-Problem)
tmpTSPECT = regexp(AUDITA,'<([>]*)>,<go>','tokens');
if ~isempty(tmpTSPECT)
    SPECT = tmpTSPECT{1}{1};
else
    SPECT = '';
end;

RAW(k).PULSES = PULSES;
RAW(k).POWERLEVEL = PL;
RAW(k).D = D;
RAW(k).SPECT = SPECT;
% stripped 32 field version
RAW(k).PULSES32 = PULSES(1:32);
RAW(k).D32 = D(1:32);
RAW(k).POWERLEVEL32 = PL(1:32);

% if TopSpin 3.0 or higher -> different Power-Handling
[PLW, PLDB] = internal_new_powerlevel(ACQUS);
RAW(k).POWERLEVELW = PLW;
RAW(k).POWERLEVELDB = PLDB;
end;

RAW(k).ByteOrder = PROCS.BYTORDP;
if isfield(PROCS,'LB')
    RAW(k).LB = PROCS.LB;
end;
if isfield(PROCS,'PHC0')
    RAW(k).PHC0 = PROCS.PHC0;
end;
if isfield(PROCS,'PHC1')
    RAW(k).PHC1 = PROCS.PHC1;
end;
RAW(k).SF = PROCS.SF;
RAW(k).file = cFileName;
```

```
% Getting the TOPSPIN-Version
if ischar(ACQUESTXT)
    v = regexp(ACQUESTXT, '[^\n]*TOPSPIN\t\tVersion (\S+)', 'tokens');
    if ~isempty(v)
        v = v{1}{1};
        tmpVersion = str2double(v);
        if isnan(tmpVersion)
            % try first 3 characters
            tmpVersion = str2double(v(1:3));
        end;
        RAW(k).TOPSPIN = tmpVersion;
    else
        RAW(k).TOPSPIN = NaN;
    end;
end;

% reading the spectrum

[si, ofs, sw, real, ncproc, imag] = internal_read_spectrum(...
    cFileName, PARS.ppmRange(1), ...
    PARS.ppmRange(2), PROCS, ACQUS, PARS);
RAW(k).Data = real.*(2^(ncproc));
RAW(k).Imag = imag.*(2^(ncproc));
RAW(k).maxppm = ofs;
RAW(k).minppm = ofs-sw;
RAW(k).sw = sw;
RAW(k).size = si;

% Reading the shim
if PARS.readShim
    shimfile = [cDir '..\..\shimvalues'];
    if exist(shimfile, 'file')
```



```
        RAW(k).SHIM = internal_read_shimvalues( shimfile);
    else
        RAW(k).SHIM = [];
    end;
else
    RAW(k).SHIM = [];
end;

% Reading scon
if PARS.readSCON
    sconfile = [cDir '..\..\scon2'];
    if exist(sconfile,'file')
        RAW(k).SCON = bbio_internal_JCAMP_read( sconfile);
    else
        RAW(k).SCON = [];
    end;
else
    RAW(k).SCON = [];
end;

% ICON-INFO
iconfile = [cDir '..\..\order.txt'];
if exist(iconfile,'file')
    RAW(k).ICON = bbio_internal_JCAMP_read( iconfile);
else
    RAW(k).ICON = [];
end;

end;

if n>=10
    fprintf('\n');
end;
```

```
% -----
% -----
% -----
% -----
% -----
```

```
function SHIM = internal_read_shimvalues( shimfile)
```

```

SHIM = [];
L = bbio_internal_loadtext(shimfile);
SL = regexp(L,'.*?\n','tokens');
StartPos = 0;
for i=1:length(SL)
    s = strtrim(SL{i}{1});
    if ~strcmp(s(1),'#')
        StartPos = i;
        break;
    end;
end;
for i=StartPos:length(SL)
    s = strtrim(SL{i}{1});
    if strcmp(s(1),'#')
        break;
    end;
    tmpV = regexp(s,'(\S*)\s+(-*\d+)','tokens');
    if ~isempty(tmpV) && length(tmpV{1})==2
        sNAME = tmpV{1}{1};
        sVALUE = str2double(tmpV{1}{2});
        sNAME = strrep(sNAME,(' ');
        sNAME = strrep(sNAME,')','');
        sNAME = strrep(sNAME,'-', 'm');
        SHIM.(['SH_' sNAME]) = sVALUE;
    end;
end;

```

```

ADDPARS = {'LOCKPOWER','LOCKPHASE','LOCKGAIN','LOCKDC','LOCKSHIFT',...
           'FIELD','LOOPGAIN','LOOPTIME','LOOPFILTER'};

```

```

for i=1:length(ADDPARS)
    tmpT = regexp(L,[ADDPARS{i} '\s+(.*?)\n'],'tokens');
    SHIM.(ADDPARS{i}) = str2double(strtrim(tmpT{1}));
end;

```

```

function [PULSES, PL, D, SPECT] = internal_add_customfields_acqus( ACQUS, ...
                        ACQUESTXT)

```

```

SPECT = [];

```

```

P = ACQUS.P;
P1 = strrep(P,sprintf('\n'),' ');
v1 = regexp(P1,'(\S+)(\s*)','tokens');
PULSES = zeros(length(v1)-1,1);
for i=2:length(v1)
    PULSES(i-1) = str2double(v1{i}{1});
end;

```

```

P = ACQUS.PL;
P1 = strrep(P,sprintf('\n'),' ');
v1 = regexp(P1,'(\S+)(\s*)','tokens');
PL = zeros(length(v1)-1,1);
for i=2:length(v1)
    PL(i-1) = str2double(v1{i}{1});
end;

```

```

P = ACQUS.D;
P1 = strrep(P,sprintf('\n'),' ');
v1 = regexp(P1,'(\S+)(\s*)','tokens');
D = zeros(length(v1)-1,1);
for i=2:length(v1)

```

```
D(i-1) = str2double(v1{i}{1});
end;

v1 = regexp(ACQUSTXT,'##OWNER=([^\n]*)\n([^\n]*)\n','tokens','once');
if length(v1)==2
    lin = (v1{2});
    atpos = strfind(lin,'@');
    if ~isempty(atpos)
        SPECT = strtrim(lin((atpos(1)+1):end));
    end;
end;
```

```
function [PLW, PLDB] = internal_new_powerlevel(ACQUS)
```

```
PLW = zeros(64,1)*NaN;
PLDB = zeros(64,1)*NaN;

if isfield(ACQUS,'PLW')

    % TopSpin 3.0 or higher
    P = ACQUS.PLW;
    P1 = strrep(P,sprintf('\n'),' ');
    v1 = regexp(P1,'(\S+)(\s*)','tokens');
    PLW = zeros(length(v1)-1,1);
    for i=2:length(v1)
        PLW(i-1) = str2double(v1{i}{1});
    end;

    PLDB = - 10 * log10(PLW);
    PLDB(isinf(PLDB)) = 1000;
end;
```

```
function [si, ofs, sw, real, ncproc, imag] = internal_read_spectrum( aFile,...
```

minppm, maxppm, INFO, INFO2,...
PARS)

```
% byteorder?
dataformat = 'l';
if INFO.BYTORDP == 1
    dataformat = 'b';
end;

% get important parameters
ofs = INFO.OFFSET;
if ~isempty(INFO2)
    sw = INFO2.SW;
else
    sw = INFO.SW_p/INFO.SF;
end;
ncproc = INFO.NC_proc;
si = INFO.SI;

minppmindex = floor((ofs-minppm)/sw*(si-1));
maxppmindex = ceil((ofs-maxppm)/sw*(si-1));

if maxppmindex < 0;
    maxppmindex = 0;
    minppmindex = si-1;
end;

if ~exist(aFile,'file')
    si = 0;
    ofs = 0;
    sw = 0;
    real = NaN;
    ncproc = 0;
    return;
end;
```

```

count = minppmindex - maxppmindex + 1;
if count>0
    f = fopen (aFile, 'r',dataformat);
    fseek(f, maxppmindex*4,0);
    [real,si2]=fread(f,count, 'int32');
    fclose (f);
    if PARS.imaginary == true
        % loading the imaginary part
        iFile = strrep(aFile,'\1r','\1i');
        f = fopen (iFile, 'r',dataformat);
        fseek(f, maxppmindex*4,0);
        imag = fread(f,count, 'int32');
        fclose (f);
    else
        imag = [];
    end;
else
    real = [];
    si2 = 0;
end;
ofs = (ofs - sw/(si-1)*maxppmindex);
sw = (ofs - sw/(si-1)*maxppmindex) - (ofs - sw/(si-1)*minppmindex);
si = si2;

```

```

function [RAW,SCALE] = bbio_spec_scale( RAWIN, varargin)

```

```

%
% -----
% ----- Bruker Biospin Toolbox -----
% -----
% bbio_spec_scale( RAWIN, Parameters...) = Scaling 1D NMR Spectra
%
% INPUT:
%     RAWIN = Struct with RAW-Data
%     Parameters:

```

```

% mode: Mode for Scaling, Possible modes are:
% int = integral
% max = maximum
% min = minimum-baseline-scale
% custom = user provides scaling factor for each spectrum
% noise = scaling to a defined noise region
% ppm: PPM-Region for Scaling
% exclusions: Exclusion, nx2-Array with ppm-ranges
% noiseregion: PPM-Range for NOISE
% minHZ: Filter-Width for Min-Filtering (mode=min)
% custom: Customized Scaling with this ARRAY
%
% -----

```

```

% Dealing with the input

```

```

p = inputParser;
p.addParamValue('mode','int');
p.addParamValue('ppm',[0.5 4.5]);
p.addParamValue('noiseregion',[]);
p.addParamValue('minHZ',10);
p.addParamValue('custom',[]);
p.parse(varargin{:});
PARS = p.Results;

```

```

if nargin==0
    RAW=PARS;
    return;
end;

```

```

if strcmpi(PARS.mode,'int')
    SCALE = internal_int(RAWIN, PARS);
end;

```

```

if strcmpi(PARS.mode,'max')
    SCALE = internal_max(RAWIN, PARS);

```

```

end;

if strcmpi(PARS.mode,'min')
    SCALE = internal_min(RAWIN, PARS);
end;

if strcmpi(PARS.mode,'custom')
    SCALE = PARS.custom;
end;

if strcmpi(PARS.mode,'noise')
    SCALE = internal_noise(RAWIN, PARS);
end;

% Apply the Scaling
RAW = RAWIN;
for i=1:length(RAW)
    RAW(i).Data = RAW(i).Data ./ SCALE(i);
end;

% -----
% -----
% -----

function idx = getppmforuse(p, PARS)

idx = find(p>PARS.ppm(1) & p<PARS.ppm(2));
if ~isempty(PARS.exclusions)
    for i=1:size(PARS.exclusions,1)
        idx = setdiff(idx, find(p>PARS.exclusions(i,1) & ...
            p<PARS.exclusions(i,2)));
    end;
end;
end;

```



```
function SCALE = internal_int(RAWIN, PARS)
% Simple integration
SCALE = ones(length(RAWIN),1);
for i=1:length(RAWIN)
    p = bbio_spec_ppm(RAWIN(i));
    delta = abs(p(2)-p(1));

    idx = getppmforuse(p, PARS);

