

**Entwicklung von SCORE-metabolite-ID
zur Identifizierung von Metaboliten aus komplexen Gemischen
durch Korrelation von NMR-, MS- und LC-Daten**

Dissertation

zur Erlangung des akademischen Grades

Doctor rerum naturalium

Dr. rer. nat.

Fakultät für Mathematik, Informatik und Naturwissenschaften

Fachbereich Chemie, Institut für Organische Chemie

Universität Hamburg

vorgelegt von

Stephanie Watermann

Hamburg, 2025

- 1. Gutachter:** Dr. Thomas Hackl
- 2. Gutachter:** Prof. Dr. Maria Buchweitz
- 3. Gutachter:** Prof. Dr. Stephan Schwarzingen

Prüfungskommission: Dr. Thomas Hackl
Prof. Dr. Ralph Holl
Prof. Dr. Wolfgang Maison

Datum der Disputation: 06.06.2025

Datum der Druckfreigabe: 11.06.2025

Die vorliegende Arbeit wurde im Zeitraum von Juli 2020 bis Dezember 2023 am Institut für Organische Chemie des Fachbereichs Chemie der Universität Hamburg unter der Anleitung von Herrn Dr. Thomas Hackl und Herrn Prof. Dr. Christian B. W. Stark angefertigt.

*„It is the set of the sails,
not the direction of the wind
that determines which way we will go.”*

– Jim Rohn

I. Publikationsliste

Publikationen:

- S. Watermann, M. Bode, T. Hackl, Identification of metabolites from complex mixtures by 3D correlation of ^1H NMR, MC and LC data using the SCORE-metabolite-ID approach, *Sci. Rep.* **2023**, *13*, 15834.
- S. Watermann, C. Schmitt, T. Schneider, T. Hackl, Comparison of Regular, Pure Shift and Fast 2D NMR Experiments for Determination of the Geographical Origin of Walnuts, *Metabolites* **2021**, *11*(1), 39.
- T. Hackl, R. Bachmann, C. Schmitt, S. Watermann, Mit Metabolic Profiling gegen Betrug, *Nachrichten aus der Chemie* **2020**, *68*, 58-61.

Tagungsbeiträge (Vorträge):

- Strategien zur Strukturaufklärung in der Analyse von Gemischen, Workshop-Beitrag, 16. Sitzung der NEXT NMR AG, 29.06.-30.06. **2022**, Karlsruhe, Deutschland.
- Vergleich von regulären, Pure Shift und schnellen HSQC-Experimenten zur geographischen Herkunftsanalyse von Walnüssen, 42. Tagung Praktische Probleme der Kernresonanzspektroskopie, **2021**, Berlin (digital), Deutschland.

Tagungsbeiträge (Poster):

- SCORE-metabolite-ID⁽²⁾ – Identification of metabolites from complex mixtures by correlation of 1D- ^1H or 2D-HSQC NMR, MS and LC data, EuroAnalysis 2023, 27.08.-31.08. **2023**, Genf, Schweiz.
- SCORE-metabolite-ID⁽²⁾ – Identification of metabolites from complex mixtures by correlation of 1D- ^1H or 2D-HSQC NMR, MS and LC data, EUROMAR 2023, The 19th European Magnetic Resonance Congress, 09.07.-13.07. **2023**, Glasgow, Schottland, Großbritannien.

- Identifizierung von Metaboliten aus Pinienkernen durch 3D-Korrelation von LC, MS und ^1H -NMR, 50. Deutscher Lebensmittelchemikertag, 19.09.-21.09. **2022**, Hamburg, Deutschland.
- SCORE-metabolite-ID: Identification of metabolites from complex mixtures by correlation of 1D or 2D NMR, MS and LC, Fachgruppe Magnetische Resonanzspektroskopie, 43rd FGMR Annual Discussion Meeting, 12.09.-15.09. **2022**, Karlsruhe, Deutschland.
- Identification of metabolites from a complex pine nut sample by 3D correlation of LC, MS and NMR, MR FOOD 2022 – 15th International Conference on the Application of Magnetic Resonance in Food Science, 07.06.-10.06. **2022**, Aarhus, Dänemark.
- Strategic approach for the identification of metabolites from complex mixtures by correlation of NMR, MS and LC, Fachgruppe Magnetische Resonanzspektroskopie, 42nd FGMR Annual Discussion Meeting, 27.09.-01.10. **2021**, Göttingen (digital), Deutschland.

II. Inhaltsverzeichnis

I.	Publikationsliste	I
II.	Inhaltsverzeichnis	III
III.	Abkürzungsverzeichnis	VII
1	Zusammenfassung	1
2	Abstract	5
3	Einleitung	9
3.1	Metabolomics-Studien	10
3.2	Multivariate Datenanalyse	12
3.2.1	Hauptkomponentenanalyse	13
3.2.2	Klassifizierungsverfahren	15
3.2.3	Validierung	16
3.3	Klassische analytische Methoden zur Identifizierung von Metaboliten	19
3.3.1	Flüssigchromatographische Trennung von Gemischen	20
3.3.2	Identifizierungsmethoden basierend auf NMR-Spektroskopie	22
3.3.3	Identifizierungsmethoden basierend auf Massenspektrometrie	30
3.4	Identifizierungsstrategien durch Kombination analytischer Methoden	32
3.4.1	Three-Dimensional Cross Correlation (3DCC)	34
3.4.2	Mathematische Korrelation von Variablen	35
4	Zielsetzung	39
5	Ergebnisse und Diskussion	41
5.1	Entwicklung von SCORE-metabolite-ID basierend auf ^1H -NMR-Spektren	41
5.1.1	Entwicklung der graphischen Benutzeroberfläche	48
5.1.2	Validierung durch das artifizielle Gemisch A	58
5.2	Anwendung von SCORE-metabolite-ID zur Identifizierung von Metaboliten	67
5.2.1	Identifizierung von Metaboliten aus polarem Pinienkernextrakt	68
5.2.2	Sortendifferenzierung von Apfelproben	118
5.2.3	Identifizierung von potenziellen Markermetaboliten in Apfelproben	141

5.3	Anwendung von SCORE-metabolite-ID zur Reaktionsverfolgung	159
5.3.1	Darstellung von 2,4,6-Tri-(6-aminocaprinsäure)-1,3,5-triazin (TACT)	159
5.3.2	NMR-spektroskopische Untersuchung von TACT	163
5.3.3	Reaktionsverfolgung durch Korrelation der Daten	165
5.4	Entwicklung von SCORE-metabolite-ID ² basierend auf 2D HSQC-NMR-Spektren	179
5.4.1	Entwicklung der graphischen Benutzeroberfläche	180
5.4.2	Vergleich verschiedener HSQC-Experimente	187
5.4.3	Validierung durch artifizielles Gemisch B	193
6	Experimenteller Teil	209
6.1	Chemikalien	209
6.2	Materialien	210
6.3	Geräte	210
6.4	Software	211
6.5	Artifizielle Gemische A und B	211
6.5.1	Phosphatpuffer (200 mM, pH 7.0, 1 mM TSP- <i>d</i> ₄ , 3 mM NaN ₃)	211
6.5.2	Phosphatpuffer (1 M, pH 7.0, 2 mM TSP- <i>d</i> ₄ , 3 mM NaN ₃)	212
6.5.3	Probenvorbereitung	212
6.5.4	NMR-Spektroskopie	217
6.6	Pinienkern-Proben	218
6.6.1	Phosphatpuffer (100 mM, pH 7.0, 1 mM TSP- <i>d</i> ₄ , 3 mM NaN ₃)	218
6.6.2	Polare Extraktionsmethode	218
6.6.3	Chromatographische Fraktionierung	219
6.6.4	Probenvorbereitung	221
6.6.5	NMR-Spektroskopie	221
6.7	Apfel-Proben für multivariate Datenanalyse	222
6.7.1	Phosphatpuffer (3.0 M, pH 8.0, 10 mM TSP- <i>d</i> ₄ , 3 mM NaN ₃)	222
6.7.2	Probenvorbereitung	223
6.7.3	NMR-Spektroskopie	223
6.8	Multivariate Datenanalyse	224
6.8.1	Hauptkomponentenanalyse	224
6.8.2	Klassifizierungsmodelle	225

6.9	Apfel-Proben für die Identifizierung von Biomarkern	225
6.9.1	Flash-Chromatographie	225
6.9.2	Probenvorbereitung	226
6.9.3	NMR-Spektroskopie	226
6.10	Synthese von 2,4,6-Tri-(6-aminocaprinsäure)-1,3,5-triazin (TACT)	227
6.10.1	Probenvorbereitung	228
6.10.2	NMR-Spektroskopie	228
6.11	Massenspektrometrie	229
6.11.1	Messungen per Direkteinlass	229
6.11.2	LC-MS-Messungen	230
6.12	Zweidimensionale und selektive NMR-Experimente	231
7	Literaturverzeichnis	233
8	Anhang	253
8.1	Verwendete Chemikalien	253
8.2	Apfelproben	259
8.3	Multivariate Datenanalyse	262
8.4	Korrelationskoeffizienten	266
8.5	Zusätzliche NMR-Spektren	277
8.6	Skripte	289
9	Danksagung	323
10	Eidesstattliche Erklärung	325

III. Abkürzungsverzeichnis

1D/2D/3D	ein-/zwei-/dreidimensional
3DCC	<i>three-dimensional cross correlation</i>
ACN	Acetonitril
Ala	Alanin
β-Ala	β-Alanin
AMP	Adenosin-5'-monophosphat
ASAP	<i>acceleration by sharing adjacent polarization</i>
Asp	Asparaginsäure
BEST	<i>band-selective excitation short-transient</i>
BMRB	<i>Biological Magnetic Resonance Data Bank</i>
BW	Baden-Württemberg
CAWG	<i>Chemical Analysis Working Group</i>
Cho.	Cholin
Cl-OH-TACT	2-(6-Aminocapronsäure)-4-chlor-6-hydroxy-1,3,5-triazin
Cl-TACT	2,4-Di-(6-aminocapronsäure)-6-chlor-1,3,5-triazin
CNN	<i>Convolutional Neural Network</i>
COSY	<i>Correlation Spectroscopy</i>
Creat.	Creatinin
δ	Chemische Verschiebung [ppm]
d1	Relaxationsdelay
DE	Deutschland
DI	<i>direct injection</i> (Direkteinlass)
DOSY	<i>Diffusion Ordered Spectroscopy</i>
DS	<i>Dummy Scans</i>
EDC	<i>Extracted Delta Chromatogram</i>
EDDC	<i>Extracted Double Delta Chromatogram</i>
EI	Elektronenionisation
EIC	<i>Extracted Ion Chromatogram</i>
EMC	<i>Extracted Mass Chromatogram</i>
ESI	Elektrospray-Ionisation

<i>et al.</i>	und andere, lat. <i>et alii</i>
F	Fraktion
FA	Ameisensäure (<i>formic acid</i>)
FID	<i>free induction decay</i>
GC	Gaschromatographie
Glc	Glucose
GNPS	<i>Global Natural Product Social Molecular Networking</i>
GSSG	Glutathion Disulfid
HH	Hamburg
HILIC	<i>Hydrophilic Interaction Chromatography</i>
Hist.	Histamin
HMBC	<i>Heteronuclear Multiple Bond Correlation</i>
HMDB	<i>Human Metabolome Database</i>
HMQC	<i>Heteronuclear Multiple Quantum Coherence</i>
HPLC	<i>High Performance Liquid Chromatography</i>
HSQC	<i>Heteronuclear Single Quantum Coherence</i>
HSQCcos	<i>Heteronuclear Single Quantum Coherence Correlation Spectroscopy</i>
IAA	Indol-3-essigsäure
Ile	Isoleucin
INEPT	<i>Insensitive Nuclei Enhanced by Polarization Transfer</i>
IR	Infrarot
IT	Italien
<i>J</i>	Kopplungskonstante [Hz]
LC	<i>Liquid chromatography</i>
Leu	Leucin
<i>m/z</i> -Verhältnis	Masse-zu-Ladungs-Verhältnis
MS	Massenspektrometrie
MS/MS	Tandem-Massenspektrometrie
NAD	Nicotinamidadenindinucleotid
NB	Nicht bekannt
NI	Niedersachsen
NMR	<i>Nuclear Magnetic Resonance</i>

NOE	<i>Nuclear Overhauser Effect</i>
NOESY	<i>Nuclear Overhauser Effect Spectroscopy</i>
NUS	<i>Nonuniform Sampling</i>
OH-TACT	2,4-Di-(6-aminocaprinsäure)-6-hydroxy-1,3,5-triazin
PARAFAC	<i>Parallel Factor Analysis</i>
PC	<i>Principal Component</i>
PCA	<i>Principal Component Analysis</i>
PDS	<i>Parallel Dynamic Spectroscopy</i>
Phe	Phenylalanin
PUREST	<i>Pure Shift Relaxation-Encoded Selective TOCSY</i>
RES	Spektrale Auflösung [ppm]
REST	<i>Relaxation-Encoded Selective TOCSY</i>
RNCV	<i>Repeated Nested Cross Validation</i>
RP	<i>Reversed Phase (Umkehrphase)</i>
SCALPEL	<i>Spectral Component Acquisition by Localized PARAFAC Extraction of Linear Components</i>
SCORE-metabolite-ID	<i>Semi-automatic CORrelation analysis for RELiable metabolite IDentification</i>
SH	Schleswig-Holstein
SHY	<i>Statistical Heterospectroscopy</i>
SI	<i>Size of real spectrum</i>
SMART (a)	<i>Small Recovery Times</i>
SMART (b)	<i>Small Molecule Accurate Recognition Technology</i>
SOFAST	<i>Band-Selective Optimized-Flip-Angle Short-Transient</i>
STOCSY	<i>Statistical Total Correlation Spectroscopy</i>
STORM	<i>Subset Optimization by Reference Matching</i>
SUMMIT	<i>Structures of Unknown Metabolomics Mixture components</i>
SVM	<i>Support Vector Machine</i>
SW	<i>Spectral width [ppm]</i>
TACT	2,4,6-Tri-(6-aminocaprinsäure)-1,3,5-triazin
TD	<i>Time domain (size of FID)</i>
TIC	<i>Total Ion Current</i>

TOCSY	<i>Total Correlation Spectroscopy</i>
TOF	<i>Time of flight</i>
Uri.	Uridin
UV	Ultraviolett
VIS	<i>visible</i> (sichtbares Licht)

1 Zusammenfassung

Metabolomics-Studien befassen sich mit der Untersuchung von Metaboliten in biologischen Systemen und gewinnen u.a. in den Bereichen der Medizin und der Lebensmittelforschung zunehmend an Bedeutung. Aufgrund der Substanzvielfalt und der großen Konzentrationsunterschiede ist die Untersuchung von komplexen Stoffgemischen aus biologischen Matrices mit besonderen Herausforderungen verbunden.

Üblicherweise werden in *Metabolomics*-Studien statistische Methoden eingesetzt, um einzelne Signale oder Signalmuster zu identifizieren, die als Biomarker für die Beantwortung verschiedener Fragestellungen dienen können. Während die statistische Analyse unabhängig von der konkreten Kenntnis der chemischen Zusammensetzung der Probe erfolgt, erfordert eine biochemische Interpretation die Identifizierung dieser Markersubstanzen. Trotz des Einsatzes moderner und hochauflösender analytischer Verfahren spiegelt sich jedoch die Komplexität der Stoffgemische in den analytischen Daten z.B. in Form von Signalüberlagerungen wider. Die direkte Identifizierung von einzelnen Substanzen aus den Gemischspektren ist häufig nicht möglich, sodass die zuverlässige Identifizierung von Metaboliten in der Regel spezifische und aufwendige Schritte der Isolierung und Reinigung erfordert. Die Bereitstellung von Technologien zur vereinfachten Identifizierung von einzelnen Substanzen aus diesen komplexen Stoffgemischen stellt daher eine zentrale Herausforderung dar.

In der vorliegenden Arbeit wurde eine Methode zur semi-automatischen Korrelation von NMR- und MS-Daten zur vereinfachten Identifizierung von Metaboliten (*SCORE-metabolite-ID* (*Semi-automatic CORrelation analysis for RELiable metabolite IDentification*)) entwickelt. Das zugrunde liegende Konzept beruht auf der Methode 3DCC (*Three-dimensional cross-correlation*), die zur Identifizierung der Stereochemie von Glycanen eingesetzt wurde. Die *SCORE-metabolite-ID* Methode wurde als MATLAB Applikation entwickelt und dient der effizienten und sicheren Identifizierung von Metaboliten verschiedener Substanzklassen aus Gemischen ohne Notwendigkeit einer individuellen Isolierung. Die Methode basiert auf der dreidimensionalen Korrelation von ^1H -NMR- und MS-Daten über die dritte Dimension des zeitlichen Verlaufs. Die zeitliche Dimension wird durch eine chromatographische Fraktionierung des Stoffgemisches generiert. Über die graphische Benutzeroberfläche können auf einfache Art und Weise NMR- und MS-Daten korreliert werden, um NMR-Signale den entsprechenden Masse-zu-Ladungs-

Verhältnissen aus ESI-Massenspektren zuzuordnen. Dies vereinfacht die schnelle Dereplikation bekannter Metaboliten und die Identifizierung neuer Verbindungen in komplexen Gemischen. In der folgenden Abbildung 1 sind die entwickelte Methode sowie die im Rahmen dieser Arbeit beschriebenen Anwendungen schematisch abgebildet.

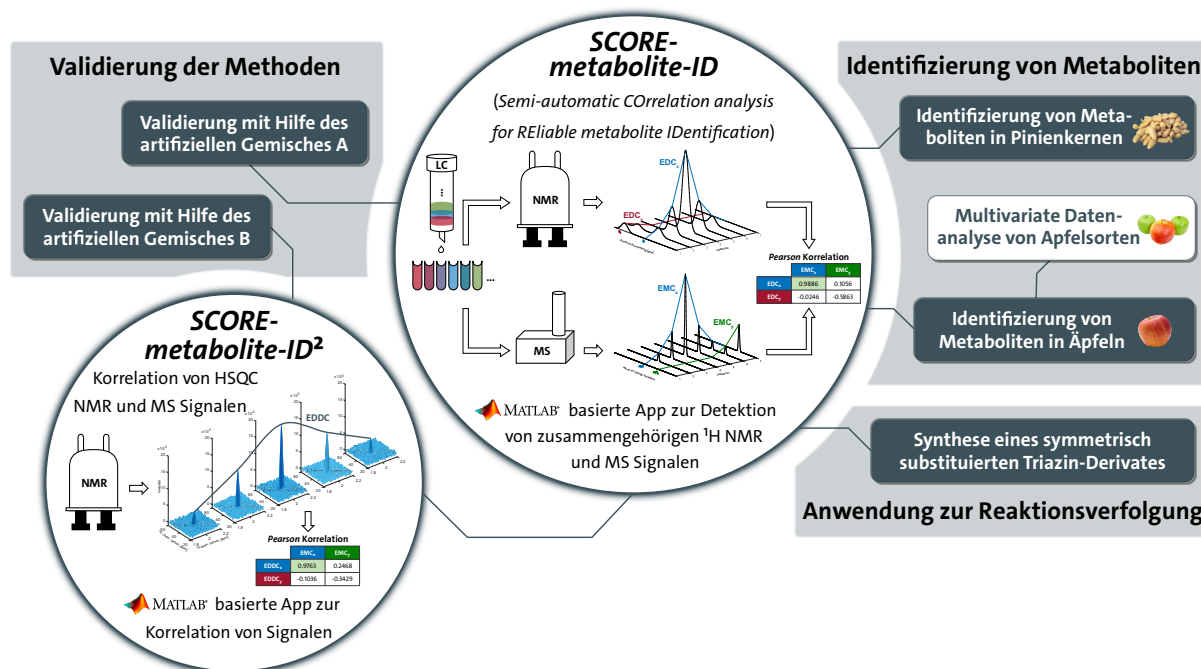


Abbildung 1: Schematische Übersicht der im Rahmen dieser Arbeit entwickelten Methoden *SCORE-metabolite-ID* und *SCORE-metabolite-ID²* sowie der in dieser Arbeit beschriebenen Anwendungen der Methoden.

Die *SCORE-metabolite-ID* Methode wurde zunächst anhand eines artifiziellen Gemisches validiert, bei dessen Zusammensetzung die Herausforderungen realer Stoffgemische simuliert wurden. Dazu gehören u.a. das Auftreten von isobaren und co-eluierenden Substanzen, wobei die Flexibilität der *SCORE-metabolite-ID* Methode auch in diesen herausfordernden Fällen die Korrelation und die Identifizierung der Analyten ermöglichte. Anschließend erfolgte die Anwendung auf verschiedene reale Extrakte biologischer Matrices.

Zunächst wurde dazu ein polares Extrakt von Pinienkernen als komplexe Matrix biologischen Ursprungs untersucht. Insgesamt konnten 45 Metaboliten verschiedener Substanzklassen (u.a. Kohlenhydrate, Betaine, Aminosäuren, organische Säuren etc.) und unterschiedlicher Konzentrationsbereiche identifiziert werden. Unter den im Pinienkernextrakt identifizierten Metaboliten sind insbesondere die sekundären Metaboliten β -D-Glucopyranosyl-1-N-indol-3-acetyl-N-L-asparaginsäure (**56**) und N^{α} -(2-Hydroxy-2-carboxymethylsuccinyl)-L-arginin (**57**)

hervorzuheben. Während zur eindeutigen Strukturaufklärung dieser Metaboliten zwar weitere Experimente notwendig waren, erleichterte und beschleunigte die *SCORE-metabolite-ID* Methode die Identifizierung durch die Korrelation der NMR- und MS-Daten jedoch erheblich.

Als weitere komplexe Matrix biologischen Ursprungs wurden Apfelproben analysiert. Die ^1H -NMR-Spektren der 86 Apfelproben der Sorten Boskoop, Elstar, Gala, Holsteiner Cox und Wellant wurden zunächst mittels multivariater Datenanalyse in Bezug auf die Sortenunterscheidung untersucht. Nach Variablenselektion durch *Boruta* wurde ein Klassifizierungsmodell auf Grundlage des *Support Vector Machine* Algorithmus zur Differenzierung der fünf Apfelsorten mit einer Klassifizierungsgenauigkeit von $90.4\% \pm 1.8\%$ erstellt. Im Anschluss wurden potenzielle Markersignale spezifisch mittels *SCORE-metabolite-ID* und weiteren selektiven und zweidimensionalen NMR-Experimenten untersucht, wodurch die sekundären Metaboliten Triandrin (**66**) sowie ein Vomifoliol-Disaccharid Glycosid (**68**) identifiziert werden konnten.

Aufgrund der Tatsache, dass komplexe Stoffgemische nicht nur in natürlichen Matrices auftreten, sondern auch das Produkt von Synthesen sein können, wurde die Anwendbarkeit der Methode *SCORE-metabolite-ID* in Bezug auf eine Reaktionsverfolgung betrachtet. Die Anwendung der Methode unterscheidet sich dabei insofern, als dass die zeitlichen Konzentrationsveränderungen nicht durch eine chromatographische Fraktionierung induziert werden müssen, sondern sich inhärent aus dem Reaktionsverlauf ergeben. Die Korrelation von NMR- und MS-Daten kann in diesem Kontext zur Identifizierung von Intermediaten und Nebenprodukten genutzt werden, um detaillierte Kenntnisse über den Reaktionspfad zu erlangen. In der vorliegenden Arbeit wurde die Synthese von 2,4,6-Tri-(6-aminocaprönsäure)-1,3,5-triazin (**78**) (TACT) ausgehend von Cyanurchlorid (**75**) und 6-Aminocaprönsäure (**77**) durch nucleophile aromatische Substitution unter alkalischen Bedingungen mittels eines dreistufigen Temperaturprofils analysiert. Durch die Korrelation der ^{13}C -NMR- und MS-Daten konnten die Zwischen- und Nebenprodukte identifiziert und die charakteristischen NMR-Signale zugeordnet werden. In Bezug auf den Reaktionspfad ergaben die Untersuchungen, dass sich der Substitutionsgrad des 1,3,5-Triazin-Rings entgegen der Annahme nicht durch die gewählten Temperaturstufen thermisch kontrollieren lässt. Vielmehr wird die Selektivität der Reaktion durch die Bildung des Mono-Hydroxy-substituierten Zwischenproduktes bestimmt, während die Hydrolyse im weiteren Verlauf der Reaktion eine untergeordnete Rolle einnimmt.

Da die in der *SCORE-metabolite-ID* Methode verwendeten eindimensionalen ^1H -NMR-Spektren trotz chromatographischer Separation teilweise starke Signalüberlagerungen aufweisen, wurde die erweiterte Methode *SCORE-metabolite-ID*² entwickelt. Bei dieser Methode können zweidimensionale HSQC-Spektren analysiert und mit den entsprechenden MS-Daten korreliert werden. Die Methode wurde anhand eines weiteren artifiziellen Gemisches validiert. Es konnte gezeigt werden, dass bei ausreichender Konzentration der Analyten eine präzisere Identifizierung der Metaboliten aus dem Gemisch ermöglicht wird und dementsprechend noch komplexere Stoffsysteme analysiert werden können.

Durch die Entwicklung von *SCORE-metabolite-ID* und *SCORE-metabolite-ID*² wurden praktische Hilfsmittel zur Identifizierung von Substanzen aus komplexen Stoffgemischen bereitgestellt. In Folge ihrer Anwendung konnten und können wertvolle Erkenntnisse gewonnen werden, wobei auf die zeit- und ressourcenintensive Isolierung einzelner Analyten verzichtet werden kann. *SCORE-metabolite-ID* und *SCORE-metabolite-ID*² ergänzen dabei die bereits etablierten Verfahren zur Untersuchung komplexer Stoffgemische und verbinden die komplementären Daten der NMR-Spektroskopie und der Massenspektrometrie.

2 Abstract

Metabolomics studies focus on the investigation of metabolites in biological systems and are becoming increasingly important for example in the fields of medicine and food research. Due to the variety of substances and the large differences in concentration, the investigation of complex mixtures of substances from biological matrices is associated with particular challenges.

Usually, statistical methods are applied in metabolomics studies to identify individual signals or signal patterns that can serve as biomarkers for addressing various research topics. While the statistical analysis is independent of the specific knowledge of the chemical composition of the sample, biochemical interpretation requires the identification of these marker substances. Despite the use of modern and high-resolution analytical methods, the complexity of the substance mixtures is reflected in the analytical data, for example, in the form of signal overlaps. The identification of individual substances from the mixture spectra is often not possible, necessitating the use of specific and complex steps of isolation and purification for the reliable identification of metabolites. The provision of technologies for the simplified identification of individual substances from these complex substance mixtures thus represents a central challenge.

In the present study, a method for the semi-automatic correlation of NMR and MS data for the simplified identification of metabolites (SCORE-metabolite-ID (Semi-automatic CORrelation analysis for RELiable metabolite Identification)) was developed for this purpose. The underlying concept is based on the method 3DCC (three-dimensional cross-correlation), which has been employed in order to determine the stereochemistry of glycans. SCORE-metabolite-ID was developed as a MATLAB application and is used for the efficient and reliable identification of metabolites of different substance classes from mixtures without the need for individual isolation. The method is based on the three-dimensional correlation of ^1H NMR and MS data over the third dimension of the time course. The temporal dimension is generated by a chromatographic fractionation of the mixture. Using the graphical user interface, NMR and MS data can be easily correlated to assign NMR signals to the corresponding mass-to-charge ratios from ESI mass spectra. This simplifies the rapid dereplication of known metabolites and the

identification of new compounds in complex mixtures. Figure 2 below provides a schematic representation of the method developed and the applications described in this thesis.

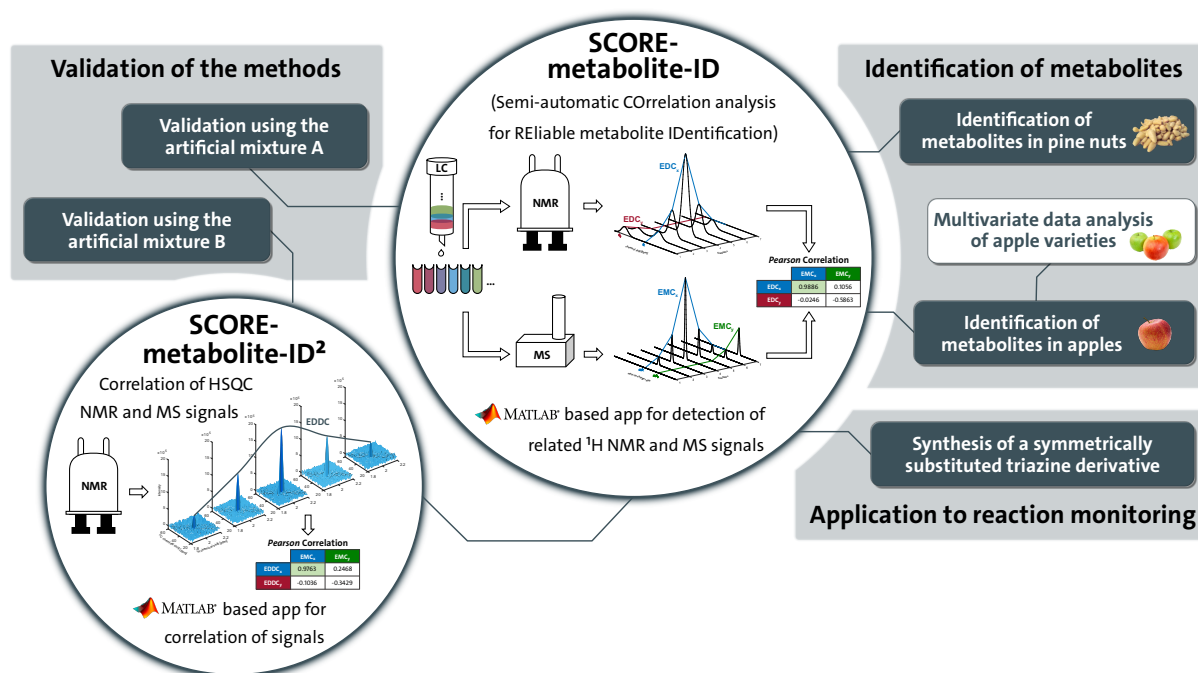


Figure 2: Schematic representation of the SCORE-metabolite-ID and SCORE-metabolite-ID² method, developed as a part of this research project, and its application as described in this thesis.

First, the SCORE-metabolite-ID method was validated using an artificial mixture whose composition simulates the challenges of real substance mixtures. These include the occurrence of isobaric and co-eluting substances, whereby the flexibility of the SCORE-metabolite-ID method made it possible to correlate and identify the analytes even in these challenging cases. Subsequently, it was applied to various real extracts of biological matrices.

Initially, a polar extract of a pine nut sample was examined as a complex matrix of biological origin. A total of 45 metabolites of different substance classes (including carbohydrates, betaines, amino acids, organic acids, etc.) and different concentration ranges were identified. Among the metabolites identified in the pine nut extract, the secondary metabolites β -D-glucopyranosyl-1-*N*-indole-3-acetyl-*N*-L-aspartic acid (**56**) and *N*^α-(2-hydroxy-2-carboxymethylsuccinyl)-L-arginine (**57**) are to be emphasized in particular. While further experiments were necessary for the unambiguous structural elucidation of these metabolites, the SCORE-metabolite-ID method facilitated and accelerated the identification considerably by correlating the NMR and MS data.

As a further complex matrix of biological origin, apple samples were analyzed. The ^1H NMR spectra of the 86 apple samples of the varieties Boskoop, Elstar, Gala, Holsteiner Cox and Wellant were first analyzed using multivariate data analysis with regard to variety discrimination. After feature selection using the *Boruta* algorithm, a classification model based on the Support Vector Machine algorithm was created for the differentiation of the apple varieties with a classification accuracy of $90.4\% \pm 1.8\%$. Subsequently, potential marker signals were specifically investigated using SCORE-metabolite-ID and further selective and two-dimensional experiments, whereby the secondary metabolites triandrin (**66**) and a vomifolioside disaccharide glycoside (**68**) could be identified.

Due to the fact that complex mixtures of substances do not only occur in natural matrices, but can also be the product of chemical syntheses, the applicability of the SCORE-metabolite-ID method was considered in relation to reaction monitoring. The application of the method differs in that the temporal concentration changes do not have to be induced by chromatographic fractionation but result inherently from the time course of the reaction. The correlation of NMR and MS data can be used in this context to identify intermediates and by-products in order to gain detailed knowledge of the reaction pathway. In the present work, the synthesis of 2,4,6-tri-(6-aminocaproic acid)-1,3,5-triazine (**78**) (TACT) starting from cyanuric chloride (**75**) and 6-aminocaproic acid (**77**) by nucleophilic aromatic substitution under alkaline conditions was analyzed using a three-step temperature profile. By correlating the ^{13}C NMR and MS data, the intermediates and by-products could be identified, and the characteristic NMR signals be assigned. With regard to the reaction pathway, the investigations revealed that, contrary to the assumption, the degree of substitution of the 1,3,5-triazine ring cannot be thermally controlled by the selected temperature levels. Rather, the selectivity of the reaction is determined by the formation of the mono-hydroxy-substituted intermediate, while hydrolysis plays a subordinate role in the further course of the reaction.

As the one-dimensional ^1H NMR spectra used in the SCORE-metabolite-ID method sometimes show strong signal overlaps despite chromatographic separation, the extended method SCORE-metabolite-ID² was developed. With this method, two-dimensional HSQC spectra can be analyzed and correlated with the corresponding MS data. The method was validated using another artificial mixture and it was shown that with sufficient concentration of the analytes, a

more precise identification of the metabolites from the mixture is possible and accordingly even more complex substance systems can be analyzed.

The development of SCORE-metabolite-ID and SCORE-metabolite-ID² has provided practical tools for the identification of substances from complex mixtures. Consequently, valuable results could and can be achieved without the necessity for time-consuming and resource-intensive isolation of individual substances. SCORE-metabolite-ID and SCORE-metabolite-ID² add to the already established methods for the analysis of complex substance mixtures and combine the complementary data of NMR spectroscopy and mass spectrometry.

3 Einleitung

Als Grundbausteine jeglicher Stoffwechselwege innerhalb eines Organismus spielen die Metaboliten, die die Edukte, Zwischenprodukte oder Produkte des Metabolismus darstellen, eine wichtige Rolle. Sowohl im medizinischen Bereich als auch im Bereich der Pflanzen und Lebensmittel gewinnen *Metabolomics*-Studien deshalb immer mehr an Bedeutung. Dabei befassen sich derartige Untersuchungen häufig mit Fragestellungen zur Authentizität, bspw. in Bezug auf Lebensmittel unterschiedlicher geographischer Herkunft oder zur Sortenunterscheidung.^[1] Nicht zuletzt weisen Vorfälle wie der Melamin-Skandal in China im Jahr 2008 oder die als Rindfleisch verkauften Pferdefleisch-Produkte im Jahr 2013 auf die Relevanz der *Metabolomics*- und Authentizität-Studien im Bereich der Lebensmittel hin.^[2,3]

Neben reinen Untersuchungen zur Klassifizierung verschiedener Probengruppen eröffnet die weiterreichende Analyse der *Metabolomics* Daten die Möglichkeit der Detektion von spezifischen Metaboliten, die sich innerhalb der Probengruppen unterscheiden. Für ein tiefergehendes Verständnis der zugrundeliegenden Stoffwechselprozesse und der biochemischen Bedeutung einzelner Moleküle ist die Identifizierung der (Marker-)Metaboliten demzufolge von großem Interesse.

Aufgrund der zahlreichen parallel ablaufenden Stoffwechselprozesse bestehen die verschiedenen Matrices im Allgemeinen allerdings aus einer Vielzahl verschiedener Metaboliten, die jeweils eine große Diversität an chemischen, strukturellen, pharmakologischen sowie funktionellen Eigenschaften aufweisen und zusätzlich innerhalb der gleichen Matrix in großen Konzentrationsunterschieden auftreten.^[4,5] Im Rahmen von *Metabolomics*-Studien werden demzufolge in der Regel hochkomplexe Gemische analysiert. Die Identifizierung einzelner Metaboliten aus derartig komplexen Gemischen gilt deshalb als große Herausforderung.^[6]

Sowohl im Rahmen von *Metabolomics*-Studien als auch zur Identifizierung von Metaboliten werden meist die fortschrittlichen Technologien NMR-Spektroskopie oder Massenspektrometrie eingesetzt. Beide Analysetechniken bergen individuelle Vorteile und die entsprechenden Daten beinhalten einen hohen Informationsgehalt. Die Komplementarität der analytischen Daten birgt außerdem ein hohes Potenzial zur Korrelation der Informationen.^[7]

3.1 Metabolomics-Studien

Metabolomics bezeichnet einen Forschungsbereich, der sich mit der umfassenden Analyse von Metaboliten, die z.B. innerhalb einer Zelle, eines Gewebes oder eines Organismus auftreten, befasst.^[8] Neben *Genomics*, *Transcriptomics* und *Proteomics* gliedert sich *Metabolomics* in die sogenannte *Omics*-Kaskade ein, die in Abbildung 3 dargestellt ist.^[9] In dieser linearen Darstellung der *Omics*-Kaskade bildet das Metabolom ausgehend vom Genom prinzipiell das Endprodukt der Transkription, Translation und der enzymatischen Katalyse, weshalb es eher dem Phänotyp des Organismus entspricht.^[10] Da allerdings sowohl die Metaboliten als auch die Proteine durch *feedback loops* und Signalkaskaden die vorangegangenen Schritte der Transkription und Translation beeinflussen, kann die *Omics*-Kaskade als bidirektional angesehen werden.^[10] Die lineare Darstellung in Abbildung 3 ist daher eine stark vereinfachte Form der *Omics*-Kaskade.

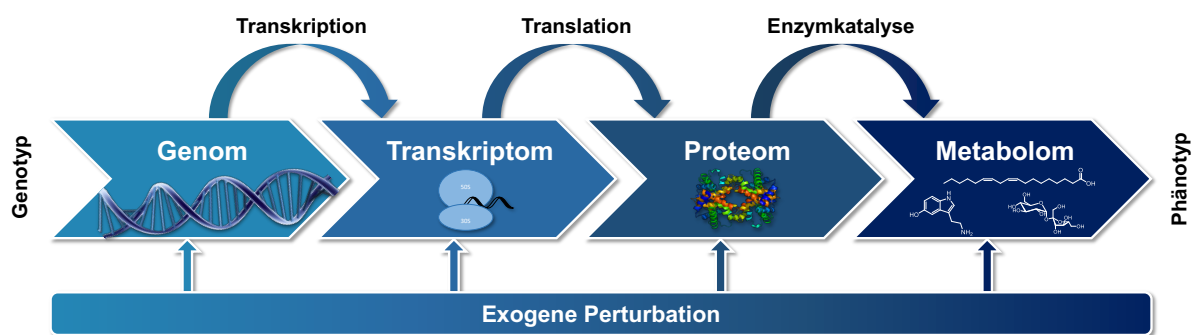


Abbildung 3: Schematische lineare Darstellung der *Omics*-Kaskade bestehend aus den vier *Omics*-Disziplinen *Genomics*, *Transcriptomics*, *Proteomics* und *Metabolomics*, die sich jeweils mit der Untersuchung des Genoms, Transkriptoms, Proteoms und des Metaboloms beschäftigen. Die Abbildung wurde nach SUSSULINI verändert.^[9]

Das Ziel von *Metabolomics*-Studien ist die Untersuchung einer Gruppe von Individuen und deren Unterscheidung, insbesondere in Hinblick auf Reaktionen bzw. Veränderungen von Biomolekülen durch exogene Einflüsse.^[9] Mögliche exogene Einflüsse können beispielsweise Pathogene (Toxine, Viren, Bakterien etc.) oder in Bezug auf pflanzliche Organismen im Speziellen die Wasser- und Nährstoffverfügbarkeit am Standort, herrschende Klimabedingungen sowie weitere Umwelteinflüsse aus der direkten Umgebung sein.^[10] *Metabolomics*-Studien befassen sich dabei sowohl mit der Identifizierung als auch mit der Quantifizierung der intra- und extrazellulären Metaboliten in dem jeweilig untersuchten biologischen System. Als Metaboliten werden dabei kleine organische Moleküle mit einem

Molekulargewicht von bis zu 1500 Da bezeichnet, die als Substrate, Zwischenprodukte und Produkte bei metabolischen Reaktionen auftreten können.^[5,11,12] Metaboliten lassen sich in primäre und sekundäre Metaboliten einteilen. Primäre Metaboliten sind essenziell für das Wachstum, die Energieversorgung und allgemeine Lebensfunktionen eines Organismus und liegen hochgradig konserviert vor. Dazu zählen beispielsweise Aminosäuren, Nukleotide, Vitamine, Fettsäuren und Kohlenhydrate.^[4,13] Zu den sekundären Metaboliten zählen beispielsweise Alkaloide, Flavonoide und Isoprene.^[12,14] Sekundären Metaboliten sind nicht direkt für die allgemeinen Lebensfunktionen verantwortlich, unterstützen den Organismus allerdings in Funktionen wie beispielsweise der antibiotischen Aktivität gegenüber Schadorganismen und der Zellkommunikation.^[13] Das Metabolom besteht dementsprechend aus einer Vielzahl primärer und sekundärer Metaboliten, die jeweils unterschiedliche chemische Eigenschaften sowie starke Konzentrationsunterschiede untereinander aufweisen, was die simultane Analyse und Identifizierung der Metaboliten erschwert.^[4,5]

Neben *Metabolomics*-Studien im pharmakologischen und klinischen Bereich stellen insbesondere Untersuchungen im Bereich der Lebensmittelanalytik einen relevanten Aspekt dar, um beispielsweise Lebensmittelbetrug (*Food Fraud*) durch Umdeklaration des Herkunftslandes oder durch Beimengungen von Lebensmitteln minderer Qualität zu analysieren.^[1,11,15–17] Im Rahmen von *untargeted* Metabolomics-Studien werden dafür eine Vielzahl von Proben mit geeigneten analytischen Verfahren wie z.B. der NMR-Spektroskopie analysiert und sogenannte *Metabolic Fingerprints* generiert.^[5,18] Diese Gemischspektren dienen als eine Art Fingerabdruck und können mit Hilfe von multivariater Datenanalyse in Bezug auf bestimmte Fragestellungen wie der Herkunft oder der Sorte untersucht werden.^[18] So gibt es beispielsweise verschiedene Studien zur Untersuchung von Äpfeln oder Apfelsäften bezüglich ihrer Sorte basierend auf unterschiedlichen analytischen Methoden und verschiedenen Ansätzen der multivariaten Datenanalyse.^[19–22] Im Gegensatz dazu zielen *targeted* Studien darauf ab, spezifische Biomarker gezielt zu identifizieren und zu quantifizieren, die beispielsweise durch die vorherige *untargeted* Analyse detektiert worden sind.^[23] So konnte beispielsweise in einer Studie von TOMITA *et al.* das Alditol L-Rhamnitol (**1**) als Biomarker zur Unterscheidung von Äpfeln aus Japan oder Neuseeland detektiert und identifiziert werden.^[24]

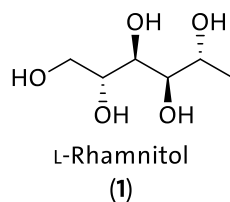


Abbildung 4: Strukturformel des Alditols L-Rhamnitol (1), das in einer Studie von TOMITA *et al.* als Biomarker zur Unterscheidung von neuseeländischen und japanischen Äpfeln identifiziert wurde.^[24]

Dabei ist hervorzuheben, dass im Rahmen von NMR-basierten *Metabolomics*-Studien bei geeigneten Messbedingungen dieselben Daten sowohl (zunächst) für *untargeted* Analysen als auch, z. B. nach erfolgreicher Identifizierung, für *targeted* Analysen verwendet werden können. Die Unterscheidung von *targeted* und *untargeted Metabolomics*-Studien ist in diesem Fall also vielmehr von der Fragestellung, der vorliegenden Informationen und der Datenauswertung abhängig.

3.2 Multivariate Datenanalyse

Da beispielsweise im Falle von NMR-Daten für *Metabolomics*-Studien große Datenmengen mit einem enormen Informationsgehalt erhalten werden, ist die Verwendung von multivariater Datenanalyse, wie z. B. einer Hauptkomponentenanalyse (PCA, *principal component analysis*), eine geeignete Methode zur Datenreduktion unter Beibehaltung der wesentlichen Informationen.^[25,26] Multivariate Datenanalyse kann sowohl als gerichtete sowie auch als ungerichtete Methode durchgeführt werden.^[25] Ein Beispiel der ungerichteten Datenanalyse ist die Hauptkomponentenanalyse, die in Abschnitt 3.2.1 genauer erläutert wird. Zusätzlich können im Rahmen von gerichteter multivariater Datenanalyse Klassifizierungsmodelle erstellt werden, wodurch die Zuordnung unbekannter Proben entsprechend einer zuvor festgelegten Fragestellung ermöglicht wird. Diese Modelle können auf unterschiedlichen Algorithmen wie Diskriminanzanalysen, *Random Forest* oder *Support Vector Machine* (SVM) Klassifikatoren beruhen. Klassifizierungsverfahren basierend auf dem SVM Algorithmus werden in Abschnitt 3.2.2 erläutert.

Während der Datenprozessierung wird für die multivariate Datenanalyse von NMR-Spektren meist ein *Bucketing* durchgeführt. Dabei wird das NMR-Spektrum in Bereiche eingeteilt, in denen die Signalintensitäten integriert werden.^[27,28] Dieser Schritt der Datenprozessierung dient

nicht nur der Datenreduktion, sondern auch der Kompensierung von Unterschieden in der exakten chemischen Verschiebung einiger Signale desselben Moleküls in unterschiedlichen Proben, die sich aufgrund von Unterschieden im pH-Wert oder der Ionenstärke der Probe ergeben können.^[27] Die für das *Bucketing* erforderliche Bereiche (*Buckets*) werden im Fall des automatischen *Bucketings* in einem immer gleichmäßig verteilten Abstand zueinander definiert (z. B. 0.03 ppm).^[27,28] Ein Nachteil des automatischen *Bucketings* ist, dass ein Signal teilweise in mehrere *Buckets* unterteilt werden kann, was anschließend zu Fehlberechnungen und Missinterpretationen der PCA führen kann.^[27,28] Alternativ können die *Buckets* in variable Bereiche unterschiedlicher Größe eingeteilt werden, die individuell für die untersuchte Matrix und die darin auftretenden Signale angepasst werden können.^[27]

Ein weiteres Ziel der multivariaten Datenanalyse ist die Ermittlung von wichtigen Variablen für das Modell der entsprechenden Fragestellung.^[26,28,29] Die Variablenselektion (*Feature Selection*) führt zusätzlich zu simpleren und robusteren Modellen, da mit der Variablenselektion eine weitere Datenreduktion einhergeht.^[26] Des Weiteren ermöglicht die Variablenselektion die Extraktion relevanter Informationen aus den NMR-Spektren, was z.B. der Detektion von Biomarkern dient.^[30] Ein möglicher Algorithmus zur Variablenselektion ist *Boruta*. Der *Boruta* Algorithmus basiert auf dem *Random Forest* Klassifizierungsalgorithmus und ermittelt die Wichtigkeit einer Variablen durch den Verlust an Klassifizierungsgenauigkeit des Modells, wenn Variablen zwischen den Proben zufällig permutiert werden.^[31] Für diese zufällige Permutation der Variablen wird die ursprüngliche Datenmatrix um Schattenvariablen erweitert, die Kopien der Variablen darstellen, dessen Variablenwerte randomisiert sind.^[31] Daraufhin wird ein Klassifizierungsmodell mit Hilfe des *Random Forest* Algorithmus erstellt und zur Evaluation der Wichtigkeit für jede Variable ein *z-Score* berechnet. Der höchste *z-Score*, der für eine der Schattenvariablen erhalten wird, wird als Schwellenwert festgelegt. Die ursprünglichen Variablen werden daraufhin anhand ihres *z-Scores* als für die jeweilige Fragestellung relevant oder irrelevant evaluiert und dementsprechend selektiert.^[31]

3.2.1 Hauptkomponentenanalyse

Die Hauptkomponentenanalyse zählt zu den ungerichteten multivariaten Datenanalyseverfahren und dient oftmals zur ersten Visualisierung des Datensatzes und zur

Untersuchung möglicher Unterscheidungen oder Verwandtschaften zwischen bestimmten Gruppen.^[28] Da es sich um eine ungerichtete Methode handelt, ist allerdings während der Analyse die Gruppenzugehörigkeit (z.B. die Herkunft oder Sorte) nicht bekannt. Stattdessen wird mit Hilfe der Hauptkomponentenanalyse die maximale Varianz der Daten abgebildet. Die Hauptkomponenten (*principal components*, PC) werden durch Linearkombination der ursprünglichen Variablen generiert.^[32] Dazu werden die Originalvariablen orthogonal transformiert, um neue, unkorrelierte, latente Variablen zu erzeugen, die bei gleichzeitiger Datenreduktion die Varianz der Originalvariablen bestmöglich abbilden.^[33] Das Prinzip der PCA ist graphisch in der folgenden Abbildung 5 dargestellt. Die schwarzen Messpunkte sind in dem zweidimensionalen Koordinatensystem aus den Originalvariablen 1 und 2 aufgetragen. Die Hauptkomponente 1 soll die größte Varianz der Daten aufzeigen und wird deshalb so in das Koordinatensystem gelegt, dass insgesamt der kürzeste Abstand zwischen den Messpunkten (schwarze Punkte) zu den Projektionen auf der PC 1 (weiße Punkte) generiert wird.^[32,34] Die Ermittlung der zweiten Hauptkomponente erfolgt analog, sodass diese die zweitgrößte Varianz der Daten abbildet.

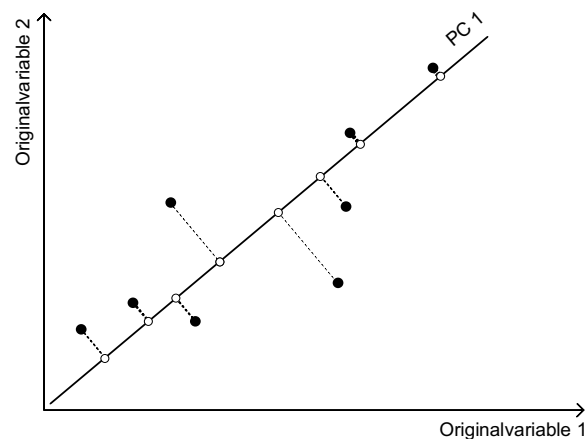


Abbildung 5: Schematische Darstellung des Prinzips der Hauptkomponentenanalyse. Die schwarzen Punkte entsprechen den Messpunkten, die im zweidimensionalen Koordinatensystem der Originalvariablen 1 und 2 vorliegen. Die weißen Punkte entsprechen der Projektion der Messpunkte auf die Hauptkomponente 1 (PC 1), die hier die maximale Varianz der Messpunkte aufweist, da so die Verbindungslinien (gestrichelte Linien) insgesamt in dieser Ausrichtung der PC 1 die kürzeste Strecke aufweisen. Die Abbildung wurde verändert nach SCHLITGEN.^[32]

In den sogenannten *Score Plots* einer PCA werden die Hauptkomponenten in einem neuen Koordinatensystem gegeneinander aufgetragen, wobei der Ursprung im Mittelwert der Datenpunkte liegt. Jeder Datenpunkt innerhalb des *Score Plots* entspricht einer Probe.^[25] Durch das Einfärben der Datenpunkte entsprechend ihrer Gruppenzugehörigkeit können so im *Score*

Plot Ähnlichkeiten (Aggregation der Datenpunkte) oder Unterschiede (Separation der Datenpunkte) visualisiert werden sowie Ausreißer erkannt werden.^[25] Eine Gruppenseparation liegt dann vor, wenn die Varianz innerhalb der Gruppe kleiner ist als die Varianz zwischen den Gruppen.^[28] Des Weiteren werden *Loading* Plots erhalten, die den Einfluss der Originalvariablen auf eine mögliche Separation oder Aggregation der Datenpunkte im *Score* Plot erklären. Jeder Datenpunkt im *Loading* Plot entspricht demnach einer Originalvariablen. Je weiter der Datenpunkt vom Ursprung entfernt liegt, desto größer ist der Einfluss dieser Variablen auf die erklärte Varianz der Hauptkomponente, die in die gleiche Richtung verläuft.^[25,28] Durch das Zusammenspiel von *Score* und *Loading* Plot können somit die Einflüsse einzelner Variablen auf die Unterscheidung bestimmter Proben untersucht werden.^[25]

3.2.2 Klassifizierungsverfahren

Als Klassifizierung wird die Vorhersage von qualitativen Variablen, wie im Fall von pflanzlichen Proben bspw. die Herkunft oder die Sorte, auf Grundlage eines zuvor erstellten Modells bezeichnet.^[35] Klassifizierungsverfahren zählen demnach zu den gerichteten multivariaten Analyseverfahren.^[36] Zur Erstellung von Klassifizierungsmodellen existieren verschiedenste Algorithmen, wie zum Beispiel die lineare Diskriminanzanalyse, *k-nearest neighbors*, *Random Forest*, *Support Vector Machine* (SVM) etc.^[35,37] Unabhängig vom Klassifizierungsalgorithmus wird zur Erstellung eines Klassifizierungsmodells der Datensatz zunächst in einen Trainingsdatensatz und einen Testdatensatz unterteilt. Mit Hilfe des Trainingsdatensatzes wird vorerst das Modell erstellt und anschließend wird die Zuordnungsgenauigkeit mit Hilfe des Testdatensatzes evaluiert.^[36] Im Folgenden ist der in dieser Arbeit verwendete Algorithmus der *Support Vector Machine* genauer erläutert.

Prinzipiell wird bei dem SVM Algorithmus die Funktion der optimalen Hyperebene ermittelt, durch die der maximale Abstand zwischen den Datenpunkten der verschiedenen Klassen erreicht werden kann.^[38] Die Hyperebene entspricht dabei jeweils einer Funktion, dessen Dimension um eine geringer ist, als die Dimensionalität des Originaldatensatzes.^[38] Die SVM ist schematisch in Abbildung 6 gezeigt.^[34,39] In diesem zweidimensionalen Datensatz entspricht die Hyperebene demnach der blau-grünen Linie.

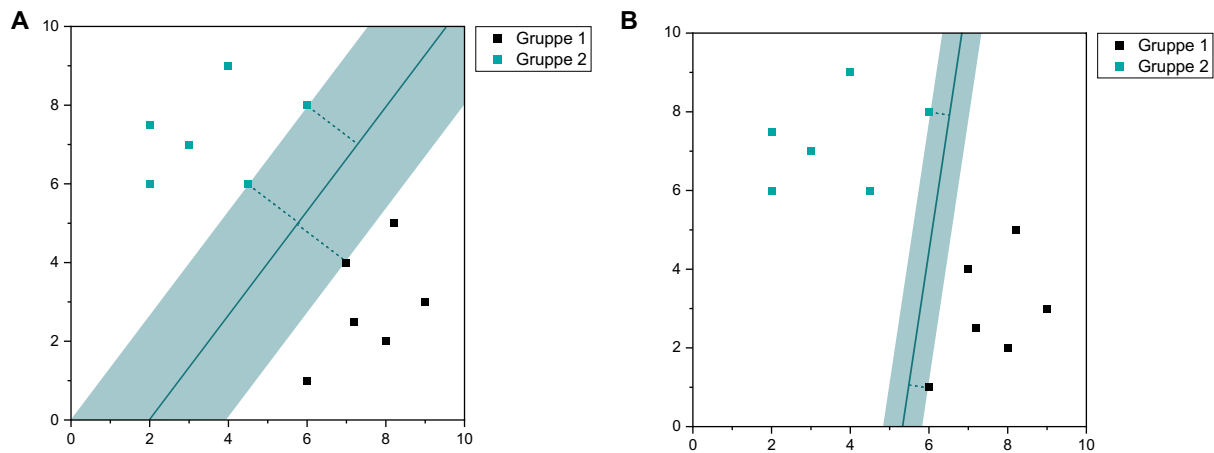


Abbildung 6: Schematische Darstellung des Klassifizierungsverfahrens mittels *Support Vector Machine* Algorithmus im zweidimensionalen Raum. Die beiden beispielhaften Gruppen 1 und 2 sind in verschiedenen Farben gekennzeichnet. Die Trennung der Gruppen ist durch die Konstruktion verschiedener Hyperebenen (dargestellt durch die dunkelblaue Linie) denkbar. Beispielhaft sind zwei der Möglichkeiten in **A** und **B** gezeigt. Das Ziel des *Support Vector Machine* Algorithmus ist die Maximierung der Fläche des Randbereichs (dargestellt durch die farblich hinterlegte Fläche). Dies wird durch die Maximierung des Abstandes der Hyperebene zu denjenigen Datenpunkten, die dieser am nächsten liegen, erreicht. Zur Ermittlung des Abstandes werden die Stützvektoren (gestrichelte Linien) verwendet, die orthogonal zur Hyperebene stehen. In diesem Fall ist die Trennung der Gruppen in **A** zu bevorzugen, da diese aufgrund der größeren Fläche des Randbereiches mit einer höheren Wahrscheinlichkeit eine korrekte Klassifizierung unbekannter Proben zulässt.^[34,39]

Die optimale Hyperebene der SVM entspricht derjenigen, bei der der Randbereich (*margin*) maximal ist. Dieser Randbereich ist größer, je größer der Abstand zwischen der Hyperebene zu den nächsten Datenpunkten, den *Support Vectors* (Stützvektoren), ist.^[34] Zur Ermittlung der optimalen Hyperebene werden demnach nur die Stützvektoren benötigt. Alle weiteren Datenpunkte, die sich von der Hyperebene im Vektorraum weiter entfernt befinden, werden für die mathematische Berechnung der Hyperebene nicht benötigt.^[40] Nach Erstellung des Modells erfolgt die Klassifizierung und Vorhersage unbekannter Proben durch eine Einordnung der Proben in den gleichen Raum und einer kategorischen Zuordnung, je nachdem auf welcher Seite der Hyperebene die Datenpunkte liegen.^[38]

3.2.3 Validierung

Insbesondere bei Verwendung von gerichteten multivariaten Datenanalyseverfahren ist eine Validierung des Modells unerlässlich. Ohne geeignete Validierungsverfahren kann beispielsweise ein *Overfitting* oder ein *Underfitting* unerkannt bleiben, was die Robustheit eines

Klassifizierungsmodells stark beeinträchtigt. Zum *Overfitting* kommt es häufig aufgrund von einer großen Anzahl an Variablen, die teilweise nur Rauschen enthalten.^[38,41] *Overfitting* äußert sich in einer sehr hohen Klassifizierungsgenauigkeit des Trainingsdatensatzes und dem gegenüberstehend einer geringen Klassifizierungsgenauigkeit des Testdatensatzes, welcher unbekannte Proben enthält.^[38] Eine Möglichkeit, um *Overfitting* zu vermeiden, ist eine Selektion der Variablen, um ein Gleichgewicht zwischen Probenanzahl zu Variablenanzahl herzustellen.^[26,29,42] *Underfitting* tritt dahingegen meist bei einer geringen Anzahl an Proben auf, wodurch ein simples Modell berechnet wird, welches die Realität nur in geringem Maße abbildet.^[26,38] Im Fall von *Underfitting* sind sowohl die Klassifizierungsgenauigkeiten des Trainings- als auch des Testdatensatzes gering.^[38]

Die häufigste Methode der internen Validierung ist die Kreuzvalidierung (*Cross-Validation*).^[26] Diese Art der Validierung kann auch schon bei einer geringen Probenanzahl verwendet werden, da alle Proben sowohl für die Erstellung des Klassifizierungsmodells als auch für die Validierung dessen verwendet werden.^[26] Dazu wird der gesamte Datensatz bei einer k -fachen internen Kreuzvalidierung in k Teile aufgeteilt. Mit $k-1$ Teilen wird das Modell erstellt und anschließend wird der für das Modell unbekannten Datenteil k zugeordnet und somit die Vorhersagegenauigkeit abgeschätzt.^[26] Nach VARMA *et al.* führt die verschachtelte Kreuzvalidierung (*Nested Cross-Validation*) im Vergleich zur Kreuzvalidierung zu einer nahezu unverzerrten Schätzung des wahren Fehlers.^[43] Bei der verschachtelten Kreuzvalidierung werden mit Hilfe der inneren Kreuzvalidierung die Modellparameter eingestellt und mit Hilfe der äußeren Kreuzvalidierung das Modell bewertet.^[43] Da zusätzlich auch die Zusammensetzung des Trainings- und Testdatensatzes einen Einfluss auf die Klassifizierungsgenauigkeit und den Vorhersagefehler hat, bietet sich die wiederholte Kreuzvalidierung (*Repeated Cross-Validation*) an.^[44] In Abbildung 7 ist die Vorgehensweise der Kombination dieser Validierungsverfahren, d.h. der wiederholten verschachtelten Kreuzvalidierung (*Repeated Nested Cross-Validation*, RNCV) gezeigt.^[45]

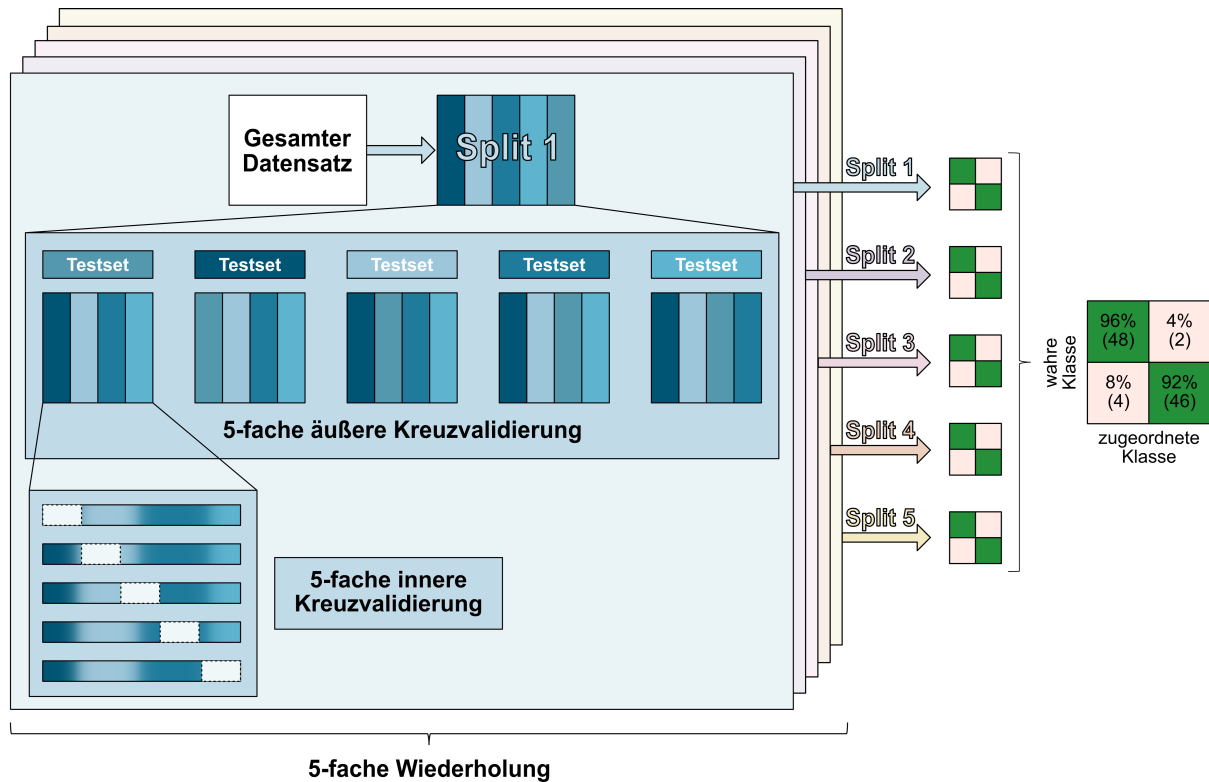


Abbildung 7: Schematische Darstellung der fünffachen *Repeated Nested Cross-Validation*. Der gesamte Datensatz wird in fünf Teile geteilt (z.B. Split 1 in blau). Jeweils vier Teile werden zu einem Trainingsdatensatz kombiniert, das Modell erstellt und mittels interner fünffacher Kreuzvalidierung optimiert. Das erhaltene Modell wird durch den übrigen fünften Teil des Split 1, der das Testset bildet, extern validiert. Diese Prozedur wird weitere vier Mal wiederholt, wobei jeweils eine andere Kombination der fünf Teile zur Erstellung des Trainingsdatensatzes verwendet wird. Das Ergebnis von Split 1 kann als Konfusionsmatrix dargestellt werden. Diese verschachtelte Kreuzvalidierung wird weitere vier Mal wiederholt, indem der gesamte Datensatz randomisiert in andere Splits eingeteilt wird (Split 2-5). Alle fünf erhaltenen Konfusionsmatrizen werden anschließend zu einer kombiniert. In diesem Fall handelt es sich beispielhaft um eine Zwei-Klassen-Klassifizierung, doch nach dem gleichen Prinzip ist auch die Validierung von Mehrklassenmodellen möglich.^[45]

Für die fünffache RNCV wird der gesamte Datensatz zunächst in fünf Teile geteilt. In diesem Split 1 werden daraufhin vier der fünf Teile zu einem Trainingsdatensatz kombiniert. Mit Hilfe dieses Trainingsdatensatzes wird das Modell erstellt und mittels fünffacher innerer Kreuzvalidierung optimiert. Der fünfte Teil des ersten Splits dient anschließend zur externen Validierung des erstellten Modells. In der Schleife der äußeren fünffachen Kreuzvalidierung wird diese Prozedur weitere vier Mal wiederholt, wobei jeweils ein anderer der fünf Teile das Testset bildet. Insgesamt wird der gesamte Prozess weitere vier Mal mit anderen Splits wiederholt, sodass die Zusammensetzung von Proben im Trainings- und Testdatensatz randomisiert variiert wird. Aus jeder verschachtelten Kreuzvalidierung wird eine Konfusionsmatrix erhalten. Nach der

fünffachen Wiederholung werden alle fünf Konfusionsmatrizen zu einer kombiniert, sodass eine Gesamtgenauigkeit der Klassifizierung inkl. Abweichung für das durch die RNCV validierte Modell angegeben werden kann.^[45]

3.3 Klassische analytische Methoden zur Identifizierung von Metaboliten

Durch Verwendung von multivariater Datenanalyse eines großen Probensatzes ist es möglich, Biomarker zu detektieren, die bspw. eine Unterscheidung von Proben unterschiedlicher Kategorien zulassen. Im Fall von pflanzlichen Proben handelt es sich dabei häufig um Eigenschaften wie Herkunft oder Sorte, im Fall von medizinischen Proben geht es häufig um die Detektion von Biomarkern, die eine (frühzeitige) Erkennung von Krankheiten zulassen.^[15,17,36] Die Identifizierung von Biomarkern ist essenziell für ein tiefergehendes Verständnis der zugrundeliegenden Stoffwechselprozesse.^[46] So erlaubt beispielsweise die Analyse von biologisch aktiven Onkometaboliten, die sich in verschiedenen Arten von Krebszellen akkumulieren können, die Erklärung der bösartigen Entartung der Zellen.^[46,47] Die Akkumulation der Metaboliten ist auf Mutationen in den für Enzyme kodierenden Genen zurückzuführen, die somit zur vermehrten Bildung von bestimmten Metaboliten führen.^[47]

Zur allgemeinen Identifizierung von chemischen Substanzen stehen zahlreiche analytische Methoden zur Verfügung. Schon seit den 1940er Jahren wird die UV/VIS-Spektroskopie sowie seit den 1950er Jahren die IR-, NMR-Spektroskopie und Massenspektrometrie (MS) verwendet.^[48] Massenspektrometrische Analysen werden zusätzlich häufig durch Kopplung mit Flüssigchromatographie (LC) oder Gaschromatographie (GC) durchgeführt. Insbesondere die NMR-Spektroskopie, aber auch die Massenspektrometrie eignen sich für die Identifizierung von bekannten sowie unbekannten Metaboliten. Im Allgemeinen können verschiedene Stufen der Metaboliten-Identifizierung von bekannten oder unbekannten Metaboliten unterschieden werden. SUMNER *et al.* nennen vier Stufen der Metaboliten-Identifizierung: 1) eindeutig identifizierte Verbindungen (nicht-neuartige Metaboliten), 2) vermeintlich annotierte Verbindungen (auf der Grundlage der spektralen Ähnlichkeit mit Spektralbibliotheken), 3) vermeintlich charakterisierte Verbindungsklassen und 4) unbekannte Verbindungen.^[49,50] Nach WISHART sollte eine weitere Unterscheidung für die Identifizierung unbekannter Metaboliten in Betracht gezogen werden.^[48,50] "Bekannte Unbekannte" („*Known unknowns*“) entsprechen Metaboliten, die bereits bekannt sind, aber in der spezifischen Probe noch nicht identifiziert

wurden, während „unbekannte Unbekannte“ („*Unknown unknowns*“) tatsächlich neuartige Metaboliten sind.^[48] Nach Angaben der *Chemical Analysis Working Group* (CAWG) erfordert die Identifizierung sowohl von nicht-neuartigen als auch von neuartigen Metaboliten die Anwendung von zwei (oder mehr) orthogonalen und unabhängigen Analysemethoden, zu denen z.B. NMR und MS gehören.^[49] Die Identifizierung einer wirklich neuartigen Verbindung erfordert zusätzlich die Isolierung und Reinigung der Verbindung sowie weitere analytische Methoden wie Elementaranalyse, exakte Massenbestimmung etc.^[49]

3.3.1 Flüssigchromatographische Trennung von Gemischen

Die Trennung von Gemischen und Isolierung einzelner Analyten kann über zahlreiche chromatographische Verfahren bewerkstelligt werden. Im Bereich der Säulenchromatographie werden häufig Säulen mit Umkehrphase (*Reversed Phase*, RP) eingesetzt. Die unpolare stationäre Phase besteht meist aus hydrophoben C₁₈-Ketten, die kovalent an Silanol-Gruppen der Silica-Partikel gebunden sind. Die Elution erfolgt nach der Polarität der Analyten. Polare Analyten werden kaum retardiert, wohingegen unpolare Analyten eine starke Retention erfahren und erst durch den steigenden unpolaren Anteil der mobilen Phase eluiert werden.

Im Rahmen von NMR-basierten *Metabolomics*-Studien werden meist polare Metaboliten analysiert. Aufgrund der hohen Polarität ist die Verwendung von RP-Säulen häufig nicht zur Trennung oder Isolierung von Analyten geeignet.^[51,52] Eine Möglichkeit zur säulenchromatographischen Trennung von polaren Analyten mit einer organisch-wässrigen mobilen Phase stellt die Methode HILIC (*Hydrophilic Interaction Chromatography*) dar.^[53,54] Die HILIC-Methode wurde von ANDREW J. ALPERT im Jahr 1990 eingeführt und kann als Variante der Normalphasen-Chromatographie verstanden werden, bei der als mobile Phase Wasser und ein mit Wasser mischbares organisches Lösungsmittel (z.B. Acetonitril oder Methanol) statt eines rein organischen Lösungsmittels verwendet werden.^[51,55] Die Elution erfolgt mit steigendem Anteil des polaren Lösungsmittels. Die stationäre Phase hat eine polare Oberfläche durch kovalente Knüpfung der Silanol-Gruppen an z.B. Amino-, Diol-, Amid- oder Cyano-Substituenten. Auch zwitterionische Phasen aus Sulfoalkylbetain-Substituenten oder aromatische Fluorphenyl-Phasen sind kommerziell erhältlich.^[51,56,57] Nach ALPERT erfolgt die Elution prinzipiell aufgrund von Verteilung der Analyten zwischen der hydrophoben mobilen

Phase und der wässrigen, hydrophilen Schicht der mobilen Phase, die sich partiell immobilisiert auf der Oberfläche der stationären Phase ausbildet (siehe Abbildung 8).^[55] Damit unterscheidet sich der HILIC-Retentionsmechanismus grundlegend von dem der Normalphasen-Chromatographie, bei der die Retention hauptsächlich auf Adsorption der Analyten an die stationäre Phase basiert. Die Elution der Analyten bei der HILIC-Chromatographie erfolgt prinzipiell in umgekehrter Reihenfolge als bei der RP-Chromatographie, d. h. die Retention bei HILIC ist proportional zur Polarität des Analyten.^[55]

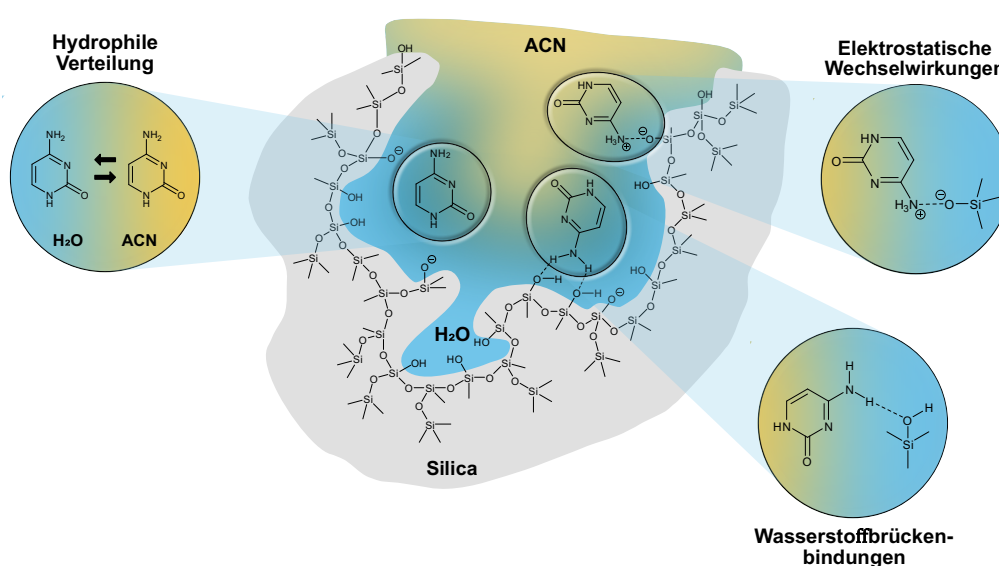


Abbildung 8: Der Retentionsmechanismus der HILIC-Methode basiert auf verschiedenen Effekten. Dazu zählen die hydrophile Verteilung des Analyten zwischen der hydrophoben mobilen Phase und der hydrophilen, wässrigen Schicht auf der polaren Oberfläche der Silica-Partikel, die Ausbildung von Wasserstoffbrückenbindungen zwischen Analyten und stationärer Phase sowie elektrostatische Wechselwirkungen. ACN: Acetonitril. Die Abbildung wurde verändert nach BUSZEWSKI *et al.*^[51]

Zusätzlich zum Retentionsmechanismus der Verteilung tragen auch die Ausbildung von Wasserstoffbrückenbindungen zwischen Analyten und der polaren stationären Phase sowie elektrostatische Wechselwirkungen zur Retention der Analyten bei (siehe Abbildung 8). Der tatsächliche Einfluss der verschiedenen Effekte ist abhängig von den chemischen Eigenschaften der stationären Phase sowie dem pH-Wert und dem Einsatz möglicher Additive in der mobilen Phase.^[52] Je nach stationärer Phase und einer Vielzahl weiterer Parameter wie dem pH-Wert, dem Puffer, der Pufferkonzentration etc. können HILIC-Säulen oftmals auch als *Mixed Mode* Phase mit partieller Nutzung in Form von Kationen- oder Anionenaustausch-Chromatographie genutzt werden.^[57,58]

Nach Trennung und Isolierung einzelner Metaboliten können diese mit klassischen analytischen Methoden wie der NMR-Spektroskopie oder der Massenspektrometrie analysiert und identifiziert werden. Die Isolierung stellt allerdings aufgrund der Vielzahl auftretender Metaboliten in biologischen Proben eine große Herausforderung dar.

3.3.2 Identifizierungsmethoden basierend auf NMR-Spektroskopie

Die NMR-Spektroskopie zählt zu den wichtigsten analytischen Methoden im Bereich der Strukturaufklärung und geht auf Arbeiten von FELIX BLOCH und EDWARD PURCELL in den 1940er Jahren zurück.^[59] Die NMR-Spektroskopie zeichnet sich durch eine hohe Reproduzierbarkeit bei minimaler Probenvorbereitung sowie durch die gleichzeitige Möglichkeit der Quantifizierung bei geeigneten Messparametern aus.^[60] NMR-Messungen sind nicht-destruktiv und im Fall von ^1H -NMR-Messungen mit einer Messdauer von einigen Minuten vergleichsweise schnell.^[60] Des Weiteren stellt die Automatisierung der NMR-Messungen einen Vorteil dar, wodurch die Messung einer großen Probenzahl einfach und schnell durchgeführt werden kann.^[60] Die Sensitivität ist dahingegen eher gering. So werden für adäquate Spektren je nach Experiment Stoffmengenkonzentrationen im Bereich von einigen Mikromol pro Liter benötigt, wodurch sich im Vergleich zur Massenspektrometrie ein mindestens 100-fach höherer Mengenbedarf des Analyten ergibt.^[60,61] Durch die Verwendung von NMR-Spektrometern bei höheren Magnetfeldstärken oder durch den Einsatz von Cryoprobenköpfen kann die Sensitivität zu einem gewissen Maße gesteigert werden.^[59,61] Die elektronischen Komponenten sind bei einem Helium-gekühlten Cryoprobenkopf auf ca. 20 K gekühlt, wodurch das elektronische Rauschen reduziert wird. Dadurch kann ein um bis zu vierfach höheres Signal-zu-Rausch-Verhältnis generiert werden.^[61] Während standardmäßig Probenköpfe mit einem Durchmesser von 5 mm eingesetzt werden, erlaubt der Einsatz von Micro-(Cryo-)Probenköpfen mit Durchmessern von 1 mm oder 1.7 mm eine weitere Möglichkeit der Sensitivitätssteigerung, die sich u. a. durch die Verringerung des Probenvolumens ergibt.^[60]

Gerade im Bereich der Strukturaufklärung ist die Verwendung der NMR-Spektroskopie essenziell, um Konnektivitäten von Atomen und Strukturelementen zu ermitteln. Neben ^1H -NMR-Spektren sind dafür insbesondere die zweidimensionalen NMR-Experimente von Bedeutung, wie z. B. die homonuklearen 2D-Experimente ^1H , ^1H -COSY (*Correlation Spectroscopy*),

^1H , ^1H -TOCSY (*Total Correlation Spectroscopy*) und ^1H , ^1H -NOESY (*Nuclear Overhauser Effect Spectroscopy*) sowie die heteronuklearen Korrelationsexperimente ^1H , ^{13}C -HSQC (*Heteronuclear Single Quantum Coherence*) und ^1H , ^{13}C -HMBC (*Heteronuclear Multiple Bond Correlation*).

Nichtsdestotrotz stellt die Identifizierung von bekannten und unbekannten Metaboliten aus einem komplexen Gemisch eine Herausforderung dar. Gerade in den für *Metabolomics*-Studien häufig verwendeten ^1H -NMR-Spektren führt die Vielzahl an Metaboliten zu starken Signalüberlagerungen, wodurch insbesondere gering konzentrierte Metaboliten unter Verwendung der klassischen Methoden oftmals nicht detektiert und identifiziert werden können. Neben den zuvor vorgestellten klassischen NMR-Experimenten, bietet sich die Verwendung von selektiven Experimenten an, wie z.B. das 1D *selective* TOCSY- und das 1D *selective* NOESY-Experiment.^[50,59,60,62] Die eindimensionale selektive Variante des TOCSY-Experiments ist insbesondere im Fall von Gemischen, in denen die Metaboliten zusätzlich in hohen Konzentrationsunterschieden auftreten, von Vorteil. So können bei selektiver Einstrahlung auf die Frequenz eines isolierten Signals die zum selben Spinsystem gehörenden Signale in einem *selective* TOCSY-Experiment detektiert werden, obwohl sie im ^1H -NMR-Spektrum des Gemisches potenziell überlagert vorliegen.^[62] Ein *selective* NOESY-Experiment ermöglicht durch Anregung eines isolierten Signals die Detektion der räumlichen Kopplung von Kernspins, deren Abstand weniger als 5 Å beträgt. Daneben wurden in den letzten Jahren vielfältige neue NMR-Experimente entwickelt, die die Herausforderung der Gemischanalytik direkt adressieren. Das NMR-Experiment REST (*Relaxation-Encoded Selective TOCSY*) ermöglicht beispielsweise die Extraktion von Teilspektren eines spezifischen Analyten aus Gemischspektren, ohne den Analyten tatsächlich zu isolieren.^[63] Dieses Experiment ähnelt prinzipiell dem DOSY (*Diffusion Ordered Spectroscopy*), bei dem ein 2D NMR-Spektrum erhalten wird, das neben der chemischen Verschiebung in der F_2 -Dimension den Diffusionskoeffizienten in der F_1 -Dimension abbildet und somit Analyten mit unterschiedlichen Diffusionskoeffizienten unterschieden werden können.^[64,65] Das REST-Experiment macht sich statt der unterschiedlichen Diffusionskoeffizienten der Analyten die unterschiedlichen Relaxationszeiten der Spins zur Differenzierung der Analyten zu Nutze. Da die Relaxationszeiten jedoch für die verschiedenen Kerne innerhalb eines Spinsystems meist unterschiedlich sind, wird nach selektiver Anregung eines Spins eine isotrope Mischungsperiode verwendet, durch die jeder Kern in einem Spinsystem die gleiche Relaxationszeit-Charakteristik aufweist.^[63] In Form des PUREST-Experimentes (*Pure Shift Relaxation-Encoded Selective TOCSY*) wird das REST-

Experiment zusätzlich mit *Pure Shift* NMR-Spektren kombiniert.^[66] Durch die Steigerung der Auflösung der homonuklear entkoppelten ^1H -NMR-Spektren (*Pure Shift* Spektren) und damit einhergehend der Reduktion von Signalüberlagerungen ermöglicht diese Kombination eine deutlichere Separation der Signale im 2D-PUREST-NMR-Spektrum.^[66] Eine weitere Entwicklung ist bspw. das Experiment SCALPEL (*Spectral Component Acquisition by Localized PARAFAC (Parallel Factor Analysis) Extraction of Linear Components*).^[67] Die Kombination verschiedener Spin-*Labelling* Methoden, z.B. auf Basis von Diffusionskoeffizienten, T_1 - oder T_2 -Relaxationszeiten, und die anschließende multivariater Datenanalyse mittels PARAFAC (*Parallel Factor Analysis*) ermöglicht die Extraktion von Teilspektren einzelner Spinsysteme aus Gemischspektren, bei denen die zuvor vorgestellten Methoden REST oder DOSY an die Grenzen stoßen.^[67]

Zur Analyse der NMR-Daten, die z.B. auf Basis der zuvor erläuterten Experimente erhalten werden können, stehen Datenbanken wie die HMDB (*Human Metabolome Database*) oder die BMRB (*Biological Magnetic Resonance Data Bank*) zur Verfügung, die den Vergleich von NMR-Daten einer Vielzahl bekannter Metaboliten mit den gemessenen NMR-Daten ermöglichen. Die geringe Dispersion von ~ 10 ppm sowie die Abhängigkeit der chemischen Verschiebung vom Lösungsmittel, dem pH-Wert oder der Ionenstärke führen allerdings dazu, dass ein Vergleich mit ^1H -NMR-Spektren aus Datenbanken oftmals nicht zu einer eindeutigen Zuordnung der Signale führt.^[68] Für kommerziell verfügbare Analyten kann, nachdem ein Strukturvorschlag erarbeitet wurde, zur Zuordnung einzelner Signale außerdem ein *Spiking* der Analyten in die NMR-Probe des Gemisches durchgeführt werden.^[48] Dabei wird die vermutete Substanz in geringen Mengen zur Probe hinzugegeben, um den selektiven Intensitätsanstieg der entsprechenden Signale im Gemischspektrum zu erreichen und damit das Vorhandensein des Metaboliten im Gemisch zu verifizieren.