    S = sum(RAWIN(i).Data(idx));
    SCALE(i) = S*delta;
end;
```

```
function SCALE = internal_max(RAWIN, PARS)
% Maximum in region = 1
SCALE = ones(length(RAWIN),1);
for i=1:length(RAWIN)
    p = bbio_spec_ppm(RAWIN(i));

    idx = getppmforuse(p, PARS);

    S = max(RAWIN(i).Data(idx));
    SCALE(i) = S;
end;
```

```
function SCALE = internal_min(RAWIN, PARS)

SCALE = ones(length(RAWIN),1);
for i=1:length(RAWIN)
    p = bbio_spec_ppm(RAWIN(i));

    idx = (p>PARS.ppm(1) & p<PARS.ppm(2));

    STEP = PARS.minHZ/(abs(p(2)-p(1))*RAWIN(i).SF);
```

```
STEP = ceil(STEP);

MIN = bbio_internal_minbase( RAWIN(i).Data(idx), STEP,0);
SCALE(i) = median(MIN); % sum or median
end;

function SCALE = internal_noise(RAWIN, PARS)
% MAD of region = 1
SCALE = ones(length(RAWIN),1);
for i=1:length(RAWIN)
    p = bbio_spec_ppm(RAWIN(i));

    idx = getppmforuse(p, PARS);

    S = bbio_mad(RAWIN(i).Data(idx));
    SCALE(i) = S;
end;
```

function RAW = bbio_spec_shift1d(RAW, shift)

```
if length(RAW)>1
    for i=1:length(RAW)
        RAW(i) = bbio_spec_shift1d(RAW(i), shift(i));
    end;
    return;
end;

RAW.minppm = RAW.minppm - shift;
RAW.maxppm = RAW.maxppm - shift;
```

% function = complete_V_coffee

```
%complete central matrix V_coffee, change empty fields with 'unknown'
```

```
S = size(V_coffee);
```

```
% 'SampleID'
for i = 1:S(1);
    b(i) = strcmp("", V_coffee{i,2});
    if b(i)> 0;
        V_coffee{i,2} = 'unknown';
    end
end

for i = 1:S(1);
    t(i) = isempty(V_coffee{i,2});
    if t(i) > 0;
        V_coffee{i,2} = 'unknown';
    end
end

% 'ProjectID'
for i = 1:S(1);
    b(i) = strcmp("", V_coffee{i,3});
    if b(i)> 0;
        V_coffee{i,3} = 'unknown';
    end
end

for i = 1:S(1);
    t(i) = isempty(V_coffee{i,3});
    if t(i) > 0;
        V_coffee{i,3} = 'unknown';
    end
end

% 'Matrix'
for i = 1:S(1);
    b(i) = strcmp("", V_coffee{i,4});
    if b(i)> 0;
```

```
V_coffee{i,4} = 'unknown';
end
end

for i = 1:S(1);
    t(i) = isempty(V_coffee{i,4});
    if t(i) > 0;
        V_coffee{i,4} = 'unknown';
    end
end

% 'Kaffeeart'
for i = 1:S(1);
    b(i) = strcmp("", V_coffee{i,5});
    if b(i) > 0;
        V_coffee{i,5} = 'unknown';
    end
end

for i = 1:S(1);
    t(i) = isempty(V_coffee{i,5});
    if t(i) > 0;
        V_coffee{i,5} = 'unknown';
    end
end

% 'KaffeeKontinent'
for i = 1:S(1);
    b(i) = strcmp("", V_coffee{i,6});
    if b(i) > 0;
        V_coffee{i,6} = 'unknown';
    end
end

for i = 1:S(1);
```

```
t(i) = isempty(V_coffee{i,6});
    if t(i) > 0;
        V_coffee{i,6} = 'unknown';
    end
end
```

```
% 'KaffeeLand'
for i = 1:S(1);
    b(i) = strcmp('', V_coffee{i,7});
    if b(i) > 0;
        V_coffee{i,7} = 'unknown';
    end
end
```

```
for i = 1:S(1);
    t(i) = isempty(V_coffee{i,7});
    if t(i) > 0;
        V_coffee{i,7} = 'unknown';
    end
end
```

```
% 'KaffeeRegion'
for i = 1:S(1);
    b(i) = strcmp('', V_coffee{i,8});
    if b(i) > 0;
        V_coffee{i,8} = 'unknown';
    end
end
```

```
for i = 1:S(1);
    t(i) = isempty(V_coffee{i,8});
    if t(i) > 0;
        V_coffee{i,8} = 'unknown';
    end
end
```

```
end

% 'KaffeeAnbau'
for i = 1:S(1);
    b(i) = strcmp('', V_coffee{i,9});
    if b(i) > 0;
        V_coffee{i,9} = 'unknown';
    end
end

for i = 1:S(1);
    t(i) = isempty(V_coffee{i,9});
    if t(i) > 0;
        V_coffee{i,9} = 'unknown';
    end
end

end

% 'KaffeeFW'
for i = 1:S(1);
    b(i) = strcmp('', V_coffee{i,10});
    if b(i) > 0;
        V_coffee{i,10} = 'unknown';
    end
end

end

for i = 1:S(1);
    t(i) = isempty(V_coffee{i,10});
    if t(i) > 0;
        V_coffee{i,10} = 'unknown';
    end
end

end

% 'KaffeeFWTonino'
for i = 1:S(1);
```

```
b(i) = strcmp("", V_coffee{i,11});
    if b(i)> 0;
        V_coffee{i,11} = 'unknown';
    end
end

for i = 1:S(1);
    t(i) = isempty(V_coffee{i,11});
    if t(i) > 0;
        V_coffee{i,11} = 'unknown';
    end
end

% 'Kaffeeinwaage'
for i = 1:S(1);
    b(i) = strcmp("", V_coffee{i,12});
    if b(i)> 0;
        V_coffee{i,12} = 'unknown';
    end
end

for i = 1:S(1);
    t(i) = isempty(V_coffee{i,12});
    if t(i) > 0;
        V_coffee{i,12} = 'unknown';
    end
end

% 'Wassergehalt'
for i = 1:S(1);
    b(i) = strcmp("", V_coffee{i,13});
    if b(i)> 0;
        V_coffee{i,13} = 'unknown';
    end
end
```

```
end

for i = 1:S(1);
    t(i) = isempty(V_coffee{i,13});
    if t(i) > 0;
        V_coffee{i,13} = 'unknown';
    end
end

% 'Roestart'
for i = 1:S(1);
    b(i) = strcmp("", V_coffee{i,14});
    if b(i) > 0;
        V_coffee{i,14} = 'unknown';
    end
end

for i = 1:S(1);
    t(i) = isempty(V_coffee{i,14});
    if t(i) > 0;
        V_coffee{i,14} = 'unknown';
    end
end

clearvars b;
clearvars i;
clearvars t;
clearvars S;
```

```
function LDAAPPLY = lda_apply( DATA, LDA, Confidence)
```

```
LDAAPPLY = [];
```

```
MODELS = LDA.MODELS;
```



```

PREDICTED_GROUPS = NaN(size(DATA,1),length(MODELS));
SCORES = cell(length(MODELS),1);

HASH = bbio_hash( DATA);
INDIST = zeros(size(DATA,1),size(LDA.CONFUSION_MATRIX,1));
multsum = zeros(size(DATA,1),1);
for i=1:length(MODELS)

    MODEL = LDA.MODELS{i};

    % die gleichen schritte auf die Testdaten
    TDATA = DATA;
    TDATA = TDATA-repmat(MODEL.modelmean, size(TDATA,1),1);
    if ~isempty(MODEL.modelstd)
        TDATA = TDATA./repmat(MODEL.modelstd, size(TDATA,1),1);
    end;
    TDATA = TDATA*MODEL.PCAcoeff; % PCA anwenden
    TDATA = TDATA(:,1:MODEL.PCADIM)*MODEL.LDA.eigenvec;

    LDADIM = MODEL.LDADIM;
    % Berechne Distanzen der Testdaten zu den Modellmittelwerten
    TESTDIST = pdist2(TDATA(:,1:LDADIM),MODEL.GRPMODELMEANS);
    [~,MINDISTGROUP] = min(TESTDIST,[],2);

    ii = find(ismember(HASH, MODEL.HASH));
    mult = ones(size(TESTDIST));
    mult(ii,:) = 0;
    multsum = multsum + mult(:,1);
    INDIST = INDIST + TESTDIST .* mult;

    PREDICTED_GROUPS(:, i) = MINDISTGROUP;
    SCORES{i} = TDATA;

    ii = find(ismember(HASH, MODEL.HASH));

```

```

    PREDICTED_GROUPS(ii,i) = NaN;
end;
INDIST = INDIST ./ repmat(multsum,1,size(INDIST,2));

LDAAPPLY.PREDICTED_GROUPS = PREDICTED_GROUPS;
LDAAPPLY.SCORES      = SCORES;
LDAAPPLY.INDIST      = INDIST;

LDAAPPLY.CONF = LDAAPPLY.INDIST;
% Betrachtung relativ zu Confidence
for i=1:size(INDIST,2)
    di = LDA.INDISTANCES(LDA.ORIGINAL_GROUPS==i,i);
    l = quantile(di,Confidence);
    LDAAPPLY.CONF(:,i) = LDAAPPLY.CONF(:,i) / l;
end;

LDAAPPLY.PVAL = NaN(size(LDAAPPLY.INDIST));
for g=1:size(INDIST,2)
    di = LDA.INDISTANCES(LDA.ORIGINAL_GROUPS==g,g);
    for i=1:size(LDAAPPLY.PVAL,1)
        LDAAPPLY.PVAL(i,g) = length(find(di>LDAAPPLY.INDIST(i,g)))/length(di);
    end;
end;
LDAAPPLY.Confidence = Confidence;

```

```

function LDARESULT = lda_create(DATA, GROUPS, MONTE_CARLO, CROSS_SETS,
PCADIM, UseAutoScaling, RANDOM_SEED)

```

```

LDARESULT = [];

s = RandStream('mt19937ar','Seed',RANDOM_SEED);
RandStream.setGlobalStream(s);

T = tabulate(GROUPS);
n = size(DATA,1);

```

```
ORIGINAL_GROUPS = NaN(n,1);
for t=1:size(T,1)
    i1 = find(ismember(GROUPS,T{t,1}));
    ORIGINAL_GROUPS(i1)=t;
end

PREDICTED_GROUPS = NaN(n,MONTE_CARLO);
MODELS = cell(MONTE_CARLO*CROSS_SETS,1);

mc = 1;
% Monte-Carlo
for MC = 1:MONTE_CARLO
    fprintf('Monte-Carlo-Run %d\n', MC);
    % Kreuzvalidierung k-fache KV mit Monte-Carlo-permutierten
    % Datensätzen
    CV_SETS = internal_cv_sets(n, CROSS_SETS);
    for CV=1:CROSS_SETS

        testset = CV_SETS{CV};
        modelset = setdiff(1:n,testset);

        modelmean = mean(DATA(modelset,:));
        MDATA = DATA(modelset,:);

        % Erzeuge Hashes für die Modelldaten
        HASH = bbio_hash( MDATA);

        MDATA = MDATA-repmat(modelmean, size(MDATA,1),1);

        % insert other scaling
        % unit-variance scaling, autoscale
        if UseAutoScaling
            modelstd = std(DATA(modelset,:));
```