3.3.2.1 Zweidimensionales HSQC-Experiment

Insbesondere in Hinblick auf einen Datenbankabgleich ist die Verwendung von ^1H , ^{13}C -HSQC-NMR-Spektren von Vorteil. Die Dispersion der ^{13}C -Verschiebung ist mit ca. 250 ppm deutlich größer als die der ^1H -NMR-Spektren, da neben der diamagnetischen Abschirmung zusätzlich die paramagnetische Abschirmung einen Einfluss auf die chemische Verschiebung der ^{13}C -Kerne

hat.^[69,70] Prinzipiell wäre dafür auch die Verwendung von 1D ^{13}C -NMR-Spektren möglich, jedoch bieten die 2D ^1H , ^{13}C -HSQC-Spektren einige Vorteile. Da der NMR-aktive Kern des Kohlenstoffs (^{13}C) nur eine natürliche Häufigkeit von 1.1 % besitzt, ist die Sensitivität von ^{13}C -NMR-Spektren deutlich geringer, was sich in einer längeren Messdauer verglichen mit ^1H -NMR-Spektren äußert.^[70] Das HSQC-Experiment wird heutzutage als inverses zweidimensionales Experiment aufgesetzt, d.h. die Aufnahme erfolgt über den im Vergleich zu ^{13}C empfindlicheren ^1H -Kern. Somit ergibt sich im Vergleich zum ^{13}C -NMR-Experiment eine höhere Sensitivität für das HSQC-Experiment, da der ^1H -Kern ein ca. viermal höheres gyromagnetisches Verhältnis aufweist.^[71] Darüber hinaus werden im HSQC-Spektrum die indirekten $^1J_{\text{C-H}}$ -Kopplungen detektiert, was einen weiteren Informationsgewinn erzielt. Die zusätzliche Dimension im HSQC-Experiment führt außerdem zu reduzierter Signalüberlagerung.^[68]

Um ein HSQC-Experiment mittels inverser Aufnahmetechnik über den ^1H -Kern zu realisieren, muss die Magnetisierung während des Experiments zeitweise auf den ^{13}C -Kern übertragen werden. Dies wird mit Hilfe der INEPT (*Insensitive Nuclei Enhanced by Polarization Transfer*) Sequenz ermöglicht, die in Abbildung 9 grün gekennzeichnet ist und den Beginn der HSQC-Pulssequenz darstellt.

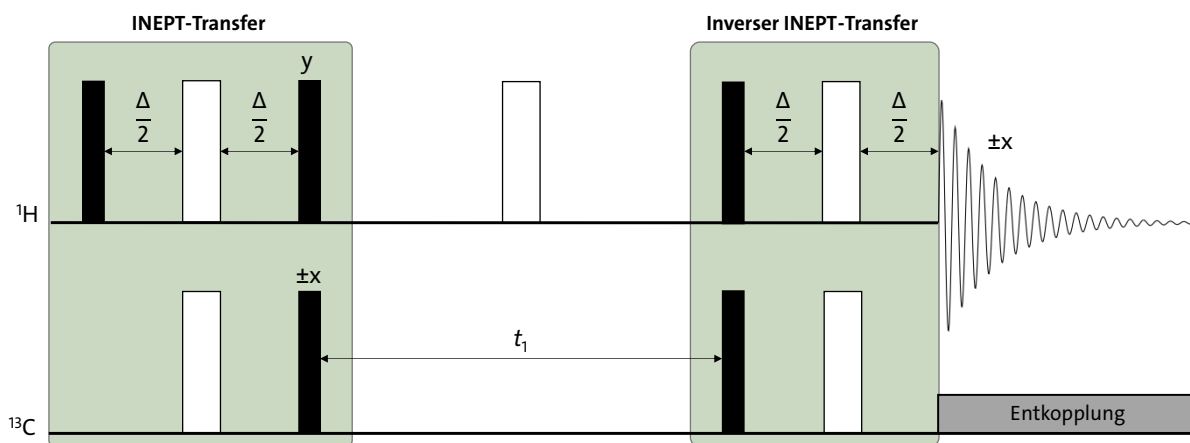


Abbildung 9: Pulssequenz des ^1H , ^{13}C -HSQC-Experimentes. Die grün hinterlegten Bereiche kennzeichnen den INEPT sowie den inversen INEPT-Transfer zum Polarisationstransfer der Magnetisierung zwischen den ^1H - und ^{13}C -Kernen. Schmale gefüllte Rechtecke kennzeichnen 90° Rechteckpuls, ungefüllte breite Rechtecke stellen 180° Rechteckpulse dar. Die Abbildung wurde nach CLARIDGE verändert.^[59]

Nach Anregung aller Protonen durch den 90° Puls entwickelt sich die chemische Verschiebung sowie die Kopplung zu den ^{13}C -Kernen. Der folgende 180° Puls führt zur Refokussierung der chemischen Verschiebung. Aufgrund des zeitgleichen 180° X-Pulses entwickelt sich die

Kopplung zwischen den Kernen während der zweiten $\frac{\Delta}{2}$ -Periode weiter. Gleichzeitig wird die Richtung der Präzession der ^1H -Kerne geändert, sodass sich nach der zweiten Periode die Magnetisierung der ^1H -Kerne auf der x-Achse in Antiphase befindet. Durch den 90° y-Puls der ^1H -Kerne wird diese wiederum auf die z-Achse übertragen.^[59]

Die Effizienz des Polarisationstransfers während der INEPT-Sequenz hängt maßgeblich von der Länge der Periode Δ ab. Ein optimaler Wert von Δ entspricht $(2J_{\text{XH}})^{-1}$.^[59] Die ^1H - ^{13}C -Kopplungskonstante über eine Bindung weist typischerweise Werte zwischen ca. 100 und 320 Hz auf. Die Kopplungskonstante ist dabei von dem s-Charakter der Hybridorbitale der Bindung und weiterhin z.B. von Substituenten abhängig.^[72] So ergeben sich für Bindungen an sp^3 -hybridisierten Kohlenstoffatomen (z.B. Alkane) Kopplungskonstanten von ca. 125 Hz, wohingegen sp^2 -hybridisierte Kohlenstoffatome (z.B. aromatische Systeme) Werte von ca. 167 Hz und sp -hybridisierte Kohlenstoffatome (z.B. Alkine) Kopplungskonstanten von ca. 250 Hz aufweisen.^[59,72] Für die Analyse von Gemischen oder in Routine-Messungen wird meist ein mittlerer Wert von 145 Hz für die 1J -Kopplungskonstante zwischen ^1H - und ^{13}C -Kernen gewählt.

Während der Evolutions-Periode t_1 entwickelt sich die ^{13}C -Magnetisierung. Durch den mittigen 180° - ^1H -Puls wird die ^1H - ^{13}C -Kopplung refokussiert, sodass in der indirekten Domäne lediglich die heteronukleare chemische Verschiebung verbleibt. Es folgt der inverse INEPT-Transfer zurück auf den ^1H -Kern und die Aufnahme des Signals in Form des *free induction decays* (FID). Die Entkopplung während der Aufnahme des FIDs eliminiert die Aufspaltung der Signale im HSQC-Spektrum. Für jedes ^1H - ^{13}C -Korrelationssignal im HSQC-Spektrum wird demzufolge ein einzelnes Kreuzsignal statt eines Dubletts erhalten, wodurch das Signal-zu-Rausch-Verhältnis im Vergleich zur Messung ohne Entkopplung erhöht wird.^[59]

3.3.2.2 Methoden zur Verkürzung der Messdauer von 2D-NMR-Experimenten

Nachteilig an der Verwendung von 2D-NMR-Spektren ist die, verglichen mit ^1H -NMR-Experimenten, lange Messdauer, die sich durch die allgemeine Aufnahmetechnik von 2D-NMR-Spektren ergibt. Dies äußert sich nicht nur nachteilig in Hinblick auf die Messung von großen Probengruppen, sondern auch in Hinblick auf die Sensitivität der Spektren in Bezug auf

Instabilitäten des Spektrometers.^[73] In den Spektren tritt teilweise das sogenannte t_1 *noise* auf, wodurch Banden von Signalen parallel zur F_1 -Achse auftreten.^[59,73]

In den letzten Jahren wurden aufgrund der beschriebenen Herausforderungen verschiedene Ansätze zur Beschleunigung der Aufnahme von 2D HSQC-Spektren entwickelt. Ein möglicher Ansatz ist die Verkürzung des Relaxationsdelays, wodurch die Wiederholrate der Spektrenaufnahme gesteigert werden kann. Dies wird beispielsweise bei der ASAP- (*acceleration by sharing adjacent polarization*), SOFAST- (*band-selective optimized-flip-angle short-transient*), BEST- (*band-selective excitation short-transient*) und der SMART- (a) (*small recovery times*) Pulssequenz angewendet.^[73–77] Ein weiterer Ansatz stellt die Entwicklung der *Ultrafast* 2D Methoden dar, bei denen die Probe prinzipiell in schmale Schichten entlang der z-Richtung unterteilt wird und das Experiment gleichzeitig in jeder Schicht mit jeweils einer anderen t_1 -Evolutionszeit akquiriert wird.^[78] So kann mit nur einem *Scan* das gesamte Spektrum innerhalb weniger Sekunden generiert werden.^[78] Eine weitere Möglichkeit der Reduktion der Gesamtmesszeit ist die Anwendung von *Nonuniform Sampling* (NUS), was prinzipiell bei jedem mehrdimensionalen NMR-Experiment angewendet werden kann.^[79,80] Während der Messung werden unter der Anwendung von NUS nicht alle der eigentlich erforderlichen Inkremente in der indirekten Dimension aufgenommen. Die Datenpunkte der fehlenden Inkremente werden im Anschluss an die Messung rekonstruiert und das 2D-Spektrum nach zweifacher Fourier-Transformation erhalten.^[81,82]

3.3.2.3 ASAP-HSQC-Experiment

Das ASAP-HSQC-Experiment ist eine beschleunigte Variante des konventionellen HSQC-Experimentes und wurde in der Gruppe um BURKHARD LUY entwickelt.^[74,83] Die Reduktion der Gesamtmessdauer beruht dabei grundsätzlich auf der Verkürzung des Relaxationsdelays (d_1) zwischen den einzelnen Messungen. Zur vollständigen Relaxation sollte der Relaxationsdelay mindestens der fünffachen Länge der longitudinalen Relaxationszeit der Kerne entsprechen, wodurch die Regenerierung von 99.3 % der Gleichgewichtsmagnetisierung erreicht wird.^[59] Bei unvollständiger Relaxation zwischen den konsekutiven Einzelaufnahmen (*Scans*) reduziert sich ansonsten das Signal-zu-Rausch-Verhältnis im Spektrum. Insbesondere bei mehrdimensionalen NMR-Experimenten führt eine Verlängerung des Relaxationsdelays jedoch zu einer mitunter vielfachen Verlängerung der Gesamtmesslänge.^[59] Um trotz Verkürzung des Relaxationsdelays

ein ausreichend hohes Polarisationsreservoir für die nächste Aufnahme zu gewährleisten, bedient sich das ASAP-HSQC-Experiment zweier Techniken.

Zum einen wird eine Variante der ERNST-Winkel-Anregung angewendet.^[74,84] Im Regelfall werden in NMR-Experimenten zur Anregung der Kerne reguläre 90°-Pulse verwendet, damit sich der Großteil der Magnetisierung in der transversalen Ebene befindet und dementsprechend eine hohe Signalintensität erreicht wird.^[59] Werden jedoch mehrere *Scans* nacheinander aufgenommen, führt ein geringerer Anregungswinkel von beispielsweise 30° zu einer schnelleren Relaxation zurück zur z-Achse und ermöglicht damit mit der Aufnahme des nächsten *Scans* schneller zu beginnen.^[59] Der optimale Anregungswinkel ist individuell für jede Probe und wird über den *Delay* Δ' während des INEPT-Transfers eingestellt.^[74,85]

Zusätzlich dazu beruht die Aufnahme der schnellen ASAP-HSQC-Spektren, basierend auf der Pulssequenz des ASAP-HMQC (*acceleration by sharing adjacent polarization heteronuclear multiple quantum coherence*), auf einem Polarisationstransfer. Bei dem ASAP-HMQC-Experiment werden selektiv all die Protonen angeregt, die an ^{13}C -Kerne gebunden sind. Dies wird durch die gezielte Nutzung der Spin-Spin-Kopplung zwischen ^1H - und ^{13}C -Kernen realisiert. Die übrigen Protonen, die an ^{12}C -Kerne gebunden sind, verbleiben entlang der z-Achse und dienen damit als Polarisationsreservoir. Während einer kurzen homonuklearen, isotropen Mischungsperiode (z.B. mittels DIPSI-2 Sequenz) kann die Magnetisierung aus dem Reservoir wiederum auf die an ^{13}C -Kerne gebundenen Protonen übertragen werden und damit als Ausgangsmagnetisierung für den nächsten *Scan* dienen.^[74,86] Im Fall des konventionellen HSQC-Experimentes findet allerdings eine Dephasierung derjenigen Protonen, die nicht an ^{13}C -Kerne gebunden sind, statt, sodass kein Polarisationsreservoir aufgebaut werden kann. In der ASAP-HSQC-Pulssequenz wird deshalb eine modifizierte Gradienten-Sequenz verwendet, sodass dennoch ein Polarisationsreservoir aufgebaut werden kann.^[74]

Das ASAP-HSQC-Experiment bedient sich außerdem Breitband-Pulsen aus der *Optimal Control Theory* (BEBOP, BURBOP und BUBI).^[74] Die Verwendung dieser Pulse statt regulärer harter Pulse dient der höheren Robustheit des Experimentes, da durch den Einsatz dieser Pulse *offset*-Effekte reduziert werden können. Dadurch werden in den ASAP-HSQC-Spektren vergleichsweise wenig bis keine Artefakte erhalten.^[83]

3.3.2.4 Chemoinformatische Methoden

Der Großteil der bisher genannten Methoden eignet sich insbesondere für die Identifizierung und Strukturaufklärung von bereits isolierten Analyten. Da im Rahmen von *Metabolomics*-Studien in der Regel große Datensätze, bestehend aus Spektren einer Vielzahl von Proben, analysiert werden, eignet sich die Anwendung der STOCSY (*Statistical Total Correlation Spectroscopy*) Methode, die 2005 von JEREMY K. NICHOLSON publiziert wurde.^[87] In Analogie zum TOCSY-Experiment werden bei STOCSY mögliche statistische Korrelationen und Antikorrelationen von Signalen in ¹H-NMR-Spektren untereinander detektiert.^[87] Die Grundlage bildet dabei ein *Metabolomics*-Datensatz basierend auf ¹H-NMR-Spektren, in denen die Multikollinearität der Intensitäten innerhalb des Datensatzes statistisch analysiert wird.^[87] Die Korrelationen können als pseudo-2D-NMR-Spektrum dargestellt werden. Die Analyse dieses Spektrums ermöglicht Aussagen über die Korrelation von ¹H-NMR-Signalen, die sich aufgrund von intramolekularer Korrelation durch tatsächliche strukturelle Konnektivität oder aufgrund von intermolekularer Korrelation durch Konnektivität biologischer Stoffwechselprozesse ergibt.^[87] Während hohe Antikorrelationen eindeutig auf intermolekulare Wechselwirkungen hindeuten, kann die Art der Korrelation im Fall von positiven Korrelationen nicht unterschieden werden. Die Weiterentwicklung STORM (*Subset Optimization by Reference Matching*) aus dem Jahr 2012 basiert auf dem gleichen Prinzip wie STOCSY, jedoch wird vor der STOCSY-Analyse die optimale Teilmenge der Spektren evaluiert, in der das spezifische Signal von Interesse vorhanden ist. Dies ermöglicht eine verbesserte Visualisierung der Korrelationsmatrix durch vermindertes Rauschen und damit die Erkennung von korrelierenden Signalen geringer Intensität.^[88] Zusätzlich wird durch das höhere Verhältnis von berücksichtigten Signalen zu Proben das statistische Rauschen durch falsch-positive Korrelationen reduziert. In Anlehnung an STOCSY basiert die von RUDD *et al.* entwickelte Methode HSQCcos (*HSQC correlation spectroscopy*) auf der Verwendung von 2D HSQC-NMR-Spektren statt ¹H-NMR-Spektren.^[89] Unter GERWICK und Mitarbeitern wurde 2017 die auf künstlichen neuronalen Netzwerken basierte Methode SMART 2.0 (b) (*Small Molecule Accurate Recognition Technology*) entwickelt.^[90,91] Durch das *Convolutional Neural Network* (CNN) können ähnliche strukturelle Motive erkannt und auf Grundlage von Datenbankspektren auf Basis von 2D HSQC-Spektren identifiziert werden. Da es sich bei der Analyse von 2D HSQC-NMR-Spektren primär um eine visuelle Analyse handelt, eignen sich CNNs als bildverarbeitende künstliche neuronale Netze besonders für diesen Zweck. Für diese Methodik wurden insgesamt 53 076 gemessene und

simulierte HSQC-Spektren von Naturstoffen und Marineorganismen in einen 180-dimensionalen Raum eingeordnet.^[91] Durch die CNN-basierte Analyse der Daten und Mustererkennung werden die Analyten, die sich in den HSQC-Spektren ähneln, in räumlicher Nähe im Ausgabe-Raum dargestellt, was sowohl die Detektion ähnlicher Struktur motive unbekannter Metaboliten als auch die Dereplikation bereits bekannter Metaboliten ermöglicht.^[90]

3.3.3 Identifizierungsmethoden basierend auf Massenspektrometrie

Neben der NMR-Spektroskopie stellt die Massenspektrometrie (MS) eine der am häufigsten verwendeten Methoden sowohl im Kontext von *Metabolomics*-Studien als auch im Bereich der Strukturaufklärung dar. Die zu untersuchenden Analyten werden in einer Ionenquelle ionisiert, daraufhin in einem Massenanalysator entsprechend ihres Masse-zu-Ladungs-Verhältnisses (m/z -Verhältnis) getrennt und anschließend wird der Ionenstrom durch einen Detektor gemessen.^[92,93] Aus den erhaltenen Massenspektren kann das Masse-zu-Ladungs-Verhältnis ermittelt werden, welches bei Verwendung von hochauflösenden Massenanalysatoren wiederum die Berechnung der monoisotopischen Masse des Moleküls und damit der Summenformel erlaubt.^[94] MS zeichnet sich durch die intrinsische hohe Sensitivität mit Nachweisgrenzen im Nanomolbereich aus, welche unter idealen Bedingungen auch bis in den Pico- und Femtomolbereich hineinreichen können.^[60,61] Aufgrund dessen ist im Vergleich zur NMR-Spektroskopie mit Hilfe von MS die Detektion einer deutlich größeren Anzahl von bis zu Tausenden Metaboliten möglich.^[60] Eine simultane Analyse einer Vielzahl von Analyten kann allerdings zur Ionensuppression führen. Dadurch kann die Signalintensität einzelner Analyten soweit reduziert werden, dass diese in der vorliegenden Mischung nur mit sehr geringer Signalintensität oder gar nicht detektiert werden können. Unter anderem aus diesem Grund wird die massenspektrometrische Analyse häufig mit chromatographischen Verfahren wie der Flüssigchromatographie (*liquid chromatography*, LC) oder der Gaschromatographie (GC) gekoppelt.^[94] Dies führt zu einer Reduktion der Komplexität des Stoffgemisches und vereinfacht somit die Auswertung der Chromatogramme und Massenspektren.^[61] Insbesondere bei einem hohen Probendurchsatz kann alternativ auch der Direkteinlass (*direct injection*, DI) eine geeignete Methode zur Injektion der Proben darstellen.^[95,96] Die Probe wird dabei entweder über eine Spritzenpumpe direkt in das Massenspektrometer injiziert oder über eine kurze

chromatographische Methode über eine (U)HPLC ((*Ultra*) *High Performance Liquid Chromatography*) gefördert.^[97]

Zur Ionisierung der Analyten stehen verschiedene Techniken zur Verfügung. Im Bereich der *Metabolomics*-Studien hat sich die Elektrospray-Ionisation (ESI) als sanfte Ionisationsmethode als vorteilhaft erwiesen, da im Gegensatz zur harten Ionisationsmethode Elektronenionisation (EI) kaum Fragmentierungen der Analyten auftreten.^[92,98] In ESI-Massenspektren werden hauptsächlich Quasimolekülionen in Form von $[M+H]^+$ bzw. $[M-H]^-$ oder Addukte z. B. mit Natrium- ($[M+Na]^+$) oder Kalium-Ionen ($[M+K]^+$), die z. B. aus der mobilen Phase, Glasflaschen, Kapillaren oder aus biologischen Matrices stammen können, detektiert.^[99–101] Im Massenanalysator werden die ionisierten Analyten entsprechend ihres m/z -Verhältnisses getrennt. Dazu müssen diese zunächst beschleunigt werden. Die Trennung der Ionen kann über verschiedene Massenanalysatoren erfolgen, wie z. B. Quadrupol, Orbitrap oder *Time of flight* (TOF). Im Fall eines Flugzeit-Analysators (*Time of flight*) werden leichte Ionen stärker beschleunigt und erreichen das Ende des Flugrohres somit schneller als schwere Ionen.^[92] TOF-Analysatoren erreichen bei akkurater Kalibrierung eine Massengenauigkeit von ca. 2 ppm.^[102]

Die Identifizierung von Metaboliten auf Grundlage von Massenspektren erfolgt meist durch einen Datenbankabgleich (z. B. HMDB) verschiedener Parameter, wie m/z -Verhältnisse, exakte Masse, Retentionszeiten, Isotopenverteilungen und Fragmentierungsmuster.^[103,104] Diese Parameter werden in einer Vielzahl von *Metabolomics*-basierten Studien meist durch LC-Tandem-Massenspektrometrie (LC-MS/MS) Experimente erhalten. Bei LC-MS/MS-Experimenten werden unterschiedliche Massenanalysatoren gekoppelt, wodurch eine weitere Fragmentierung von Vorläufer-Ionen stattfindet.^[105] Durch Interpretation der erhaltenen MS/MS-Spektren können Rückschlüsse auf Fragmentierungsreaktionen gezogen werden und sowohl die Vorläufer- als auch die Produkt-Ionen mit Datenbankspektren abgeglichen werden, was die strukturelle Identifizierung vereinfacht.^[106] Zusätzlich ist ebenso die Verwendung von leistungsfähigen Ansätzen mittels neuronaler Netzwerke möglich. Die LC-MS/MS-basierte *Global Natural Products Social Molecular Networking* Methode (GNPS) basiert auf einem neuronalen Netzwerk und erkennt ähnliche strukturelle Motive, wodurch eine Vielzahl von Strukturen auf Grundlage von Datenbankspektren charakterisiert werden kann.^[107,108]

Nichtsdestotrotz geht die Identifizierung von bisher unbekannten Metaboliten meist mit der Isolierung des Produktes einher.

3.4 Identifizierungsstrategien durch Kombination analytischer Methoden

Zur eindeutigen Identifizierung von bekannten wie auch unbekannten Metaboliten ist die Kenntnis von möglichst vielen analytischen Parametern vonnöten. Da die NMR-Spektroskopie und die Massenspektrometrie komplementäre Methoden sind, deren Experimente einander ergänzende Informationen liefern, ist eine Kombination beider Analysemethoden eine vielversprechende Methode. Die Kombination von NMR-Spektroskopie und Massenspektrometrie kann auf unterschiedliche Art und Weise erfolgen. BINGOL *et al.* klassifizieren die Methoden entsprechend der folgenden fünf Kategorien: Hardware-basierte Kopplung, Kopplung durch chemische Modifizierung, Kopplung über statistische Methoden, stabile Isotopen-basierte Kopplung und kombinatorische Chemoinformatik-basierte Methoden.^[7] Die Kopplung durch chemische Modifizierung und die *Metabolomics*-Analyse basierend auf Verteilung von stabilen Isotopen sind in diesem Kontext jedoch eher als spezielle Anwendungen zu sehen. Für die Kopplung durch chemische Modifizierung wird ein spezifisches *Labeling* von Metaboliten mit bestimmten Strukturmotiven durchgeführt. Beispielsweise können Metaboliten mit Carboxylat-Funktion über eine Kondensationsreaktion mit ¹⁵N-Aminocholin verknüpft werden, wodurch sowohl in den Massenspektren durch die bereits vorhandene positive Ladung als auch in den NMR-Spektren über ¹H-¹⁵N-HSQC-Spektren eine erhöhte Sensitivität erreicht wird und die Kopplung der Methoden dementsprechend erleichtert wird.^[109] Im Fall von *Stable Isotope-resolved Metabolomics* wird z. B. ¹³C-gelabelte Glucose in den zu untersuchenden Organismus eingebracht, um anschließend über Isotopenmuster den Metabolismus und den Einfluss auf biochemische Stoffwechselwege zu untersuchen.^[110]

Eine Möglichkeit der Hardware-basierten Kopplung ist beispielsweise die Online LC-MS-NMR-Kopplung, die jedoch in der praktischen Anwendung mit einigen Nachteilen und Hindernissen behaftet ist. So ist für die direkte Kopplung die Nutzung von deuterierten Lösungsmitteln für die Flüssigchromatographie nötig. Wie bereits zuvor erläutert, unterscheiden sich die NMR-Spektroskopie und die Massenspektrometrie wesentlich bzgl. ihrer Sensitivität, was die Kopplung der Hardware erschwert.^[6]

Beispiele der Kopplung der Methoden auf Basis von Chemoinformatik sind die Methoden *NMR/MS Translator* und *SUMMIT MS/NMR (Structures of Unknown Metabolomic Mixture components by MS/NMR)*.^[111,112] Die Methode *NMR/MS Translator* beruht auf der Erkennung potenziell vorhandener Metaboliten durch einen Datenbankabgleich der NMR-Spektren. Anschließend werden von den in Frage kommenden Metaboliten die m/z -Werte der Molekülionen, möglicher Addukte und Fragment-Ionen berechnet und mit den gemessenen Massenspektren verglichen.^[112] Diese Methode ist aufgrund des Datenbankabgleiches dementsprechend nur zur Identifizierung bereits bekannter Metaboliten geeignet. Die Methode *SUMMIT MS/NMR* kann dahingegen auch zur Identifizierung von bisher unbekannten Metaboliten genutzt werden. Durch die Aufnahme von hoch-auflösenden Massenspektren können für ein Signal eines bestimmten m/z -Wertes alle möglichen Summenformeln berechnet werden, die anschließend in mögliche Strukturformeln übersetzt werden. Für jede der Strukturformeln wird anschließend das jeweilige NMR-Spektrum berechnet, das daraufhin mit den experimentellen NMR-Spektren verglichen wird.^[111]

Eine auf Statistik basierende Methode zur Kombination von NMR-Spektroskopie und Massenspektrometrie ist *Statistical Heterospectroscopy (SHY)*.^[113] Diese Methode ist ein analoger Ansatz zu den bereits unter Abschnitt 3.3.2.4 erläuterten Methoden STOCSY und STORM, bei denen statistische Zusammenhänge in *Metabolomics*-Datensätzen basierend allein auf NMR-Spektren evaluiert werden.^[87,88] Im Unterschied dazu werden bei SHY die Datensätze verschiedener Analyseverfahren, d.h. beispielsweise NMR- und MS-Daten, statistisch co-analysiert.^[113] Durch die Analyse von Korrelationen und Antikorrelationen zwischen Signalen der Datensätze können somit sowohl strukturelle Konnektivitäten als auch Zusammenhänge aufgrund von biochemischen Stoffwechselwegen untersucht werden.^[113] Wie auch bei STOCSY ist es allerdings im Fall von positiven Korrelationen nicht möglich, die Art der Korrelation zu unterscheiden. Des Weiteren werden für diese Art der Korrelation große Datensätze derselben Proben mit beiden Analyseverfahren benötigt.

Neben den bereits vorgestellten Methoden kann die Kombination von NMR- und MS-Daten auch nach einer unvollständigen Trennung der Substanzen durch Flüssigchromatographie erreicht werden. Ein Beispiel stellt die von Dai *et al.* eingeführte Methode *NMR/LC-MS parallel dynamic spectroscopy (NMR/LC-MS PDS)* dar.^[114] Nach unvollständiger Separation eines Extraktes werden NMR- und LC-MS-Daten der Fraktionen generiert und anschließend visuell

analysiert und manuell miteinander korreliert.^[114] Diese Methode wurde für die Identifizierung von verschiedenen Flavonoiden und organischen Säuren aus Pflanzenextrakten verwendet.^[114,115] Der unter MEYER *et al.* entwickelten Methode *Three-Dimensional Cross Correlation* (3DCC) zur Identifizierung von Glycanen liegt ebenfalls eine unvollständige chromatographische Trennung zu Grunde.^[116–118] Die 3DCC-Methodik ist im folgenden Abschnitt 3.4.1 genauer erläutert.

3.4.1 Three-Dimensional Cross Correlation (3DCC)

Die 3DCC-Methode wurde von MEYER *et al.* zur Identifizierung von Glycanen entwickelt. Als Glycane werden die in Glycoproteinen vorhandenen Oligo- und Polysaccharide bezeichnet, die in Zellen und Geweben vorkommen und eine wichtige Rolle bei verschiedenen Krankheiten und vielen biochemischen Stoffwechselprozessen einnehmen. Zur genaueren Untersuchung der Krankheiten und der zugrundeliegenden Stoffwechselprozesse ist deshalb die Kenntnis der Glycan-Struktur erforderlich.^[116]

Im Rahmen der 3DCC-Methode erfolgt im ersten Schritt eine chromatographische Separation der Glycane mittels LC-MS und anschließender Fraktionierung. Aus den LC-MS-Chromatogrammen können die Intensitätsprofile der individuellen Ionen in Form von *Extracted Ion Chromatograms* (EICs) generiert werden. Von jeder der durch die Fraktionierung erhaltenen Fraktionen wird im nächsten Schritt ein NMR-Spektrum aufgenommen. In Analogie zu den EICs können für die NMR-Spektren sogenannte EDCs (*Extracted Delta Chromatograms*) generiert werden, die die zweidimensionalen Signalamplitudenprofile bei einer spezifischen chemischen Verschiebung über die Fraktionen zeigen. Schematisch ist dies in Abbildung 10 gezeigt.

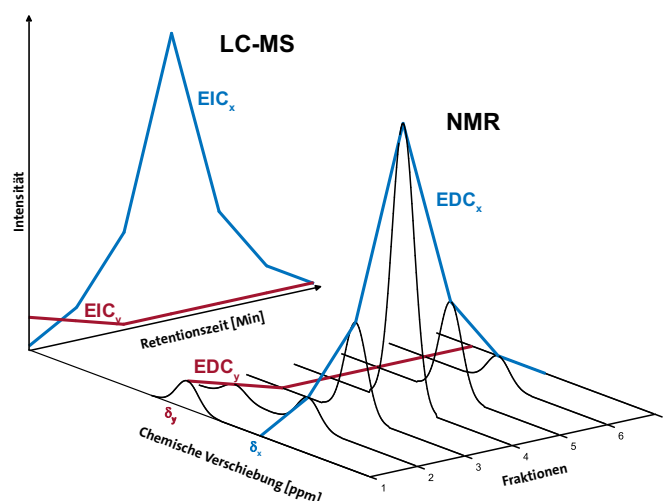


Abbildung 10: Schematische Darstellung der für die 3DCC-Methode verwendeten Zeit-Intensitätsprofile. Für die Massenspektren werden die EICs aus den LC-MS-Chromatogrammen generiert. Für die NMR-Spektren, die von jeder einzelnen Fraktion aufgenommen werden, werden analog sogenannte EDCs als Zeit-Intensitätsprofile bei einer bestimmten chemischen Verschiebung δ [ppm] generiert. Die Abbildung wurde nach BEHNKEN *et al.* verändert.^[116]

Die beiden spezifischen EICs des LC-MS-Chromatogramms (EIC_x und EIC_y) in Abbildung 10 weisen unterschiedliche Verläufe bzgl. der chromatographischen Separation auf. In den NMR-Spektren, die von den einzelnen Fraktionen akquiriert werden, d.h. zu spezifisch definierten Zeitpunkten, zeigen EDC_x und EDC_y die gleichen Verläufe. Das NMR-Signal bei der chemischen Verschiebung δ_x gehört demnach zu dem EIC_x und das NMR-Signal bei δ_y zu EIC_y . Über verschiedene Methoden der Korrelation, beispielsweise basierend auf einer Gewichtung der errechneten Korrelationskoeffizienten, durch Optimierung der kleinsten Quadrate (*least squares optimization*) oder durch Singulärwertzerlegung (*singular value decomposition*), können die Reinspektren der einzelnen Substanzen berechnet werden.^[116,118]

3.4.2 Mathematische Korrelation von Variablen

Zur Evaluation des linearen Zusammenhangs zweier Variablen x und y kann der Korrelationskoeffizient r_{x_j, x_k} berechnet werden. Der Korrelationskoeffizient nach PEARSON wird über den Quotienten der Kovarianz $cov(x_j, x_k)$ der beiden Variablen mit dem Produkt der Standardabweichungen s_j und s_k berechnet (Gleichung (1)).^[119,120] Die Division durch die beiden Varianzen, die die jeweilige Streuung der Variablen angeben, dient der Normierung der Daten.

$$r_{x_j, x_k} = \frac{\text{cov}(x_j, x_k)}{s_j s_k} \quad (1)$$

In Gleichung (1) entspricht $\text{cov}(x_j, x_k)$ der Kovarianz, die gemäß Gleichung (2) berechnet wird.^[119]

$$\text{cov}(x_j, x_k) = \frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k) \quad (2)$$

Gleichung (2) gilt für $j, k = 1, \dots, p$ und

$$j \neq k$$

mit

n = Anzahl der Objekte und

p = Anzahl der Variablen.

Die Varianz $s_k^2 = \text{var}(x_k)$ entspricht der Kovarianz $\text{cov}(x_j, x_k)$ für den Fall $j = k$, d.h. $\text{cov}(x_1, x_1) = \text{var}(x_1)$. Die Varianz wird entsprechend Gleichung (3) berechnet.^[119] Die Standardabweichungen s_j bzw. s_k entsprechen der Quadratwurzel der Varianz s_k^2 bzw. s_j^2 .

$$\text{var}(x_k) = s_k^2 = \frac{1}{n-1} \sum_{i=1}^n (x_{ik} - \bar{x}_k)^2 \quad (3)$$

Für den Korrelationskoeffizienten r_{x_j, x_k} nach PEARSON ergibt sich demnach für den Fall $j = k$ ein Wert von 1. Im Falle von mehreren Variablen p ergibt sich eine Korrelationsmatrix $\mathbf{R}$ mit folgendem Aufbau (siehe Gleichung (4)).

$$\mathbf{R} = (r_{j,k}) = \begin{pmatrix} 1 & r_{1,2} & \cdots & r_{1,p} \\ r_{2,1} & 1 & \cdots & r_{2,p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p,1} & r_{p,2} & \cdots & 1 \end{pmatrix} \quad (4)$$

Der Korrelationskoeffizient r_{xy} kann Werte zwischen -1 und $+1$ annehmen und beschreibt dadurch die Stärke des Zusammenhangs zwischen den Variablen. Ein Korrelationskoeffizient von $+1$ entspricht demnach einem stark positiven linearen Zusammenhang, wohingegen ein Wert von 0 eine fehlende Korrelation und ein Wert -1 eine Antikorrelation der Daten anzeigt.

Für ordinal skalierte Daten kann der Korrelationskoeffizient nach PEARSON nicht berechnet werden, da dieser nur für metrisch skalierte Daten verwendbar ist. Im Fall von ordinal skalierten Daten können stattdessen Rangkorrelationskoeffizienten z.B. nach SPEARMAN oder KENDALL errechnet werden.^[121] Zur Berechnung von Rangkorrelationskoeffizienten werden zunächst die verschiedenen Ausprägungen der Variablen in zu- oder abnehmender Weise der Größe nach

geordnet und daraus die Rangfolge erstellt, wodurch jeder Ausprägung eine Rangzahl r zugeordnet werden kann. Somit entsprechen die Ausprägungen für diese Art der Korrelationsanalyse lediglich den Rangzahlen und nicht tatsächlich interpretierbaren Zahlenwerten. Der SPEARMAN Rangkorrelationskoeffizient r_s kann über Gleichung (5) berechnet werden.^[122]

$$r_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \quad (5)$$

In Gleichung (5) entspricht der Term d_i den Rangdifferenzen, der entsprechend in Gleichung (6) berechnet werden kann.^[122]

$$d_i = r(x_i) - r(y_i) \quad (6)$$

Der Rangkorrelationskoeffizient nach KENDALL berechnet sich zu KENDALLS τ über Gleichung (7). Dabei entspricht Q_i der Anzahl derjenigen Rangzahlen $R(y_i)$, die kleiner oder gleich und damit in der Rangreihenfolge hinter $R(y_i)$ stehen.^[121]

$$\tau = 1 - \frac{4 \sum_{i=1}^n Q_i}{n(n-1)} \quad (7)$$

Unabhängig von der gewählten Art der Berechnung des Korrelationskoeffizienten muss in allen Fällen evaluiert werden, ob es sich um tatsächliche und sinnvolle Korrelationen handelt, oder ob Schein- oder Nonsenskorrelationen vorliegen. Eine Scheinkorrelation entsteht, wenn zwei nicht-korrelierte Variablen jeweils mit einer dritten Variable korrelieren. Die Nonsenskorrelation beschreibt dahingegen Korrelationen, die nicht in einem sachlogischen Zusammenhang zueinander stehen, d. h. inhaltlich unsinnig sind.^[121]

4 Zielsetzung

Im Rahmen von *Metabolomics*-Studien stellt die Identifizierung von Metaboliten auf der einen Seite einen wichtigen Faktor, auf der anderen Seite aber auch eine große Herausforderung dar. Mittels NMR-Spektroskopie lassen sich in Extrakten biologischer Proben bis zu 200 Metaboliten detektieren. Die Vielzahl der Metaboliten und die zusätzlich starken Konzentrationsunterschiede zwischen einzelnen Metaboliten innerhalb der zu untersuchenden Matrix erschweren die direkte und eindeutige Identifizierung von Metaboliten aus komplexen Gemischen auf alleiniger Basis von eindimensionalen NMR-Spektren.

Das Ziel dieser Arbeit ist deshalb die Entwicklung eines strategischen Ansatzes zur einfachen, sicheren und effizienten Identifizierung von Metaboliten aus komplexen Gemischen ohne die Notwendigkeit der individuellen Isolierung einzelner Metaboliten. Das grundlegende Konzept basiert auf der unter B. MEYER entwickelten Methode 3DCC (*Three-dimensional cross correlation*), die zur Identifizierung der Stereochemie von Glycanen durch Korrelation von LC-MS- und ^1H -NMR-Daten angewendet wurde. Im Rahmen dieser Arbeit soll die Korrelation komplementärer ^1H -NMR- und MS-Daten nach flüssigchromatographischer Separation eines komplexen Gemisches zur Identifizierung von Metaboliten aus biologischen Extrakten entwickelt werden. Der Fokus der Flüssigchromatographie liegt dabei nicht primär auf der Isolierung einzelner Analyten, sondern es wird die Verteilung dieser über mehrere Fraktionen angestrebt. Zur Auswertung der Daten soll eine MATLAB Applikation mit graphischer Benutzeroberfläche entwickelt werden, die eine Strukturaufklärung aus komplexen Gemischen wesentlich vereinfacht. Das entwickelte Konzept soll zunächst anhand eines artifiziellen Gemisches validiert und anschließend anhand verschiedener biologischer Matrices angewendet werden. Dabei sollen Metaboliten in Extrakten biologischer Proben sowie potenzielle Biomarker nach multivariater Datenauswertung eines *Metabolomics*-Datensatzes identifiziert werden. Weiterhin soll untersucht werden, ob die generelle Methode auch außerhalb von *Metabolomics*-Studien angewendet werden kann, indem bei einer Reaktionsverfolgung der zeitliche Konzentrationsverlauf der Signale nicht durch eine Flüssigchromatographie induziert wird, sondern direkt aus dem Verlauf einer Reaktion untersucht werden kann. Um der Herausforderung der starken Signalüberlagerungen in eindimensionalen ^1H -NMR-Spektren zu begegnen, soll die Methode außerdem erweitert werden, sodass auch zweidimensionale HSQC-NMR-Daten verarbeitet und mit den entsprechenden MS-Daten korreliert werden können.

5 Ergebnisse und Diskussion

Die Identifizierung von Metaboliten aus komplexen Gemischen ohne individuelle Isolierung jedes einzelnen Metaboliten stellt trotz der sich immer weiter entwickelnden Varietät verschiedener Analysetechniken, v.a. im Bereich der NMR-Spektroskopie und der Massenspektrometrie, eine der wesentlichen Herausforderungen dar. Die unter MEYER *et al.* entwickelte Methode *Three-dimensional cross correlation* (3DCC) ermöglicht, wie in Abschnitt 3.4.1 beschrieben, die Bestimmung der exakten Stereochemie von Glycanen aus Gemischen und somit deren Identifizierung.^[116] Dabei werden ¹H-NMR-Daten mit LC-MS-Daten korreliert, um NMR-Spektren der jeweiligen Analyten zu berechnen, ohne diese isolieren zu müssen. Für die Identifizierung von Glycanen reicht, entsprechend des *structural reporter group* Konzeptes, allerdings die Analyse von einzelnen, spezifischen Signalen der Glycane aus.^[123] Zur Identifizierung von Metaboliten verschiedenster Substanzklassen ist dahingegen die Kenntnis aller NMR-Signale von Relevanz. Im Rahmen dieser Arbeit wurde deshalb die Methode *SCORE-metabolite-ID* (*Semi-automatic CORrelation analysis for RELiable metabolite IDentification*) entwickelt, die grundlegend auf dem Prinzip der 3DCC-Methode basiert.