```

    MDATA = MDATA./repmat(modelstd, size(MDATA,1),1);
else
    modelstd = [];
end

PCAccoeff = pca(MDATA);
MDATA = MDATA*PCAccoeff; % PCA anwenden
MDATA = MDATA(:,1:PCADIM);

% LDA
[~,~,LDA] = manova1(MDATA,GROUPS(modelset));
MDATA = MDATA*LDA.eigenvec;

% die gleichen schritte auf die Testdaten
TDATA = DATA(testset,:);
TDATA = TDATA-repmat(modelmean, size(TDATA,1),1);
if UseAutoScaling
    TDATA = TDATA./repmat(modelstd, size(TDATA,1),1);
end
TDATA = TDATA*PCAccoeff; % PCA anwenden
TDATA = TDATA(:,1:PCADIM)*LDA.eigenvec;

LDADIM = size(T,1)-1;
GRPMODELMEANS = NaN(size(T,1), LDADIM);
for t=1:size(T,1)
    i1 = find(ismember(GROUPS(modelset),T{t,1}));
    GRPMODELMEANS(t,:) = mean(MDATA(i1,1:LDADIM));
end

% Berechne Distanzen der Testdaten zu den Modellmittelwerten
TESTDIST = pdist2(TDATA(:,1:LDADIM),GRPMODELMEANS);
[~,MINDISTGROUP] = min(TESTDIST,[],2);

PREDICTED_GROUPS(testset, MC) = MINDISTGROUP;

% Model speichern

```

```
MODELS{mc} = [];  
MODELS{mc}.modelmean = modelmean;  
MODELS{mc}.modelstd = modelstd;  
MODELS{mc}.PCAcoeff = PCAcoeff;  
MODELS{mc}.LDA = LDA;  
MODELS{mc}.GRPMODELMEANS = GRPMODELMEANS;  
MODELS{mc}.MDATA = MDATA;  
MODELS{mc}.TDATA = TDATA;  
MODELS{mc}.testset = testset;  
MODELS{mc}.modelset = modelset;  
MODELS{mc}.PCADIM = PCADIM;  
MODELS{mc}.LDADIM = LDADIM;  
MODELS{mc}.HASH = HASH;  
MODELS{mc}.TESTDIST = TESTDIST;  
  
mc = mc + 1;  
end  
  
end  
  
LDARESULT.MODELS = MODELS;  
LDARESULT.PREDICTED_GROUPS = PREDICTED_GROUPS;  
LDARESULT.ORIGINAL_GROUPS = ORIGINAL_GROUPS;  
LDARESULT.GROUP_LEGEND = T(:,1);  
  
LDARESULT.CORRECT_PREDICTION = NaN(n,1);  
for i=1:n  
    l = length(find(PREDICTED_GROUPS(i,:)==ORIGINAL_GROUPS(i)));  
    LDARESULT.CORRECT_PREDICTION(i) = l/size(PREDICTED_GROUPS,2);  
end  
  
CONFUSION_MATRIX = zeros(size(T,1),size(T,1));  
for m=1:MONTE_CARLO
```

```
for t=1:size(T,1)
    i1 = find(ismember(GROUPS,T{t,1}));
    for k=1:size(T,1)
        ii = find(PREDICTED_GROUPS(i1,m)==k);
        CONFUSION_MATRIX(t,k) = CONFUSION_MATRIX(t,k) + length(ii);
    end
end
end

for t=1:size(T,1)
    i1 = find(ismember(GROUPS,T{t,1}));
    CONFUSION_MATRIX(t,:) = CONFUSION_MATRIX(t,+)/length(i1)*100;
end
CONFUSION_MATRIX = CONFUSION_MATRIX/MONTE_CARLO;
LDARESULT.CONFUSION_MATRIX = CONFUSION_MATRIX;

% Sammeln der Distanz-Informationen
DISTANCES = zeros(size(DATA,1),size(T,1));
for i=1:length(LDARESULT.MODELS)
    M = LDARESULT.MODELS{i};
    M.TESTDIST
    whos DISTANCES
    length(M.testset)
    DISTANCES(M.testset,:) = DISTANCES(M.testset,) + M.TESTDIST;
end
DISTANCES = DISTANCES/MONTE_CARLO;

LDARESULT.INDISTANCES = DISTANCES;
```

function lda_visu_clouds(LDA, MODELNR, confalpha, ShowPoints)

```
if nargin<4
    ShowPoints = true;
end;
```

```

VMODEL = LDA.MODELS{MODELNR};
MDATA = VMODEL.MDATA;
TDATA = VMODEL.TDATA;

cla;
hold on;

lp=perms([-1 0 1]);
for i=1:4
    light('Position',lp(i,:),'Style','infinite');
end;

CM = jet(length(LDA.GROUP_LEGEND));
for t=1:length(LDA.GROUP_LEGEND)
    patch([0 0 0],[0 0 0],'k','EdgeColor','none','FaceColor',CM(t:),'FaceAlpha',0.3,
'FaceLighting','Phong');
end;
for t=1:length(LDA.GROUP_LEGEND)
    i1 = find(LDA.ORIGINAL_GROUPS(VMODEL.modelset)==t);
    if ShowPoints
        plot3(MDATA(i1,1),MDATA(i1,2),MDATA(i1,3),'ko','MarkerFaceColor',CM(t:),'Markersize',3)
    end;
    mp = bbio_conf3dellipsoid(MDATA(i1,1:3),confalpha,30);
    patch(mp,'EdgeColor','none','FaceColor',CM(t:),'FaceAlpha',0.3, 'Face-
Lighting','Phong');
end;

for t=1:length(LDA.GROUP_LEGEND)
    i1 = find(LDA.ORIGINAL_GROUPS(VMODEL.testset)==t);
    if ShowPoints
        plot3(TDATA(i1,1),TDATA(i1,2),TDATA(i1,3),'ks','MarkerFaceColor',CM(t:))
    end;
end;

```

```

box on;
grid on;
axis vis3d;
view(30,30);
legend(LDA.GROUP_LEGEND)
set(legend,'Location','best');

```

function lda_visu_confusion(LDA)

```

T = LDA.GROUP_LEGEND;
CONFUSION_MATRIX = LDA.CONFUSION_MATRIX;
ng = length(T);

cla;
set(gcf,'Color','W');
CM = flipud(gray(101));
for t=1:size(T,1)
    for k=1:size(T,1)
        c = floor(CONFUSION_MATRIX(t,k))+1;
        col = CM(c,:);
        patch([t-0.5 t+0.5 t+0.5 t-0.5],[k-0.5 k-0.5 k+0.5 k+0.5],col); %patch -> Funktion
        zeichnet Quadrate und unterteilt es in die einzelne Felder
        text(t,k, sprintf('%0.1f',CONFUSION_MATRIX(t,k)),'HorizontalAlignment','center',
        'Color',1-col);
    end;
end;
set(gca,'xlim',[0.5 size(T,1)+0.5]);
set(gca,'ylim',[0.5 size(T,1)+0.5]);
set(gca,'xtick',1:ng,'ytick',1:ng,'yticklabel',T,'xticklabel',T,'xticklabelrotation',45);

```


9.12 MATLAB®-Skript zur Quantifizierung ausgewählter Kaffee- metaboliten durch ¹H-NMR-Spektroskopie und PULCON

9.11.1 Verwendetes MATLAB®-Hauptskript

```
%% Quantifizierung wasserlöslicher Analyten in Kaffee
% NMP N-Methylpyridinium
% CQA Chlorogensäure quantifiziert als 5-Caffeoylchinasäure
% HMF 5-Hydroxymethylfurfural

%% Konstante Parameter
%Molare Masse
MW_HMF      = (126.11);
MW_Trigonellin = (137.14);
MW_NMP      = (94.134);
MW_Ameisensaure = (46.03);
MW_Coffein  = (194.19);
MW_Essigsäure = (60.052);
MW_Milchsäure = (90.08);
MW_CQA      = (354.31);
MW_Citronensäure = (192.12);

%Anzahl der Protonen fuer entsprechendes Signal
NH_HMF      = (1);
NH_Trigonellin = (1);
NH_NMP      = (1);
NH_Ameisensaure = (1);
NH_Coffein  = (3);
NH_Essigsäure = (3);
NH_Milchsäure = (3);
NH_CQA      = (1);
NH_Citronensäure = (2);

% ppm-Bereich des entsprechenden Signals
ppm_HMF      = [9.44 9.48];
ppm_Trigonellin = [9.08 9.17];
```

```
ppm_NMP      = [8.49 8.60];
ppm_Ameisensaure = [8.42 8.49];
ppm_Coffein   = [3.315 3.362];
ppm_Essigsaeure = [1.915 1.94];
ppm_Milchsaeure1 = [1.32 1.49];    % ppm Bereich für QuantRef
ppm_Milchsaeure = [1.315 1.36];
ppm_CQA       = [6.89 7.01];
ppm_Citronensaure = [2.920 3.026]; % ppm Bereich für QuantRef
```

% Korrekturfaktoren bezueglich T1 Zeit Bestimmung und HPLC-Vergleichsmessung

```
CF_HMF      = (1);
CF_Trigonellin = (1.1);
CF_NMP      = (1);
CF_Ameisensaure = (1.2);
CF_Coffein   = (1.2);    %Faktor berechnet anhand HPLC-Vergleichsmessung
CF_Essigsaeure = (1.1);
CF_Maleinsaure = (1);
CF_Milchsaeure = (1);
CF_CQA       = (1);
CF_CQA1      = (1);
CF_Citronensaure = (1);
```

V = 0.008; % Volumen der Extraktionslösung in L

E = 0.0002; % Einwaage der Kaffeesprobe in kg

VE = V/E;

SI = 262144;

DF = 0.78; % Verdünnungsfaktor da 600 µL Probe mit 100 µL Puffer und 70 µL TSP
(600/770 = 0.78)

ppm = [-1 10];

minHz = 15;

Path = 'N:\Praktikanten\Vera Gottstein\Kaffee_wässrig\Quantifizierung\Neue Daten';
%Datenpfad der Spektren

```
%% Importieren der spektralen Daten und Basislinienkorrektur

FileList = bbio_all_1r_files(Path, 10, 1);
mvm      = bbio_spec_read1d(FileList);
xc       = bbio_spec_filter_blk(mvm, minHz, ppm);

V_Coffee_H2O = cell(length(mvm),17);

%% Import der Metadaten in die Matrix V_Coffee_H2O

CoffeeID = bbio_spec_info(xc,'SAMPLEID');

for i=1:length(mvm)
    V_Coffee_H2O{i,1} = mvm(i).file;
    V_Coffee_H2O{i,2} = xc(i).PULSES(1);
    V_Coffee_H2O{i,3} = xc(i).ERETIC(1);
end

%% Importieren von SW (spektrale Breite) aus den spektralen Daten
SW = xc.sw;

%% Berechnung des ERETIC-Faktors aus der QuantRef

V_QR = cell(length(xc),2);

tf=find(cellfun('length',strfind({xc.file},'QuantRef_a')));

ppmi=bbio_spec_ppm(xc(tf));
i1 = (ppmi>ppm_Milchsaeure1(1) & ppmi<ppm_Milchsaeure1(2));
SubData = xc(tf).Data(i1);
Integral_Milchsaeure = trapz(SubData);
V_QR{i,1} = Integral_Milchsaeure;
```

```

ppmi=bbio_spec_ppm(xc(tf));
i1 = (ppmi>ppm_Citronensaeure(1) & ppmi<ppm_Citronensaeure(2));
SubData = xc(tf).Data(i1);
Integral_Citronensaeure = trapz(SubData);
V_QR{i,2} = Integral_Citronensaeure;

e(1,1) = (V_QR{i,1}*SW*MW_Milchsaeure)/(SI*923.04*NH_Milchsaeure*DF);
e(1,2) = (V_QR{i,2}*SW*MW_Citronensaeure)/(SI*4236.27*NH_Citronensaeure*DF);
eretic = (e(1,1) + e(1,2))/2;
eretic = round(eretic,2);
Pulse = xc(tf).PULSES(1);

%% Integration des HMF-Siganls mittels Trapez-Integration
for i=1:length(xc)
    ppmi=bbio_spec_ppm(xc(i));
    i1 = (ppmi>ppm_HMF(1) & ppmi<ppm_HMF(2));
    SubData = xc(i).Data(i1);
    Integral = trapz(SubData);
    V_Coffee_H2O{i,4} = Integral;
end

%% Integration des Trigonellin-Siganls mittels Trapez-Integration
for i=1:length(xc)
    ppmi=bbio_spec_ppm(xc(i));
    i1 = (ppmi>ppm_Trigonellin(1) & ppmi<ppm_Trigonellin(2));
    SubData = xc(i).Data(i1);
    Integral = trapz(SubData);
    V_Coffee_H2O{i,5} = Integral;
end

%% Integration des NMP-Siganls mittels Trapez-Integration
for i=1:length(xc)
    ppmi=bbio_spec_ppm(xc(i));
    i1 = (ppmi>ppm_NMP(1) & ppmi<ppm_NMP(2));

```

```
SubData = xc(i).Data(i1);
Integral = trapz(SubData);
V_Coffee_H2O{i,6} = Integral;
end

%% Integration des Ameisensäure-Siganls mittels Trapez-Integration
for i=1:length(xc)
    ppmi=bbio_spec_ppm(xc(i));
    i1 = (ppmi>ppm_Ameisensaure(1) & ppmi<ppm_Ameisensaure(2));
    SubData = xc(i).Data(i1);
    Integral = trapz(SubData);
    V_Coffee_H2O{i,7} = Integral;
end

%% Integration des Koffein-Siganls mittels line-fit Algorithmus

V_Coffee_H2O{i,8} = NaN(length(xc),1);

for i = 1:length(xc)

    CY = xc(i);

    P = bbio_sig_FitSignal();
    P.TOLFUN = 1E-6;
    P.BADMIN = -0.01;
    P.BADMINF = 80;
    P.MinWidth = 0.6;
    P.MaxWidth = 3;
    P.UseExponents = false;
    P.DoPlot = true;

    %Methoden Basislinienkorrektur
    P.SubstractBaseLine = true;    %alter Wert: false
    P.DoMinFiltering = false;
    %P.MinFilterWidth = 20;
```

```
P.AdjustBaseLine = false;    %alert Wert: true

%Beschreibung Signalform
P.SIG_ppm = ppm_Coffein;
P.SIG_center = 3.344;
P.SIG_shifftol = 0.05;
P.SIG_coupltol = 0;
P.SIG_factors = [1];
P.SIG_couplings = [0];

FIT = bbio_sig_FitSignal(CY,P);
INT = FIT.INTEGRAL;
V_Coffee_H2O{i,8} = INT
end

%% Integration des Essigsäure-Siganls mittels line-fit Algorithmus
V_Coffee_H2O{i,9} = NaN(length(xc),1);

for i = 1:length(xc)

    CY = xc(i);

    P = bbio_sig_FitSignal();
    P.TOLFUN = 1E-8;
    P.BADMIN = -0.005;
    P.BADMINF = 10;
    P.MinWidth = 0.5;
    P.MaxWidth = 4;
    P.UseExponents = false;
    P.DoPlot = true;
    % P.MinGauss = 0.5;
    % P.MaxGauss = 0.9;
```

```
%Methoden Basislinienkorrektur
P.SubstractBaseLine = false;
P.DoMinFiltering = false;
P.MinFilterWidth = 20;
P.AdjustBaseLine = true;

%Beschreibung Signalform
P.SIG_ppm = ppm_Essigsaeure;
P.SIG_center = 1.9273;
P.SIG_shifftol = 0.05;
P.SIG_coupltol = 0;
P.SIG_factors = [1];
P.SIG_couplings = [0];

FIT = bbio_sig_FitSignal(CY,P);
INT = FIT.INTEGRAL;
V_Coffee_H2O{i,9} = INT;

end

%% Integration des Milchsäure-Siganls mittels line-fit Algorithmus
V_Coffee_H2O{i,10} = NaN(length(xc),1);

for i = 1:length(xc)

    CY = xc(i);

    P = bbio_sig_FitSignal();
    P.TOLFUN = 1E-6;
    P.BADMIN = -0.001;
    P.BADMINF = 80;
    P.MinWidth = 0.5;
    P.MaxWidth = 3;
    P.UseExponents = false;
    P.DoPlot = true;
```

```
%Methoden Basislinienkorrektur
P.SubstractBaseLine = true;
P.DoMinFiltering = false;
P.MinFilterWidth = 20;
P.AdjustBaseLine = false;

%Beschreibung Signalform
P.SIG_ppm = ppm_Milchsaeure;
P.SIG_center = 1.3349;
P.SIG_shifftol = 0.02;
P.SIG_coupltol = 0.2;
P.SIG_factors = [1.02 1.00];
P.SIG_couplings = [3.48 -3.48];

FIT = bbio_sig_FitSignal(CY,P);
INT = FIT.INTEGRAL;
V_Coffee_H2O{i,10} = INT
end

%% Integration des CQA-Siganls mittels line-fit Algorithmus
V_Coffee_H2O{i,11} = NaN(length(xc),1);

for i = 1:length(xc)

    CY = xc(i);

    P = bbio_sig_FitSignal();
    P.TOLFUN = 1E-8; % Genauigkeit des Fits: Je größer der
Wert desto genauer des Fit jedoch wird auch mehr Zeit für Fit benötigt
    P.BADMIN = -0.00025;
    P.BADMINF = 80; % Fit nach Least Squares. Je höher der
Wert, desto unwahrscheinlicher überschreitet Fit die die Linien des "wahren" Signas
    P.MinWidth = 4;
    P.MaxWidth = 6;
```

```

P.UseExponents = false;           % Spllte immer deaktiviert (=false) sein
P.DoPlot = true;                  % Fit wird als Abbildung dargestellt

%Methoden Basislinienkorrektur
P.SubstractBaseLine = true;       % Basislinie wird von dem Fir des
Signals abgezogen
P.DoMinFiltering = false;         % ermittelt Untergrund mit einem
Fenster, welches so groß ist wie MinFilterWidth
P.MinFilterWidth = 8;            % Angabe in Hz
P.AdjustBaseLine = false;        % Basislinie wird mitgefitet (wichtig,
wenn Basislinie offset hat)

%Beschreibung Signalform
P.SIG_ppm = ppm_CQA;
P.SIG_center = 6.9425;           % Mitte des Signals
P.SIG_shifftol = 0.02;          % Toleranz für Signalshift in Hz
P.SIG_coupltol = 0.5;           % Toleranz für Kopplungskonstanze in
Hz
P.SIG_factors = [1.4 2.3 0.8 2.0]; % Beschreibt Signalform [1 1] -
-> Duplett
P.SIG_couplings = [5.2 3.9 -2.7 -4.0]; % Koplungskonstante/2
ausgehend von der Mitte des Signals

FIT = bbio_sig_FitSignal(CY,P);
INT = FIT.INTEGRAL;
V_Coffee_H2O{i,11} = INT
end

%% Berechnung der Halbwertbreite des TSP-Signals

ResultList_TSP = cell(length(FileList)+1,3);
ResultList_TSP{1,1} = 'Datei';
ResultList_TSP{1,2} = 'Integral';
ResultList_TSP{1,3} = 'Error';
ResultList_TSP{1,4} = 'Position';

```

```
ResultList_TSP{1,5} = 'LineWidth';
```

```
ResultList_TSP{1,6} = 'P1';
```

```
for i=1:length(FileList)
```

```
    figure(1);
```

```
    clf;
```

```
    [Integral, Error, Position, LineWidth, SPEC] = FitTSP(FileList{i});
```

```
    ResultList_TSP{i+1,1} = FileList{i};
```

```
    ResultList_TSP{i+1,2} = Integral;
```

```
    V_Coffee_H2O{i,12} = LineWidth;
```

```
    ResultList_TSP{i+1,3} = sprintf('%.2f',Error*100);
```

```
    ResultList_TSP{i+1,4} = sprintf('%.4f', Position);
```

```
    ResultList_TSP{i+1,5} = sprintf('%.2f',LineWidth);
```

```
    ResultList_TSP{i+1,6} = sprintf('%.2f',SPEC.PULSES32(2));
```

```
    drawnow;
```

```
end
```

```
ResultList_TSP;
```

```
%% Berechnung der Konzentration in mg/kg
```

```
for i=1:length(mvm)
```

```
    C(i,1) = ((V_Coffee_H2O{i,4}*SW*MW_HMF*xc(i).PULSES(1))/(SI*eretic*NH_HMF*DF*Pulse))*CF_HMF*VE;
```

```
    C(i,2) = ((V_Coffee_H2O{i,5}*SW*MW_Trigonellin*xc(i).PULSES(1))/(SI*eretic*NH_Trigonellin*DF*Pulse))*CF_Trigonellin*VE;
```

```
    C(i,3) = ((V_Coffee_H2O{i,6}*SW*MW_NMP*xc(i).PULSES(1))/(SI*eretic*NH_NMP*DF*Pulse))*CF_NMP*VE;
```

```
    C(i,4) = ((V_Coffee_H2O{i,7}*SW*MW_Ameisensaure*xc(i).PULSES(1))/(SI*eretic*NH_Ameisensaure*DF*Pulse))*CF_Ameisensaure*VE;
```

```
    C(i,5) = ((V_Coffee_H2O{i,8}*MW_Coffein*xc(i).PULSES(1))/(eretic*NH_Coffein*DF*Pulse))*CF_Coffein*VE;
```

```
    C(i,6) = ((V_Coffee_H2O{i,9}*MW_Essigsaeure*xc(i).PULSES(1))/(eretic*NH_Essigsaeure*DF*Pulse))*CF_Essigsaeure*VE;
```

```
    C(i,7) = ((V_Coffee_H2O{i,10}*MW_Milchsaeure*xc(i).PULSES(1))/(eretic*NH_Milchsaeure*DF*Pulse))*CF_Milchsaeure*VE;
```

```
C(i,8) = ((V_Coffee_H2O{i,11}*MW_CQA*xc(i).PULSES(1))/(eretic*NH_CQA*DF*Pulse))*CF_CQA*VE;
```