5.1 Entwicklung von SCORE-metabolite-ID basierend auf ¹H-NMR-Spektren

Das generelle Konzept von *SCORE-metabolite-ID* ist in Abbildung 11 schematisch dargestellt. Im ersten Schritt wird ein beliebiges Extrakt säulenchromatographisch fraktioniert. Der Fokus der säulenchromatographischen Methode liegt dabei nicht primär auf der Isolierung einzelner Metaboliten. Stattdessen wird eine Verteilung der Analyten über mehrere Fraktionen angestrebt, die infolge einer beabsichtigten unvollständigen Trennung realisiert wird. Die Optimierung der flüssigchromatographischen Methode verfolgt das Ziel, individuelle Elutionsprofile für die Metaboliten zu generieren. Von jeder erhaltenen Fraktion werden NMR-Spektren und Massenspektren aufgenommen. Entsprechend des Elutionsprofils der Metaboliten werden sowohl in den NMR-Spektren als auch in den Massenspektren Verläufe der Signalintensitäten erhalten, die proportional zur relativen Konzentration der Analyten sind. Der zeitliche Verlauf (entsprechend der säulenchromatographischen Fraktionen) dieser Signalamplitudenprofile wird im Fall der Massenspektren als *Extracted Mass Chromatogram* (EMC) und im Fall der NMR-Spektren in Analogie zur 3DCC-Methode als *Extracted Delta*

Chromatogram (EDC) bezeichnet. Die EDCs und EMCs können, z. B. mittels PEARSON Korrelation, untereinander korreliert werden, um Informationen über die NMR- und MS-Daten zu erlangen, die zu demselben Metaboliten gehören. Der Korrelationskoeffizient kann Werte von -1 bis $+1$ annehmen. Je höher der Korrelationskoeffizient ist, desto wahrscheinlicher ist eine Zusammengehörigkeit der entsprechenden EDCs und EMCs, d. h. der entsprechenden NMR- und MS-Signale, und damit ergibt sich eine hohe Wahrscheinlichkeit, dass die Signale vom gleichen Molekül stammen.

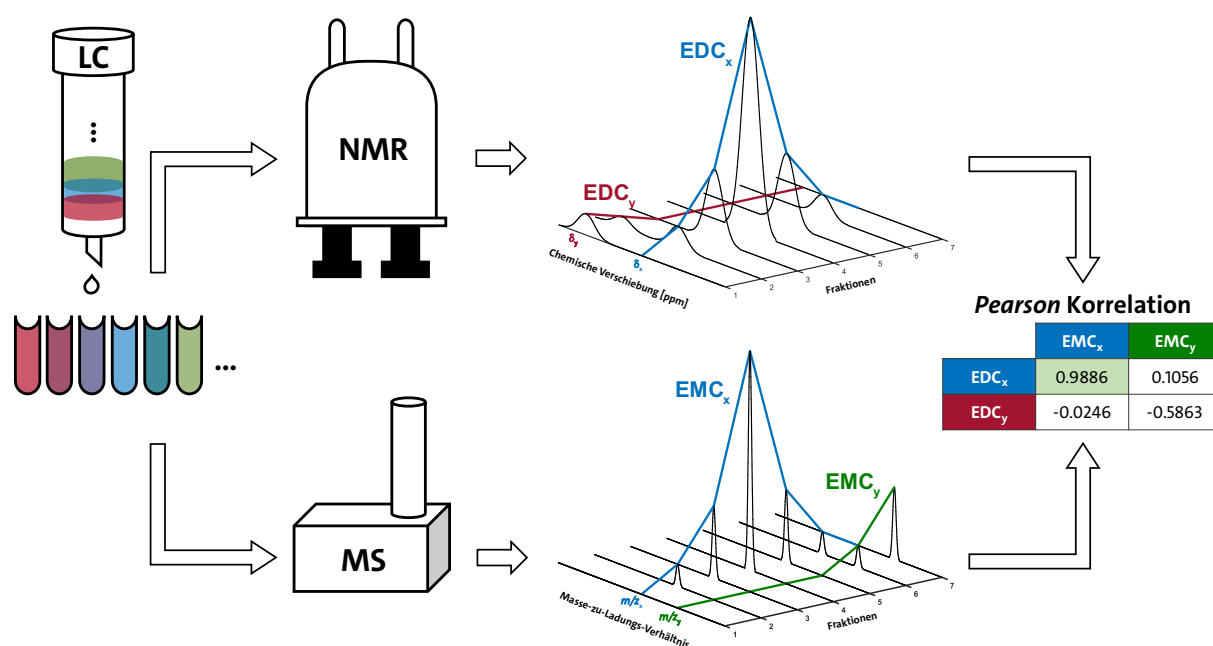


Abbildung 11: Schematische Darstellung des Konzeptes von *SCORE-metabolite-ID*. Nach Säulenchromatographie wird von jeder Fraktion sowohl ein NMR-Spektrum als auch ein Massenspektrum akquiriert. Entsprechend des Elutionsprofils der Analyten wird die zu- und abnehmende Signalintensität sowohl in den NMR-Spektren als auch in den Massenspektren detektiert. Die aus den NMR-Spektren generierten EDCs (*Extracted Delta Chromatograms*) und aus den Massenspektren generierten EMCs (*Extracted Mass Chromatograms*) entsprechen demnach dem jeweiligen relativen Konzentrationsverlauf und können z. B. über PEARSON Korrelation miteinander korreliert werden. Dadurch können NMR-Signale den m/z -Verhältnissen aus Massenspektren zugeordnet werden.

Im ersten Schritt wurde das generelle Konzept von *SCORE-metabolite-ID* evaluiert. Dazu wurde zunächst die Linearität der verwendeten Methoden untersucht, da zur Verwendung der PEARSON Korrelation ein linearer Zusammenhang zwischen Konzentration und Signalintensität in beiden analytischen Methoden unterstellt wird. Für die eindimensionalen NMR-Spektren wurde im Rahmen dieser Arbeit das 1D ^1H -NOESY-Experiment verwendet. Dieses Experiment wird im Rahmen von *Metabolomics*-Studien aufgrund der effektiven Lösungsmittelunterdrückung

häufig eingesetzt. Vorteilhaft ist dabei vor allem die wirksame Unterdrückung des Wassersignals bei relativ geringem Optimierungsbedarf.^[124] Die Unterdrückung des Wassersignals beruht hauptsächlich auf der T_1 -Diskriminierung, d.h. der unterschiedlich langen longitudinalen Relaxationszeiten (Spin-Gitter-Relaxation) des Lösungsmittels und der gelösten Analyten.^[124] Durch den Effekt des *Radiation Damping* relaxiert das Lösungsmittel (in diesem Fall Wasser) deutlich schneller zur +z-Achse als die gelösten Analyten. *Radiation Damping* beschreibt den Effekt, dass das Lösungsmittel ein so großes NMR-Signal induziert, dass dieses Signal selbst ein sekundäres Radiofrequenzfeld generiert. Dies bewirkt, dass die Magnetisierung des Lösungsmittels schneller auf die +z-Achse relaxiert, als es durch natürliche Relaxation der Fall wäre. Neben dem *Radiation Damping* führen zusätzlich die Anwendung von Phasenzyklen und der optimalen Mischzeit zu einer effektiven Wasserunterdrückung.^[124]

In Abbildung 12 ist die Abhängigkeit der Signalintensität von der Konzentration in einem Konzentrationsbereich von 0.05 mg/mL bis 100 mg/mL am Beispiel von Saccharose (**2**) in den 1D ^1H -NOESY-Spektren gezeigt. Die lineare Regression in Abbildung 12 A mit einem korrigierten R-Quadrat-Wert von 0.99845 legt einen linearen Zusammenhang zwischen Konzentration und Signalintensität nahe.

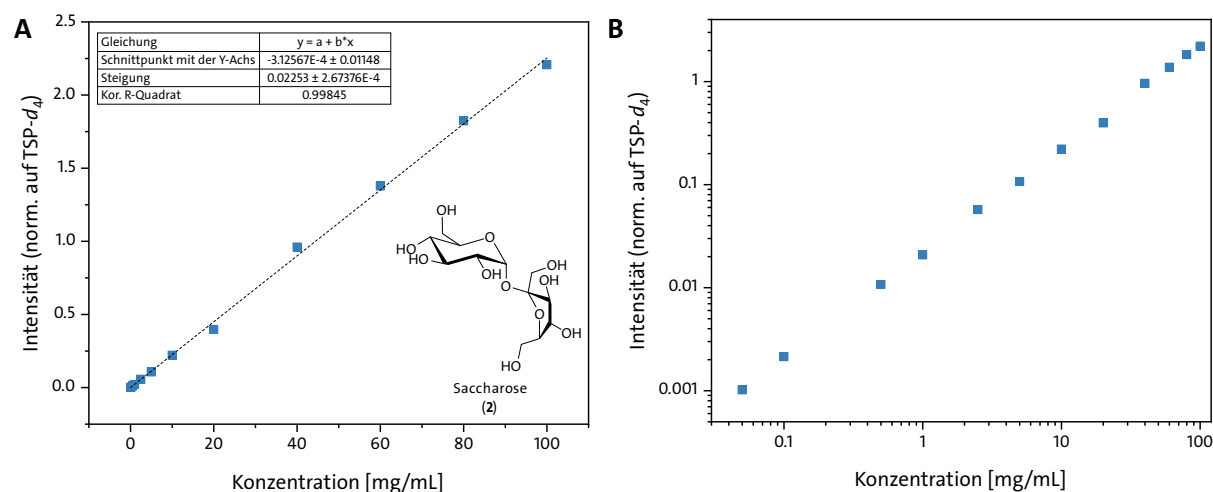


Abbildung 12: Abhängigkeit der Signalintensität von der Konzentration in 1D ^1H -NOESY-Spektren. Es wurde die Signalintensität von H-1 bei einer chemischen Verschiebung von 5.42 ppm von Saccharose (**2**) in D_2O (+ 5 mM TSP- d_4) in einem Konzentrationsbereich von 0.05 mg/mL bis 100 mg/mL untersucht. Es ist jeweils die auf das TSP- d_4 -normierte Signalintensität aufgetragen. In **A** sind die Achsen linear skaliert, in **B** sind die Achsen logarithmisch skaliert.

Im Gegensatz zur NMR-Spektroskopie ist der lineare Bereich in massenspektrometrischen Analysen deutlich kleiner. In Abbildung 13 ist die Abhängigkeit der absoluten Signalintensität des Natrium-Addukts des Analyten Saccharose (**2**) von der Konzentration im Bereich von 0.05 µg/mL bis 100 µg/mL gezeigt.

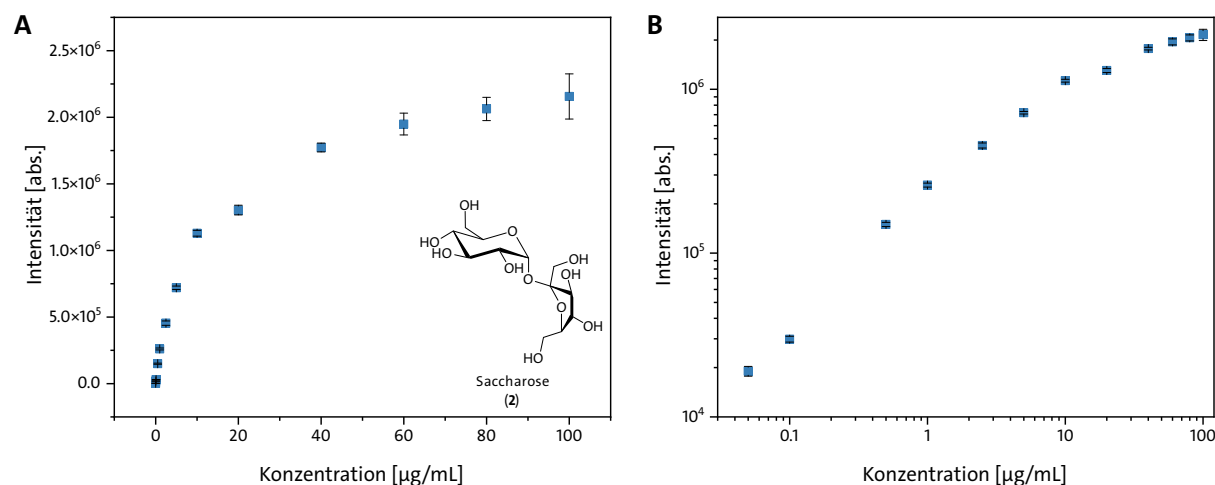


Abbildung 13: Abhängigkeit der Signalintensität von der Konzentration in der massenspektrometrischen Analyse basierend auf DI-Massenspektren. Es wurde die absolute Signalintensität des $[M+Na]^+$ -Signals von Saccharose (**2**) ($m/z = 365.11$) in einem Konzentrationsbereich von 0.05 µg/mL bis 100 µg/mL untersucht. Es sind die gemittelten Werte aus jeweils vier Messwerten sowie die daraus berechnete Standardabweichung als Fehlerbalken dargestellt. In **A** sind die x- und y-Achse linear und in **B** logarithmisch skaliert.

Wie der Abbildung 13 zu entnehmen ist, ist im Fall von Saccharose (**2**) nicht über den gesamten untersuchten Konzentrationsbereich ein linearer Zusammenhang zwischen der Signalintensität und der Konzentration des Analyten gegeben. Im niedrigen Konzentrationsbereich zwischen 0.05 bis 10 µg/mL ist eine hohe Linearität gegeben. Im Konzentrationsbereich ab 20 µg/mL ist eine Abhängigkeit in Form einer Sättigungsfunktion erkennbar. Dabei handelt es sich nicht um die Detektorsättigung, sondern um eine Signalsättigung, die bei hohen Konzentrationen eines Analyten auftreten kann. Mögliche Gründe für eine Signalsättigung in der Massenspektrometrie sind bspw. die begrenzte Menge der zur Verfügung stehenden Ladung, um den Analyten zu ionisieren.^[100] Ein weiterer Grund ist der begrenzte Raum auf der Oberfläche der Tröpfchen, die im ESI-Spray gebildet werden. Nach KOSTIAINEN *et al.* werden die Tröpfchen bei hohen Analytkonzentrationen so sehr mit Analyt gesättigt, sodass die Oberflächen vollständig gefüllt sind und dadurch der Ionenaustritt aus den Tröpfchen heraus gehemmt wird.^[100,125] Da insgesamt dennoch ein direkter Zusammenhang zwischen Konzentration und Signalintensität vorliegt, wurde exemplarisch der PEARSON Korrelationskoeffizient zwischen den

Konzentrationsreihen der Saccharose-Proben der ^1H -NMR-Spektren (Abbildung 12) und der Massenspektren (Abbildung 13) berechnet. Über den gesamten Konzentrationsbereich beträgt der Korrelationskoeffizient 0.9073. Über einen Konzentrationsbereich von 0.05 mg/mL bis 10 mg/mL (für ^1H -NMR) bzw. 0.05 $\mu\text{g/mL}$ bis 10 $\mu\text{g/mL}$ (für MS), d.h. innerhalb des linearen Bereiches beider Analysemethoden im Fall von Saccharose (2), beträgt der Korrelationskoeffizient 0.9815. Während bei der NMR-Spektroskopie eine Linearität über die verschiedenen Moleküle hinweg angenommen werden kann, ist die Abhängigkeit der Signalintensität von der Konzentration in der Massenspektrometrie zusätzlich vom jeweiligen Molekül abhängig. Dies ist beispielhaft an den Analyten Chinasäure (3) (Abbildung 14 A) sowie Phenylalanin (4) (Abbildung 14 B) gezeigt.

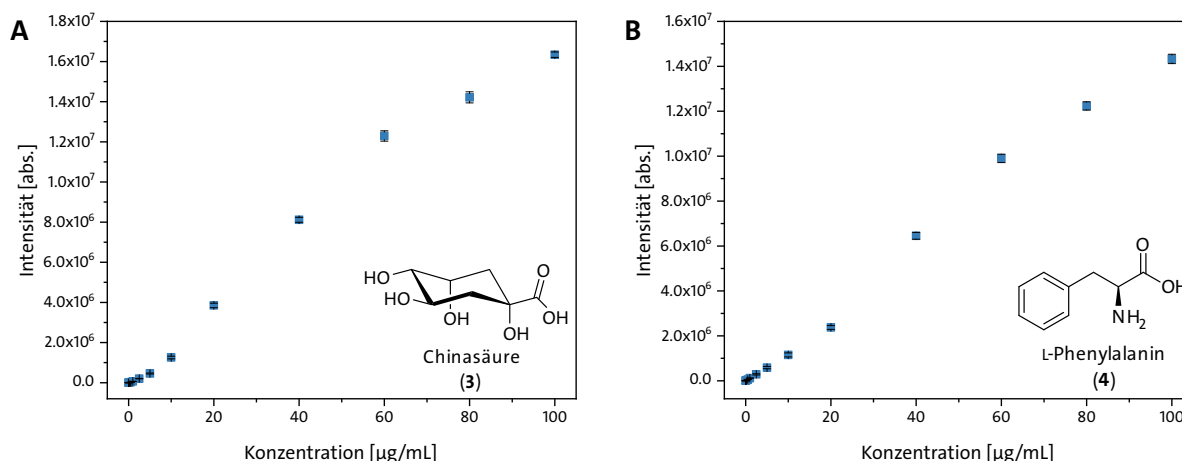


Abbildung 14: Abhängigkeit der Signalintensität von der Konzentration in Massenspektren von Chinasäure (3) (A) und Phenylalanin (4) (B) in einem Konzentrationsbereich von 0.05 $\mu\text{g/mL}$ bis 100 $\mu\text{g/mL}$. Für Chinasäure (3) wurde das $[\text{M}+\text{Na}]^+$ -Signal bei $m/z = 215.05$ und für Phenylalanin (4) das $[\text{M}+\text{H}]^+$ -Signal bei $m/z = 166.09$ analysiert. Es sind jeweils die Mittelwerte aus drei Messungen sowie deren Standardabweichung in Form von Fehlerbalken dargestellt.

Diese beiden Analyten zeigen eine andere Abhängigkeit der Signalintensität von der Konzentration im Vergleich zu Saccharose (2). Es ist keine Sättigung bei einer hohen Konzentration erkennbar, dennoch ist ebenfalls keine exakt lineare Proportionalität zwischen Konzentration und Signalintensität gegeben. Zum Vergleich ist in Abbildung 15 die Konzentrationsreihe von Chinasäure (3) im NMR in einem 1000-fach höheren Konzentrationsbereich gezeigt. Der Korrelationskoeffizient zwischen den beiden Konzentrationsreihen von Chinasäure (3) im MS und im NMR wurde exemplarisch berechnet und weist mit 0.9935 eine starke Korrelation auf.

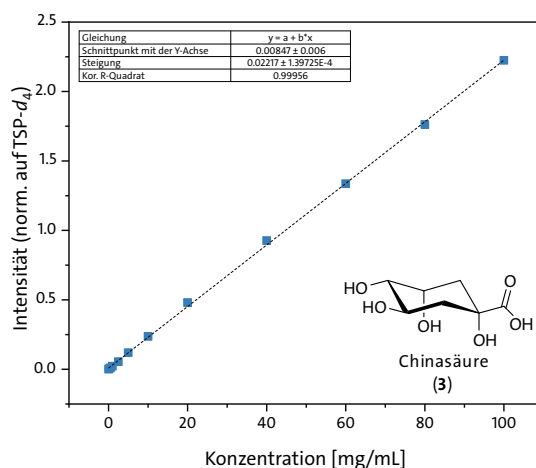


Abbildung 15: Abhängigkeit der Signalintensität von der Konzentration in 1D ^1H -NOESY-Spektren am Beispiel von Chinasäure (3). Es wurde die Signalintensität des Dublett vom Dublett bei 1.95 ppm von Chinasäure (3) in D_2O (+ 5 mM TSP- d_4) in einem Konzentrationsbereich von 0.05 mg/mL bis 100 mg/mL untersucht. Es ist jeweils die auf das TSP- d_4 -normierte Signalintensität aufgetragen.

Während die Probenvorbereitung für die Analytik von Einzelsubstanzen so gestaltet werden kann, dass die Konzentrationen im linearen Bereich liegen, ist dies bei der Untersuchung von komplexen Gemischen nicht ohne Weiteres möglich. Einerseits muss davon ausgegangen werden, dass die einzelnen Substanzen und damit die Grenzen des jeweiligen linearen Bereiches nicht bekannt sind und andererseits führen die großen Konzentrationsunterschiede in der Probe dazu, dass es unwahrscheinlich ist, eine Konzentration zu erreichen, bei der die Konzentrationen aller Metaboliten innerhalb des linearen Bereiches liegen, ohne dass die Intensität der Signale unter die Nachweisgrenze fällt.

Bei der verwendeten Methode erfolgt die Korrelation über einen zeitlichen Konzentrationsverlauf, welcher durch eine chromatographische Verteilung über mehrere Fraktionen erreicht wird. Jede Substanz hat dabei einen individuellen Elutionsverlauf, welcher verschiedene Formen annehmen kann. Durch eine größere Anzahl der Fraktionen steigt dabei die Wahrscheinlichkeit, dass ein Metabolit über mehrere Fraktionen mit niedrigerer Konzentration vorliegt und damit den linearen Bereich nicht überschreitet. Beispielfhaft lässt sich der Effekt an einem idealen, GAUß-förmigen Elutionsprofil zeigen.

Obwohl der lineare Konzentrationsbereich im Rahmen des Elutionsprofils nach oben möglicherweise verlassen wird, verteilt sich die Anzahl der betroffenen Fraktionen nur auf einen verhältnismäßig geringen Teil bezogen auf die Gesamtzahl der Proben bzw. Fraktionen. Zu Beginn und am Ende der Elution eines Analyten sind die Konzentrationen relativ gering, was die

Wahrscheinlichkeit erhöht, dass die Konzentration innerhalb des linearen Bereichs liegt. Die Ungenauigkeiten der EMCs beschränken sich deshalb auf wenige Fraktionen am Maximum der EMCs. Der Effekt ist in Abbildung 16 beispielhaft anhand von simulierten Daten dargestellt. Die Anzahl der betroffenen Datenpunkte ist dabei von der Konzentration abhängig, bei der der lineare Bereich überschritten wird. In Abbildung 16 sind daher verschiedene Fälle dargestellt, bei denen der lineare Bereich bei unterschiedlichen Schwellenwerten überschritten wird.

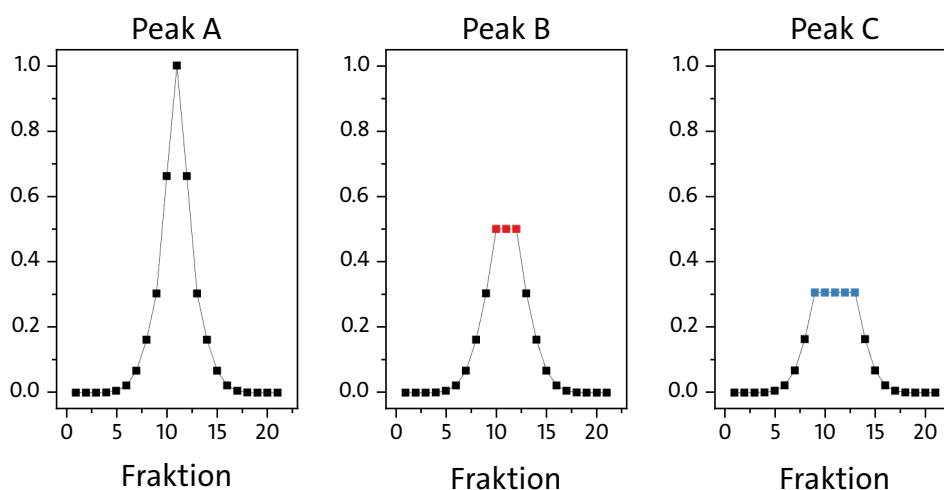


Abbildung 16: Simulierte, normalverteilte Chromatographie-Peaks, deren Verläufe sich lediglich in den mittleren drei (Peak B) bzw. fünf (Peak C), farblich markierten Datenpunkten voneinander unterscheiden.

In der späteren Anwendung könnte der Peak A beispielsweise einem EDC-Verlauf und der Peak B oder Peak C einem EMC-Verlauf entsprechen, falls der Analyt bei hohen Konzentrationen keine lineare Abhängigkeit mehr zeigt. In dem exemplarischen Fall von Peak B in Abbildung 16 tritt eine Sättigung bereits bei 50 % bezogen auf die maximale Intensität des Analyten in Peak A in Fraktion 11 ein. Über die Fraktionen 10-12 wird deshalb bei Peak B ein Plateau des EMC-Verlaufes detektiert. Trotz dieses visuell stark unterschiedlichen Verlaufes, beträgt der PEARSON Korrelationskoeffizient zwischen den beiden Peaks 0.9582. Im Fall des Peaks C tritt die Sättigung bereits bei 30 % der maximalen Intensität bezogen auf Peak A ein, weshalb hier fünf Datenpunkte (Fraktion 9-13) betroffen sind. In diesem Fall beträgt der Korrelationskoeffizient zwischen Peak A und Peak C 0.8672. Je mehr Datenpunkte im Bereich der Sättigung liegen, desto geringer ist der Korrelationskoeffizient. Die geeignete Verdünnung der Proben für die massenspektrometrische Analyse ist deshalb äußerst relevant. Eine zu hohe Verdünnung der Proben kann dazu führen, dass die Signale der niedrig konzentrierten Metaboliten bzw. der Metaboliten mit niedriger Ionisierungseffizienz unter die Nachweisgrenze fallen und damit

nicht mehr detektiert werden können. Andererseits sollte die Verdünnung hoch genug sein, damit weder der Detektor, noch die Analyten eine Sättigung erfahren. Daraus ergibt sich für den Anwender eine individuelle Ermessensfrage, wobei je nach spezifischer Anwendung, Komplexität der Proben, Sensitivität des Massenspektrometers etc. eine individuelle Abwägung erfolgen muss.

Im Allgemeinen kann bei geeigneten Verdünnungen der Proben für die massenspektrometrischen Analysen in den meisten Fällen ab einem Schwellenwert von 0.95 von einer tatsächlichen Korrelation der NMR- und MS-Daten ausgegangen werden. Unter Umständen kann der Korrelationskoeffizient allerdings auch geringer sein, beispielsweise wenn die Intensität in den Massenspektren einzelner Fraktionen zu hoch ist, sodass der lineare Bereich verlassen wird und damit eine höhere Abweichung zum erwarteten Intensitätswert auftritt. Im Fall der NMR-Daten können pH-Wert-Änderungen oder Unterschiede in der Ionenstärke zwischen den verschiedenen Fraktionen zu Verschiebungen der Signale im Verlauf der Fraktionen führen. Auch in diesem Fall wird u. U. ein geringerer Korrelationskoeffizient erhalten, da der EDC nicht mehr den realen Verlauf darstellt. Zur bestmöglichen Vermeidung von Signalverschiebungen im Verlauf der Fraktionen bietet es sich deshalb schon während der Probenvorbereitung an, ein gepuffertes Lösungsmittelsystem zu verwenden (beispielsweise Phosphatpuffer), um Änderungen der pH-Werte in den Fraktionen zu vermeiden. Zusätzlich können *Shifts* in der chemischen Verschiebung zwischen verschiedenen Proben teilweise über *Alignment*-Methoden reduziert bzw. eliminiert werden. Dazu wird bei NMR-Daten häufig der Algorithmus *icoshift* verwendet.^[126] Daneben ist auch die Verwendung von MestReNova geeignet, um ein *Alignment* graphisch durchzuführen.

5.1.1 Entwicklung der graphischen Benutzeroberfläche

Entsprechend des zuvor dargestellten Prinzips der Korrelation von NMR- und MS-Daten über eine dritte Dimension der Flüssigchromatographie wurde eine graphische Benutzeroberfläche entwickelt, um eine allgemein anwendbare Applikation zu konzipieren und eine graphische und semi-automatische Analyse der Daten auf simple Weise zu ermöglichen. Die Entwicklung der Applikation inkl. der graphischen Benutzeroberfläche erfolgte in MATLAB. In Abbildung 17 ist die graphische Benutzeroberfläche nach dem Starten der App gezeigt. Grundsätzlich ist die Oberfläche in drei Teile unterteilt. Der linke Teil dient dem Import und der Prozessierung der

NMR-Daten (links oben) und der MS-Daten (links unten). Im mittleren Teil werden die importierten Daten in einem dreidimensionalen Plot visualisiert und können graphisch analysiert werden. Der rechte Teil der App dient der Korrelation der Daten und der Berechnung der entsprechenden Korrelationskoeffizienten.

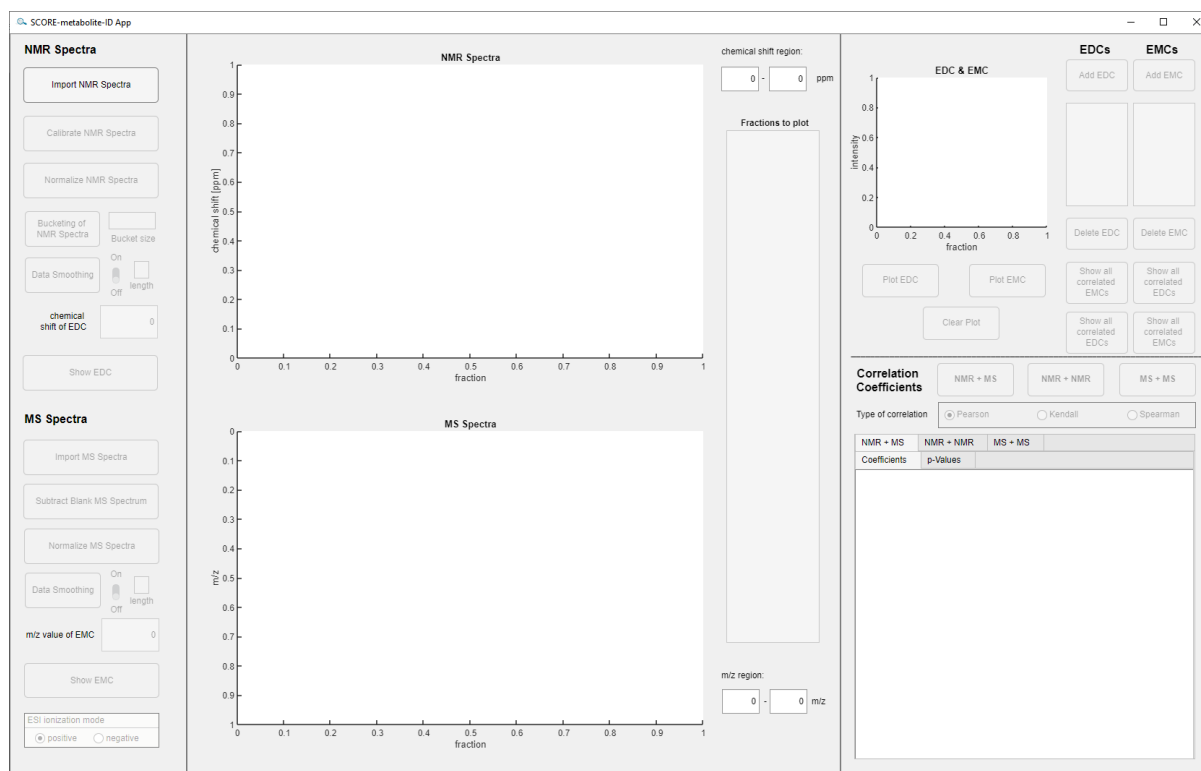


Abbildung 17: Graphische Benutzeroberfläche der in MATLAB entwickelten App *SCORE-metabolite-ID* (Semi-automatic Correlation analysis for REliable metabolite IDentification) vor dem Import und der Analyse von Daten.

5.1.1.1 Import und Prozessierung der NMR-Daten

Das Klicken der Schaltfläche „*Import NMR Spectra*“ öffnet ein Auswahlfenster, in dem ausgewählt werden kann, ob die NMR-Daten im Standard BRUKER Format oder als Text-Datei (.txt) vorliegen. In beiden Fällen müssen die NMR-Daten in einem separaten Ordner gespeichert sein, in dem die jeweiligen Dateien entsprechend der aufsteigenden Fraktionsnummer nummeriert sind. Für den Fall, dass die NMR-Daten im Standard BRUKER Format vorliegen, erfolgt der Import über die Funktion „*rbnmr*“.^[127] Die Verwendung von Text-Dateien ist insbesondere dann sinnvoll, wenn die NMR-Daten zuvor außerhalb von TopSpin (BRUKER) prozessiert wurden, z.B. nachdem ein *Alignment* der Spektren durchgeführt wurde. Für die Erstellung der EDCs ist es unabdingbar, dass sich die chemische Verschiebung des für die EDC-Erstellung ausgewählten

Signals über die konsekutiven Fraktionen nicht ändert, was allerdings aufgrund von pH-Wert-Änderungen, Unterschieden in der Ionenstärke etc. auftreten könnte.^[128] Sofern beispielsweise ein *Alignment* in MestReNova durchgeführt wurde, können die Daten anschließend als Text-Datei des Formates chemische Verschiebung [ppm] vs. Intensität exportiert werden und anschließend in *SCORE-metabolite-ID* importiert werden. Die importierten NMR-Daten werden daraufhin in der graphischen Benutzeroberfläche als dreidimensionaler Plot dargestellt, wobei die chemische Verschiebung [ppm] gegen die Intensität gegen die Fraktion aufgetragen wird.

Nach dem Import findet eine Kalibrierung bzgl. der chemischen Verschiebung der NMR-Daten auf das Signal eines Referenzstandards (z. B. TSP- d_4 bei $\delta = 0.00$ ppm) statt. Auch wenn bereits zuvor eine Kalibrierung innerhalb von bspw. TopSpin stattgefunden hat, ist dieser Schritt erforderlich, um für jede Probe die gleiche x-Achse (entspr. chemische Verschiebung [ppm]) zu erzeugen. Die Kalibrierung innerhalb von *SCORE-metabolite-ID* erfordert die Eingabe eines Bereiches der chemischen Verschiebung, in dem sich das Signal des Referenzstandards befindet (z. B. -0.20 bis 0.20 ppm), sowie die Eingabe der tatsächlichen chemischen Verschiebung, auf die die Kalibrierung erfolgen soll (z. B. $\delta = 0.00$ ppm für TSP- d_4). Nach der Eingabe wird das intensivste Signal innerhalb dieses Bereiches automatisch ermittelt und dieses anschließend zur Kalibrierung genutzt. Die Kalibrierung erfolgt simultan für alle Fraktionen. Anschließend wird über die Schaltfläche „*Normalize NMR Spectra*“ eine Normalisierung der Intensität des Referenzstandards durchgeführt.

Über die Schaltfläche „*Bucketing of NMR Spectra*“ kann ein *Bucketing* durchgeführt werden, wobei dies optional ist. Neben der zuvor erwähnten *Alignment* Methode kann auch ein *Bucketing* genutzt werden, um kleine Änderungen der exakten chemischen Verschiebung von Signalen über verschiedene Proben zu kompensieren. Des Weiteren wird durch das *Bucketing* die Anzahl der Datenpunkte reduziert, was sich positiv auf die Rechenzeit auswirkt. In *SCORE-metabolite-ID* wird ein automatisches *Bucketing* durchgeführt, d. h. es werden Bereiche (*Buckets*) gleicher Größe generiert, innerhalb derer die Signalintensitäten summiert werden.^[27,28] Innerhalb der App kann die Größe der *Buckets* individuell eingestellt werden. Diese kann auch nachträglich über die Schaltfläche „*Change the Bucket size*“ variiert werden. Zu beachten ist dabei, dass die Größe der *Buckets* nicht kleiner als die spektrale Auflösung des Spektrums (RES [ppm]) sein darf, die sich entsprechend Gleichung (8) berechnet.

$$\text{RES [ppm]} = \frac{\text{SW [ppm]}}{\text{SI}} \quad (8)$$

mit SW = *Spectral width* [ppm] und SI = *Size of real spectrum*.

Um den Einfluss möglicher Ausreißer in einzelnen Messungen zu kompensieren, kann optional eine Datenglättung über die Schaltfläche „*Data Smoothing*“ durchgeführt werden. Die Datenglättung erfolgt mit Hilfe der MATLAB-Funktion „*smoothdata*“ mittels *Moving Average* Algorithmus („*movmean*“). Dabei wird der gleitende Durchschnitt in einem Bereich einer festgelegten Anzahl von Datenpunkten berechnet. Bezogen auf die Art der Daten in *SCORE-metabolite-ID* entspricht die Anzahl der Datenpunkte, über die die Glättung stattfinden soll, den konsekutiven Fraktionen. Dieser Wert kann individuell über die Eingabe im sich öffnenden Fenster „*Sliding window length*“ eingestellt und ebenso im Nachhinein angepasst werden.

Im Anschluss besteht die Möglichkeit, EDCs im dreidimensionalen Plot der NMR-Spektren zu visualisieren. Dazu ist die entsprechende chemische Verschiebung [ppm] in das Eingabefeld „*chemical shift of EDC*“ einzugeben. Der Verlauf dieses Signals wird anschließend über die Schaltfläche „*Show EDC*“ im mittleren Teil der App angezeigt. In den beiden Eingabefeldern „*chemical shift region*“ kann ein Bereich der chemischen Verschiebung angegeben werden, um spezifische Bereiche vergrößert darzustellen. Abbildung 18 zeigt die Oberfläche der App nach dem Import und der Prozessierung der NMR-Daten.

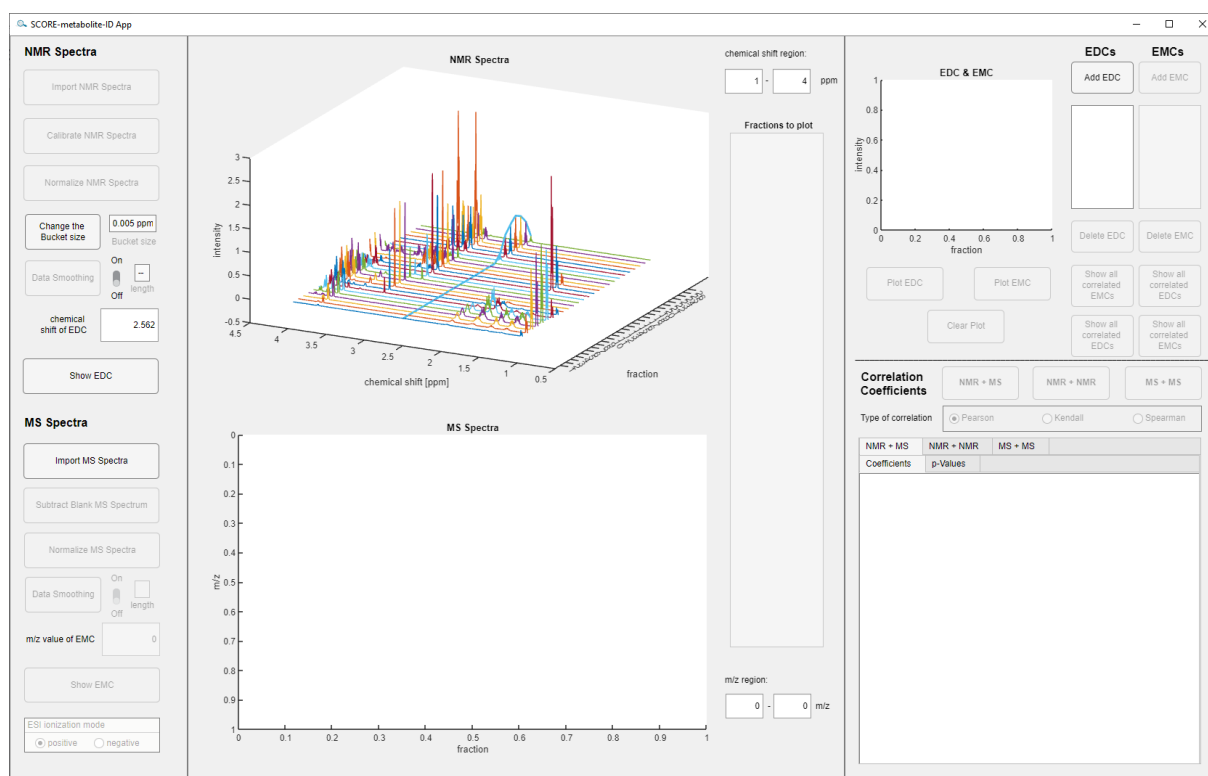


Abbildung 18: Nach dem Import und der Prozessierung der NMR-Daten werden diese im mittleren Teil der App dargestellt. Es wurde eine *Bucket* Größe von 0.005 ppm gewählt und keine Datenglättung durchgeführt. In diesem Fall werden die NMR-Spektren im Bereich zwischen 1 und 4 ppm angezeigt, wie in den Eingabefeldern „chemical shift region“ angegeben ist. Außerdem wird der EDC bei der chemischen Verschiebung 2.562 ppm in blau visualisiert.

5.1.1.2 Import und Prozessierung der MS-Daten

Nach dem Import der NMR-Daten können anschließend die MS-Daten importiert werden. Die Massenspektren können z. B. in Compass DataAnalysis (BRUKER) prozessiert werden und als Text-Datei im Format m/z -Verhältnis vs. Intensität exportiert werden. Damit in der nachfolgenden Analyse der Daten die zu analysierenden EMCs innerhalb der *SCORE-metabolite-ID* App korrekt erkannt und dargestellt werden können, werden die m/z -Werte mit lediglich zwei Nachkommastellen angegeben. Diese Dateien müssen entsprechend ihrer Fraktionsnummer nummeriert sein und in einem separaten Ordner gespeichert sein. Über die Schaltfläche „Import MS Spectra“ kann der Ordner, in dem die Dateien gespeichert sind, ausgewählt werden. Weiterhin ist zum Import die Eingabe der Anzahl der in diesem Ordner befindlichen Massenspektren, d.h. die Anzahl der Fraktionen, erforderlich. Während des Imports werden gleiche x-Achsen (entspr. m/z -Verhältnis) für alle Proben generiert. Über die Schaltfläche

„*Subtract Blank MS Spectrum*“ kann optional eine Subtraktion der bereits im Massenspektrum einer Leermessung (*Blank*) vorhandenen Signale erfolgen. Dazu wird die Text-Datei der Leermessung im Format m/z -Verhältnis (mit zwei Nachkommastellen) vs. Intensität ausgewählt. Die Subtraktion erfolgt für alle bereits importierten Massenspektren der Fraktionen. Außerdem erfolgt die Abfrage des Ionisationsmodus. Diese Angabe hat keinen Einfluss auf die weiteren Schritte der Berechnung, sondern dient lediglich der Übersichtlichkeit. Sofern eine Normalisierung der Massenspektren erfolgen soll und die entsprechende Referenzsubstanz bei den Messungen hinzugegeben wurde, kann dies über die Schaltfläche „*Normalize MS Spectra*“ und die Eingabe des entsprechenden Masse-zu-Ladungs-Verhältnisses erfolgen. In Analogie zur Prozessierung der NMR-Spektren kann auch bei den Massenspektren optional eine Datenglättung über den gleitenden Durchschnitt erfolgen. Im Eingabefeld „*m/z value of EMC*“ kann nun ein Signal vorgegeben werden, dessen Verlauf in Form des EMCs anschließend über die Schaltfläche „*Show EMC*“ im dreidimensionalen Plot der Massenspektren visualisiert werden kann. Dabei muss das m/z -Verhältnis mit zwei Nachkommastellen angegeben werden. Über die Eingabefelder „*m/z region*“ kann auch bei den Massenspektren der darzustellende Bereich angegeben werden. In Abbildung 19 ist ein Überblick der App nach Import der NMR- und MS-Daten gegeben.