```
end
```

```
%% Anwendung der Bestimmungsgrenze
```

```
% HMF
```

```
for i=1:length(mvm)
```

```
if C(i, 1) < 0.22
```

```
    C(i, 1)=0;
```

```
end
```

```
end
```

```
% Trigonellin
```

```
for i=1:length(mvm)
```

```
if C(i, 2) < 1.64
```

```
    C(i, 2)=0;
```

```
end
```

```
end
```

```
% NMP
```

```
for i=1:length(mvm)
```

```
if C(i, 3) < 0.77
```

```
    C(i, 3)=0;
```

```
end
```

```
end
```

```
% Ameisensäure
```

```
for i=1:length(mvm)
```

```
if C(i, 4) < 1.27
```

```
    C(i, 4)=0;
```

```
end
```

```
end
```

```
% Koffein
for i=1:length(mvm)
    if C(i, 5) < 6.6
        C(i, 5)=0;
    end
end

% Essigsäure
for i=1:length(mvm)
    if C(i, 6) < 6.6
        C(i, 6)=0;
    end
end

% Milchsäure
for i=1:length(mvm)
    if C(i, 7) < 0.43
        C(i, 7)=0;
    end
end

% CQA
for i=1:length(mvm)
    if C(i, 8) < 4.87
        C(i, 8)=0;
    end
end

%% Excel Report
cd 'N:\Praktikanten\Vera Gottstein\Kaffee_wässrig\Quantifizierung'
filename = [datestr(now,'dd-mm-yyyy'),'_Coffee_H2O_MF.xlsx'];

A= {'HMF','Trigonellin','N Methylpyridinium','Ameisensäure','Coffein','Essigsäure',
'Milchsäure','Chlorogensäure'};
A1= {'mg/kg','mg/kg','mg/kg','mg/kg','mg/kg','mg/kg','mg/kg','mg/kg'};
```

```
xlswrite(filename,{'Kaffee waessrig'},1,'A1');
```

```
xlswrite(filename,CoffeeID,1, 'A3');
```

```
xlswrite(filename,C,1, 'B3');
```

```
xlswrite(filename,A1,1, 'B2');
```

```
xlswrite(filename,A,1, 'B1');
```

```
xlswrite(filename,{'eretic'},1,'N1');
```

```
xlswrite(filename,eretic,1,'N2');
```

```
xlswrite(filename,{'ppm*L/mmol'},1,'O2');
```

```
xlswrite(filename,{'Integrals'},2,'A1');
```

```
xlswrite(filename,CoffeeID,2,'A3');
```

```
for i=1:length(mvm)
```

```
    Int(i,1) = V_Coffee_H2O{i,4};
```

```
    Int(i,2) = V_Coffee_H2O{i,5};
```

```
    Int(i,3) = V_Coffee_H2O{i,6};
```

```
    Int(i,4) = V_Coffee_H2O{i,7};
```

```
    Int(i,5) = V_Coffee_H2O{i,8};
```

```
    Int(i,6) = V_Coffee_H2O{i,9};
```

```
    Int(i,7) = V_Coffee_H2O{i,10};
```

```
    Int(i,8) = V_Coffee_H2O{i,11};
```

```
end
```

```
xlswrite(filename,Int,2,'B3');
```

```
xlswrite(filename,A,2,'B1');
```

```
for i=1:length(mvm)
```

```
    TSP(i,1) = round(V_Coffee_H2O{i,12},2); % TSPhalf width
```

```
end
```

```
xlswrite(filename,TSP,1, 'P3');
```

```
xlswrite(filename,{'TSP half width'},1,'P1');
xlswrite(filename,{'Hz'},1,'P2')

for i=1:length(mvm)
    ER(i,1) = round(V_Coffee_H2O{i,3},2); % ERETIC
end

xlswrite(filename,ER,1, 'Q3');
xlswrite(filename,{'ERETIC'},1,'Q1');
xlswrite(filename,{' '},1,'Q2')
```

9.11.2 Verwendete MATLAB®-Unterskripte

```
function FILELIST = bbio_all_1r_files(START_DIR, ExpNo, ProcNo)
```

```
if nargin<2
    ExpNo = [];
    ProcNo = [];
end;

FILELIST = recursive_reading(START_DIR, ExpNo, ProcNo);
```

```
function FILELIST = recursive_reading(BASEDIR, ExpNo, ProcNo)
```

```
FILELIST = {};

D = dir(BASEDIR);

for i=1:length(D)
    s = D(i).name;
    c = [BASEDIR '\' s];
    if strcmp(s, '.') || strcmp(s, '..')
        continue;
    end;
```

```
if isdir(c)
    tmpL = recursive_reading(c, ExpNo, ProcNo);
    FILELIST = union(FILELIST, tmpL);
end;
if strcmp(s,'1r') && (isempty(findstr(c,'\99999\pdata\')))

    c = strrep(c, '/', '\');
    %c = strrep(c, '\\', '\');

    % check the filter
    try
        tokens = regexp(c, '(\d+)\p{data}\p{data}(\d+)\1r', 'tokens');
        tokens = tokens{1};
        E = str2double(tokens{1});
        P = str2double(tokens{2});
    catch %#ok<CTCH>
        continue;
    end;

    if ~isempty(ExpNo)
        if (E~=ExpNo) || (P~=ProcNo)
            continue;
        end;
    end;

    FILELIST = union(FILELIST, c);
end;
end;
```

function FITSTRUCT = bbio_sig_FitSignal(SPEC, varargin)

```
% Dealing with the input
p = inputParser;
p.addParamValue('MSTEPS', 15000);
```

```

p.addParamValue('TOLFUN', 1e-8);
p.addParamValue('BADMIN', -0.025);
p.addParamValue('BADMINF', 5);
p.addParamValue('MinGauss',0);
p.addParamValue('MaxGauss',1);
p.addParamValue('MinWidth',0.5);
p.addParamValue('MaxWidth',2);
p.addParamValue('StartGauss',0.5);
p.addParamValue('StartWidth',1);
p.addParamValue('UseExponents',true);
p.addParamValue('StartExp',1);
p.addParamValue('MinExp', 0.75);
p.addParamValue('MaxExp', 1.5);
p.addParamValue('DoPlot', false);    % Boolean
p.addParamValue('DoMinFiltering', false); % Boolean
p.addParamValue('MinFilterWidth', 20); % Hz
p.addParamValue('SubstractBaseLine', true); % Boolean
p.addParamValue('InterpolateFactor', 1); % Double (>=1)
p.addParamValue('AdjustBaseLine', false); % Boolean

% Parameters for Signal (example is Sucrose in juice)
p.addParamValue('SIG_ppm', [5.38 5.42]); %ppm
p.addParamValue('SIG_center', 5.403); %ppm
p.addParamValue('SIG_shifftol', []); % when empty use the whole region,ppm
p.addParamValue('SIG_coupltol', 1); % Hz
p.addParamValue('SIG_factors', [1 1]);
p.addParamValue('SIG_couplings', [1.92 -1.92]); %Hz

p.addParamValue('IsMinor', false);
p.addParamValue('MinorMaxSigZero', 2);
p.addParamValue('IterStartPos', 50);
p.addParamValue('ApplyLB', 0);

p.parse(varargin{:});
PARS = p.Results;

```



```
if nargin==0
    FITSTRUCT = PARS;
    return;
end;

FITSTRUCT = [];

VISUREG = PARS.SIG_ppm;
MINREG = VISUREG;
CENTER = PARS.SIG_center;
SHIFTTOL = PARS.SIG_shifftol;
COUPLTOL = PARS.SIG_coupltol;
FACTORS = PARS.SIG_factors;
COUPLS = PARS.SIG_couplings;

FITPARS = [];
FITPARS.FIELD = SPEC.SF;
FITPARS.USEEXP = PARS.UseExponents;
FITPARS.BADMIN = PARS.BADMIN;
FITPARS.BADMINFACTOR = PARS.BADMINF;
FITPARS.AdjustBaseLine = PARS.AdjustBaseLine;
FITPARS.IsMinor = PARS.IsMinor;
FITPARS.MinorMaxSigZero = PARS.MinorMaxSigZero;

ppm = bbio_spec_ppm(SPEC);

if PARS.ApplyLB ~= 0
    SPEC = bbio_spec_ApplyLB(SPEC, PARS.ApplyLB);
end;

% Min-Filtering?
```

```

if PARS.DoMinFiltering==1
    SPEC = bbio_spec_filter(SPEC,'filter','min','minHz',PARS.MinFilterWidth,...
        'ppm',[-3 11]);
    DATA = SPEC.Data(:);
    NOISE = median(DATA(ppm<-1 & ppm>-2)); %#ok<UDIM>
    SPEC.Data = SPEC.Data - NOISE;
end;

% Preparing the data
DATA = SPEC.Data(:);
idx = find(ppm>VISUREG(1) & ppm<VISUREG(2));

% Center of Signal
if isempty(CENTER)
    CENTER = mean(VISUREG);
end;

% Subtract constant Baseline
if PARS.SubstractBaseLine
    idxmin = ppm>MINREG(1) & ppm<MINREG(2);
    MMIN = min(DATA(idxmin));

    DATA = DATA - MMIN;
    % adjust for noise
    %NOISE = bbio_mad(DATA(ppm<-1 & ppm>-1.5))*2;
    %DATA = DATA - NOISE;
end;

MMAX = max(DATA(idx));
DATA = DATA./MMAX;

mDATA = DATA(idx);

```

```

mppm = ppm(idx);

if PARS.InterpolateFactor~=1
    s = (length(mDATA)-1)*PARS.InterpolateFactor;
    d = (max(mppm)-min(mppm))/s;

    nppm = min(mppm):d:max(mppm);
    mDATA = interp1(mppm,mDATA,nppm,'pchip');
    mppm = nppm;
end;

% Starting Vector
START = [ PARS.StartWidth ... % Width
         PARS.StartGauss ... % Gauss
         ];

% Using Exponents?
if FITPARS.USEEXP
    START = [START PARS.StartExp];
end;

% Adjust Baseline
StartBaseOffset = 0; %#ok<NASGU>
if FITPARS.AdjustBaseLine
    StartBaseOffset = min(mDATA);
    START = [ START StartBaseOffset 0];
end;

% Adding Null-Positions
index_position = length(START)+1;
START = [START 0 zeros(1,length(FACTORS))];

searchOptions = struct(...
    'Display','on',...

```

```
'MaxIter',PARS.MSTEPS,...
'MaxFunEvals',PARS.MSTEPS,...
'TolX',PARS.TOLFUN,...
'TolFun',PARS.TOLFUN,...
'FunValCheck','off',...
'OutputFcn',[]];

POS = [CENTER COUPLS(:)'];

LOW = [PARS.MinWidth PARS.MinGauss];
UP = [PARS.MaxWidth PARS.MaxGauss];

if FITPARS.USEEXP
    LOW = [LOW PARS.MinExp ];
    UP = [UP PARS.MaxExp ];
end;

if FITPARS.AdjustBaseLine
    LOW = [LOW 0 -0];
    UP = [UP 1 0];
end;

% Calculating Shift-Tolerances
if isempty(SHIFTTOL)
    M1 = max(VISUREG)-CENTER;
    M2 = -(CENTER-min(VISUREG));
else
    M1 = SHIFTTOL;
    M2 = -SHIFTTOL;
end;

LOW = [LOW M2 ones(1,length(COUPLS))*(-COUPLTOL)];
UP = [UP M1 ones(1,length(COUPLS))*(COUPLTOL)];

% Finding the optimal starting position
```

```

% -----
optpos = M2:(M1-M2)/PARS.IterStartPos:M1;
if ~isempty(optpos)
    optval = zeros(length(optpos),1);
    for k=1:length(optpos)
        START(index_position) = optpos(k);
        d = BuildSignal(START, POS, FACTORS, mppm, mData, FITPARS);
        optval(k) = d;
    end;
    [optvalue,optindex] = min(optval); %#ok<ASGLU>
    % No we have optimized the starting value
    START(index_position) = optpos(optindex);
end;
[d, ystart] = BuildSignal(START, POS, FACTORS, mppm, mData, FITPARS);
%#ok<ASGLU,NASGU>

% Calculation
END = fminsearchbnd(@(x, varargin) ...
    BuildSignal(x, POS, FACTORS, mppm, mData, FITPARS),...
    START, LOW, UP, searchOptions);

% Integral
[FINERROR, y, ENDPAR, BASELINE, SCALE] = BuildSignal(END, POS,...
    FACTORS, mppm, mData, FITPARS);

% Build complete signal over large ppm-area
delta = abs(mppm(2)-mppm(1));
tppm = mean(ENDPAR.Centers) + (-0.25:delta:0.25);
iSIGNAL = bbio_sig_BuildNMRSignal( tppm, ENDPAR);
iINTEGRAL = sum(iSIGNAL*SCALE)*MMAX*abs(tppm(2)-tppm(1));

%iINTEGRAL = sum(y)*MMAX*abs(mppm(2)-mppm(1));

FITSTRUCT.ENDERROR = FINERROR;

```

```

FITSTRUCT.ENDPAR = ENDPAR;
FITSTRUCT.FIT    = y;
FITSTRUCT.IFACTOR = MMAX;
FITSTRUCT.ODATA  = mData;
FITSTRUCT.PPM    = mppm;
FITSTRUCT.INTEGRAL = iINTEGRAL;
FITSTRUCT.PARS   = PARS;
FITSTRUCT.BASELINE = BASELINE;

if PARS.DoPlot
    cla;
    hold on;
    plot(ppm(idx),DATA(idx),'k--','LineWidth',1);
    plot(mppm,mDATA,'k','LineWidth',1);
    %plot(mppm,ystart + StartBaseOffset,'m--');

    FinalSignal = y + BASELINE.OFFSET + (1:length(y))*BASELINE.SLOPE/length(y);

    plot(mppm,FinalSignal,'b--');
    %mf = min([FinalSignal(1) FinalSignal(end)]);
    fy = [FinalSignal 0 0];
    patch([mppm mppm(end) mppm(1)], fy , zeros(size(fy))-1,'b','FaceAlpha',1,...
        'EdgeColor','none','FaceColor',[0.75 0.75 1]);
    %plot(mppm,mDATA-y,'r-');

    s = sprintf('%s: e=%f',cell2mat(bbio_spec_info(SPEC,'SAMPLEID')), FINERROR);

    title(s,'interpreter','none');
    set(gca,'xdir','reverse');
    set(gca,'ylim',[-0.1 1.2]);
    set(gca,'xlim',[min(mppm) max(mppm)]);
    grid on;
    box on;
    drawnow;
end

```

```
function [D,y, SPAR, BASELINE, SCALE] = BuildSignal(X, POS, FACTORS, ppm, Signal,
P)
```

```
    % Converting Parameters from X
```

```
    Width = X(1);
```

```
    Gauss = X(2);
```

```
    ci = 3;
```

```
    EXPL = 1;
```

```
    SLOPE = 0;
```

```
    BL = 0;
```

```
    if P.USEEXP
```

```
        EXPL = X(3);
```

```
        ci = 4;
```

```
    end;
```

```
    if P.AdjustBaseLine
```

```
        BL = X(ci);
```

```
        SLOPE = X(ci+1);
```

```
        ci = ci + 2;
```

```
    end;
```

```
    PT = X(ci:end);
```

```
    POSIS = zeros(1,length(PT)-1);
```

```
    for i=1:length(POSI)
```

```
        POSIS(i) = POS(1) + POS(i+1)/P.FIELD + PT(1) + PT(i+1)/P.FIELD;
```

```
    end;
```

```
    SPAR = bbio_sig_BuildNMRSignal;
```

```
    SPAR.Width = Width;
```

```
    SPAR.Gauss = Gauss;
```

```

SPAR.ExpL = EXPL;
SPAR.Centers = POSIS;
SPAR.Factors = FACTORS;
SPAR.Field = P.FIELD;
SIGNAL = bbio_sig_BuildNMRSignal( ppm, SPAR);

```

```

BASELINE = [];
BASELINE.OFFSET = BL;
BASELINE.SLOPE = SLOPE;

```

```

y = SIGNAL;
Signal = Signal - BL - (1:length(y))*SLOPE/length(y);

```

```

% Adjust for largest Peak

```

```

SV = interp1(ppm,Signal,POISIS);
RV = interp1(ppm,y,POISIS);
SV = (SV./RV);
SCALE = min(SV);
if SCALE<0 || ~isempty(find(isnan(SV),1))
    SCALE = 0;
end;
y = y.*SCALE;

```

```

% Calculate Error

```

```

%smean = mean(SPAR.Centers);
%idx = ppm>=(smean-0.1) & ppm<=(smean+0.1);
%y(~idx) = 0;

```

```

ERROR = Signal(:)-y(:);

```

```

BADMIN = P.BADMIN;

```

```

% Dealing with minor signals

```

```

if P.IsMinor

```

```

    BADMIN = BADMIN * max(y); % Adjust to real Signal height
    igr = Signal(:)>max(y(:))*P.MinorMaxSigZero;
    ERROR(igr) = ERROR(igr)/10;
end;

% Old Version
im1 = (ERROR<BADMIN);
im2 = (ERROR>=BADMIN);

D = sum(ERROR(im1).^2)*P.BADMINFACTOR + sum(ERROR(im2).^2);

% New Version
%EFactor = ones(size(ERROR));
%idx = find(ERROR<BADMIN);
%EFactor(idx) = 1 + abs( P.BADMINFACTOR*(ERROR(idx)-BADMIN));
%EFactor(Signal<0) = 0;
%D = sum((ERROR.*EFactor).^2);

D = D./length(ERROR);

```

%Anhand der Bereiche -2ppm bis -1ppm und 11ppm bis 12ppm den Median der
%Basislinie bilden und (als Bias) vom Gesamtspektrum subtrahieren

function RAW = bbio_spec_filter_blk(RAWIN, minHz, ppm)

```
RAW = internal_min(RAWIN, minHz, ppm);
```

```
function RAW = internal_min(RAWIN, minHz, ppm)
```

```
RAW = RAWIN;
```

```
minHZ = abs(minHz);
```

```
for i=1:length(RAWIN)
```

```

p = bbio_spec_ppm(RAWIN(i));

for j=1:size(ppm,1)

    idx = (p>ppm(j,1) & p<ppm(j,2));

    STEP = minHZ/(abs(p(2)-p(1))*RAWIN(i).SF);
    STEP = ceil(STEP);

    D = RAWIN(i).Data(idx);
    MIN = calcminbase( D, STEP);
    DNew = D - MIN;
    if minHZ<0
        DNew = MIN;
    end;
    RAW(i).Data(idx) = DNew;
end;
end;

idx1 = (p>-2 & p<-1);
idx2 = (p>11 & p<12);
idx = logical(sum([idx1;idx2],1)); %Indexvektor mit Indices für Position -2 bis -1 ppm
und 11 bis 12 ppm als Logical

```

```

for i = 1:length(RAWIN)
    Medi = median(RAW(i).Data(idx));
    RAW(i).Data = RAW(i).Data - Medi;
end

```

```

function MIN = calcminbase(X, STEP)

```

```

MIN = zeros(length(X),1);

```

```

for i=(STEP+1):length(X)-STEP
    MIN(i) = min(X(i-STEP:i+STEP));
end;

```

```

function INFO = bbio_spec_info(RAWData, Parameter)

```

```

INFO = [];

```

```

DirectNumbers = {'LB','SI','Date','O1','PHC0','PHC1','SF','RG','NS','TD','BF1','SW'};

```

```

Arrays        = {'PULSES','DELAYS','POWERLEVELS', 'POWERLEVELSW', 'POW-
ERLEVELSDB'};

```

```

Singles       = {'P1','D1','SR','EXPNO','PROCNO', 'SAMPLEID','PROBHD'};

```

```

DirectStrings = {'PULPROG','SPECT','TITLE','file'};

```

```

if nargin==0

```

```

    INFO = [];

```

```

    INFO = union(INFO, Arrays);

```

```

    INFO = union(INFO, DirectNumbers);

```

```

    INFO = union(INFO, Singles);

```

```

    INFO = union(INFO, DirectStrings);

```

```

    return;

```

```

end;

```

```

if ~isempty(find(ismember(DirectStrings, Parameter), 1))

```

```

    INFO = ({RAWData.(Parameter)});

```

```

    return;

```

```

end;

```

```

if ~isempty(find(ismember(DirectNumbers, Parameter), 1))

```

```

    if strcmp(Parameter,'SI')

```

```

        Parameter = 'size';

```

```

    end;

```

```

    if strcmp(Parameter,'SW')

```

```

        Parameter = 'sw';

```

```
end;
INFO = cell2mat({RAWData.(Parameter)});
return;
end;

PULSES = cell2mat({RAWData.PULSES})';
DELAYS = cell2mat({RAWData.D})';
POWERLEVEL = cell2mat({RAWData.POWERLEVEL})';
POWERLEVELW = cell2mat({RAWData.POWERLEVELW})';
POWERLEVELDB = cell2mat({RAWData.POWERLEVELDB})';

if strcmp(Parameter,'PL1');
    INFO = POWERLEVEL(:,2);
    if RAWData(1).TOPSPIN>=3
        INFO = POWERLEVELDB(:,2);
    end;
    return;
end;

if strcmp(Parameter,'PL9');
    INFO = POWERLEVEL(:,10);
    if RAWData(1).TOPSPIN>=3
        INFO = POWERLEVELDB(:,10);
    end;
    return;
end;

if strcmp(Parameter,'PULSES');
    INFO = PULSES;
    return;
end;

if strcmp(Parameter,'DELAYS');
    INFO = DELAYS;
    return;
end;
```

```
end;

if strcmp(Parameter,'POWERLEVELS');
    INFO = POWERLEVEL;
    return;
end;

if strcmp(Parameter,'POWERLEVELSW');
    INFO = POWERLEVELW;
    return;
end;

if strcmp(Parameter,'POWERLEVELSDB');
    INFO = POWERLEVELDB;
    return;
end;

if strcmp(Parameter,'P1');
    INFO = PULSES(:,2);
    return;
end;

if strcmp(Parameter,'D1');
    INFO = DELAYS(:,2);
    return;
end;

files = {RAWData.file};
files = strrep(files,'/','\');

if strcmp(Parameter,'EXPNO');
    INFO = cellfun(@(x) str2double(x{1}),regexp(files,'(\d+)\p{data}\\\d+\\\d+r','tokens'));
    return;
end;
```

```
if strcmp(Parameter,'PROBHD');
    INFO = cell(length(RAWData),1);
    for i=1:length(RAWData)
        s = strtrim(RAWData(i).ACQUS.PROBHD(2:end-1));
        INFO{i} = s;
    end;
    return;
end;
```