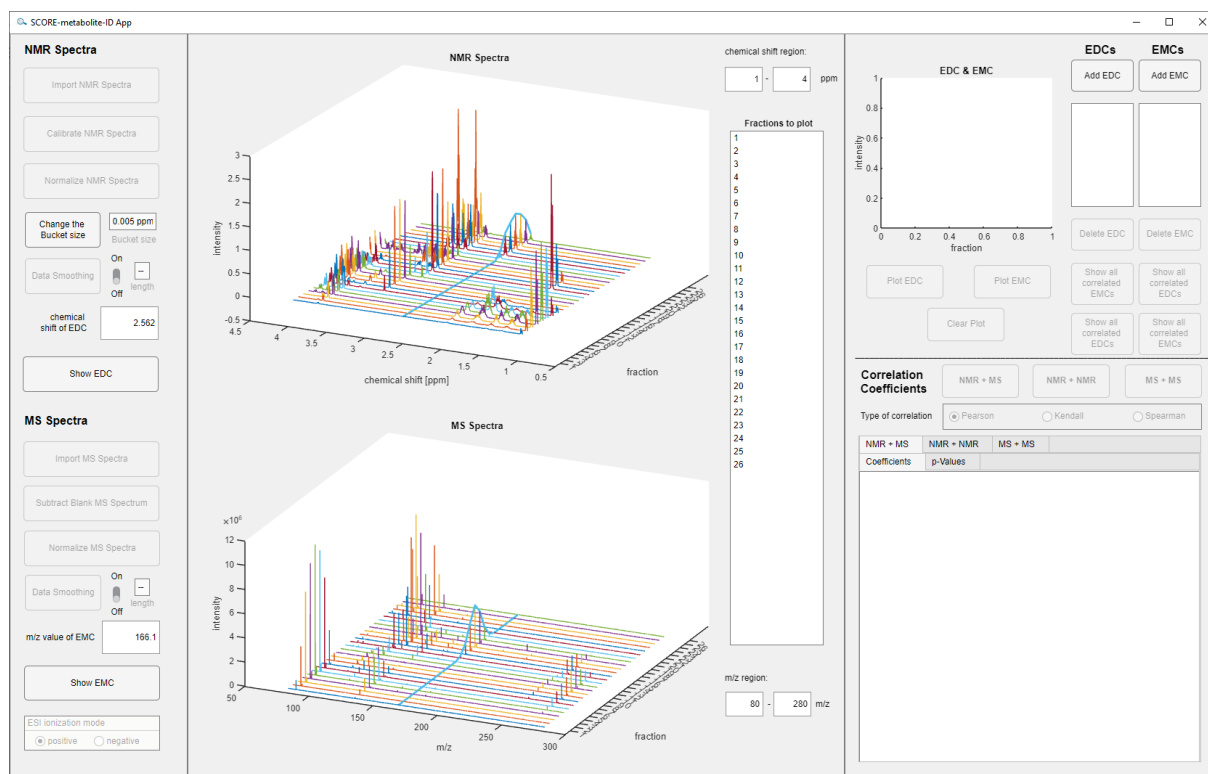


Abbildung 19: Die Massenspektren der Fraktionen werden im unteren mittleren Teil der App dargestellt. In Analogie zu den NMR-Spektren können hier die EMCs über die Eingabe im Feld „*m/z value of EMC*“ im dreidimensionalen Plot visualisiert werden. Nach Import von NMR- und MS-Daten ist die Liste über die zu betrachtenden Fraktionen („*Fractions to plot*“) aktiviert, worüber individuell ausgewählt werden kann, welche Fraktionen geplottet und zur Berechnung der Korrelationskoeffizienten herangezogen werden sollen.

Die aus den hochaufgelösten Massenspektren erhaltenen Ergebnisse erlauben die Angabe der *m/z*-Werte mit einer höheren Genauigkeit von üblicherweise vier Nachkommastellen, als den zwei Nachkommastellen, die im Rahmen der Auswertung mittels der *SCORE-metabolite-ID* App verwendet werden. Um aus den Ergebnissen der *SCORE-metabolite-ID* App z.B. mögliche Summenformeln zu bestimmen, ist es daher sinnvoll die Daten im Anschluss zu präzisieren, indem die genaueren *m/z*-Werte der entsprechenden EMCs aus den Massenspektren herangezogen werden.

5.1.1.3 Fraktionsauswahl

Nachdem sowohl die NMR- als auch die MS-Daten importiert wurden, wird die Liste „*Fractions to plot*“ aktiviert. Darüber kann eine Auswahl getroffen werden, welche Fraktionen angezeigt werden sollen. Die folgende Berechnung der Korrelationskoeffizienten erfolgt nur auf der in

dieser Liste ausgewählten Fraktionen. Diese individuelle Auswahl der zu betrachtenden Fraktionen ist insbesondere dann relevant, wenn innerhalb eines EDCs oder EMCs im zeitlichen Verlauf zwei (oder mehr) getrennte Verläufe zu erkennen sind. Bei NMR-Spektren tritt dieser Fall beispielsweise ein, wenn Signale von verschiedenen Metaboliten, die in verschiedenen Fraktionen vorkommen, die gleiche chemische Verschiebung haben. Insbesondere bei der Verwendung von ^1H -NMR-Spektren mit einer vergleichsweise schmalen Dispersion der Signale von ca. 10 ppm, tritt dieser Fall häufig ein. Auch in den Massenspektren ist diese individuelle Fraktionsauswahl von großer Relevanz, wenn in unterschiedlichen Fraktionen ein Signal mit dem gleichen m/z -Verhältnis erscheint. Dieser Fall tritt z. B. für isobare Verbindungen ein, d. h. Substanzen mit derselben Summenformel, die zu Signalen mit gleichem Masse-zu-Ladungs-Verhältnis führen. Wenn beispielsweise zwei isobare Verbindungen zu unterschiedlichen Zeiten bei der Säulenchromatographie eluieren, führt dies zum Auftreten von zwei Maxima im EMC-Verlauf dieses m/z -Wertes. Über die individuelle Auswahl der Fraktionen können die zu diesem Zeitpunkt irrelevanten Bereiche von der jeweiligen Berechnung ausgeschlossen werden, sodass der EMC nur das Maximum des gerade zu analysierenden Signals enthält. Würde die Korrelation der Daten über alle Fraktionen durchgeführt werden, könnte sich dies negativ auf den berechneten Korrelationskoeffizienten auswirken und dazu führen, dass relevante Ergebnisse nicht wahrgenommen werden.

5.1.1.4 Berechnung von Korrelationskoeffizienten

Die rechte Seite der Benutzeroberfläche dient der Analyse der Daten und der Berechnung der Korrelationskoeffizienten (siehe Abbildung 20). Die zu analysierenden EDCs und EMCs können in die entsprechenden Listen auf der rechten Seite über die Schaltfläche „Add EDC“ bzw. „Add EMC“ hinzugefügt werden. Eine zweidimensionale Darstellung aller in der Liste gespeicherten EDCs bzw. EMCs erfolgt über die entsprechenden Schaltflächen „Plot EDC“ bzw. „Plot EMC“. Insbesondere im Fall von großen Unterschieden in Signalintensitäten bietet sich eine Normalisierung der geplotteten Daten an, um die Darstellung zu vereinfachen. Die Normalisierung kann über einen Rechtsklick in den EDC- bzw. EMC-Plot durchgeführt werden.

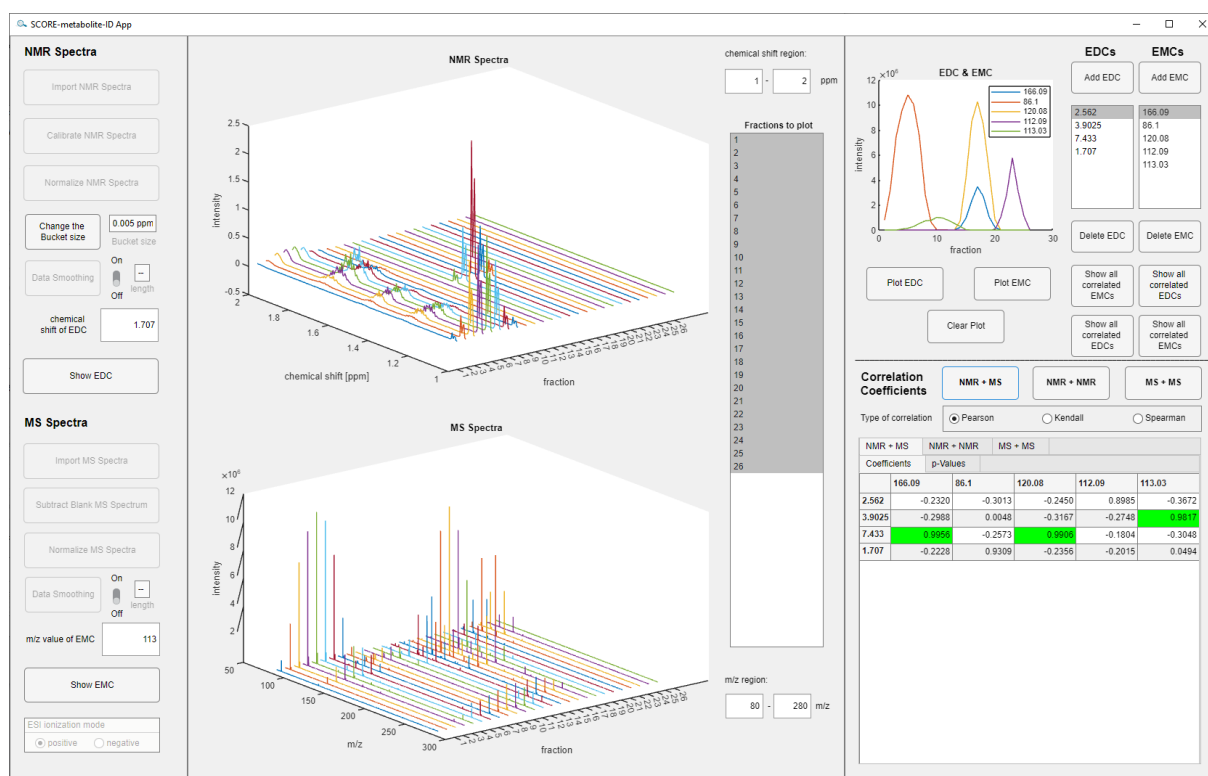


Abbildung 20: Im rechten Teil der App erfolgt über die Schaltfläche „NMR + MS“ die Berechnung der Korrelationskoeffizienten und der p -Werte entsprechend der ausgewählten Korrelationsmethode (PEARSON, KENDALL oder SPEARMAN) über die EDCs und EMCs, die in den obigen Listen aufgeführt sind, und über die Fraktionen, die in der Liste „Fractions to plot“ ausgewählt sind. Analog ist auch die Berechnung der Korrelationskoeffizienten zwischen den EDCs untereinander über die Schaltfläche „NMR + NMR“ sowie zwischen den EMCs über „MS + MS“ möglich. Im oberen Teil können die EDCs und EMCs über die Schaltflächen „Plot EDC“ bzw. „Plot EMC“ im Bereich der ausgewählten Fraktionen zweidimensional dargestellt werden.

Im unteren rechten Teil der App werden die Korrelationskoeffizienten zwischen allen EDCs und EMCs berechnet, die in den obigen Listen aufgeführt sind. Dabei kann unter „Type of correlation“ ausgewählt werden, ob die Korrelationskoeffizienten gemäß PEARSON, KENDALL oder SPEARMAN berechnet werden sollen. Über die Schaltfläche „NMR + MS“ bzw. „NMR + NMR“ oder „MS + MS“ werden die Korrelationskoeffizienten sowie deren p -Werte der jeweilig angegebenen Daten berechnet und im unteren Teil tabellarisch dargestellt. Korrelationskoeffizienten > 0.95 werden grün hinterlegt dargestellt.

5.1.1.5 Semi-automatische Nutzung

Neben der schnellen und einfachen Berechnung von Korrelationskoeffizienten zwischen zuvor ausgewählten EDCs und EMCs, liegt der Vorteil von *SCORE-metabolite-ID* in der semi-automatischen Nutzung, wodurch bspw. MS-Signale detektiert werden können, deren Verlauf zu dem eines definierten NMR-Signals korreliert. Dazu dienen die Schaltflächen „*Show all correlated EMCs*“ bzw. „*Show all correlated EDCs*“ unter den EDC- bzw. EMC-Listen. Wird beispielsweise ein EDC vorgegeben, kann über die sich unter der EDC-Liste befindliche Schaltfläche „*Show all correlated EMCs*“ eine Tabelle erzeugt werden, die alle zu diesem NMR-Signal korrelierenden EMCs innerhalb der ausgewählten Fraktionen mit einem Korrelationskoeffizienten oberhalb eines gewählten Schwellenwertes enthält. Standardmäßig ist dieser Schwellenwert auf 0.95 eingestellt, kann jedoch über die Eingabe im sich öffnenden Fenster geändert werden. Die Liste, die sich wiederum in einem neuen Fenster öffnet, beinhaltet Informationen über den *m/z*-Wert, den Korrelationskoeffizienten, den *p*-Wert sowie die maximale Intensität des MS-Signals, das in einer der ausgewählten Fraktionen detektiert wurde. Die 15 MS-Signale mit maximaler Intensität werden in Grüntönen abnehmender Intensität hervorgehoben. In gleicher Art und Weise können zu dem vorgegebenen EDC auch weitere EDCs detektiert werden, die einen mindestens so hohen Korrelationskoeffizienten wie der Schwellenwert zu dem vorgegebenen EDC aufweisen. Die Tabelle enthält Informationen über die chemische Verschiebung [ppm], den Korrelationskoeffizienten, den *p*-Wert sowie die maximale Intensität in einer der ausgewählten Fraktionen. Für einen vorgegeben EMC in der EMC-Liste können in analoger Weise semi-automatisch korrelierende EDCs und EMCs detektiert werden.

Neben der tabellarischen Analyse können die erhaltenen Korrelationen außerdem visualisiert werden. Der Plot kann über die Schaltfläche „*Create plot of correlations*“ im Kontextmenü der Tabelle hoch-korrelierender EDCs oder EMCs erstellt werden. Zur Visualisierung wird die Funktion *VisLoadings1D_NMR* bzw. *VisLoadings1D_MS* verwendet, welche eine Abwandlung der Funktion *VisLoadings1D* aus der *Edison Lab Metabolomics Toolbox* von ARTHUR EDISON darstellt.^[129] Dabei wird die Kovarianz graphisch dargestellt, die entsprechend des ermittelten Korrelationskoeffizienten gefärbt ist. Korrelationskoeffizienten > 0.95 werden schwarz dargestellt. Eine beispielhafte Darstellung von korrelierenden EDCs zu dem vorgegebenen EMC bei *m/z* = 166.09 ($[M+H]^+$ -Signal von Phenylalanin (**4**)) über 26 Fraktionen im artifiziellen

Gemisch A ist in Abbildung 21 B gezeigt. Zum Vergleich ist ein Referenzspektrum von L-Phenylalanin (**4**) in Abbildung 21 A dargestellt.^[130]

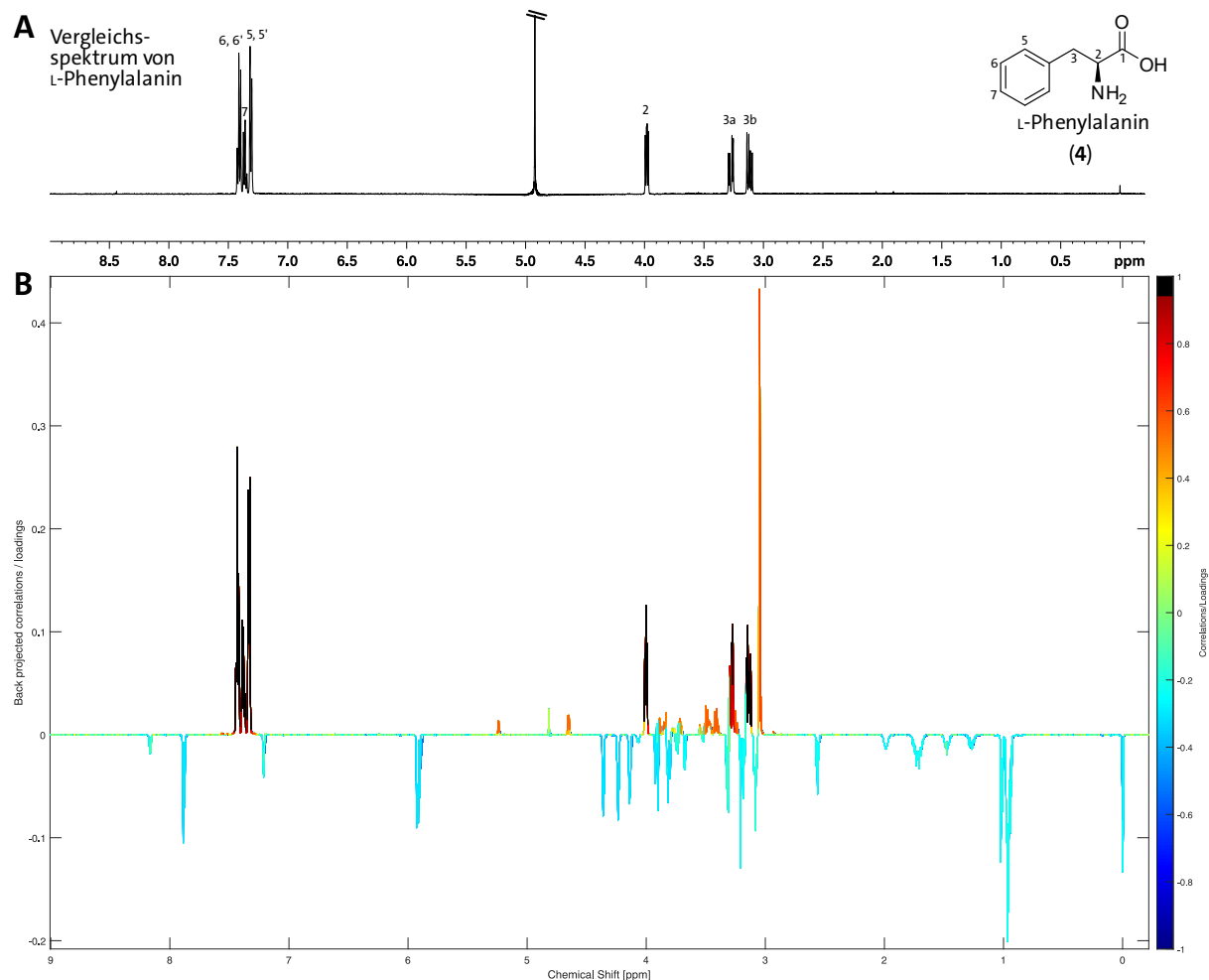


Abbildung 21: A: Datenbank-Vergleichsspektrum von L-Phenylalanin (**4**) in D₂O.^[130] B: Visualisierung der korrelierenden EDCs zu dem vorgegeben EMC bei $m/z = 166.09$ ($[M+H]^+$ -Signal von Phenylalanin (**4**)) im artifizierten Gemisch A durch die Verwendung der Funktion *VisLoadings1D_NMR*, abgewandelt nach *VisLoadings1D*.^[129]

In Analogie können auch die korrelierenden m/z -Verhältnisse in Form eines Massenspektrums dargestellt werden. Diese Art der Darstellung dient der einfachen visuellen Analyse der korrelierenden Signale.

5.1.2 Validierung durch das artifizielle Gemisch A

Zur Evaluierung und Validierung des generellen Konzeptes wurde das artifizielle Gemisch A verwendet, das artifizielle Elutionsverläufe verschiedener Substanzen abbilden soll. Dazu wurden zehn Analyten in jeweils unterschiedlichen Konzentrationen in 26 synthetische

Fraktionen gegeben, um ein säulenchromatographisches Elutionsprofil für jede der zehn Substanzen zu imitieren. Die verwendeten Analyten sind in der folgenden Abbildung 22 dargestellt.

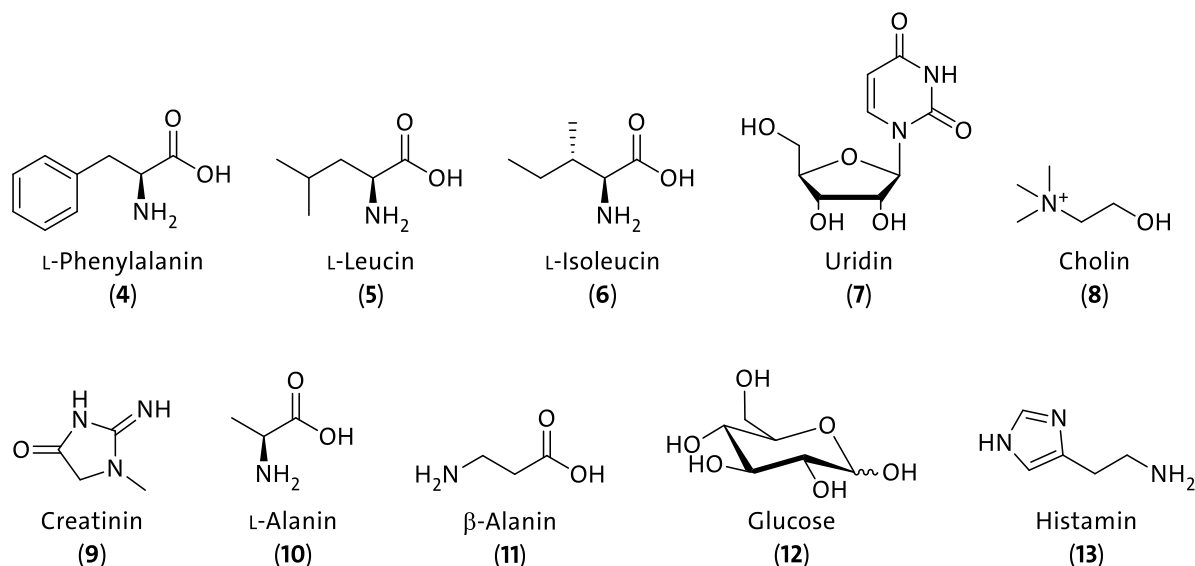


Abbildung 22: Übersicht der zehn Analyten, die zur Erstellung des artifiziellen Gemisches A eingesetzt wurden.

Um ein möglichst realitätsnahes Gemisch zu erhalten, wurden Verbindungen unterschiedlicher Substanzklassen in verschiedenen Konzentrationen eingesetzt. Weiterhin wurde auch die absolute Anzahl der Proben, in denen eine Substanz vorkommt, variiert, um verschiedene Elutionsverläufe zu simulieren. Auf Basis dieses Vorgehens wird eine chromatographische Trennung simuliert und die zeitliche Dimension durch die verschiedenen Proben mit bekannter Zusammensetzung dargestellt. Deshalb entsprechen die Proben dem theoretischen Ergebnis einer Fraktionierung und werden im Folgenden als Fraktion bezeichnet. Bei der Erstellung des Zusammensetzungsprofils wurde dabei versucht, eine Vielzahl von möglichen Problemstellungen realer Fallbeispiele abzubilden. Dazu gehören zum Beispiel der Einsatz isobarer Substanzen und der Einsatz von Substanzen, die in den NMR-Spektren Signale bei der gleichen chemischen Verschiebung erzeugen, was in beiden Fällen aufgrund der unter Abschnitt 5.1.1.3 beschriebenen Eigenschaften die Korrelation der Daten erschwert. Nach der Herstellung der einzelnen Fraktionen wurde jeweils ein Massenspektrum sowie ein NMR-Spektrum aufgenommen und die Daten entsprechend dem zuvor beschriebenen Vorgehen in die Applikation *SCORE-metabolite-ID* importiert. In Abbildung 23 A sind die NMR-Spektren der 26 Fraktionen dargestellt. Von jeder der zehn Verbindungen ist beispielhaft ein EDC farblich

hervorgehoben. In Abbildung 23 B ist jeweils die zweidimensionale Darstellung der EDCs sowie der EMCs gegen die Fraktionen gezeigt.

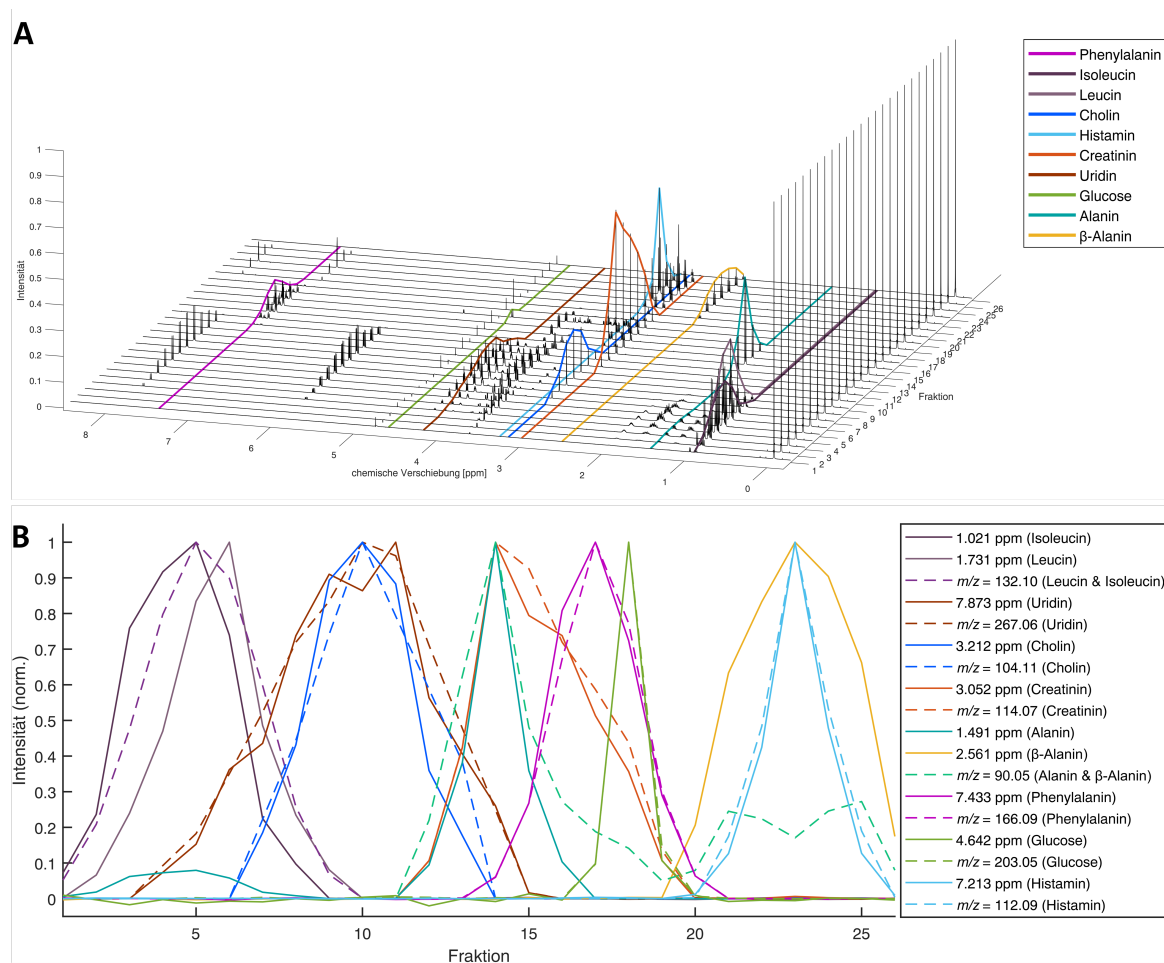


Abbildung 23: **A:** NMR-Spektren des artifiziiellen Gemisches A, das zehn Substanzen verteilt über 26 artifiziielle Fraktionen enthält. Für jede Substanz wurde jeweils ein EDC farblich hervorgehoben. **B:** Zweidimensionale Darstellung der normalisierten EDCs (durchgehende Linien) und EMCs (gestrichelte Linien) der Substanzen des artifiziiellen Gemisches A über alle 26 Fraktionen. Für die NMR-Spektren wurde ein *Bucketing* mit einer *Bucket*-Größe von 0.01 ppm angewendet.

Um adäquate EDCs und EMC-Profile generieren zu können, müssen die Signale hinsichtlich der chemischen Verschiebung bzw. des m/z -Wertes über alle Fraktionen hinweg konsistent sein. Für die NMR-Spektren wurde neben der Verwendung eines Phosphatpuffers, der TSP- d_4 zur Referenzierung der Spektren enthält, in diesem Fall in der *SCORE-metabolite-ID* App ein *Bucketing* mit einer Größe der *Buckets* von 0.01 ppm angewendet. Um die Massengenauigkeit in den Massenspektren zu gewährleisten, wurde sowohl eine externe als auch eine interne Massenkalisierung mit Hilfe eines Natrium-Formiat-Clusters durchgeführt. Zu den in

Abbildung 23 hervorgehobenen EDCs und EMCs sind in Tabelle 1 die berechneten Korrelationskoeffizienten nach PEARSON angegeben.

Tabelle 1: PEARSON Korrelationskoeffizienten zwischen ausgewählten EDCs (*Bucketing*: 0.01 ppm) und EMCs (positiver Ionisationsmodus) des artifiziellen Gemisches A über alle 26 artifiziellen Fraktionen. Die Korrelationskoeffizienten sind entsprechend der Farbskala hervorgehoben.

	<i>m/z</i>	132.10 [M+H] ⁺	267.06 [M+Na] ⁺	104.11 [M] ⁺	114.07 [M+H] ⁺	90.05 [M+H] ⁺	166.09 [M+H] ⁺	203.05 [M+Na] ⁺	112.09 [M+H] ⁺
[ppm]	Analyt	Ile/ Leu	Uri.	Cho.	Creat.	Ala/ β-Ala	Phe	Glc	Hist.
1.021	Ile	0.9412	-0.0882	-0.2360	-0.2791	-0.3640	-0.2311	-0.1268	-0.2024
1.731	Leu	0.9616	0.1014	-0.1352	-0.2699	-0.3535	-0.2253	-0.1227	-0.1998
7.873	Uri.	0.0318	0.9904	0.9340	-0.1785	-0.1593	-0.3024	-0.1672	-0.2755
3.212	Cho.	-0.1583	0.9173	0.9731	-0.2287	-0.2549	-0.2273	-0.1205	-0.1980
3.052	Creat.	-0.2987	-0.1657	-0.1943	0.9956	0.7903	0.5553	0.1577	-0.2129
1.491	Ala	-0.0869	0.0163	-0.0961	0.7343	0.8752	-0.0327	-0.1027	-0.1683
7.433	Phe	-0.2494	-0.3094	-0.2474	0.6011	0.1157	0.9948	0.5148	-0.1810
4.642	Glc	-0.1474	-0.1697	-0.1286	0.2087	-0.0291	0.5728	0.9985	-0.0979
2.561	β-Ala	-0.2926	-0.3722	-0.2899	-0.2895	0.1127	-0.2319	-0.1302	0.8985
7.213	Hist.	-0.2088	-0.2685	-0.2097	-0.2079	0.0687	-0.1692	-0.0943	0.9965

In vielen Fällen wird eine eindeutige Korrelation zwischen EDC und EMC der jeweiligen Substanz mit einem Korrelationskoeffizienten > 0.99 erreicht, wie z.B. im Fall von Creatinin (9), Phenylalanin (4) und Glucose (12). Die co-eluierenden Verbindungen Cholin (8) und Uridin (7) sowie Histamin (13) und β-Alanin (11) weisen ebenfalls relativ hohe Korrelationskoeffizienten zwischen EDC und EMC der jeweils anderen Verbindung auf. Dennoch zeigen die tatsächlich zusammengehörigen EDCs und EMCs deutlich höhere Korrelationskoeffizienten.

Das artifizielle Gemisch A enthält zwei Paare isobarer Verbindungen (Alanin (10) und β-Alanin (11) sowie Leucin (5) und Isoleucin (6)). Im Fall von Alanin (10) und β-Alanin (11) ist eine eindeutige Korrelation zwischen den entsprechenden EDCs und dem EMC hierbei nicht möglich, wenn die Berechnung des Korrelationskoeffizienten über die Gesamtheit aller 26 Fraktionen erfolgt. In Abbildung 24 ist der Verlauf der EDCs von Alanin (10) und β-Alanin (11) sowie des EMCs im negativen Ionisationsmodus bei einem Masse-zu-Ladungsverhältnis von 88.04 ([M-H]⁻) der beiden Verbindungen gezeigt. Im simulierten Elutionsverlauf befindet sich das Maximum des

EDCs von Alanin (**10**) in Fraktion 14 und das Maximum des EDCs von β -Alanin (**11**) in Fraktion 23. Da die isobaren Analyten allerdings das gleiche Quasimolekülion erzeugen, weist der EMC bei $m/z = 88.04$ zwei Maxima in Fraktion 14 und in Fraktion 23 auf. Zusätzlich zeigt auch der EDC von Alanin (**10**) eine geringe Signalintensität in den Fraktionen 1-7, was durch NMR-Signale von Isoleucin (**6**) hervorgerufen wird, die bei der gleichen chemischen Verschiebung auftreten.

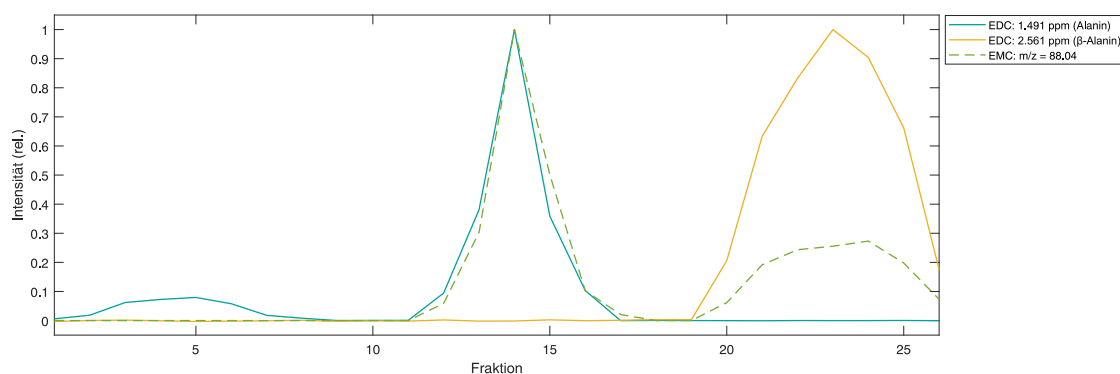


Abbildung 24: Zweidimensionale Darstellung der normalisierten EDCs von Alanin (**10**) (1.491 ppm) und β -Alanin (**11**) (2.561 ppm) sowie des EMC beider Analyten bei $m/z = 88.04$ ($[M-H]^-$).

Wie bereits unter Abschnitt 5.1.1.3 erläutert, erlaubt die *SCORE-metabolite-ID* App es jedoch in solchen Fällen nur einen Teil der Fraktionen für die Berechnung der Korrelationskoeffizienten auszuwählen. Über die graphische Benutzeroberfläche von *SCORE-metabolite-ID* kann dies je nach Einzelfall individuell ausgewählt werden und der zu diesem Zeitpunkt relevante Fraktionsbereich somit einfach und schnell evaluiert werden. In Tabelle 2 sind die berechneten Korrelationskoeffizienten bei Verwendung aller 26 Fraktionen (Spalte 1) oder unter Verwendung einer Teilmenge (Spalten 2 und 3) gezeigt. Somit ist in diesem Fall die eindeutige Zuordnung von NMR- zu MS-Signalen sogar bei isobaren Verbindungen möglich.

Tabelle 2: Korrelationskoeffizienten nach PEARSON zwischen den EDCs bei 1.491 ppm (Alanin (**10**)) und 2.561 ppm (β -Alanin (**11**)) zu dem Quasimolekülion bei $m/z = 88.04$ ($[M-H]^-$) von Alanin (**10**) und β -Alanin (**11**), berechnet über alle 26 Fraktionen (Spalte 1) oder nur einem Teil der Fraktionen (Fraktionen 10-19 in Spalte 2 und Fraktionen 19-26 in Spalte 3).

	Fraktion 1-26	Fraktion 10-19	Fraktion 19-26
Alanin (10) (1.491 ppm)	0.8765	0.9855	0.1590
β-Alanin (11) (2.561 ppm)	0.2348	-0.2912	0.9895

Diese Flexibilität der *SCORE-metabolite-ID* Methode ist von großer Bedeutung. In Abbildung 24 und Tabelle 2 ist dies anhand von isobaren Analyten veranschaulicht, die zu denselben m/z -Werten in den Massenspektren führen. Wie unter Abschnitt 5.1.1.3 beschrieben, ist diese Flexibilität der Auswahl einer Teilmenge von Fraktionen jedoch auch bei der Verwendung der ^1H -NMR-Spektren sehr wichtig, da aufgrund der schmalen Dispersion der ^1H -NMR-Spektren oftmals NMR-Signale unterschiedlicher Moleküle bei der gleichen chemischen Verschiebung auftreten. Wie in Abbildung 24 gezeigt, ist dies beispielsweise für den EDC von Alanin (**10**) bei 1.491 ppm der Fall, da in den Fraktionen 1-7 ein Signal von Isoleucin (**6**) bei der gleichen chemischen Verschiebung auftritt.

Wie in Abbildung 25 zu sehen ist, weisen die Verbindungen Leucin (**5**) und Isoleucin (**6**) im artifiziellen Gemisch A ein nur minimal verschobenes Co-Elutionsprofil auf. Da es sich bei Leucin (**5**) und Isoleucin (**6**) zusätzlich um isobare Verbindungen handelt, führt die Co-Elution dazu, dass der EMC bei $m/z = 132.10$ ($[\text{M}+\text{H}]^+$) prinzipiell dem summierten Elutionsprofil von Leucin (**5**) und Isoleucin (**6**) entspricht. Demzufolge ist der ermittelte Korrelationskoeffizient zwischen dem EMC zu den exemplarischen EDCs der beiden Analyten in Tabelle 1 mit 0.9412 (für Isoleucin (**6**)) und 0.9616 (für Leucin (**5**)) im Vergleich zu anderen Verbindungen des artifiziellen Gemisches A etwas niedriger. Bei einer Co-Elution von isobaren Analyten ist somit keine eindeutige Korrelation zwischen EDCs und dem EMC möglich. Um einem solchen Problem zu begegnen, kann *SCORE-metabolite-ID* jedoch auch dazu genutzt werden, ausschließlich EDCs miteinander zu korrelieren.

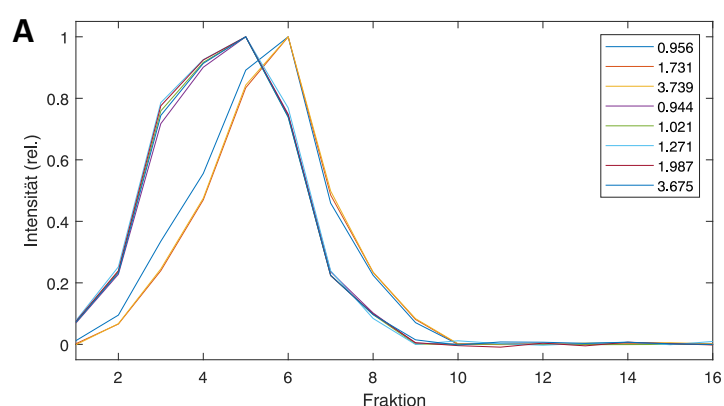


Abbildung 25: Verlauf der verschiedenen EDCs von Isoleucin (**6**) und Leucin (**5**) in Fraktionen 1-16 des artifiziellen Gemisches.

Tabelle 3 zeigt die berechneten Korrelationskoeffizienten zwischen den EDCs von Isoleucin (6) und Leucin (5). Trotz nur minimalem Unterschied des Elutionsprofils ist so eine eindeutige Zuordnung aller NMR-Signale der beiden Aminosäuren möglich.

Tabelle 3: Korrelationskoeffizienten berechnet nach PEARSON zwischen den EDCs von Isoleucin (6) und Leucin (5) in Fraktionen 1-16 des artifiziellen Gemisches A.

[ppm]	0.956	1.731	3.739	0.944	1.021	1.271	1.987	3.675
0.956	1.0000	0.9949	0.9950	0.8886	0.8763	0.8764	0.8738	0.8788
1.731	0.9949	1.0000	0.9999	0.8381	0.8234	0.8236	0.8205	0.8264
3.739	0.9950	0.9999	1.0000	0.8390	0.8243	0.8245	0.8214	0.8273
0.944	0.8886	0.8381	0.8390	1.0000	0.9996	0.9990	0.9992	0.9997
1.021	0.8763	0.8234	0.8243	0.9996	1.0000	0.9997	0.9999	0.9999
1.271	0.8764	0.8236	0.8245	0.9990	0.9997	1.0000	0.9996	0.9994
1.987	0.8738	0.8205	0.8214	0.9992	0.9999	0.9996	1.0000	0.9998
3.675	0.8788	0.8264	0.8273	0.9997	0.9999	0.9994	0.9998	1.0000

In Analogie dazu lassen sich auch EMCs untereinander korrelieren. Dies ist besonders hilfreich, um beispielsweise weitere Addukt- oder Fragment-Ionen-Signale der jeweiligen Substanz in den Massenspektren zu detektieren. Über die semi-automatische Korrelation können nach Vorgabe eines spezifischen EMCs somit weitere EMCs detektiert werden, die einen hohen Korrelationskoeffizienten (z.B. > 0.95) zum *Input* aufweisen. Beispielhaft ist dies für den EMC von $m/z = 132.10$ ($[M+H]^+$ von Isoleucin (6) bzw. Leucin (5)) in Abbildung 26 gezeigt, wobei die Ergebnisse graphisch über die Funktion *VisLoadings1D_MS* dargestellt sind. Die schwarzen Signale entsprechen EMCs, die einen Korrelationskoeffizienten > 0.95 zu dem EMC bei $m/z = 132.10$ aufweisen.