```
if strcmp(Parameter,'PROCNO');
    INFO = cellfun(@(x) str2double(x{1}),regexp(files,'\d+\\pdata\\(\d+)\\d+r','tokens'));
    return;
end;
```

```
if strcmp(Parameter,'SAMPLEID');
    INFO = cellfun(@(x) (x{1}),regexp(files,'\\([^\d]*)\\d+\\pdata\\d+\\','tokens'));
    return;
end;
```

```
if strcmp(Parameter,'SR');
    SF = bbio_spec_info(RAWData, 'SF');
    BF = bbio_spec_info(RAWData, 'BF1');
    INFO = (SF-BF)*1E6;
    return;
end;
```

```
if strcmp(Parameter,'SWH');
    SW = bbio_spec_info(RAWData, 'SW');
    BF = bbio_spec_info(RAWData, 'BF1');
    O1 = bbio_spec_info(RAWData, 'O1');
    INFO = SW .* (BF + 1E-6*O1);
    return;
```

```
end;
```

```
function [ppm] = bbio_spec_ppm(Spec)
```

```
    % results ppm axis for spectrum
```

```
    step = Spec.sw / (Spec.size-1);
```

```
    ppm = Spec.maxppm:-step:Spec.minppm;
```

```
function RAW = bbio_spec_read1d( FileList, varargin)
```

```
%
```

```
% -----
```

```
% ----- Bruker Biospin Toolbox -----
```

```
% -----
```

```
% bbio_spec_read1d = Read 1R-NMR files
```

```
%
```

```
% This function reads the spectra of all given filenames
```

```
% and returns a struct RAW with the Data and almost all
```

```
% available parameters
```

```
%
```

```
% INPUT:
```

```
%   FileList:
```

```
%       is a cell-array containing the filenames
```

```
%       of the files to be read (with 1r)
```

```
% Parameters
```

```
%   ppmRange: [min max]
```

```
%       is a 1x2-array defining the ppm-Range for
```

```
%       reading (optional)
```

```
%   imaginary: (true/false)
```

```
%       boolean, "true" for reading imaginary-part
```

```
%
```

```
% -----
```

```
% Dealing with the input
```

```
p = inputParser;
```

```
p.addParamValue('imaginary',false);
```

```
p.addParamValue('ppmRange',[-1E99 1E99]);
```

```
p.addValue('readShim',true);
p.addValue('readSCON',true);
p.addValue('readICON',true);
p.addValue('readERETIC',true);
p.addValue('PrepDilution',1);
p.addValue('OnlyMetaFast',false);
p.addValue('CACHEDIR','');

p.parse(varargin{:});
PARS = p.Results;

if ischar(FileList)
    FileList = {FileList};
end;

n = length(FileList);
RAW = struct('Data',cell(n,1),'Imag',cell(n,1));

if n>=10
    fprintf('1D-Reading (%d spectra)...\n', n);
end;
for k=1:n

    FileList{k} = strrep(FileList{k}, '/' , '\');
    cFileName = FileList{k};
    cDir      = cFileName(1:(end-2));

    if mod(k,100)==0 && k>2
        fprintf(' %d (%.02f%%)\n', k, 100*k/n);
    end;

    if mod(k,10)==1 && n>=10
        fprintf('.');
    end;
end;
```



```
if ~isempty(PARS.CACHEDIR) && k>1
    tmpap = bbio_internal_loadtext([cDir 'auditp.txt']);
    h = bbio.tools.hash(tmpap);
    h = h{1};
    if exist([PARS.CACHEDIR '\ ' h '.mat'],'file')
        tmpL = load([PARS.CACHEDIR '\ ' h '.mat']);
        RAW(k) = tmpL.tmpL.S;
        continue;
    end;
end;

if exist(cFileName, 'file')~=2
    fprintf('\nFile "%s" does not exist!\n', cFileName);
    continue;
end;

% loading jcamp-info files acqu, procs

[ACQUS, ACQUSTXT] = bbio_internal_JCAMP_read( [cDir '..\..\acqus']);
[PROCS, PROCSTXT] = bbio_internal_JCAMP_read( [cDir 'procs']);

AUDITA = bbio_internal_loadtext([cDir '..\..\audita.txt']);
AUDITP = bbio_internal_loadtext([cDir 'auditp.txt']);

titlefile = [cDir 'title'];
if exist(titlefile,'file')
    myTitle = strtrim(bbio_internal_loadtext(titlefile));
else
    myTitle = '';
end;

RAW(k).TITLE = myTitle;
RAW(k).ACQUS = ACQUS;
RAW(k).PROCS = PROCS;
```

```

RAW(k).ACQUESTXT = ACQUESTXT;
RAW(k).PROCSTXT = PROCSTXT;
RAW(k).AUDITA = AUDITA;
RAW(k).AUDITP = AUDITP;

if PARS.OnlyMetaFast
    continue;
end;

% save important parameters at top-level
if ~isempty(ACQUS)
    RAW(k).O1 = ACQUS.O1;
    RAW(k).RG = ACQUS.RG;
    RAW(k).PULPROG = ACQUS.PULPROG;
    RAW(k).BF1 = ACQUS.BF1;
    RAW(k).TD = ACQUS.TD;
    RAW(k).NS = ACQUS.NS;
    RAW(k).Date = bbio_internal_UnixToMatLabDate(ACQUS.DATE);

    % Real Aquisition Date with correct TimeZone
    v = regexp(RAW(k).AUDITA,'1,<[^\>]*>,<[^\>]*>,<[^\>]*>,<(go)>|<(go4)>','tokens');
    if ~isempty(v)
        ds = v{1}{1};
        RAW(k).DateReal = datenum(ds(1:19),'yyyy-mm-dd HH:MM:SS');
        RAW(k).TimeZone = str2double(ds(25:27));
    else
        RAW(k).DateReal = NaN;
        RAW(k).TimeZone = NaN;
    end;

    % extract lists like pulses, powerlevels, delays
    [PULSES, PL, D] = internal_add_customfields_acqus(ACQUS,...
        ACQUESTXT);

```

```

% getting the spectrometer is complicated (WRPA-Problem)
tmpTSPECT = regexp(AUDITA,'<([>]*)>,<go','tokens');
if ~isempty(tmpTSPECT)
    SPECT = tmpTSPECT{1}{1};
else
    SPECT = '';
end;
% For AvNeo-Systems
v = regexp(AUDITA,'created by zg on "([^\n]*)" \\.([^\n]*)\ by user','tokens','once');
if length(v)==2
    SPECT = v{2};
end;

RAW(k).PULSES = PULSES;
RAW(k).POWERLEVEL = PL;
RAW(k).D = D;
RAW(k).SPECT = SPECT;
% stripped 32 field version
RAW(k).PULSES32 = PULSES(1:32);
RAW(k).D32 = D(1:32);
RAW(k).POWERLEVEL32 = PL(1:32);

% if TopSpin 3.0 or higher -> different Power-Handling
[PLW, PLDB] = internal_new_powerlevel(ACQUS);
RAW(k).POWERLEVELW = PLW;
RAW(k).POWERLEVELDB = PLDB;
end;

RAW(k).ByteOrder = PROCS.BYTORDP;
if isfield(PROCS,'LB')
    RAW(k).LB = PROCS.LB;
end;
if isfield(PROCS,'PHC0')
    RAW(k).PHC0 = PROCS.PHC0;

```

```
end;
if isfield(PROCS,'PHC1')
    RAW(k).PHC1    = PROCS.PHC1;
end;
RAW(k).SF        = PROCS.SF;
RAW(k).file      = cFileName;

% Getting the TOPSPIN-Version
if ischar(ACQUESTXT)
    v = regexp(ACQUESTXT, '[^\n]*TOPSPIN\t\tVersion (\S+)', 'tokens');
    if isempty(v)
        v = regexp(ACQUESTXT, '[^\n]*TopSpin (\S+)', 'tokens');
    end;
    if ~isempty(v)
        v = v{1}{1};
        tmpVersion = str2double(v);
        if isnan(tmpVersion)
            % try first 3 characters
            tmpVersion = str2double(v(1:3));
        end;
        RAW(k).TOPSPIN = tmpVersion;
    else
        RAW(k).TOPSPIN = NaN;
    end;
end;

% reading the spectrum

[si, ofs, sw, real, ncproc, imag] = internal_read_spectrum(...
    cFileName, PARS.ppmRange(1), ...
    PARS.ppmRange(2), PROCS, ACQUS, PARS);
RAW(k).Data = real.*(2^(ncproc));
RAW(k).Imag = imag.*(2^(ncproc));
```

```
RAW(k).maxppm = ofs;
RAW(k).minppm = ofs-sw;
RAW(k).sw = sw;
RAW(k).size = si;

% Reading the shim
if PARS.readShim
    shimfile = [cDir '..\..\shimvalues'];
    if exist(shimfile,'file')
        RAW(k).SHIM = internal_read_shimvalues( shimfile);
    else
        RAW(k).SHIM = [];
    end;
else
    RAW(k).SHIM = [];
end;

% Reading scon
if PARS.readSCON
    sconfile = [cDir '..\..\scon2'];
    if exist(sconfile,'file')
        RAW(k).SCON = bbio_internal_JCAMP_read( sconfile);
    else
        RAW(k).SCON = [];
    end;
else
    RAW(k).SCON = [];
end;

% ICON-INFO
if PARS.readSCON
    iconfile = [cDir '..\..\order.txt'];
    if exist(iconfile,'file')
        RAW(k).ICON = bbio_internal_JCAMP_read( iconfile);
    else
```

```
        RAW(k).ICON = [];  
    end;  
else  
    RAW(k).ICON = [];  
end;  
  
% Read Eretic from File EreticFactorSample or eretic.xml  
if PARS.readERETIC  
  
    RAW(k).ERETIC = bbio_spec_ERETICfromFile( cFileName);  
    if ~isnan(RAW(k).ERETIC)  
        RAW(k).ERETICUNIT = 'mmol/L';  
    else  
        RAW(k).ERETICUNIT = '';  
    end;  
else  
    RAW(k).ERETIC = NaN;  
    RAW(k).ERETICUNIT = '';  
end;  
RAW(k).PREPDILUTION = PARS.PrepareDilution;  
  
if ~isempty(PARS.CACHEDIR)  
    tmpap = bbio_internal_loadtext(['cDir 'auditp.txt']);  
%    h = bbio_hash2(tmpap);  
    h = bbio.tools.hash(tmpap);  
    h = h{1};  
    tmpL = [];  
    tmpL.S = RAW(k);  
    save(['PARS.CACHEDIR \' h'],'tmpL');  
end;  
  
end;  
  
if n>=10  
    fprintf('\n');
```

```

end;

% -----
% -----
% -----
% -----
% -----

function SHIM = internal_read_shimvalues( shimfile)

SHIM = [];
L = bbio_internal_loadtext(shimfile);
SL = regexp(L,'.*?\n','tokens');
StartPos = 0;
for i=1:length(SL)
    s = strtrim(SL{i}{1});
    if ~strcmp(s(1),'#')
        StartPos = i;
        break;
    end;
end;
for i=StartPos:length(SL)
    s = strtrim(SL{i}{1});
    if strcmp(s(1),'#')
        break;
    end;
    tmpV = regexp(s,{'\S*'}\s+({'*\d+'),'tokens');
    if ~isempty(tmpV) && length(tmpV{1})==2
        sNAME = tmpV{1}{1};
        sVALUE = str2double(tmpV{1}{2});
        sNAME = strrep(sNAME, '(', '');
        sNAME = strrep(sNAME, ')', '');
        sNAME = strrep(sNAME, '-', 'm');
        SHIM.(['SH_' sNAME]) = sVALUE;
    end;
end;

```