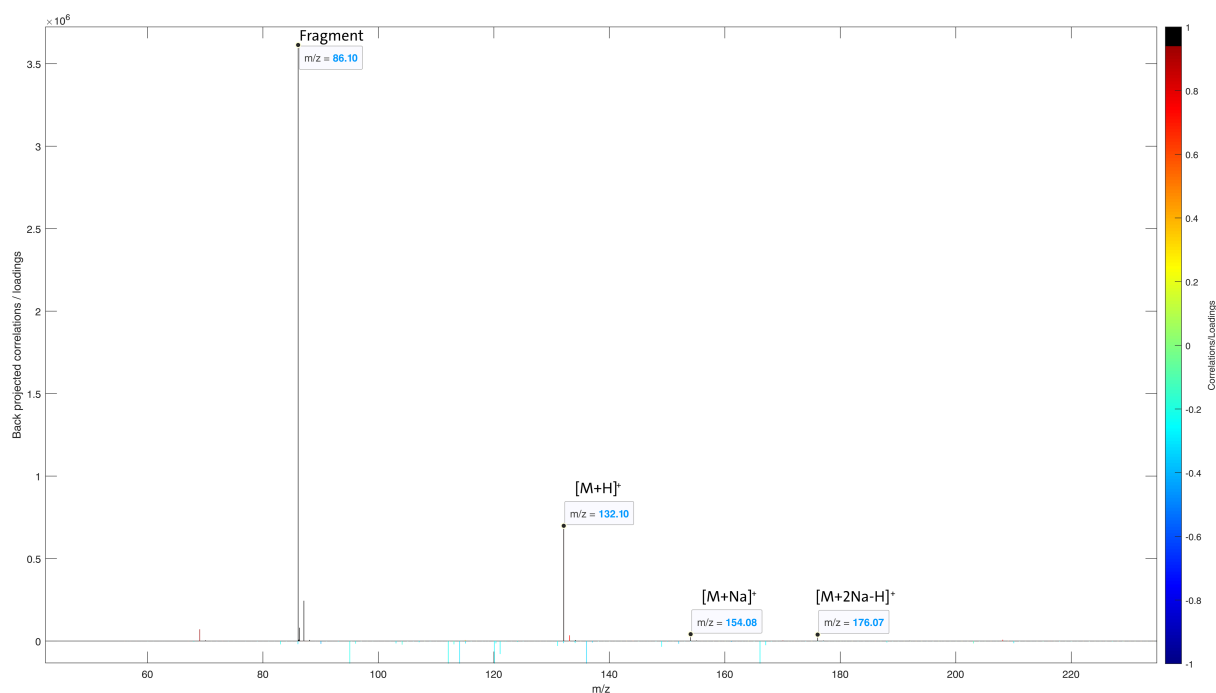


Abbildung 26: Graphische Darstellung der korrelierenden EMCs zu dem vorgegebenen EMC bei $m/z = 132.10$ ($[M+H]^+$ von Leucin (5) bzw. Isoleucin (6)) in Fraktionen 1-26. Neben den Natrium-Addukten mit $m/z = 154.08$ und 176.07 wird das Iminium-Ion als Fragment-Ion mit $m/z = 86.10$ mit einer hohen Intensität detektiert.

Neben der protonierten Form des Analyten bei $m/z = 132.10$ treten weitere Signale mit hohen Korrelationskoeffizienten bei $m/z = 86.10$, $m/z = 154.08$ und $m/z = 176.07$ auf. Diese Signale können zum einen dem Natrium-Addukt ($m/z = 154.08$) sowie der deprotonierten Form des zweifachen Natrium-Adduktes ($m/z = 176.07$) und zum anderen dem Iminium-Ion der Aminosäuren als Fragment-Ion ($m/z = 86.10$) zugeordnet werden. Wie in Abbildung 27 am Beispiel von L-Leucin (5) gezeigt ist, wird durch konsekutive Abspaltung von Wasser und Kohlenstoffmonoxid das Iminium-Ion gebildet, dessen massenspektrometrisches Signal charakteristisch für den Großteil der proteinogenen α -Aminosäuren ist.^[131]

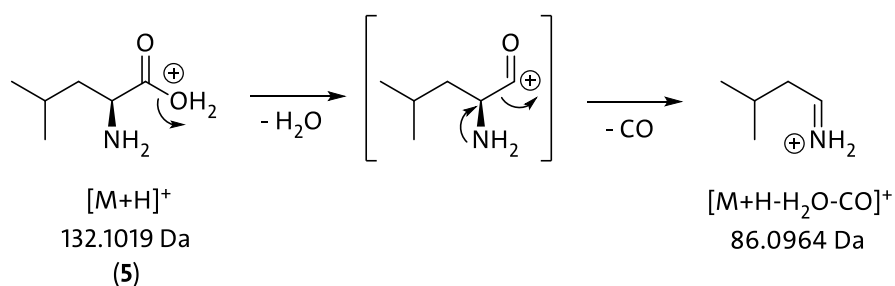


Abbildung 27: Häufig auftretende Fragmentierungsreaktion von Aminosäuren im positiven Ionisationsmodus in ESI-Massenspektren am Beispiel von L-Leucin (5). Durch konsekutive Abspaltung von Wasser und Kohlenstoffmonoxid wird das Iminium-Ion erhalten, welches ein charakteristisches Fragment-Ion vieler α -Aminosäuren ist.^[131]

Die Analyse des artifiziellen Gemisches A zeigt, dass eine Korrelation von NMR- und MS-Signalen, die jeweils variierende Signalintensität in einer Reihe von Fraktionen aufweisen, möglich ist. Durch die Korrelation der Daten wird die Assoziation von NMR- und MS-Signalen derselben Verbindungen ermöglicht und aufgrund der Komplementarität der Informationen die Auswertung von Spektren komplexer Gemische enorm erleichtert. Darüber hinaus können auch NMR-Signale untereinander korreliert werden, was den Nachweis und die Identifizierung isobarer Verbindungen erleichtern kann. Analog können auch die MS-Signale miteinander korreliert werden, um bspw. weitere Addukt- oder Fragment-Ionen-Signale zu detektieren. Die Kenntnis möglichst vieler NMR- und MS-Signale zu einer Verbindung ist insbesondere bei der Identifizierung von unbekannten oder in der untersuchten Matrix nicht erwarteten Verbindungen von großem Vorteil. Ein weiterer Vorteil der *SCORE-metabolite-ID* Methode und dessen graphischer Benutzeroberfläche ist die Flexibilität, z.B. nur eine Teilmenge von Fraktionen für die Analyse zu verwenden. Dies ist insbesondere für die hier verwendeten ^1H -NMR-Spektren mit einer vergleichsweise geringen Dispersion der chemischen Verschiebung notwendig, da in diesen Spektren oftmals Signale bei der gleichen chemischen Verschiebung auftreten, die aber von verschiedenen Substanzen in unterschiedlichen Fraktionen hervorgerufen werden.

5.2 Anwendung von SCORE-metabolite-ID zur Identifizierung von Metaboliten

Nachdem die Methode *SCORE-metabolite-ID* anhand des artifiziellen Gemisches A validiert wurde, wurde diese anschließend zur Identifizierung von Metaboliten in realen Gemischen eingesetzt. Die Analyse realer, komplexer Gemische birgt dabei zusätzliche Herausforderungen, die nicht in Gänze bei der Zusammenstellung eines artifiziellen Gemisches berücksichtigt werden können. Neben der generellen Komplexität der Spektren, die sich aus der Vielzahl an Analyten im Extrakt ergibt, steht der geringen Anzahl bereits bekannter Metaboliten eine Vielzahl unbekannter Analyten gegenüber. Wie bereits unter Abschnitt 3.3 erläutert, lassen sich diese unbekannten Metaboliten weiter in die folgenden beiden Kategorien unterteilen. Zum einen gibt es die „bekannten Unbekannten“ Metaboliten („*Known unknowns*“), die z. B. bereits in anderen Matrices beschrieben, aber in der spezifischen Probe noch nicht bestätigt sind, und damit z. B. durch den Vergleich von Spektren aus Datenbanken oder durch *Spike-In* Experimente vergleichsweise schnell verifiziert werden können. Zum anderen gibt es „unbekannte Unbekannte“ („*Unknown unknowns*“) Metaboliten, deren Struktur gänzlich neuartig ist.^[48] Weitere Herausforderungen ergeben sich aus den teilweise erheblichen Konzentrationsunterschieden zwischen den Metaboliten sowie durch potenzielle Signalüberlagerungen in den ¹H-NMR-Spektren, die sich aus der Komplexität der Matrices ergeben. Da zur Korrelation der Daten auf Basis der EDCs außerdem vorausgesetzt wird, dass die Signale über die fortfolgenden Fraktionen bei jeweils der gleichen chemischen Verschiebung auftreten, stellen potenzielle pH-Wert-Unterschiede oder Unterschiede der Ionenstärke zwischen den Fraktionen eine Herausforderung dar.

Die entwickelte Methode *SCORE-metabolite-ID* soll einen strategischen Ansatz zur Bewältigung einiger dieser Herausforderungen bieten. Durch die flüssigchromatographische Separation wird nicht nur die Komplexität der Gemische reduziert, sondern auch die Grundlage zur Korrelation der analytischen Daten gebildet. Die Korrelation der orthogonalen analytischen Methoden NMR-Spektroskopie und Massenspektrometrie ermöglicht anschließend die schnelle und einfache Ermittlung einer Vielzahl relevanter Informationen zur Identifizierung der Analyten. Die prinzipielle Vorgehensweise zur Analyse der Daten beginnt zunächst meist mit der Auswahl eines Signals von Interesse in den NMR-Spektren oder Massenspektren. Zur Vermeidung von Änderungen der chemischen Verschiebung im Verlauf der Fraktionen wird neben der Verwendung eines Phosphatpuffers bei der Probenvorbereitung zusätzlich ein *Bucketing* mit

frei zu wählender *Bucket*-Größe in der *SCORE-metabolite-ID* App durchgeführt. Durch die semi-automatische Korrelation können anschließend zugehörige NMR- oder MS-Signale detektiert werden. Im Fall von unbekannten Metaboliten können die Signale der Massenspektren daraufhin häufig zur Ermittlung einer wahrscheinlichen Summenformel genutzt werden. Die Analyse der chemischen Verschiebungen, Multiplizitäten, Integrale und Kopplungskonstanten der ^1H -NMR-Daten ermöglicht weiterhin meist die Ermittlung einzelner Struktur motive. Insbesondere die Dereplikation bekannter Metaboliten ist somit prinzipiell durch die Kombination der Daten besonders effizient möglich.

Sofern z.B. durch Überlagerung von Signalen oder durch starke Konzentrationsunterschiede unter Anwendung von *SCORE-metabolite-ID* nicht alle relevanten NMR- oder MS-Signale eindeutig detektiert werden können, ist zusätzlich die Erweiterung der analytischen Daten um bspw. weitere NMR-Spektren (wie z.B. selektive oder zweidimensionale NMR-Experimente) oder LC-MS-Messungen möglich. Auch ein gezielter Datenbankvergleich oder ein *Spiking* der Analyten zur Verifizierung kann systematisch eingesetzt werden. Durch die vorherige Anwendung von *SCORE-metabolite-ID* können jedoch häufig bereits Strukturfragmente identifiziert oder Strukturvorschläge generiert werden, wodurch anschließend gezielte Experimente zweckgerichtet durchgeführt werden können.

In den folgenden Kapiteln wird die Identifizierungsstrategie auf Basis der *SCORE-metabolite-ID* Methode zunächst zur Identifizierung von Metaboliten aus dem polaren Extrakt der Pinienkernen verwendet. Außerdem wurden Apfelproben verschiedener Sorten mittels multivariater Datenanalyse untersucht und potenzielle Markersignale mit Hilfe von *SCORE-metabolite-ID* analysiert.

5.2.1 Identifizierung von Metaboliten aus polarem Pinienkernextrakt

Nüsse und nussähnliche Lebensmittel gelten als nährstoffreiche und hochwertige Lebensmittel und weisen einen hohen Gehalt an ungesättigten Fettsäuren auf.^[132] Daneben beinhalten Nüsse und Kerne eine Vielzahl sekundärer Pflanzenstoffe, die multifunktionelle Eigenschaften aufweisen können.^[132] Strukturell handelt es sich dabei z.B. um Flavonoide, Tannine oder Alkaloide. Im Allgemeinen handelt es sich demzufolge bei Nüssen und nussähnlichen Lebensmitteln um komplexe Matrices biologischen Ursprungs mit einer Vielzahl von Signalen

in den entsprechenden ^1H -NMR-Spektren. Aufgrund der hohen Preise und des weltweiten Handels sind Nüsse und Kerne besonders anfällig für Lebensmittelbetrug (*Food Fraud*).^[133] Im Rahmen der von Europol und Interpol koordinierten Operation OPSON VI im Jahr 2016/2017 wurde bspw. die Vermischung von 1.3 Tonnen gerösteten Haselnüssen mit ca. 8 % Erdnüssen festgestellt.^[134] Neben wirtschaftlichen Vorteilen für die Produzenten in Form von Streckung der Ware durch Beimischungen der kostengünstigeren Erdnüsse, ergibt sich durch derartigen Lebensmittelbetrug auch eine mögliche Gesundheitsgefahr, z.B. durch Allergien.^[133] Ebenso werden teilweise die bei Pesto deklarierten Pinienkerne durch die kostengünstigeren Cashew-Kerne ersetzt.^[135] Zum Nachweis von möglichen Beimischungen von z. B. Nüssen oder Kernen ist deshalb die Detektion und Identifizierung von Markersignalen erforderlich. In Abbildung 28 sind die ^1H -NMR-Spektren der polaren Extrakte nach Zwei-Phasen-Extraktion mit Methanol, Chloroform und Wasser (Verhältnis 4:5:6) verschiedener Nüsse und Kerne (Mandel, Haselnuss, Erdnuss, Walnuss, Pinienkerne, Sonnenblumenkerne, Cashewkerne und Kürbiskerne) im Vergleich gezeigt.

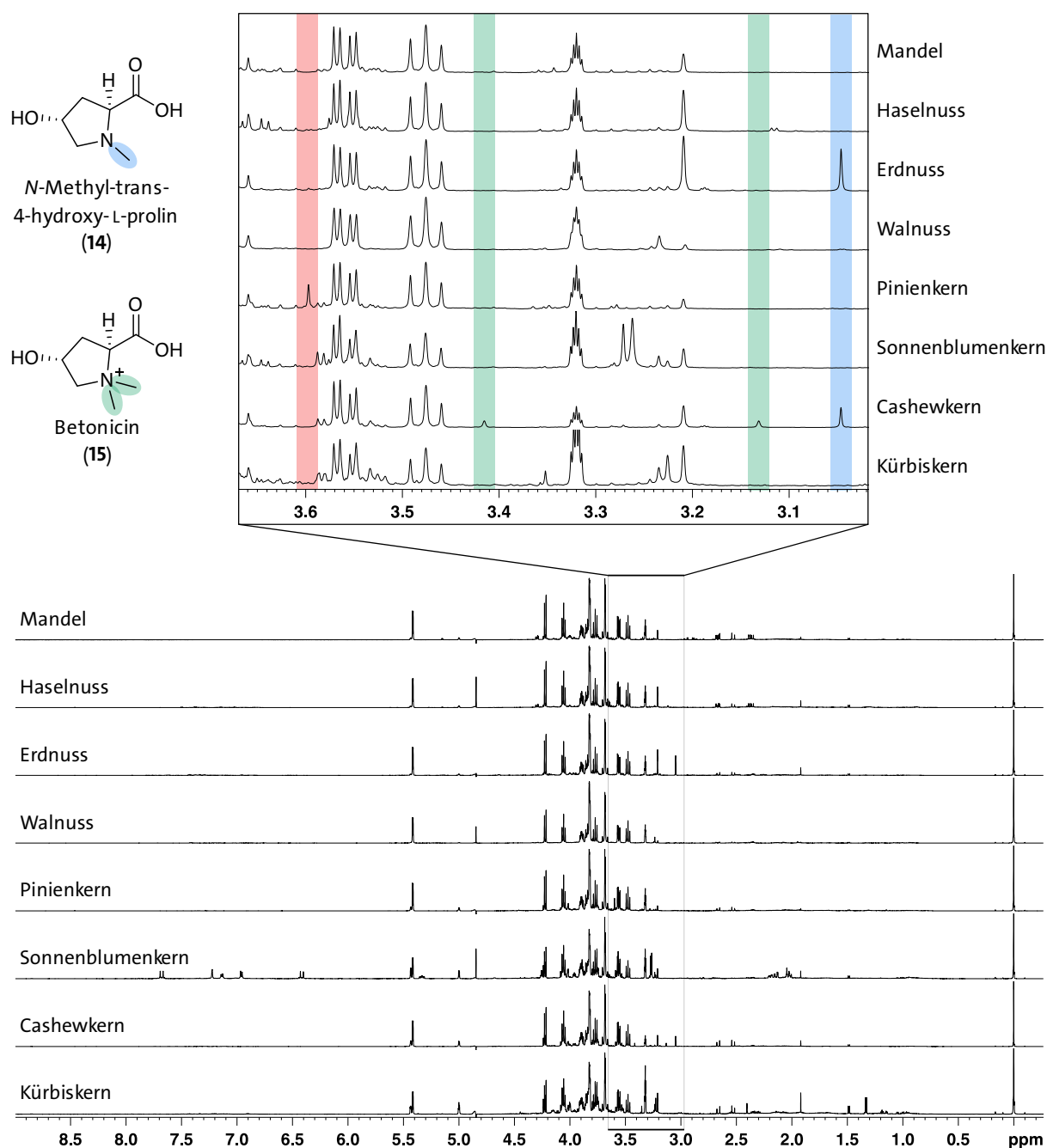


Abbildung 28: ¹H-NMR-Spektren der polaren Extrakte verschiedener Nüsse und Kerne. Der Bereich zwischen 3.0 und 3.7 ppm ist vergrößert dargestellt. Signale, die sich u.U. als Marker eignen, sind farblich hervorgehoben. Der blau hinterlegte Bereich zeigt ein Signal des von DR. CAROLINE SCHMITT in Erdnüssen identifizierten Metaboliten *N*-Methyl-*trans*-4-Hydroxy-L-Prolin (**14**) bei 3.05 ppm und die grün hinterlegten Bereiche zeigen Signale des von NATHALIE HOLZ in Cashewkernen identifizierten Metaboliten Betonicin (**15**) bei 3.13 und 3.41 ppm.^[45,136] Das rot hinterlegte Signal tritt lediglich im ¹H-NMR-Spektrum der Pinienkerne auf. Es wurde die *noesygprr1d*-Pulssequenz mit einer Wasserunterdrückung bei 4.808 ppm verwendet und jeweils 128 *Scans* an einem BRUKER Avance III HD 600 MHz Spektrometer aufgenommen.

In der vergrößerten Darstellung des Bereichs zwischen 3.0 und 3.7 ppm sind einzelne Signale farbig hinterlegt. Diese Signale können als Marker für die verschiedenen Nüsse und Kerne dienen. Das blau hinterlegte Signal (3.05 ppm) in Erdnuss-Proben wurden von DR. CAROLINE SCHMITT bereits als (2*S*,4*R*)-*N*-Methyl-4-Hydroxy-L-Prolin (**14**) identifiziert.^[45] Dieser Metabolit kann beispielsweise als Indikator zur Detektion von Erdnuss-Beimengungen in Mandeln, Haselnüssen oder Walnüssen dienen.^[137] In Cashewkernen konnten von NATHALIE HOLZ die grün hinterlegten Signale bei 3.13 und 3.41 ppm als Betonicin (**15**) identifiziert werden.^[136] Das NMR-Spektrum der Pinienkerne zeigt ein Markersignal bei 3.60 ppm, das in keinem der Spektren der anderen Nüsse und Kerne auftritt. Zusätzlich weist die Pinienkern-Probe eine Vielzahl überlagerter Signale im aliphatischen und aromatischen Bereich auf und stellt deshalb eine vielversprechende Matrix für die Untersuchung mittels *SCORE-metabolite-ID* dar.

Die gemahlene und lyophilisierten Pinienkerne wurden in einer Zweiphasen-Extraktion mit Chloroform, Methanol und Wasser extrahiert, die polare Phase anschließend separiert und das Lösungsmittel entfernt. Das ¹H-NMR-Spektrum des in Deuteriumoxid gelösten Extraktes ist in Abbildung 29 abgebildet.

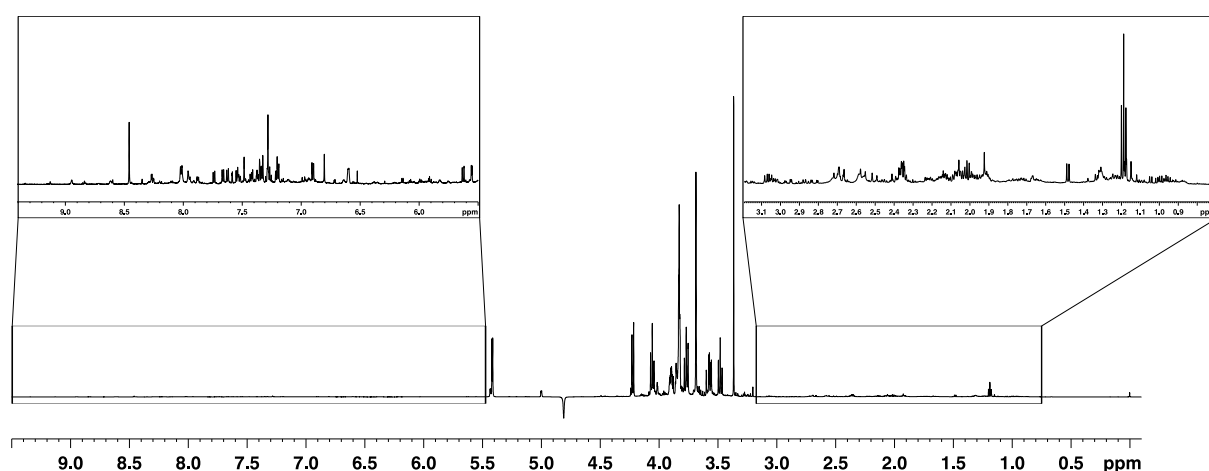


Abbildung 29: ¹H-NMR-Spektrum in D₂O des polaren Gesamtextraktes der Pinienkerne. Die Signale im aliphatischen und aromatischen Bereich sind vergrößert dargestellt. Es wurde die *noesygppr1d*-Pulssequenz mit einer Wasserunterdrückung bei 4.699 ppm verwendet und 128 *Scans* an einem BRUKER Avance III HD 600 MHz Spektrometer aufgenommen.

Das ¹H-NMR-Spektrum zeigt eine Vielzahl von überlagerten Signalen, die eine Identifizierung erschweren. Die Signale treten im gesamten Bereich der chemischen Verschiebung von ca. 0.7 bis 9.5 ppm auf und stammen von verschiedenen Verbindungen unterschiedlicher Substanzklassen, wie Kohlenhydraten, organischen Säuren, Aminosäuren, aromatischen

Verbindungen etc. Zusätzlich weisen die starken Intensitätsunterschiede auf bedeutende Konzentrationsunterschiede der Substanzen hin. Deutlich wird dies unter anderem durch den Vergleich der intensiven Signale der Kohlenhydrate im Bereich zwischen 3 und 4.5 ppm und den vergleichsweise geringen Intensitäten der Signale im aromatischen Bereich. Eine direkte Identifizierung von Metaboliten aus dem Spektrum des Gesamtextraktes ist daher kaum möglich, was in besonderem Maße für die gering konzentrierten Metaboliten gilt.

Der erste Schritt der Korrelation der NMR- und MS-Daten mittels *SCORE-metabolite-ID* ist die Generierung der dritten Dimension, über die eine Korrelation erfolgt. Dies wird über die flüssigchromatographische Verteilung der Metaboliten erreicht. Zur Entwicklung der optimalen Fraktionierungsmethode wurden verschiedene Ansätze getestet. Die Methoden sind unter Abschnitt 6.6.3.1 aufgeführt. Zunächst wurde eine reguläre RP-Säule mit einer mit C₁₈-Ketten und hydrophilem *Endcapping* modifizierten stationären Phase eingesetzt. Die Fraktionierung nach klassischer Umkehrphase (Methoden A, B und C) mit einer stark hydrophoben stationären Phase ist jedoch für die Verteilung der polaren Analyten ungeeignet. Der Großteil der Metaboliten wird kaum retardiert und eluiert simultan in den Fraktionen 4 bis 10, sodass kaum individuelle Konzentrationsverläufe beobachtet werden können und diese zusätzlich auf eine sehr begrenzte Anzahl an Fraktionen beschränkt sind. Diejenigen Analyten, die eine gewisse Retention erfahren, zeigen eine ausgeprägte Verzögerung der vollständigen Elution über die späteren Fraktionen, was sich in einem langen *Tailing* äußert. Um die Retention der Analyten zu optimieren, wurde daraufhin eine Amino-funktionalisierte Säule unter RP-Bedingungen getestet (Methode D). Die Verteilung der Analyten erstreckte sich auf ca. 20 Fraktionen, sodass sich die Verwendung dieser Säule als deutlich geeigneter im Vergleich zur klassischen RP-Säule erwies. Zur Steigerung der Verteilung über mehr Fraktionen wurde die gleiche Säule erneut unter HILIC-Bedingungen eingesetzt (Methode E). Es zeigte sich eine weitere Verbesserung der Verteilung der Analyten auf ca. 42 Fraktionen. Die eingesetzte Säule der Firma BÜCHI zeigte jedoch sowohl bei Methode D als auch bei Methode E erhebliche Instabilitäten der stationären Phase („Säulenbluten“). Da sich dies auch durch wiederholtes Konditionieren und Spülen nicht beheben ließ und sich die Verschleppung der Silica-Partikel nachteilig auf die NMR-Spektren der Fraktionen auswirkten, wurde der Einsatz dieses speziellen Säulenmaterials nicht weiterverfolgt. Prinzipiell erwies sich die Amino-Funktionalisierung der stationären Phase jedoch als geeignet, sodass vergleichbares Säulenmaterial von der Firma MACHEREY-NAGEL bezogen und getestet wurde. Auch diese stationäre Phase zeigt leichte Instabilitäten bei hohen

Anteilen wässriger Phase, jedoch ist der Effekt deutlich geringer als zuvor. Da das Säulenmaterial separat zu erwerben ist, bietet dies zusätzlich die Möglichkeit der Adsorption des Extraktes an das Säulenmaterial. Dies ist unter den verwendeten HILIC-Bedingungen besonders sinnvoll, da das polare Extrakt nicht in den Startbedingungen des Gradienten bestehend aus 90 % Acetonitril und lediglich 10 % Wasser löslich ist. Die Methode F erwies sich als besonders geeignete Methode zur Fraktionierung des polaren Extraktes.

Zur Untersuchung des polaren Extraktes der Pinienkerne wurde als stationäre Phase demnach eine HILIC-Phase mit Aminopropyl-Modifikation eingesetzt. Im Vergleich zur optimierten Methode F wurde der Zeitraum der Fraktionierung um weitere 10 Minuten verlängert, um die vollständige Elution der Analyten zu gewährleisten. Entsprechend dieser Fraktionierungsmethode (siehe Abschnitt 6.6.3.2) wurden somit 80 Fraktionen erhalten, von denen nach Probenaufarbeitung jeweils ein 1D ^1H -NOESY-Spektrum und ein Massenspektrum, sowohl im positiven als auch im negativen Ionisationsmodus, aufgenommen wurden. Nach der Datenprozessierung wurden die NMR- und MS-Daten mit Hilfe der *SCORE-metabolite-ID* App analysiert. Für die NMR-Daten wurde, wenn nicht anders angegeben, ein *Bucketing* mit einer Größe von 0.005 ppm angewendet.

5.2.1.1 Identifizierung von Kohlenhydraten, Cyclitolen und Alditolen

Zunächst wurde das Markersignal bei einer chemischen Verschiebung von 3.60 ppm analysiert. Durch die semi-automatische Korrelation nach Vorgabe des EDCs bei 3.599 ppm konnten spezifische EMCs im positiven Ionisationsmodus zu dem NMR-Signal zugeordnet werden. In Tabelle 4 sind die vier intensivsten EMCs, die einen Korrelationskoeffizient > 0.95 zu dem EDC bei 3.599 ppm aufweisen, aufgelistet.

Tabelle 4: Intensivste EMCs, die einen Korrelationskoeffizient > 0.95 zu dem vorgegebenen EDC bei 3.599 ppm in Fraktion 1-80 aufweisen. Neben dem m/z -Wert und dem Korrelationskoeffizient ist die ermittelte Zuordnung der EMCs angegeben.

EMC [m/z]	Korrelationskoeffizient	Zuordnung
195.09	0.9511	$[\text{M}+\text{H}]^+$
217.07	0.9822	$[\text{M}+\text{Na}]^+$
218.07	0.9834	$[\text{M}+\text{Na}]^+$ Isotopenpeak
411.15	0.9736	$[2\text{M}+\text{Na}]^+$

Anhand der erhaltenen EMCs, die einen hohen Korrelationskoeffizienten zu dem vorgegebenen EDC aufweisen, kann eine Zuordnung der m/z -Werte getroffen werden, indem charakteristische Massendifferenzen ermittelt werden, welche sich verschiedenen Addukten zuordnen lassen. Es ergab sich eine chemisch sinnvolle Zuordnung, wenn der EMC bei $m/z = 195.09$ als $[M+H]^+$ -Signal zu Grunde gelegt wird. Die weiteren m/z -Werte können unter dieser Annahme als Natrium-Addukt ($m/z = 217.07$) sowie als ^{13}C -Isotopenpeak des Natrium-Adduktes ($m/z = 218.07$) und als Dimer des Natrium-Adduktes ($m/z = 411.15$) zugeordnet werden. Anhand der Ergebnisse und Schlussfolgerungen kann die monoisotopische Masse des Metaboliten errechnet werden, wobei sich ein Wert von 194.08 Da (bzw. 194.0788 ± 0.0005 Da) ergibt. Eine mögliche Summenformel wäre somit $\text{C}_7\text{H}_{14}\text{O}_6$ mit einer monoisotopischen Masse von 194.0790 Da. Auf Basis der geringen Abweichung der gemessenen und der berechneten monoisotopischen Masse von 1.0 ppm und der sinnvoll erscheinenden Anzahl von einem Doppelbindungsäquivalent wird diese Summenformel als wahrscheinlichste angesehen. Nach der erfolgreichen Ermittlung der Summenformel des Metaboliten wurden die weiteren NMR-Signale untersucht. In Abbildung 30 sind die hoch-korrelierenden EDCs zu dem vorgegebenen EMC bei $m/z = 217.07$ in den Fraktionen 1 bis 37 gezeigt. Die Anzahl der Fraktionen wurde verringert, da die EDCs in den späteren Fraktionen erneut Signalintensität aufweisen, die durch das Auftreten von Signalen bei der gleichen chemischen Verschiebung von anderen Metaboliten hervorgerufen wird.

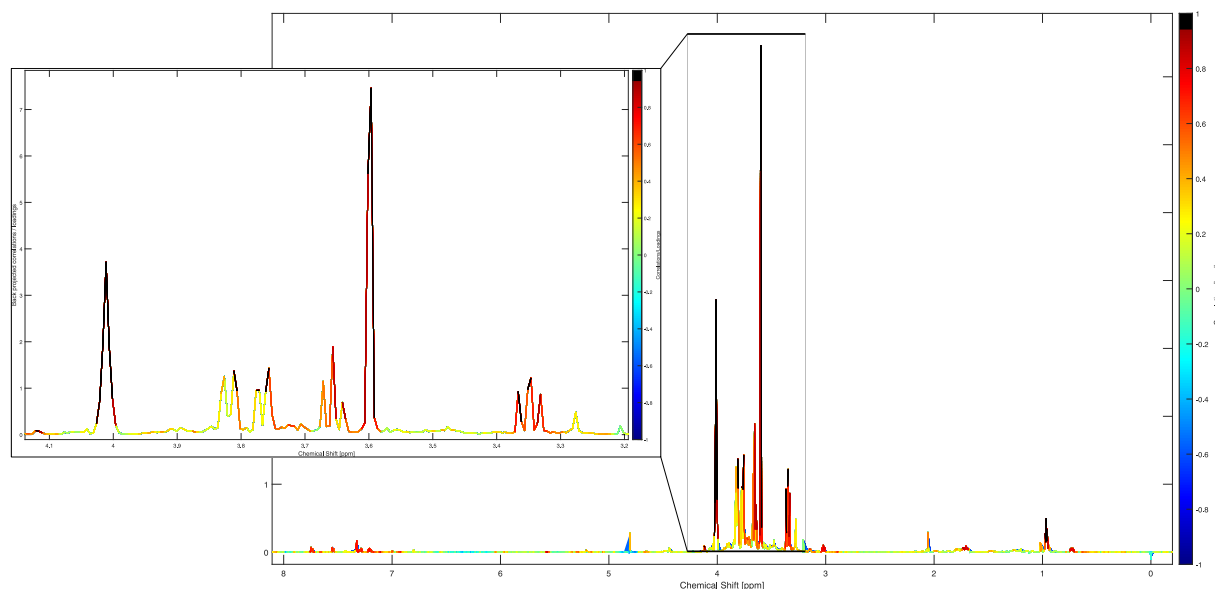


Abbildung 30: Graphische Darstellung der EDCs, die einen Korrelationskoeffizienten > 0.95 zu dem EMC bei $m/z = 217.07$ in den Fraktionen 1-37 aufweisen.

Es werden einige Signale im Bereich zwischen 3.3 und 4.1 ppm detektiert, die einen hohen Korrelationskoeffizienten zu dem EMC des vermeintlichen Natrium-Adduktes des Metaboliten aufweisen. Zusätzlich weist ein Signal bei einer chemischen Verschiebung von 0.97 ppm hohe Korrelationskoeffizienten auf. Über weitere 2D- und selektive NMR-Experimente kann die Zugehörigkeit des Signals bei 0.97 ppm zu den anderen Signalen ausgeschlossen werden. Das Singulett bei einer chemischen Verschiebung von 3.599 ppm lässt auf einen Methoxy-Substituenten schließen. Auf Grundlage aller vorhandenen Daten wurde darauf geschlossen, dass es sich um D-Pinitol (3-*O*-Methyl-D-*chiro*-Inositol) (**16**) mit einer monoisotopischen Masse von 194.0790 Da ($C_7H_{14}O_6$) handelt. Zur Überprüfung wurde ein *Spike-In* Experiment durchgeführt. Es wurden zwei Mal je 10 μ L einer 0.1 M Lösung von kommerziell erworbenem D-Pinitol (**16**) zu der Fraktion 33 dazugegeben (siehe Abbildung 31).

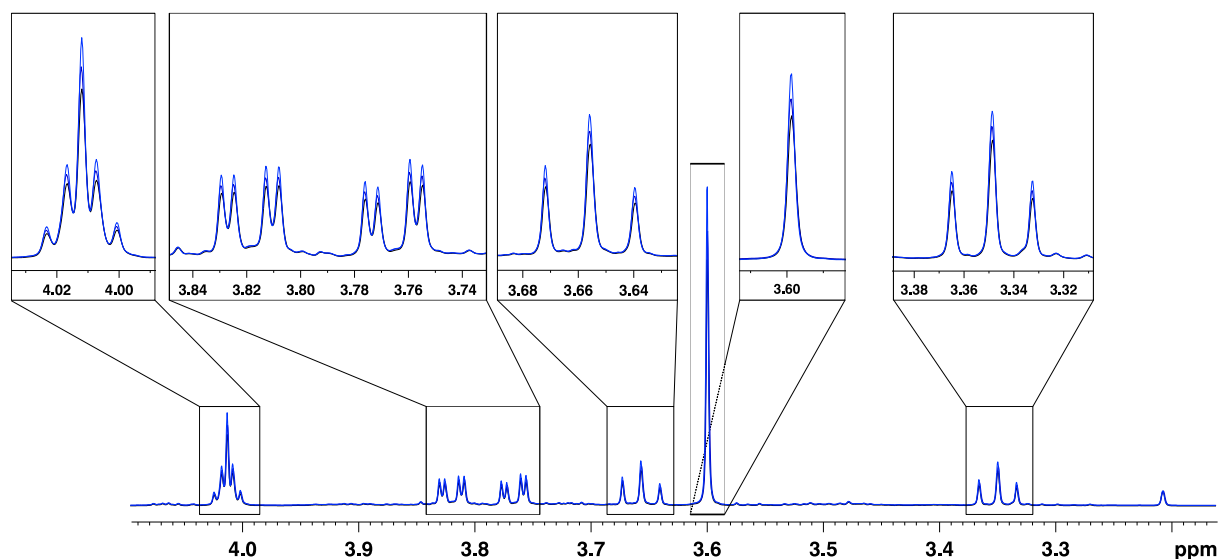
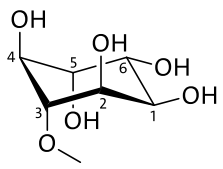


Abbildung 31: 1D-NOESY-Spektrum der Fraktion 33 nach *Spike-In* Experiment mit einer 0.1 M Lösung von D-Pinitol (**16**) in D_2O . Es wurden sukzessiv jeweils 10 μ L der Lösung zugegeben (blaue Spektren).

Nach Zugabe der D-Pinitol-Stammlösung erfolgt ein deutlicher Anstieg der Signale zwischen 3.3 und 4.1 ppm, wodurch die Signale verifiziert wurden und der Metabolit im polaren Extrakt der Pinienkerne eindeutig nachgewiesen wurde. Die Korrelationskoeffizienten zwischen den NMR-Signalen und MS-Signalen von D-Pinitol (**16**) sind in Tabelle 38 aufgeführt. In Tabelle 5 sind die 1H -NMR-Daten sowie die Zuordnung der Signale des Metaboliten aufgeführt.

Tabelle 5: ^1H -NMR-Daten (600.13 MHz, D_2O , pH-Wert = 7.0, kalibriert auf TSP- d_4 ($\delta = 0.00$ ppm)) des Metaboliten D-Pinitol (3-O-Methyl-D-*chiro*-Inositol) (**16**), der im polaren Extrakt der Pinienkerne identifiziert werden konnte.

Metabolit	δ [ppm]	Multiplizität	J [Hz]	Integral	Zuordnung
 D-Pinitol (16)	3.99-4.03	m	-	2	H-1, H-6
	3.82	dd	10.0, 2.8	1	H-2
	3.77	dd	10.0, 2.8	1	H-5
	3.66	dd	9.7, 9.7	1	H-4
	3.60	s	-	3	H-7
	3.35	dd	9.7, 9.7	1	H-3

Pinitol (**16**) gehört zur Gruppe der Cyclitole und ist ein Derivat des *chiro*-Inositols. Der Metabolit konnte in großen Mengen aus dem Exsudat der Zucker-Kiefer (*Pinus lambertiana* DOUGL.) isoliert werden, wodurch im Jahr 1952, nachdem die Stereochemie des Inositol-Derivats bereits 1936 von POSTERNAK ermittelt wurde, zusätzlich die Position der Methylgruppe von ANDERSON *et al.* durch Derivatisierung bestimmt werden konnte.^[138,139] In Pinienkernen wurde der Metabolit erstmals 2012 von RUIZ-ACEITUNO durch Derivatisierung zu den entsprechenden Trimetylsilyl-Oximen mittels GC-MS identifiziert.^[140] Die Biosynthese von Pinitol (**16**) verläuft entsprechend des Biosynthesewegs nach LOEWUS ausgehend von Glucose über *myo*-Inositol und D-Ononitol zu D-Pinitol (**16**).^[141,142] NMR-spektroskopisch kann nicht zwischen den Enantiomeren D-Pinitol und L-Pinitol unterschieden werden. Aufgrund des Biosynthese-Weges nach LOEWUS kann jedoch davon ausgegangen werden, dass es sich ausgehend von der Biosynthese aus D-Glucose um die natürlich vorkommende Variante des D-Pinitols (**16**) handelt. Die relative Konfiguration der in Pinienkernen als Markersubstanz dienende Metabolit Pinitol (**16**) kann somit durch die Anwendung der *SCORE-metabolite-ID* Methode ohne Isolierung aus dem Gemisch heraus eindeutig identifiziert werden.

Zur weiteren Evaluierung der *SCORE-metabolite-ID* Methode wird das Gemisch bzgl. der Identifizierung weiterer Metaboliten untersucht. Im Folgenden werden die identifizierten Metaboliten entsprechend ihrer Substanzklassen zusammengefasst und die Identifizierung ausgewählter Metaboliten im Detail diskutiert.

Neben Pinitol (**16**) konnten im polaren Extrakt der Pinienkerne die Alditole Sorbitol (**17**) und Glycerol (**18**) sowie das Disaccharid Saccharose (**2**), das Trisaccharid Raffinose (**19**) und das Tetrasaccharid Stachyose (**20**) identifiziert werden (siehe Abbildung 32).

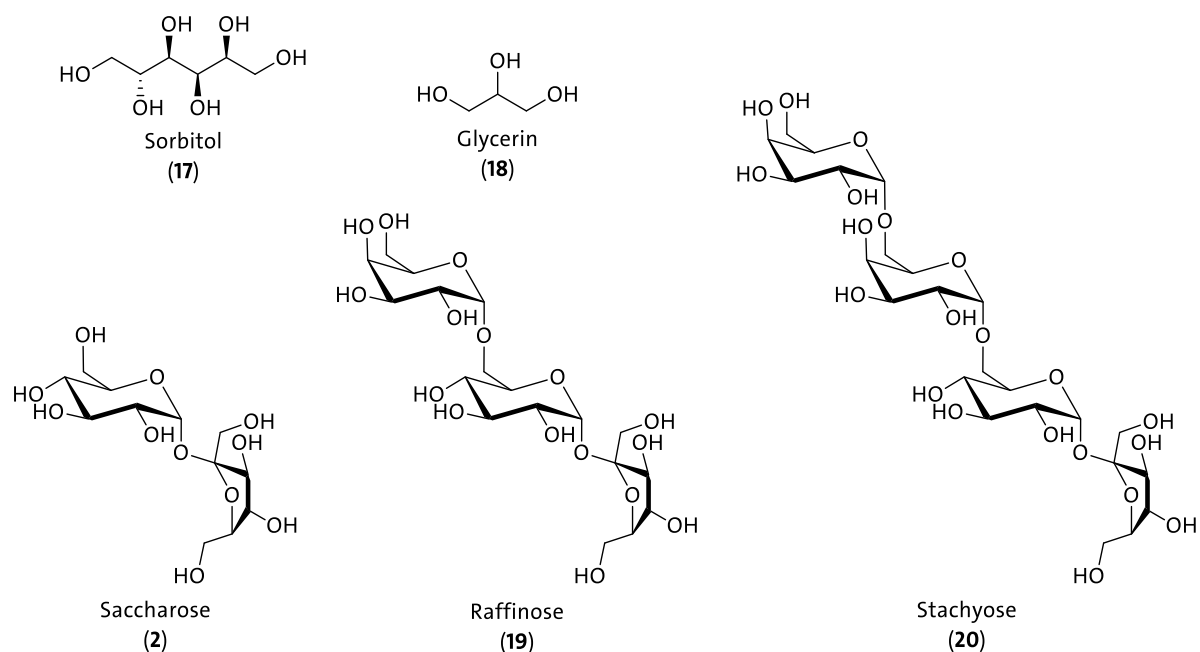


Abbildung 32: Strukturformeln der in Pinienkernen identifizierte Alditole Sorbitol (17) und Glycerin (18) sowie die Kohlenhydrate Saccharose (2), Raffinose (19) und Stachyose (20).