```

end;
end;

ADDPARS = {'LOCKPOWER','LOCKPHASE','LOCKGAIN','LOCKDC','LOCKSHIFT',...
           'FIELD','LOOPGAIN','LOOPTIME','LOOPFILTER'};

for i=1:length(ADDPARS)
    tmpT = regexp(L,[ADDPARS{i} '\s+(.*?)\n'],'tokens');
    SHIM.(ADDPARS{i}) = str2double(strtrim(tmpT{1}));
end;

```

```

function [PULSES, PL, D, SPECT] = internal_add_customfields_acqus( ACQUS, ...
                                                                    ACQUESTXT)

```

```

SPECT = [];

P = ACQUS.P;
P1 = strrep(P,sprintf('\n'),' '); %#ok<*SPRINTFN>
v1 = regexp(P1,('\S+)(\s*)','tokens');
PULSES = zeros(length(v1)-1,1);
for i=2:length(v1)
    PULSES(i-1) = str2double(v1{i}{1});
end;

P = ACQUS.PL;
P1 = strrep(P,sprintf('\n'),' ');
v1 = regexp(P1,('\S+)(\s*)','tokens');
PL = zeros(length(v1)-1,1);
for i=2:length(v1)
    PL(i-1) = str2double(v1{i}{1});
end;

P = ACQUS.D;
P1 = strrep(P,sprintf('\n'),' ');

```



```
v1 = regexp(P1, '(\\S+)(\\s*)', 'tokens');
D = zeros(length(v1)-1,1);
for i=2:length(v1)
    D(i-1) = str2double(v1{i}{1});
end;

v1 = regexp(ACQUSTXT, '##OWNER=([\\^\\n]*)\\n([\\^\\n]*)\\n', 'tokens', 'once');
if length(v1)==2
    lin = (v1{2});
    atpos = strfind(lin, '@');
    if ~isempty(atpos)
        SPECT = strtrim(lin((atpos(1)+1):end));
    end;
end;
```

```
function [PLW, PLDB] = internal_new_powerlevel(ACQUS)
```

```
PLW = zeros(64,1)*NaN;
PLDB = zeros(64,1)*NaN;

if isfield(ACQUS, 'PLW')

    % TopSpin 3.0 or higher
    P = ACQUS.PLW;
    P1 = strrep(P, sprintf('\\n'), ' ');
    v1 = regexp(P1, '(\\S+)(\\s*)', 'tokens');
    PLW = zeros(length(v1)-1,1);
    for i=2:length(v1)
        PLW(i-1) = str2double(v1{i}{1});
    end;

    PLDB = - 10 * log10(PLW);
    PLDB(isinf(PLDB)) = 1000;
```

```
end;
```

```
function [si, ofs, sw, real, ncproc, imag] = internal_read_spectrum( aFile,...  
    minppm, maxppm, INFO, INFO2,...  
    PARS)
```

```
% byteorder?
```

```
dataformat = 'l';
```

```
if INFO.BYTORDP == 1
```

```
    dataformat = 'b';
```

```
end;
```

```
% get important parameters
```

```
ofs = INFO.OFFSET;
```

```
if ~isempty(INFO2)
```

```
    sw = INFO2.SW;
```

```
else
```

```
    sw = INFO.SW_p/INFO.SF;
```

```
end;
```

```
ncproc = INFO.NC_proc;
```

```
si = INFO.SI;
```

```
minppmindex = floor((ofs-minppm)/sw*(si-1));
```

```
maxppmindex = ceil((ofs-maxppm)/sw*(si-1));
```

```
if maxppmindex < 0;
```

```
    maxppmindex = 0;
```

```
    minppmindex = si-1;
```

```
end;
```

```
if ~exist(aFile,'file')
```

```
    si = 0;
```

```
    ofs = 0;
```

```
    sw = 0;
```

```
    real = NaN;
```

```
ncproc = 0;
return;
end;

count = minppmindex - maxppmindex + 1;
if count>0
    f = fopen (aFile, 'r',dataformat);
    fseek(f, maxppmindex*4,0);
    [real,si2]=fread(f,count, 'int32');
    fclose (f);
    if PARS.imaginary == true
        % loading the imaginary part
        iFile = strrep(aFile,'\1r','\1i');
        f = fopen (iFile, 'r',dataformat);
        fseek(f, maxppmindex*4,0);
        imag = fread(f,count, 'int32');
        fclose (f);
    else
        imag = [];
    end;
else
    real = [];
    imag = [];
    si2 = 0;
end;
ofs = (ofs - sw/(si-1)*maxppmindex);
sw = (ofs - sw/(si-1)*maxppmindex) - (ofs - sw/(si-1)*minppmindex);
si = si2;
```

function [Integral, Error, Position, LineWidth, S] = FitTSP(afile)

```
S = bbio_spec_read1d( afile);
S = bbio_spec_filter(S, 15,[3 4]);
```

```
TSPpos = [-0.10 0.10]; % Position des Signals
```

```

ppmref = [-0.10 0.10]; % Region des Fits

ppm = bbio_spec_ppm(S);

idx = find(ppm>ppmref(1) & ppm<ppmref(2));
D = S.Data(idx);
ppm2 = ppm(idx);

% Fit this signal, using a start
START = [mean(TSPpos);... % start-position
1.5;... % start width
0.1]; % Gauss/Lorentz
UP = [max(TSPpos) 2.5 1]; % obere Grenzen, maxwidth = 2.5
LOW = [min(TSPpos) 0.5 0]; % untere Grenzen, minwidth = 0.5
END = fminsearchbnd(@(x, varargin) ...
FitSignal(x, ppm2, D', S.SF),...
START, LOW, UP);
FittedSignal = InternalNMRSignal(ppm2, END(1), END(2), END(3), S.SF);

fprintf('Position: %.4f ppm\n', END(1));
fprintf('Linienbreite: %.2f Hz\n', END(2));
fprintf('Gauss/Lorentz: %.2f \n', END(3));

intensity = interp1(ppm2, D, END(1));
FittedSignal = FittedSignal* intensity;
Integral = sum(FittedSignal)*abs(ppm2(2)-ppm2(1));

% Fehler bei Doppelter Linienbreite
ppmerror = [END(1)-(END(2)/S.SF)*2 END(1)+(END(2)/S.SF)*2];
ierror = find(ppm2>ppmerror(1) & ppm2<ppmerror(2));
Error = median(abs((D(ierror)')-FittedSignal(ierror))./D(ierror)'));

Position = END(1);
LineWidth = END(2);

```

```

clf;hold on;
plot(ppm2,D,'-');
plot(ppm2,FittedSignal,'-r');
set(gca,'xdir','reverse');
set(gca,'xlim',ppmref);
legend('Messung','Fit TSP');
box on;
title(sprintf('Integral TSP: %.1f, Halbwertsbreite [Hz]: %.1f%', Integral, LineWidth));

end

```

```

function d = FitSignal(x, ppm, ReferenzSignal, Field)
    Position = x(1);
    Width   = x(2);
    Gauss   = x(3);

    intensity = interp1(ppm, ReferenzSignal, Position);

    y1 = InternalNMRSignal(ppm, Position, Width, Gauss, Field);
    y1 = y1 * intensity;

    d = sum((y1-ReferenzSignal).^2);

    % Oberhalb des Spektrums wird mit Faktor 3 gewichtet
    ii = find(y1>ReferenzSignal);
    d = d + sum((y1(ii)-ReferenzSignal(ii)).^2)*2;

end

```

```

function y1 = InternalNMRSignal(ppm, SPos, Width, Gauss, Field)
    % Lorentz
    y1 = (1./(1+((ppm-SPos)/(Width/Field)*2).^2));
    % Gauss

```

```
y1 = y1*(1-Gauss) +...  
      exp(-log(2)*((ppm-SPos).^2/((0.5*Width/Field)^2)).*Gauss;  
end
```

10 Publikationsliste

Publikationen in Fachzeitschriften

Teipel, J., Gottstein, V., Hölzle, E., Kaltenbach, K., Lachenmeier & D., Kuballa, T. (2022). An easy and reliable method for the mitigation of deuterated chloroform decomposition to stabilise susceptible NMR samples. *Chemistry* **4**, 776–785

Segatz, V., Steger, M. C., Blumenthal, P., Gottstein, V., Rigling, M., Schwarz, S., Zhang, Y. & Lachenmeier, D. W. (2021). Evaluation of analytical methods to determine regulatory compliance of coffee leaf tea. *Biology and Life Sciences Forum* **6**, 45.

Gottstein, V., Bernhardt, M., Dilger, E., Keller, J., Breitling-Utzmann, C., Schwarz, S., Kuballa, T., Lachenmeier, D. W. & Bunzel, M. (2021). Coffee silver skin: chemical characterization with special consideration of dietary fiber and heat-induced contaminants. *Foods* **10**, 1705.

Claassen, L., Rinderknecht, M., Porth, T., Röhnisch, J., Seren, H. Y., Scharinger, A., Gottstein, V., Noack, D., Schwarz, S., Winkler, G. & Lachenmeier, D. W. (2021). Cold brew coffee - pilot studies on definition, extraction, consumer preference, chemical characterization and microbiological hazards. *Foods* **10**, 865.

Klingel, T., Kremer, J. I., Gottstein, V., Rajcic de Rezende, T., Schwarz, S. & Lachenmeier, D. W. (2020). A review of coffee by-products including leaf, flower, cherry, husk, silver skin, and spent grounds as novel foods within the European Union. *Foods* **9**, 665.

Gottstein, V., Müller, M., Günther, J., Kuballa, T. & Vetter, W. (2019). Direct ¹H NMR quantitation of valuable furan fatty acids in fish oils and fish oil fractions. *Journal of Agricultural and Food Chemistry* **67**, 11788-11795.

Vorträge auf Fachtagungen

Gottstein, V., Bunzel, M., Lachenmeier, D. W. & Kuballa, T. (2023). NMR-spectroscopic applications to control coffee authenticity, *ACS Spring 2023 – Before the Coffee Break: The Rich and Complex Chemistry of Coffee*, digital.

Gottstein, V. (2021). Coffee silver skin: chemical characterization with special consideration of dietary fiber and heat-induced contaminants, *Live session of the Foods2021 - "Coffee by-products as sustainable novel foods"*.

Gottstein, V. (2020). NMR applications for coffee authenticity control, *Coffee Symposium*, Messe Stuttgart.

Posterbeiträge auf Fachtagungen

Gottstein, V., Bunzel, M., Lachenmeier, D. W. & Kuballa, T. (2022). Quantifizierung ausgewählter Kaffeeinhaltsstoffe mittels Protonen-Kernspinresonanzspektroskopie, 50. *Deutscher Lebensmittelchemikertag*, Hamburg.

Gottstein, V., Müller, M., Kuballa, T. & Vetter, W. (2019). Quantifizierung von monomethylierten und dimethylierten Furanfettsäuren mittels ^1H -NMR, 48. *Deutscher Lebensmittelchemikertag*, Dresden.