Die Korrelationskoeffizienten der Metaboliten sind im Anhang in Tabelle 38 aufgeführt. Im Fall von Stachyose (20) erfolgte der Nachweis über den EMC des Tetrasaccharids ($m/z = 665.21$ ($[M-H]^-$)). Dass es sich bei dem Tetrasaccharid um Stachyose (20) handelt, wurde zunächst aufgrund des häufigen natürlichen Vorkommens dieses Kohlenhydrats in pflanzlichen Matrices postuliert. Ein Datenbankvergleich der NMR-Spektren sowohl von Stachyose (20) als auch von Raffinose (19) erhärteten diesen vorläufigen Verdacht. Für Stachyose (20) ist eine Korrelation mit den NMR-Daten jedoch in diesem Fall nicht möglich, da dieses Kohlenhydrat in vergleichsweise geringen Konzentrationen auftritt und die EDCs aufgrund von Signalüberlagerungen anderer Metaboliten (v. a. Raffinose (19)) in früheren Fraktionen nicht den adäquaten Verlauf des Metaboliten widerspiegeln. Wie die Abbildung 33 anhand der EMC-Verläufe der drei Saccharide veranschaulicht, eluieren die Metaboliten zwar in unterschiedlichen Fraktionen, jedoch überlagern die Verläufe von Raffinose (19) und Stachyose (20) bis einschließlich Fraktion 45.

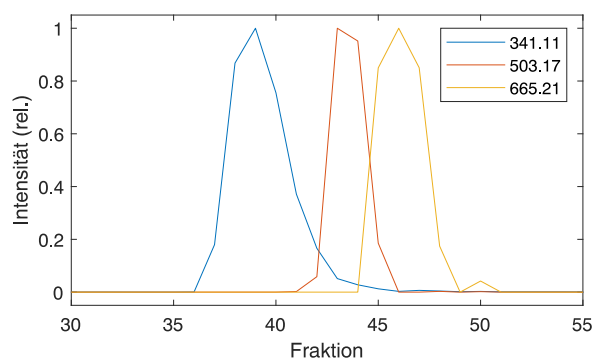


Abbildung 33: Die normierten EMC-Verläufe der Kohlenhydrate Saccharose (**2**) ($m/z = 341.11$ ($[M-H]^-$)), Raffinose (**19**) ($m/z = 503.17$ ($[M-H]^-$)) und Stachyose (**20**) ($m/z = 665.21$ ($[M-H]^-$)) im negativen Ionisationsmodus zeigen die Elutionsreihenfolge dieser Kohlenhydrate.

Aufgrund der strukturell starken Ähnlichkeit von Raffinose (**19**) und Stachyose (**20**) zeigen beide Metaboliten NMR-Signale bei sehr ähnlichen chemischen Verschiebungen, sodass die NMR-Signale des gering konzentrierten Metaboliten Stachyose (**20**) von denen des höher konzentrierten Metaboliten Raffinose (**19**) im Großteil der Fraktionen überlagert sind. Aufgrund dessen und aufgrund der zu Beginn getroffenen Annahme wurden *Spike-In* Experimente mit den beiden Metaboliten durchgeführt (siehe Abbildung 34), wodurch beide Metaboliten eindeutig in den Pinienkernen nachgewiesen werden konnten.

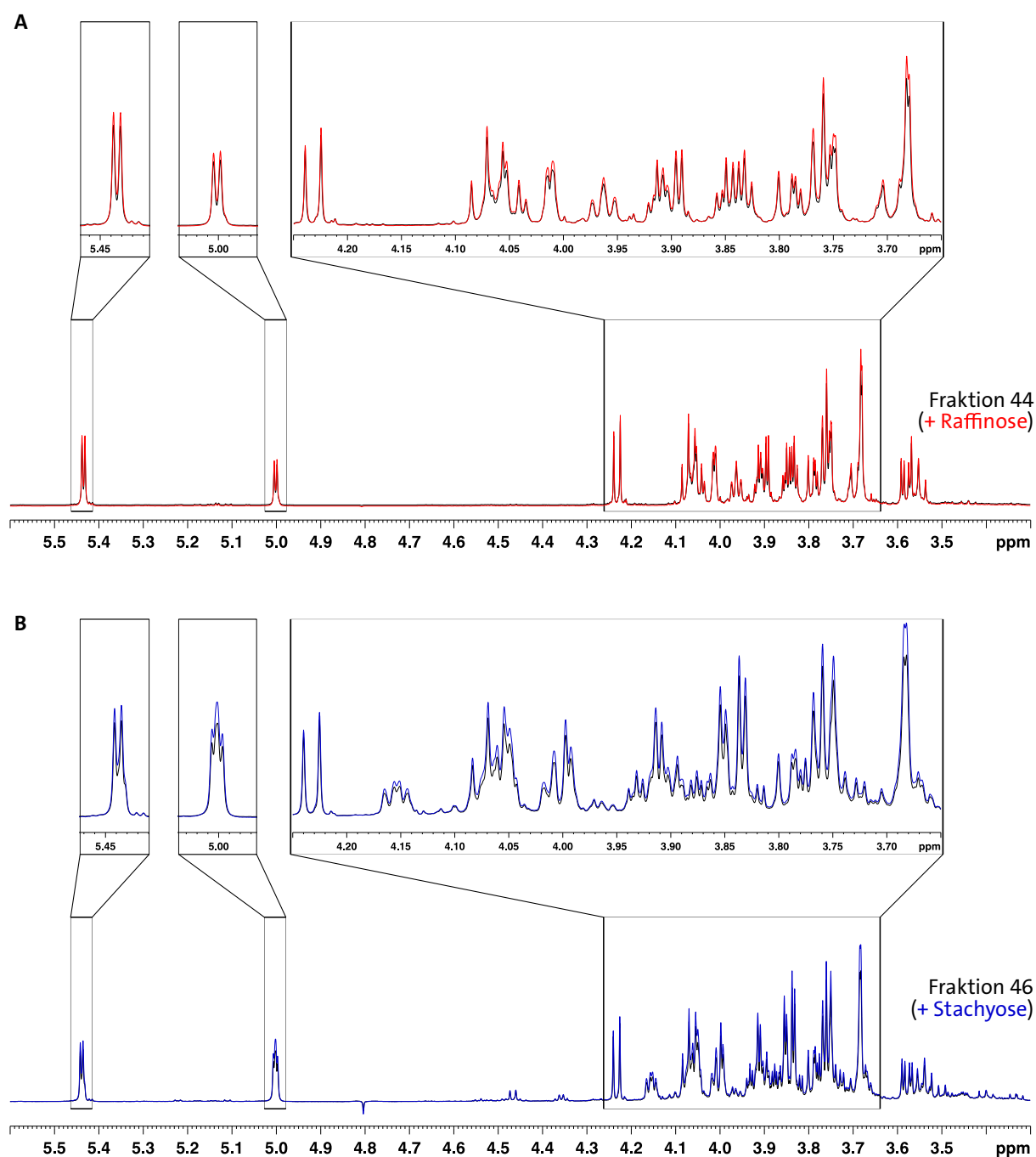


Abbildung 34: *Spike-In* Experimente der Metaboliten Raffinose (**19**) und Stachyose (**20**). **A** zeigt das ^1H -NMR-Spektrum der Fraktion 44 in schwarz und das Spektrum derselben Fraktion nach Zugabe von 15 μL Raffinose (**19**) in D_2O (0.1 M) in rot. In **B** ist das ^1H -NMR-Spektrum der Fraktion 46 in schwarz gezeigt sowie in blau das Spektrum derselben Fraktion nach Zugabe von 3 μL Stachyose (**20**) in D_2O (0.1 M).

Auch im Fall von Sorbitol (**17**) erfolgte die Identifizierung zunächst ausgehend von EMCs. Im positiven und negativen Ionisationsmodus wurde jeweils ein EMC mit gleichem Verlauf detektiert (siehe Abbildung 35). Der Korrelationskoeffizient zwischen beiden Verläufen beträgt 0.9923. Unter der Annahme, dass es sich bei dem Signal im positiven Ionisationsmodus um das

Natrium-Addukt ($m/z = 205.07$) und im negativen Ionisationsmodus um das deprotonierte Molekölion ($m/z = 181.07$) des Metaboliten handelt, kann die Summenformel $C_6H_{14}O_6$ postuliert werden.

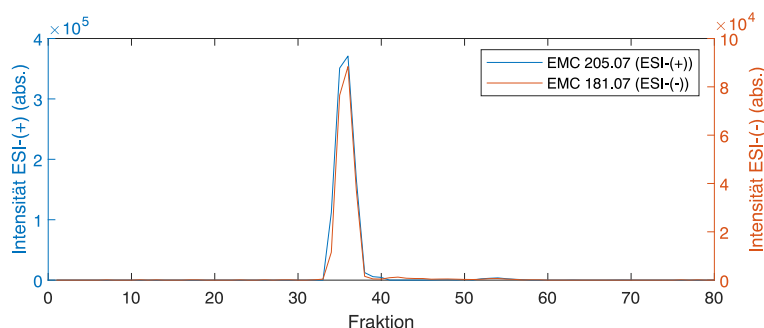


Abbildung 35: Zweidimensionale Darstellung der EMCs bei $m/z = 205.07$ im positiven Ionisationsmodus und bei $m/z = 181.07$ im negativen Ionisationsmodus. Der Korrelationskoeffizient zwischen den beiden EMCs beträgt 0.9923.

Zu den möglichen Metaboliten mit dieser Summenformel gehören die Alditole, die ausgehend von Hexosen durch Reduktion erhalten werden. Da reguläre Alditole ausschließlich zu NMR-Signalen im Bereich zwischen 3 und 4 ppm führen und in den Fraktionen 30-40 eine Vielzahl anderer Kohlenhydrate eluieren, die dementsprechend viele Signale in diesem Bereich und in diesen Fraktionen hervorrufen, ist erneut keine eindeutige Korrelation mit den NMR-Daten möglich. Es wurde deshalb ein selektives TOCSY-Experiment der Fraktion 36 aufgenommen, bei dem auf ein Signal bei 3.85 ppm selektiv eingestrahlt wurde, wodurch die Kopplung zu Protonen desselben Spinsystems detektiert werden können (siehe Abbildung 36 Mitte).

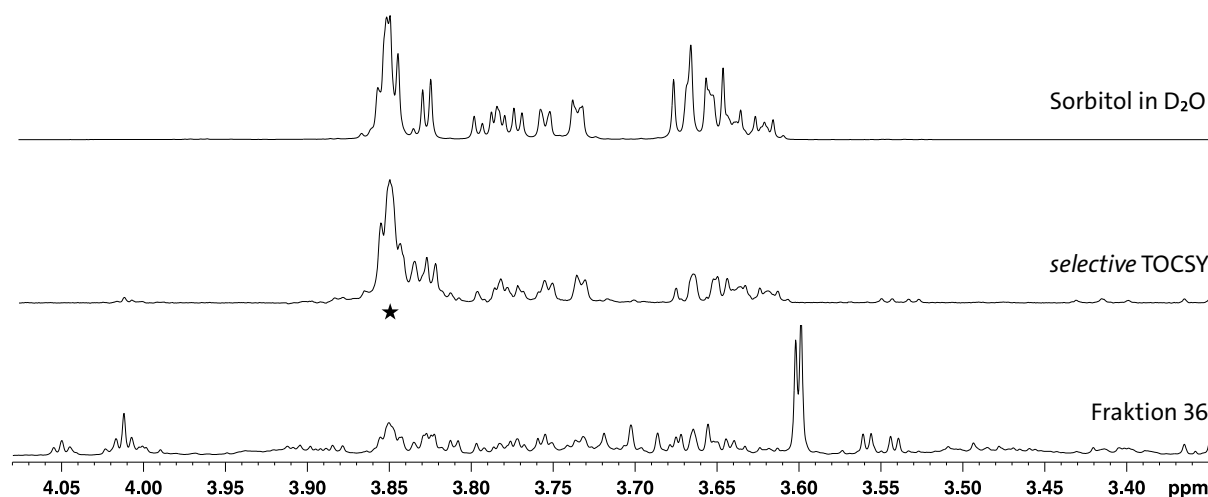


Abbildung 36: 1H -NMR-Spektrum der Fraktion 36 des polaren Pinienkernextraktes (unten). In der Mitte ist ein selektives TOCSY-Spektrum dargestellt, bei dem auf das mit einem Sternchen markierte Signal bei 3.85 ppm eingestrahlt wurde. Oben ist ein Referenzspektrum des Metaboliten Sorbitol (**17**) in Deuteriumoxid gezeigt.

Zusätzlich wurde ein Referenzspektrum des Alditols Sorbitol (**17**) in Deuteriumoxid aufgenommen, das in Abbildung 36 oben abgebildet ist. Trotz starker Überlagerung der NMR-Signale im Spektrum der Fraktion 36, können die Signale über das selektive TOCSY-Spektrum detektiert werden. Das *Spike-In* Experiment verifiziert den Strukturvorschlag (siehe Abbildung 137 im Anhang unter Abschnitt 8.5). Sorbitol (**17**) kann metabolisch durch Reduktion von α -D-Glucose (**12**), D-Fructose (**20**) oder L-Sorbose (**21**) gebildet werden (vgl. Abbildung 37).^[143]

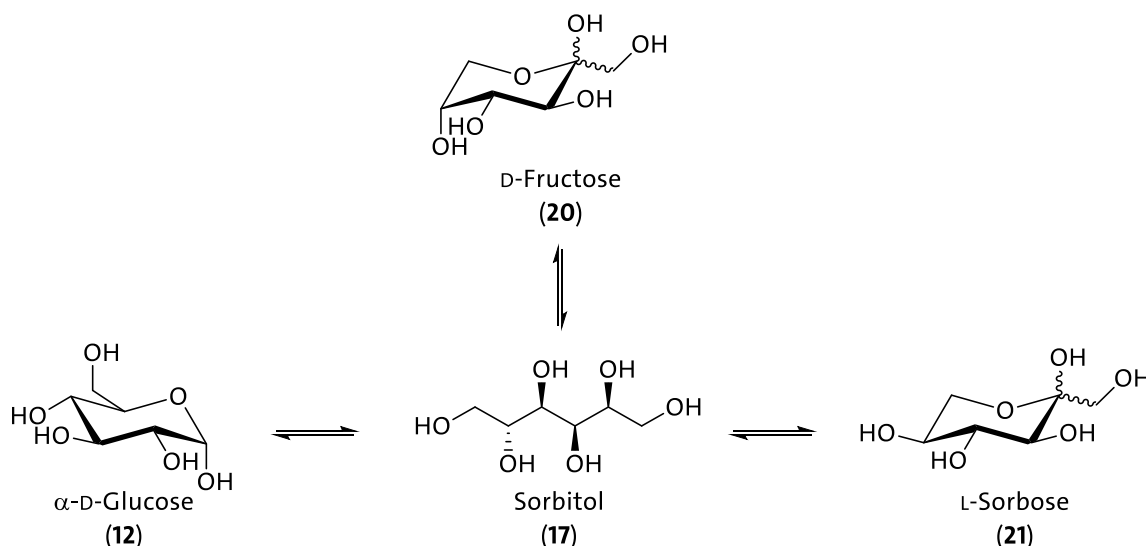


Abbildung 37: Sorbitol (**17**) kann biosynthetisch ausgehend von den natürlich vorkommenden Kohlenhydraten α -D-Glucose (**12**), D-Fructose (**20**) oder L-Sorbose (**21**) gebildet werden.^[143]

Anhand der dargestellten Ergebnisse wird eine besondere Herausforderung bei der Identifizierung von Kohlenhydraten aus komplexen Gemischen deutlich. Dabei wirkt sich vor allem die strukturelle Ähnlichkeit der verschiedenen Kohlenhydrate erschwerend auf die Auswertung aus. Zum einen führt die Ähnlichkeit bereits bei der Fraktionierung des Gemisches dazu, dass sich die Elutionsverläufe der verschiedenen Kohlenhydrate kaum unterscheiden. Weiterhin führen die ähnlichen Strukturmodule in den erhaltenen ^1H -NMR-Spektren zu starken Signalüberlagerungen im Bereich von 3 bis 4 ppm. Damit werden gleich zwei Dimensionen, anhand derer eine Unterscheidung und Identifizierung von Substanzen mittels *SCORE-metabolite-ID* erfolgt, nachteilig beeinflusst. Trotzdem lassen sich durch die Anwendung besonders auf Grundlage der m/z -Werte Rückschlüsse auf die Kohlenhydrat-Klasse (z. B. Mono-, Di- oder Oligosaccharid) ziehen. Im Rahmen der Auswertung mit spezieller Berücksichtigung der Kohlenhydrate reduziert der Einsatz von *SCORE-metabolite-ID* trotzdem die Anzahl der relevanten Signale in Bezug auf das gesamte Gemisch. Dadurch wird ein effizienterer Vergleich

mit Datenbanken möglich und speziell die Signale der anomeren Protonen lassen einen Strukturvorschlag oder eine Zuordnung zu. Die direkte eindeutige Zuordnung aller NMR-Signale ist jedoch nicht in allen Fällen möglich, sodass es sinnvoll sein kann, den Einsatz von *SCORE-metabolite-ID* durch weitere Methoden zu ergänzen.

5.2.1.2 Identifizierung von Aminosäuren

Im polaren Extrakt der Pinienkerne konnten 17 der 20 proteinogenen Aminosäuren identifiziert werden, die in Abbildung 38 dargestellt sind. Ein Großteil der proteinogenen α -Aminosäuren zeigt in den Massenspektren ein Signal, das durch die zuvor gezeigte charakteristische Fragmentierungsreaktion der konsekutiven Abspaltung von Wasser und Kohlenstoffmonoxid unter Bildung eines Iminium-Ions hervorgerufen wird (siehe Abbildung 27).^[131] Das Auftreten von Signalen bei m/z -Werten des protonierten Moleküls der entsprechenden Aminosäure und dem jeweiligen Iminium-Ion ($\Delta m/z = 46.01$) wurde deshalb im Folgenden häufig als Grundlage zur Erstellung des Strukturvorschlags genutzt.

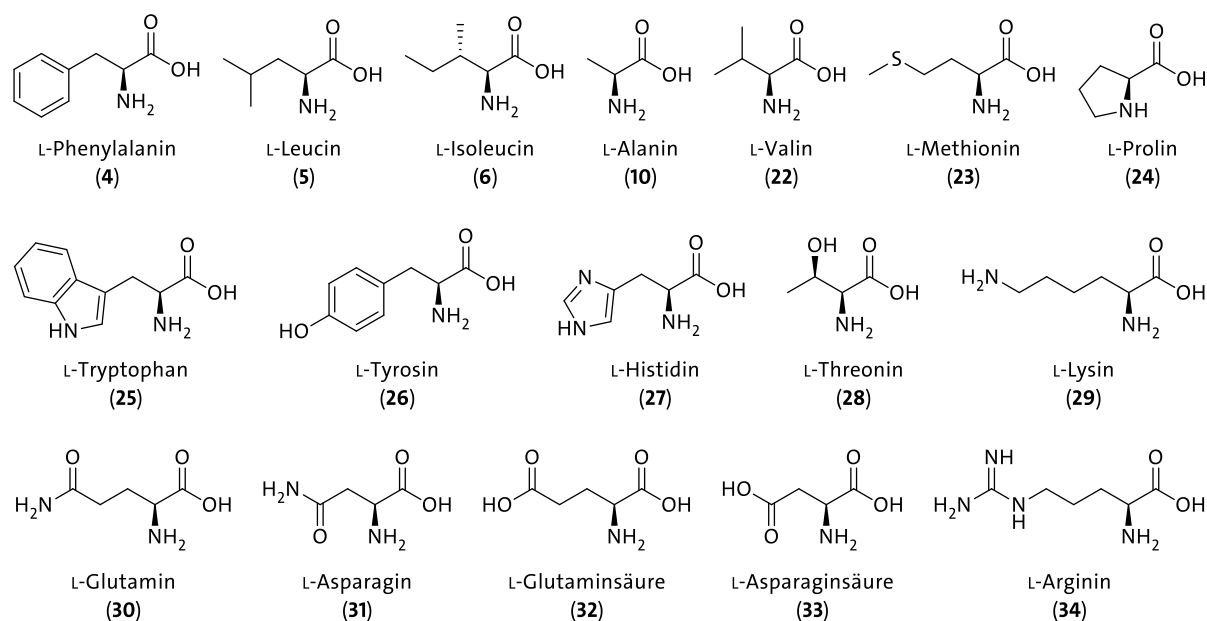


Abbildung 38: Übersicht der im polaren Extrakt der Pinienkerne identifizierten Aminosäuren.

Die Aminosäuren L-Serin (35), L-Cystein (36) und Glycin (37) wurden nicht detektiert. Diese drei Aminosäuren weisen jedoch ausschließlich ^1H -NMR-Signale im Bereich von 3 bis 4 ppm auf. In diesem Bereich treten trotz Fraktionierung des Extraktes viele Signale auf, die zusätzlich oft überlagert sind. Es ist zu erwarten, dass auch die Aminosäuren L-Serin (35), L-Cystein (36) und

Glycin (**37**) im Extrakt der Pinienkerne vorliegen, jedoch in den ^1H -NMR-Spektren z.B. aufgrund von Überlagerungen nicht detektiert werden können.

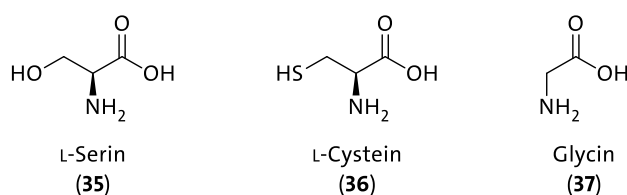


Abbildung 39: Strukturformeln der drei proteinogenen Aminosäuren L-Serin (**35**), L-Cystein (**36**) und Glycin (**37**), die im Rahmen der durchgeführten Untersuchungen im polaren Extrakt der Pinienkerne auf Grundlage nicht detektiert werden konnten.

In Abbildung 40 ist der Elutionsverlauf aller 17 identifizierten Aminosäuren anhand jeweils eines EMCs dargestellt.

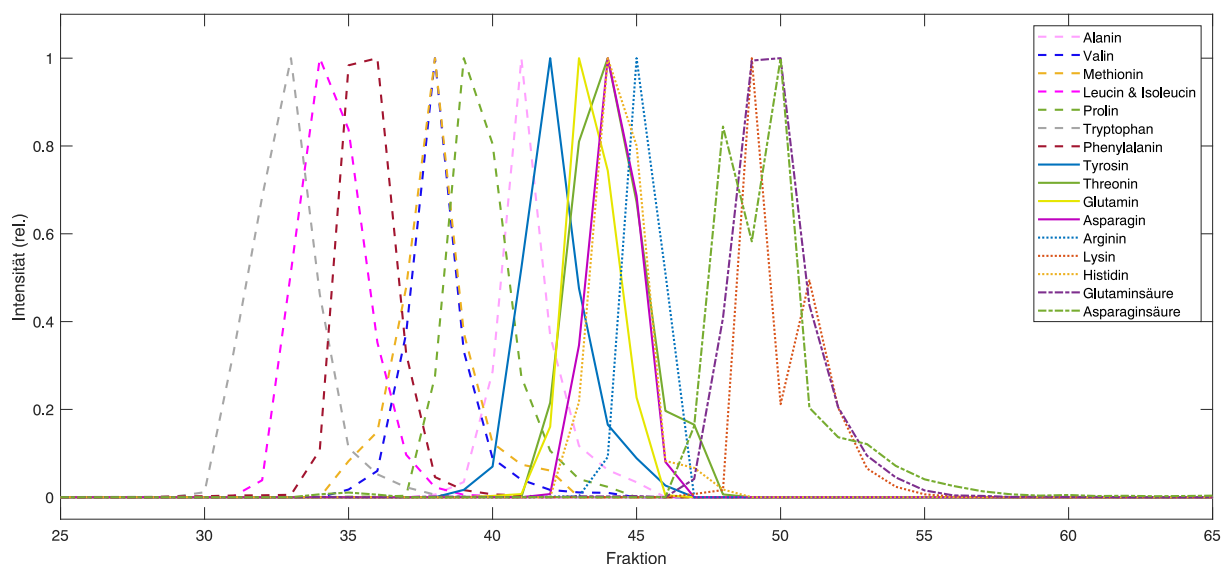


Abbildung 40: Elutionsverlauf der identifizierten Aminosäuren anhand jeweils eines EMCs. Die Intensitäten der EMCs wurden normiert. Die EMCs sind entsprechend ihrer Einteilung in unpolare, polare, saure und basische Aminosäure-Reste durch unterschiedliche Linien gekennzeichnet.

Die 17 Aminosäuren zeigen eine chromatographische Verteilung über insgesamt ca. 25 Fraktionen. Prinzipiell erfolgt nach ALPERT die Elution an HILIC-Phasen proportional zur Polarität der Analyten.^[55] Grundsätzlich ist dies entsprechend der Einteilung der Aminosäuren in polare, unpolare, saure und basische Aminosäuren gegeben. So eluieren zunächst die unpolaren Aminosäuren, gefolgt von den polaren und zuletzt die sauren und basischen. Die EMCs der unpolaren und polaren Aminosäuren weisen meist symmetrische, GAUß-förmige Profile auf, wohingegen die EMCs der sauren und teilweise der basischen Aminosäuren deutlich

breitere Profile aufweisen. Diese Elution über mehrere Fraktionen mit teilweise asymmetrischen Verläufen ist vermutlich auf weitere Retentionsmechanismen der HILIC-Phasen zurückzuführen. So hängt die Elution zusätzlich zur Hydrophilie auch von elektrostatischen Wechselwirkungen und dem Auftreten bzw. der Stärke von Wasserstoffbrückenbindungen ab. Insbesondere die sauren und basischen Aminosäuren können je nach Ladungszustand vermehrt auch durch diese Mechanismen retardiert werden.

Wie in Abbildung 40 außerdem zu sehen ist, weisen einzelne Aminosäuren einen sehr ähnlichen Verlauf der EMCs auf. Die EMCs der Aminosäuren Valin (**22**) und Methionin (**23**) (orange und blau gestrichelte Verläufe) korrelieren beispielsweise miteinander, wodurch eine Unterscheidung der NMR- und MS-Signale der beiden Aminosäuren ohne weitere Analyse nicht möglich ist. Es wurden deshalb selektive TOCSY- und NOESY-Experimente durchgeführt, um die NMR-Signale zuordnen zu können. Selektive TOCSY-Spektren ermöglichen die Detektion von Kopplungen der Protonen innerhalb eines Spinsystems. Durch selektive NOESY-Spektren können Kopplungen in Folge von räumlicher Nähe der Kerne detektiert werden. Diese NMR-Spektren und die Zuordnung der Signale zu den beiden Aminosäuren Valin (**22**) und Methionin (**23**) sind in Abbildung 41 dargestellt. Eine besondere Schwierigkeit ergibt sich hierbei durch die hohen Konzentrationsunterschiede der Metaboliten innerhalb der Fraktion 38. Das Kohlenhydrat Saccharose (**2**) weist in dieser Fraktion das Maximum auf und liegt im Vergleich zu Methionin (**23**) in ca. 360-fach höherer Stoffmengenkonzentration vor. Diese hohen Konzentrationsunterschiede wirken sich nachteilig auf die Anwendbarkeit der selektiven Experimente aus. In Abbildung 41 äußert sich dies bspw. in den Artefakten, die an den Positionen der Signale von Saccharose (**2**) auftreten. Durch die in Abbildung 41 gezeigten selektiven NMR-Experimente können die Spinsysteme der beiden Aminosäuren dennoch eindeutig nachgewiesen werden. Am Beispiel von Valin (**22**) und Methionin (**23**) wird gezeigt, dass durch die Anwendung von *SCORE-metabolite-ID* somit zunächst die relevanten NMR- und MS-Signale eingeschränkt werden können. Durch die Erweiterung der analytischen Daten um die gezielt eingesetzten selektiven Experimente ist anschließend die eindeutige Zuordnung der Signale möglich.

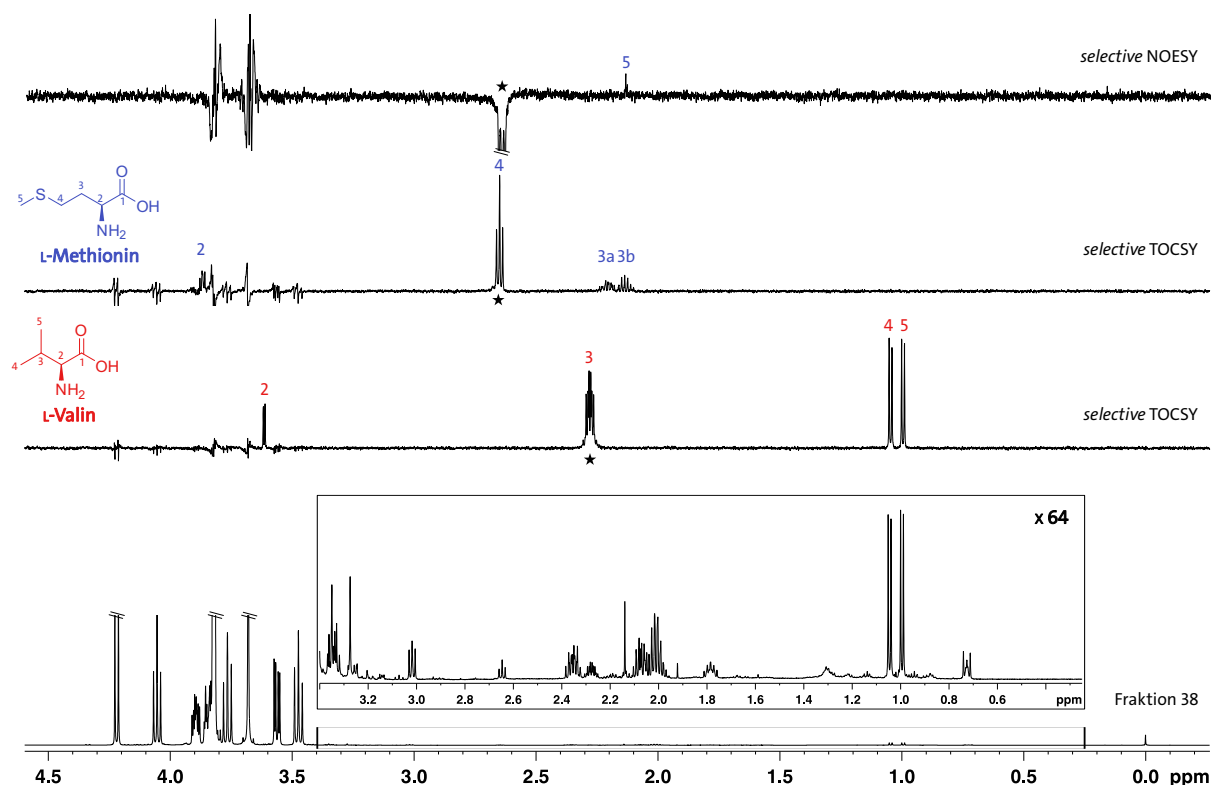


Abbildung 41: ¹H-NMR-Spektrum der Fraktion 38 des polaren Pinienkernextraktes (unten) sowie selektive TOCSY- und NOESY-NMR-Spektren (oben). Die Signale im Bereich zwischen 0.25 und 3.4 ppm der Fraktion 38 sind aufgrund der starken Konzentrationsunterschiede der Analyten in dieser Fraktion in 64-facher Vergrößerung gezeigt. Es wurde jeweils auf das mit einem Sternchen markierte Signal eingestrahlt. Die selektiven Experimente zeigen die Aminosäuren Valin (**22**) (zweites Spektrum von unten, Einstrahlfrequenz bei 2.28 ppm) sowie die Aminosäure Methionin (**23**) (drittes und viertes Spektrum von unten, Einstrahlfrequenz jeweils bei 2.65 ppm).

Ebenso weisen auch die EMCs der Aminosäuren Histidin (**27**) und Asparagin (**31**) einen sehr ähnlichen Verlauf auf und korrelieren miteinander. Der Strukturvorschlag dieser beiden Aminosäuren wurde auf Grundlage der NMR-Signale und der *m/z*-Werte der korrelierten EDCs und EMCs sowie durch den Vergleich mit NMR-Spektren aus Datenbanken getroffen. Im Anschluss wurden zur Verifizierung der Annahme *Spike-In* Experimente mit den Aminosäuren Histidin (**27**) und Asparagin (**31**) in dem ¹H-NMR-Spektrum der Fraktion 44 durchgeführt.

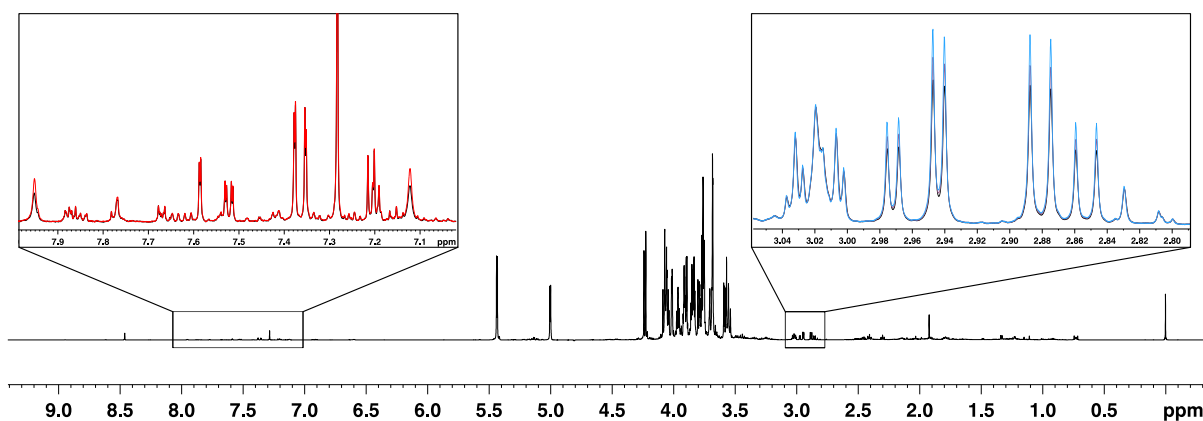


Abbildung 42: *Spike-In* Experiment der Metaboliten Histidin (**27**) und Asparagin (**31**). Das schwarze Spektrum zeigt das ^1H -NMR-Spektrum der Fraktion 44. In rot ist das Spektrum nach Zugabe von 0,5 μL Histidin (**27**) (0,1 M in D_2O) und in blau das Spektrum nach Zugabe von jeweils 4 μL Asparagin (**31**) (0,1 M in D_2O) gezeigt.

Anhand der in Abbildung 42 gezeigten *Spike-In* Experimente können die Strukturen Histidin (**27**) (*Spike-In* Experiment dargestellt in rot im linken Bereich der Abbildung) und Asparagin (**31**) (*Spike-In* Experiment dargestellt in blau im rechten Bereich der Abbildung) verifiziert werden. Diese Beispiele zeigen, dass bei Anwendung weiterer analytischer Methoden, wie beispielsweise zusätzliche (selektive oder mehrdimensionale) NMR-Experimente oder durch *Spike-In* Experimente, somit auch Metaboliten, die bei chromatographischer Separation ein ähnliches Elutionsverhalten aufweisen, nach einem vorläufigen Strukturvorschlag eindeutig nachgewiesen und identifiziert werden können. Dabei bildet die Anwendung von *SCORE-metabolite-ID* durch die Korrelation der NMR- und MS-Daten die Grundlage zur Erstellung eines Strukturvorschlags.

Wie bereits zuvor anhand des artifiziellen Gemisches A gezeigt wurde, birgt die Analyse von isobaren Verbindungen besondere Herausforderungen. Die isobaren Metaboliten Leucin (**5**) und Isoleucin (**6**) eluieren in den Fraktionen 32 bis 39 (vgl. Abbildung 40). Beide Metaboliten weisen eine ähnliche Struktur und Polarität auf. Dennoch unterscheiden sich die Elutionsverläufe minimal. Dies ist anhand jeweils eines EDCs jedes Metaboliten in Abbildung 43 gezeigt. Während das Maximum des EDCs von Leucin (**5**) bei einer chemischen Verschiebung von 0,966 ppm (blaue, durchgezogene Linie) in Fraktion 34 liegt, befindet sich das Maximum des EDCs von Isoleucin (**6**) bei einer chemischen Verschiebung von 0,946 ppm (blaue, gestrichelte Linie) in Fraktion 35.

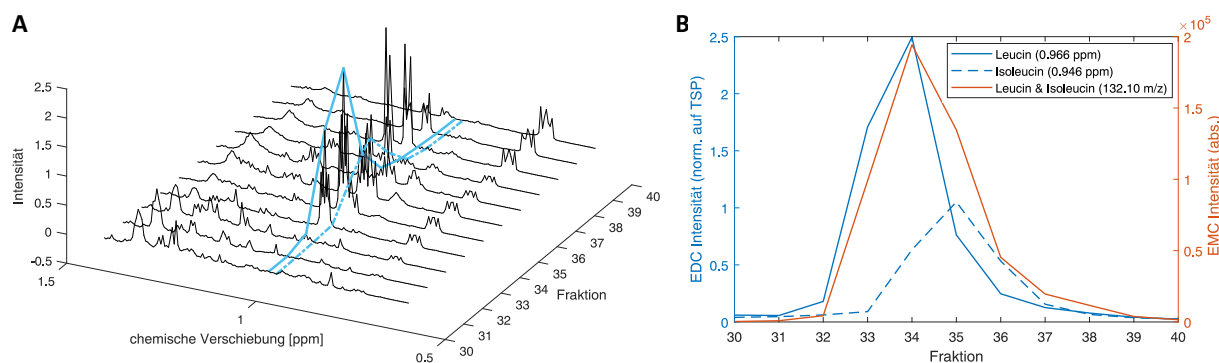


Abbildung 43: A: Dreidimensionale Darstellung der EDC-Verläufe von Leucin (5) (durchgezogene Linie) und Isoleucin (6) (gestrichelte Linie) in den NMR-Spektren in den Fraktionen 30-40. B: Zweidimensionale Darstellung der gleichen EDC-Verläufe in blau (linke y-Achse) sowie des EMC-Verlaufes des [M+H]⁺-Signals ($m/z = 132.10$) beider Metaboliten in orange (rechte y-Achse).

Da die beiden Metaboliten isobar sind und dementsprechend zu dem gleichen [M+H]⁺-Signal bei einem m/z -Wert von 132.10 führen, entspricht der EMC-Verlauf (orange Linie in Abbildung 43 B) der Summe beider EDC-Verläufe. Eine Korrelation zwischen den NMR- und MS-Signalen ist in diesem Fall deshalb nicht sinnvoll. In Analogie zum artifiziellen Gemisch A wurden deshalb die NMR-Signale untereinander korreliert. Die entsprechenden Korrelationskoeffizienten sind in Tabelle 6 angegeben.

Tabelle 6: PEARSON Korrelationskoeffizienten der EDCs von Isoleucin (6) und Leucin (5) berechnet über die Fraktionen 30-37 des polaren Pinienkernextraktes.

EDC [ppm]	0.966	1.707	0.944	1.001	1.477	1.987
0.966	1.0000	0.9967	0.3599	0.2926	0.2518	0.2377
1.707	0.9967	1.0000	0.3341	0.2663	0.2291	0.2163
0.944	0.3599	0.3341	1.0000	0.9956	0.9836	0.9721
1.001	0.2926	0.2663	0.9956	1.0000	0.9942	0.9858
1.477	0.2518	0.2291	0.9836	0.9942	1.0000	0.9977
1.987	0.2377	0.2163	0.9721	0.9858	0.9977	1.0000

Trotz der Tatsache, dass die Metaboliten um nur eine Fraktion versetzt eluieren, lassen sich die EDCs von Leucin (5) und Isoleucin (6) eindeutig unterscheiden. Die ¹H-NMR-Signale der α-Protonen lassen sich nicht durch die Korrelation detektieren, da die Signale in diesem Bereich überlagert sind. Zusätzlich wurden deshalb selektive TOCSY-Experimente durchgeführt (siehe Abbildung 44).

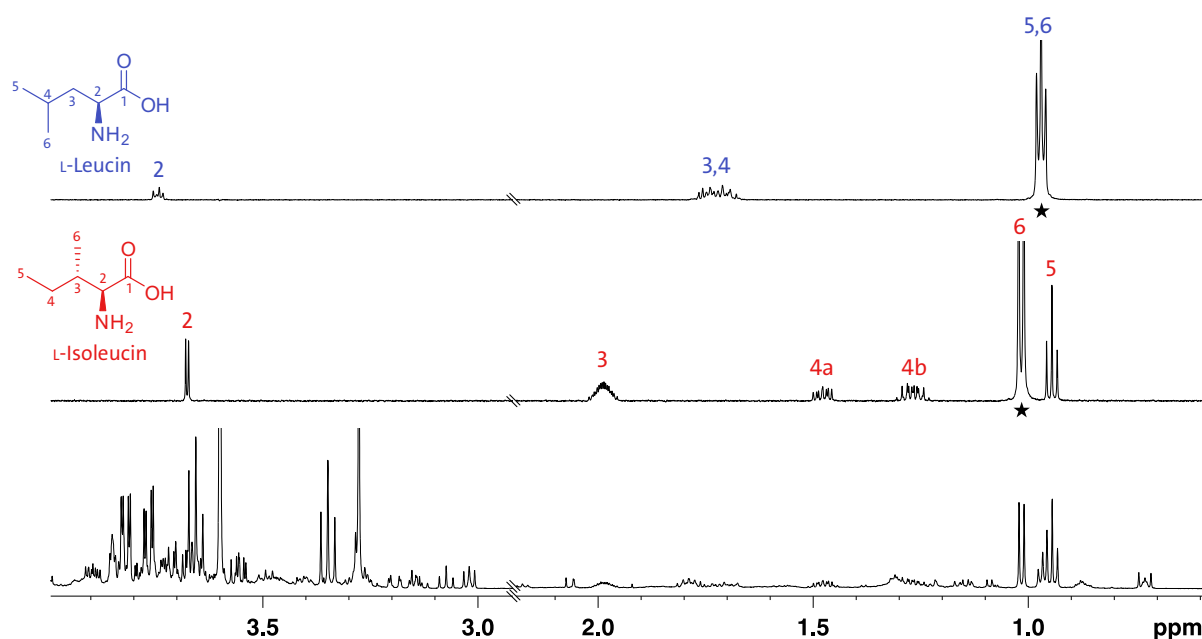


Abbildung 44: Selektive TOCSY-Experimente der Fraktion 35 (unteres Spektrum) des polaren Pinienkernextraktes. Die Sternchen markieren die eingestrahlte Frequenz. Die selektiven Experimente zeigen die Aminosäuren Isoleucin (**6**) (zweites Spektrum von unten, Einstrahlfrequenz bei 1.02 ppm) sowie die Aminosäure Leucin (**5**) (drittes Spektrum von unten, Einstrahlfrequenz bei 0.97 ppm).

Die selektiven TOCSY-Experimente verifizieren die Zusammengehörigkeit der EDCs in Tabelle 6. Zusätzlich können über die selektiven TOCSY-Experimente weitere ^1H -NMR-Signale detektiert werden, die über die Korrelation nicht detektiert wurden. Dazu zählen die Signale bei 3.68 und 3.74 ppm, die den α -Protonen der beiden Aminosäuren zugeordnet werden können, sowie das Multiplett bei 1.22-1.31 ppm, das einem der diastereotopen Protonen des C4 von Isoleucin (**6**) zugeordnet werden kann. Für alle anderen Aminosäuren wurden Korrelationen zwischen spezifischen EDCs und den EMCs des protonierten Moleküls, des Natrium-Addukts oder des Iminium-Ions detektiert. Diese sind im Anhang in Tabelle 39 aufgeführt. Diese Ergebnisse, die durch die *SCORE-metabolite-ID* App generiert wurden, wurden als Grundlage zur Zuordnung der Signale herangezogen. Zusätzlich wurden selektive NMR-Experimente durchgeführt, um überlagerte Signale detektieren zu können. Dies betrifft in vielen Fällen die α - und β -Protonen mancher Aminosäuren. Die selektiven Experimente sind im Anhang in den Abbildungen 138 bis 146 aufgeführt.

Neben den zuvor gezeigten Aminosäuren konnte in den Fraktionen 65-72 außerdem das oxidierte Derivat des Tripeptids Glutathion (Glutathion Disulfid, GSSG) (**38**) nachgewiesen werden. Ausgehend vom EDC bei einer chemischen Verschiebung von 2.976 ppm wurden im

positiven Ionisationsmodus hoch-korrelierende EMCs bei einem Masse-zu-Ladungs-Verhältnis von 307.08 und 307.58 ermittelt. Aufgrund der Massendifferenz der Isotopenpeaks kann darauf geschlossen werden, dass es sich um das Signal eines zweifach-geladenes Moleküls handelt, d.h. die monoisotopische Masse des Metaboliten beträgt 612.15 Da. Im negativen Ionisationsmodus wurde zu demselben EDC ein hoch-korrelierender EMC bei $m/z = 611.15$ detektiert, wobei es sich demnach um das deprotonierte Molekölion des Metaboliten handelt.

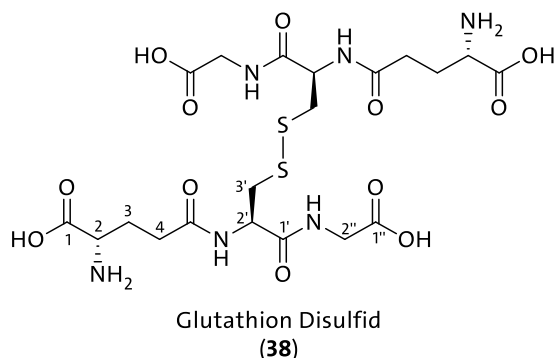


Abbildung 45: Strukturformel des im polaren Extrakt der Pinienkerne identifizierten Metaboliten Glutathion Disulfid (38).

Weitere NMR-Spektren der Fraktion 71 zeigen zum einen ein Glutamin-ähnliches Spinsystem im selektiven TOCSY-Spektrum (siehe Abbildung 46) sowie die Kopplung zweier Dubletts vom Dublett bei chemischen Verschiebungen von 2.99 ppm und 4.18 ppm zu einem weiteren Signal bei einer chemischen Verschiebung von 4.77 ppm im 2D TOCSY-Spektrum (siehe Abbildung 147 im Anhang). Das Signal bei 4.77 ppm ist aufgrund der Wasserunterdrückung in dem ^1H -NMR-Spektrum der Fraktion 71 bei einer chemischen Verschiebung von 4.80 ppm nicht zu sehen. Letztlich können die MS- und NMR-Signale dem Metaboliten Glutathion Disulfid (38) mit einer monoisotopischen Masse von 612.15 Da zugeordnet werden. Es handelt sich dabei um das Disulfid des Tripeptids Glutathion, das aus den Aminosäuren Glutaminsäure (32), Cystein (36) und Glycin (37) aufgebaut ist. Die selektiven TOCSY-Spektren in Abbildung 46 zeigen die Spinsysteme von Glutaminsäure (32) (zweites Spektrum von unten) und die diastereotopen Protonen 3'a und 3'b von Cystein (36). Das α -Proton wurde über das 2D-TOCSY-Spektrum (siehe Abbildung 147 im Anhang) zugeordnet. Das Signal des Glycin-Restes (37) des Tripeptids ist im ^1H -NMR-Spektrum von dem Signal des α -Protons des Glutaminsäure-Restes (32) bei 3.78 ppm überlagert. In Abbildung 46 ist ein Vergleichsspektrum des Metaboliten (38) inkl. Zuordnung der Signale dargestellt.^[144]

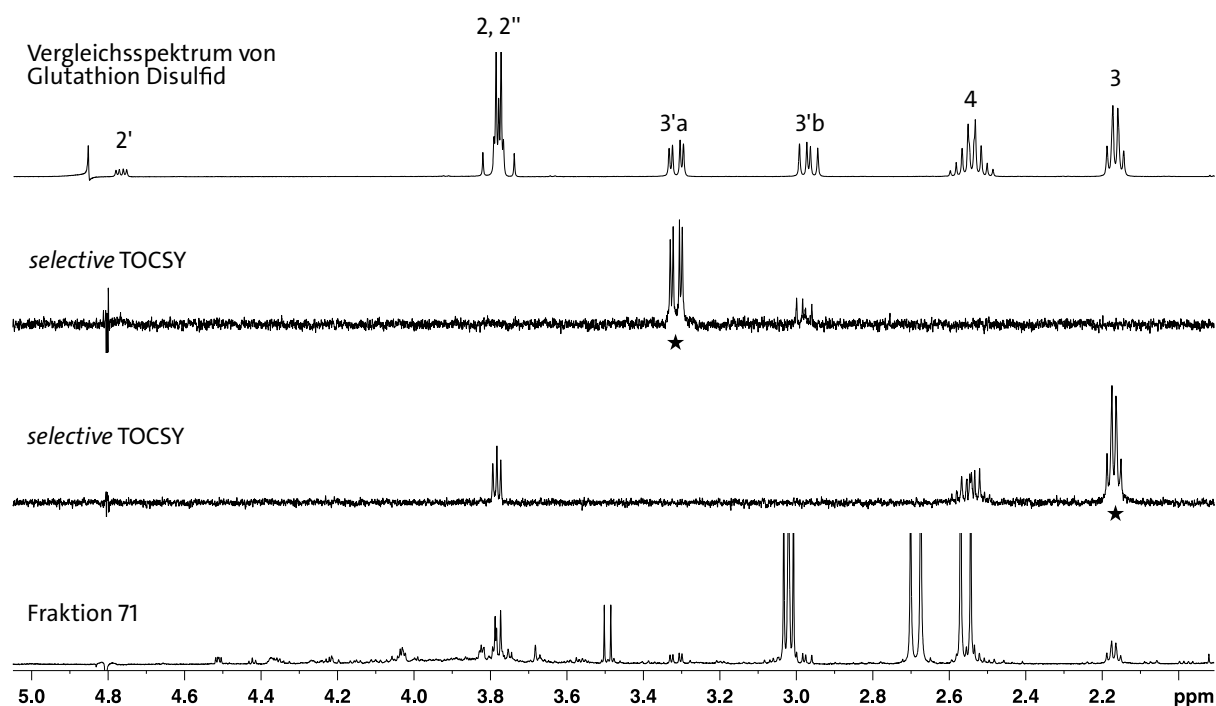


Abbildung 46: ^1H -NMR-Spektrum der Fraktion 71 des polaren Pinienkernextraktes (unten) sowie selektive TOCSY-Spektren (zweites und drittes Spektrum von unten). Die selektive Einstrahlungsfrequenz ist mit einem Sternchen gekennzeichnet. Das oberste Spektrum zeigt ein Datenbank-Vergleichsspektrum von Glutathion Disulfid (**38**) in D_2O .^[144]

5.2.1.3 Identifizierung von Betainen und Derivaten

Neben Betain (*N,N,N*-Trimethylglycin) (**39**) konnten im polaren Extrakt der Pinienkerne die Betain-Derivate Cholin (**8**), Phosphorylcholin (**40**), Trigonellin (**41**) und Ergothionein (**42**) identifiziert werden, deren Strukturen in Abbildung 47 dargestellt sind. Betaine dienen oftmals als Methyl-Gruppen-Donoren bei Transmethylierungen in biochemischen Stoffwechselwegen.^[145–147]

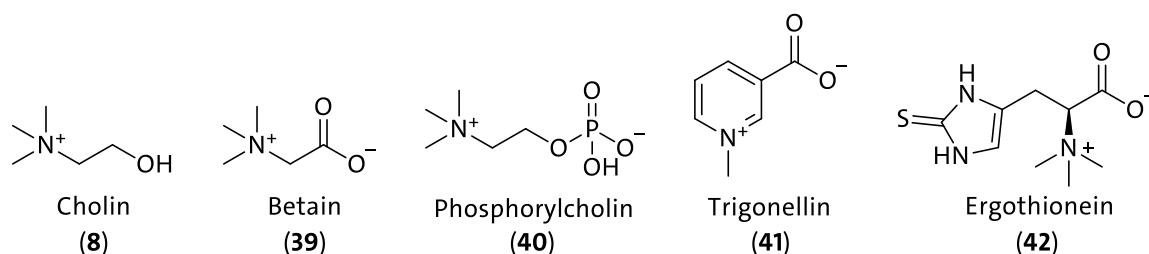


Abbildung 47: Im polaren Extrakt der Pinienkerne identifizierte Betain-Derivate.

Mit der Ausnahme von Trigonellin (**41**) führen die *N*-Methylgruppen der anderen identifizierten Betaine in den NMR-Spektren zu jeweils einem Singulett bei einer chemischen Verschiebung von 3.2 bis 3.3 ppm. In Abbildung 48 ist dies für die in den Fraktionen 24 bis 44 eluierenden Metaboliten Betain (**39**) (oranger EDC bei 3.269 ppm), Cholin (**8**) (blauer EDC bei 3.206 ppm) und Ergothionein (**42**) (gelber EDC bei 3.276 ppm) gezeigt. Die *N*-Methylgruppen des Phosphorylcholins (**40**) führen ebenfalls zu einem Singulett bei einer chemischen Verschiebung von 3.23 ppm, allerdings eluiert der Metabolit erst in den Fraktionen 49 bis 56, weshalb der EDC hier nicht abgebildet ist.

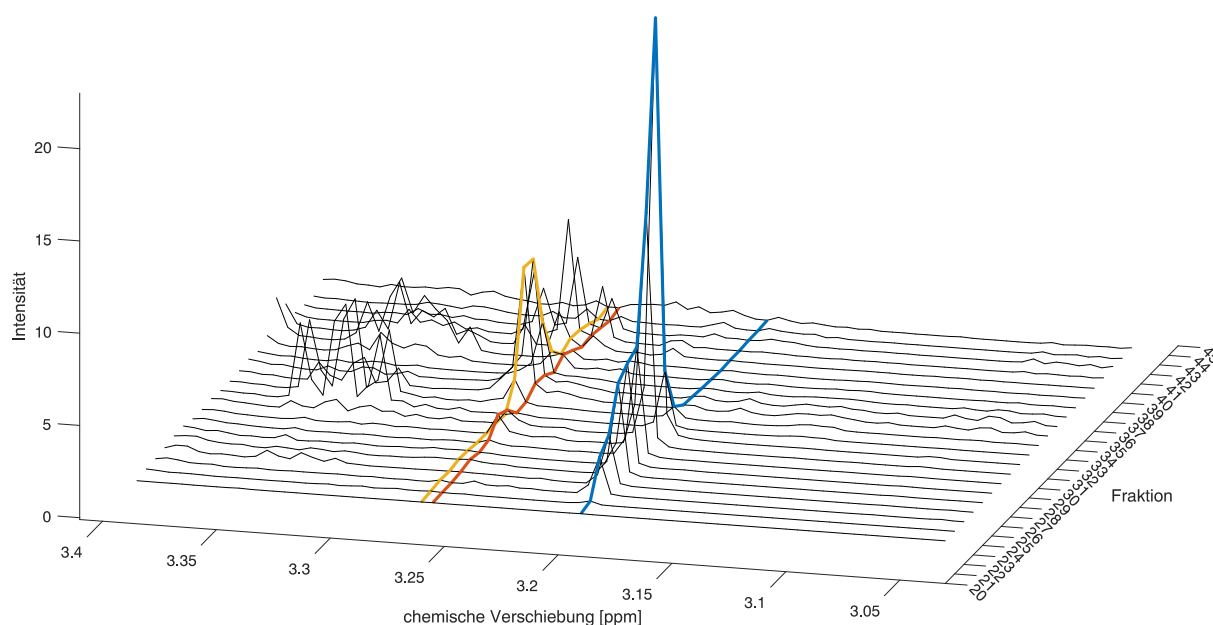


Abbildung 48: Ausschnitt der ^1H -NMR-Spektren der Fraktionen 24-44 im Bereich von 3.0-3.4 ppm. In diesem Bereich treten die Signale der *N*-Methylgruppen von Betain (**39**), Cholin (**8**) und Ergothionein (**42**) auf. Die EDCs dieser Methylgruppen sind farblich hervorgehoben. Der EDC von Cholin (**8**) bei 3.206 ppm ist in blau, der von Betain (**39**) bei 3.269 ppm in orange und der von Ergothionein (**42**) bei 3.276 ppm in gelb gezeigt.

Aufgrund der sehr ähnlichen chemischen Verschiebungen der Singulettts der Methylgruppen ist ohne weitere Analyse keine sichere Zuordnung möglich. Auch ein Vergleich mit Spektren aus einer Datenbank ist nur bedingt von Nutzen, da sich die exakte chemische Verschiebung durch einen anderen pH-Wert, ein anderes Lösungsmittel, eine andere Ionenstärke oder ähnliche Einflussfaktoren unterscheiden kann. Die Korrelation dieser EDCs mit den MS-Daten über die *SCORE-metabolite-ID* Methode ermöglicht jedoch die eindeutige Zuordnung zu spezifischen Masse-zu-Ladungs-Verhältnissen. Für die Metaboliten Betain (**39**), Cholin (**8**) und Phosphorylcholin (**40**) sind die relevanten Korrelationen in Tabelle 7 gezeigt. Weitere

Korrelationskoeffizienten sowie die Korrelationskoeffizienten des Metaboliten Trigonellin (41) sind im Anhang in Tabelle 40 angegeben.

Tabelle 7: Korrelationskoeffizienten der Metaboliten Betain (39), Cholin (8) und Phosphorylcholin (40) zwischen spezifischen EDCs und jeweils einem EMC. Für die Metaboliten ist jeweils die monoisotopische Masse sowie der Fraktionsbereich, über den die Korrelationskoeffizienten berechnet wurden, angegeben.

Betain (39) 117.0790 Da F21-34		Cholin (8) 104.1070 Da F1-36		Phosphorylcholin (40) 183.0660 Da F49-80	
ESI-(+)	118.09 ([M+H] ⁺)	ESI-(+)	104.11 ([M] ⁺)	ESI-(+)	184.07 ([M+H] ⁺)
3.269 ppm	0.9503	4.069 ppm	0.9983	4.176 ppm	0.9792
		3.531 ppm	0.9695	3.601 ppm	0.9754
		3.206 ppm	0.9894	3.226 ppm	0.9874

Die Metaboliten Betain (39), Cholin (8), Phosphorylcholin (40) und Trigonellin (41) konnten somit durch die Zuordnung von NMR- zu MS-Signalen eindeutig identifiziert werden. Im Fall des Metaboliten, der durch die im Folgenden dargestellte Analyse als Ergothionein (42) identifiziert werden konnte, werden durch die semi-automatische Korrelation des EDCs bei einer chemischen Verschiebung von 3.276 ppm eine Reihe von EMCs mit einem hohen Korrelationskoeffizienten detektiert. Die Signale mit einem hohen Korrelationskoeffizienten und gleichzeitig höchster Signalintensität sind in Tabelle 8 angegeben.

Tabelle 8: Intensivste EMCs mit einem Korrelationskoeffizienten > 0.95 zu dem vorgegebenen EDC bei einer chemischen Verschiebung von 3.276 ppm berechnet über die Fraktionen 1-80.

EMC [m/z]	Korrelationskoeffizient
120.08	0.9871
166.09	0.9887
230.10	0.9958
252.08	0.9834

Nicht alle der EMCs, die einen hohen Korrelationskoeffizienten zu dem EDC bei 3.276 ppm aufweisen, lassen sich aus chemischer Sicht sinnvoll einem einzelnen Metaboliten zuordnen. Bei Betrachtung der 80 Fraktionen weisen jedoch alle Verläufe der in Tabelle 8 aufgeführten EMCs ein ununterscheidbares Profil auf (siehe Abbildung 49 A). Aus diesem Grund wurde von der

Fraktion 35 ein LC-MS-Chromatogramm aufgenommen. Es wurde eine NUCLEOSHELL HILIC-Säule von MACHEREY-NAGEL verwendet. In Abbildung 49 B sind ausgewählte EICs im Retentionszeit-Bereich von 9.25 bis 11.75 Minuten dargestellt. Es sind die EICs derjenigen Masse-zu-Ladungs-Verhältnisse gezeigt, die durch die zuvor gezeigte semi-automatische Korrelation als EMCs mit hohem Korrelationskoeffizienten zu dem NMR-Signal detektiert wurden.

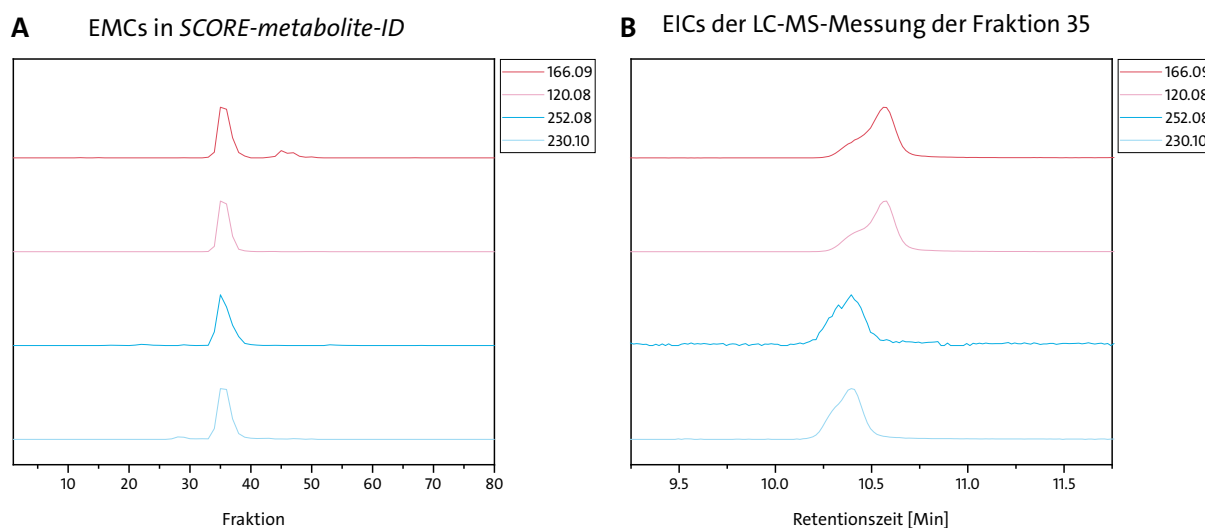
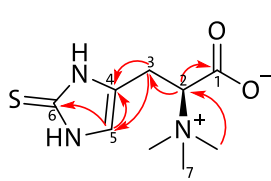


Abbildung 49: A: EMCs der in der Legende angegebenen Masse-zu-Ladungs-Verhältnisse, dargestellt über alle 80 Fraktionen. B: EICs der LC-MS-Messung der Fraktion 35 des polaren Pinienkernextraktes der gleichen Masse-zu-Ladungs-Verhältnisse. Es ist der Ausschnitt im Retentionszeitfenster zwischen 9.25 und 11.75 Minuten gezeigt.

Bei der erneuten säulenchromatographischen Trennung der Fraktion 35 mit Hilfe der analytischen HILIC-Säule ist deutlich erkennbar, dass sich die Verläufe der EICs teilweise unterscheiden. Die jeweils in einer Farbrichtung dargestellten EICs weisen im LC-MS-Chromatogramm den gleichen Verlauf auf und gehören dementsprechend mit hoher Wahrscheinlichkeit zum gleichen Molekül, wie z. B. die EICs der m/z -Werte 230.10 und 252.08 in Blautönen. Genauso gehören die EICs mit m/z = 166.09 und 120.08 in Rottönen zur selben Verbindung. Die roten EICs können der bereits zuvor analysierten Aminosäuren Phenylalanin (4) zugeordnet werden. Somit lässt sich der EDC bei 3.276 ppm eindeutig den EMCs bei m/z = 230.10 und 252.08 zuweisen. Bei Betrachtung der Differenz der EMC-Signale lassen sich diese als $[M+H]^+$ (m/z = 230.10) und als $[M+Na]^+$ -Signal (m/z = 252.08) zuordnen. Für den Metaboliten ergibt sich demnach eine monoisotopische Masse von 229.09 Da (bzw. 229.0885 ± 0.0001 Da). Auf Grundlage dessen wurde die Summenformel $C_9H_{15}N_3O_2S$ postuliert. In Analogie zur semi-automatischen Analyse der EMCs wurden semi-automatisch diejenigen EDCs ermittelt, die einen hohen Korrelationskoeffizienten zu dem EDC bei 3.276 ppm in den Fraktionen 30-40

aufweisen. Auch hier werden die aromatischen Signale der Aminosäure Phenylalanin (**4**) detektiert. Die Zusammengehörigkeit dieser Signale kann jedoch aufgrund der zuvor gewonnenen Ergebnisse ausgeschlossen werden. Auch unabhängig von der LC-MS-Messung kann über die Bestimmung der Integrale, z.B. in dem ^1H -NMR-Spektrum der Fraktion 35, die Zusammengehörigkeit dieser Signale aufgrund des Protonenverhältnisses ausgeschlossen werden. Zusätzlich werden EDCs mit einem hohen Korrelationskoeffizienten bei einer chemischen Verschiebung von 6.80 ppm und im Bereich von 3.1 bis 3.2 ppm detektiert. Ein selektives NOESY-Spektrum bestätigt die Korrelation des Signals bei 3.276 ppm zu den Signalen im Bereich von 3.1 bis 3.2 ppm (siehe Abbildung 148 im Anhang). Zusätzlich weisen die Protonen des Singulett eine räumliche Kopplung zu einem Dublett vom Dublett bei einer chemischen Verschiebung von 3.89 ppm auf. Basierend auf der gemeinsamen Betrachtung der einzelnen Ergebnisse kann auf die Struktur von Ergothionein (**42**) geschlossen werden. Die 2D HSQC- und HMBC-Spektren der Fraktion 35 (siehe Abbildung 149 im Anhang) bestätigen den Strukturvorschlag. Die NMR-Daten sind in Tabelle 9 aufgeführt.

Tabelle 9: ^1H - und ^{13}C -NMR-Daten des im polaren Extrakt der Pinienkerne identifizierten Metaboliten Ergothionein (**42**). Die ^{13}C -Daten wurden aus 2D HSQC- und HMBC-Daten der Fraktion 35 ermittelt. Die Aufspaltungen und Kopplungskonstanten der mit * gekennzeichneten Signale der diastereotopen Protonen H-3a und H-3b konnten aufgrund von Überlagerung mit dem Singulett von H-7 und der geringen Intensität der Signale nicht eindeutig bestimmt werden. Jedoch ergibt sich die Lage der Signale eindeutig aus den entsprechenden 2D NMR-Spektren. Die roten Pfeile zeigen die analysierten HMBC-Kopplungen an. 600.13 MHz (^1H -NMR) bzw. 150.90 MHz (^{13}C -NMR), D_2O , pH-Wert = 7.0, kalibriert auf TSP- d_4 .

Metabolit	$\delta_{13\text{C}}$ [ppm]	$\delta_{1\text{H}}$ [ppm]	Multiplizität	J [Hz]	Zuordnung
 Ergothionein (42)	173.0	-	-	-	1
	156.7	-	-	-	6
	126.7	-	-	-	4
	118.2	6.80	s	-	5
	80.1	3.89	dd	11.6, 3.5	2
	55.0	3.28	s	-	7
	25.7	3.26 3.18	-*	-*	3

5.2.1.4 Identifizierung von Nucleosiden, Nucleotiden und Derivaten

In Abbildung 50 sind die Nucleoside, Nucleotide, Nucleobasen und Derivate, die im Rahmen dieser Arbeit im polaren Extrakt der Pinienkerne identifiziert werden konnten, gezeigt. Die Relevanz der Korrelation von NMR- und MS-Daten ist anhand der Metaboliten Adenin (**43**), Adenosin (**44**), Adenosin-5'-monophosphat (**45**) (AMP) und Nicotinamidadenindinucleotid (**49**) (NAD) im Folgenden gezeigt. Die Korrelationskoeffizienten der anderen identifizierten Metaboliten dieser Klasse sind im Anhang in Tabelle 41 aufgeführt.

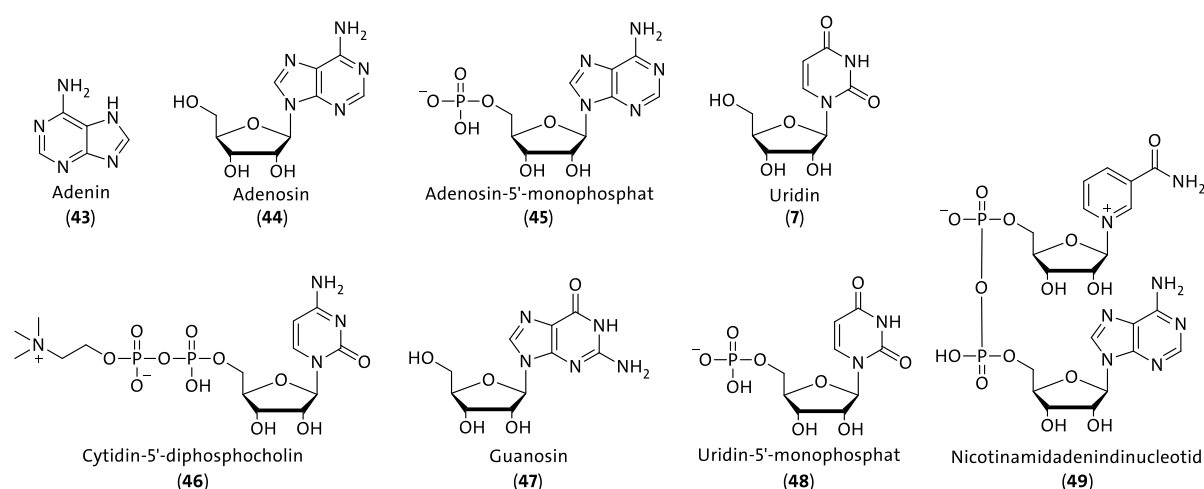


Abbildung 50: Im polaren Extrakt der Pinienkerne identifizierte Nucleoside, Nucleotide, Nucleobasen und Derivate.

Die Metaboliten der Adenin-Familie (Adenin (**43**), Adenosin (**44**), AMP (**45**) und NAD (**49**)) beinhalten das Strukturmotiv der Adenin-Nucleobase, das im ^1H -NMR-Spektrum lediglich zu zwei Singuletts im Bereich zwischen 8.1 und 8.7 ppm führt. Die chemische Verschiebung der Singuletts unterscheidet sich zwischen den verschiedenen Metaboliten, doch ein reiner Vergleich mit Spektren aus Datenbanken ist aufgrund der bereits zuvor erwähnten pH-Wert- und Konzentrations-Abhängigkeit nur bedingt von Nutzen. Der Ribose-Substituent des Adenosins (**44**), AMPs (**45**) und NADs (**49**) führt zusätzlich zu Signalen im Bereich zwischen 3 und 4 ppm, die in den Gemischspektren häufig von Signalen anderer Kohlenhydrate überlagert sind. Durch die Korrelation der NMR-Daten mit den Daten der Massenspektren ist es jedoch möglich, den verschiedenen Singuletts spezifische m/z -Werte zuzuordnen zu können. Für die EDCs der aromatischen Singuletts der Nucleobase sind beispielhaft die Korrelationen zu den Quasimolekülen in Tabelle 10 gezeigt.

Tabelle 10: Korrelationskoeffizienten der Metaboliten Adenin (**43**), Adenosin (**44**), AMP (**45**) und NAD (**49**) zwischen den spezifischen EDCs der Singulets der Nucleobase und jeweils einem EMC. Für die Metaboliten ist jeweils die monoisotopische Masse sowie der Fraktionsbereich, über den die Korrelationskoeffizienten berechnet wurden, angegeben.

Adenin (43) 135.0545 Da F1-37		Adenosin (44) 267.0968 Da F1-28		AMP (45) 347.0631 Da F65-75		NAD (49) 663.1091 Da F45-65	
ESI-(+)	136.06 [M+H] ⁺	ESI-(+)	268.10 [M+H] ⁺	ESI-(+)	348.07 [M+H] ⁺	ESI-(+)	664.12 [M+H] ⁺
8.254 ppm	0.9506	8.347 ppm	0.9943	8.602 ppm	0.9738	8.432 ppm	0.9501
8.201 ppm	0.9509	8.267 ppm	0.9750	8.272 ppm	0.9856	8.179 ppm	0.9381

Somit lässt sich durch die Korrelation der Daten in allen Fällen ein MS-Signal eindeutig zuordnen. Für die Nucleoside und Nucleotide werden zusätzlich auch Korrelationen mit weiteren NMR-Signalen der anderen Protonen detektiert. Diese sind im Anhang in Tabelle 41 aufgeführt. Trotz der ähnlichen Struktur motive lassen sich somit die Metaboliten eindeutig unterscheiden und ohne individuelle Isolierung identifizieren. Auch die Identifizierung der Metaboliten Adenosin (**44**) und AMP (**45**), die ¹H-NMR-spektroskopisch nur im direkten Vergleich der exakten chemischen Verschiebungen unterscheidbar sind, ist durch die Korrelation der Daten möglich.

5.2.1.5 Identifizierung organischer Säuren

Im polaren Extrakt der Pinienkerne konnten die in Abbildung 51 gezeigten organischen Säuren Nicotinsäure (**50**), Äpfelsäure (**51**), Zitronensäure (**52**), Fumarsäure (**53**) und Milchsäure (**54**) identifiziert werden. Die Korrelationskoeffizienten sind im Anhang in Tabelle 42 angegeben.

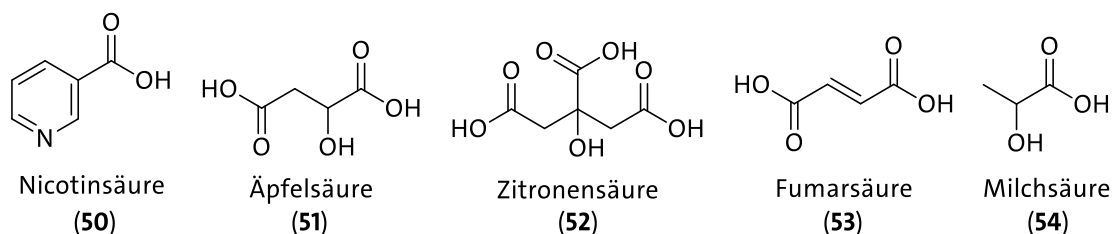


Abbildung 51: Im polaren Extrakt der Pinienkerne identifizierte organische Säuren.

In den Fraktionen 41 bis 45 wurden im aromatischen Bereich des NMR-Spektrums zwischen 7.5 und 9 ppm vier Signale detektiert, die untereinander eine hohe Korrelation aufweisen. Die Multiplizitäten und Verschiebungen der Signale lassen auf ein Pyridin-Derivat mit einem Substituenten in Position 3 schließen. Daher kommen beispielsweise Nicotinsäure (**50**) oder Nicotinamid (**55**) als natürlich vorkommende Metaboliten mit diesem Strukturmotiv in Frage (vgl. Abbildung 52).

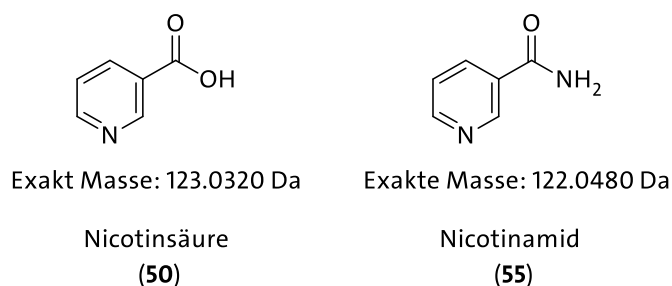


Abbildung 52: Strukturformeln der natürlich vorkommenden Metaboliten Nicotinsäure (**50**) und Nicotinamid (**55**), die das Strukturmotiv eines in Position 3 substituierten Pyridin-Derivats aufweisen, unter Angabe der monoisotopischen Masse.

Aufgrund des sehr ähnlichen Strukturmotivs lassen sich die beiden Metaboliten lediglich auf Grundlage der ^1H -NMR-Spektren nicht unterscheiden. Allerdings weisen die vier EDCs zwischen 7.5 und 9 ppm eine hohe Korrelation zu den EMCs bei $m/z = 124.04$ im positiven Ionisationsmodus und bei $m/z = 122.03$ im negativen Ionisationsmodus auf. Die Korrelationskoeffizienten sind im Anhang in Tabelle 42 angegeben.

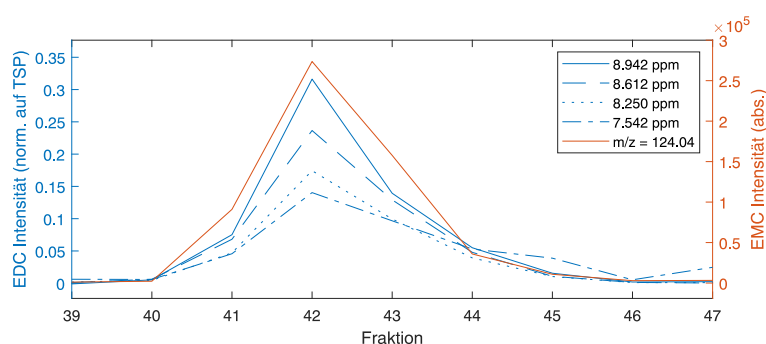


Abbildung 53: Zweidimensionale Darstellung der EDC-Verläufe in blau (linke y-Achse) sowie des EMC-Verlaufes des $[\text{M}+\text{H}]^+$ -Signals ($m/z = 124.04$) von Nicotinsäure (**50**) in orange (rechte y-Achse).

Unter der Annahme, dass es sich um das Signal des protonierten ($m/z = 124.04$) und des deprotonierten ($m/z = 122.03$) Moleküls handelt, lässt sich die monoisotopische Masse des Metaboliten zu 123.03 Da bestimmen. Demzufolge handelt es sich bei dem Metaboliten um

Nicotinsäure (**50**) mit einer monoisotopischen Masse von 123.03 Da und nicht um Nicotinamid (**55**) mit einer monoisotopischen Masse von 122.05 Da. Trotz der sehr ähnlichen Strukturformeln kann somit durch die Korrelation der NMR-Daten mit den MS-Daten eindeutig zwischen den Metaboliten unterschieden werden und die Signale eindeutig der Nicotinsäure (**50**) zugeordnet werden. Die Zuordnung wurde über ein *Spike-In* Experiment (siehe Abbildung 54) verifiziert.

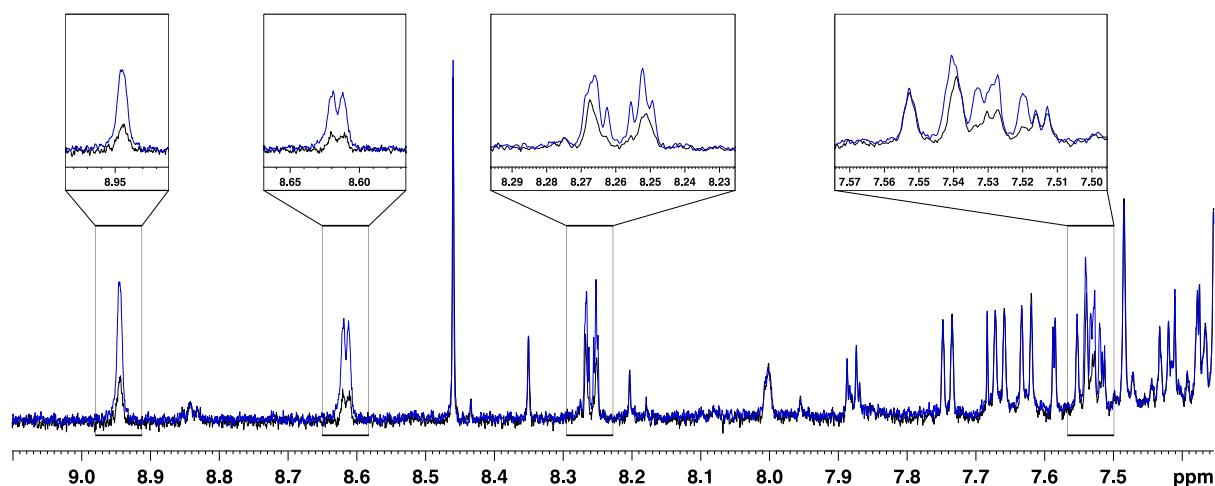


Abbildung 54: *Spike-In* Experiment des Metaboliten Nicotinsäure (**51**), der im polaren Extrakt der Pinienkerne identifiziert werden konnte, in eine Probe des Gesamtextraktes. Es wurden 0.5 μ L einer 0.1 M Lösung von Nicotinsäure (**50**) in D_2O hinzugegeben.

Der Metabolit Fumarsäure (**53**) eluiert in den Fraktionen 49 bis 58 und zeigt lediglich ein Singulett bei einer chemischen Verschiebung von 6.52 ppm. Für den Metaboliten konnte lediglich eine Korrelation zwischen dem EDC bei 6.52 ppm zu einem EMC bei $m/z = 99.01$ im positiven Ionisationsmodus detektiert werden. Dieses MS-Signal weist eine sehr geringe Intensität auf und kann dem dehydroxylierten Molekülion zugeordnet werden. Die Ionisierungseffizienz von Fumarsäure (**53**) scheint im Vergleich zu anderen Metaboliten sowohl im positiven als auch im negativen Ionisationsmodus vergleichsweise gering zu sein. Über ein *Spike-In* Experiment (siehe Abbildung 55) konnte der Metabolit verifiziert werden.

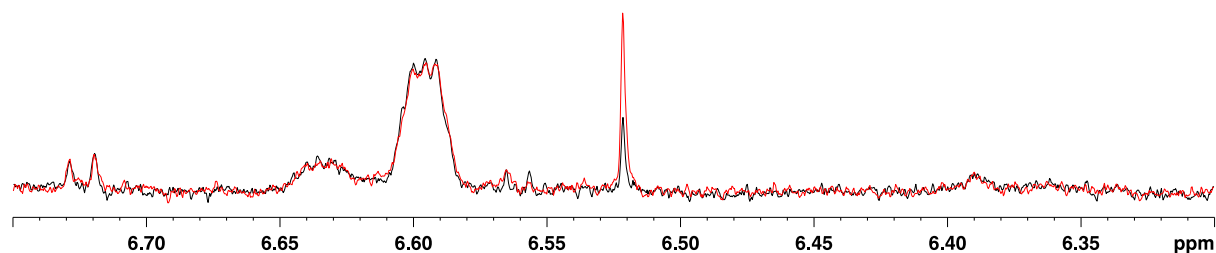


Abbildung 55: *Spike-In* Experiment mit Fumarsäure (**53**) (Singulett bei einer chemischen Verschiebung von 6.52 ppm). Es wurden 0.5 μ L einer 0.1M Lösung von Fumarsäure (**53**) in D₂O zur Gesamtextrakt-Probe der Pinienkerne hinzugegeben.

Bei der vorangegangenen Analyse der Aminosäure Threonin (**28**) wurde festgestellt, dass das Dublett bei einer chemischen Verschiebung von 1.33 ppm von einem weiteren Dublett überlagert ist. Das überlagernde Signal ist dabei leicht ins Hochfeld verschoben. Das Signal in der Fraktion 43 ist in Abbildung 56 dargestellt.

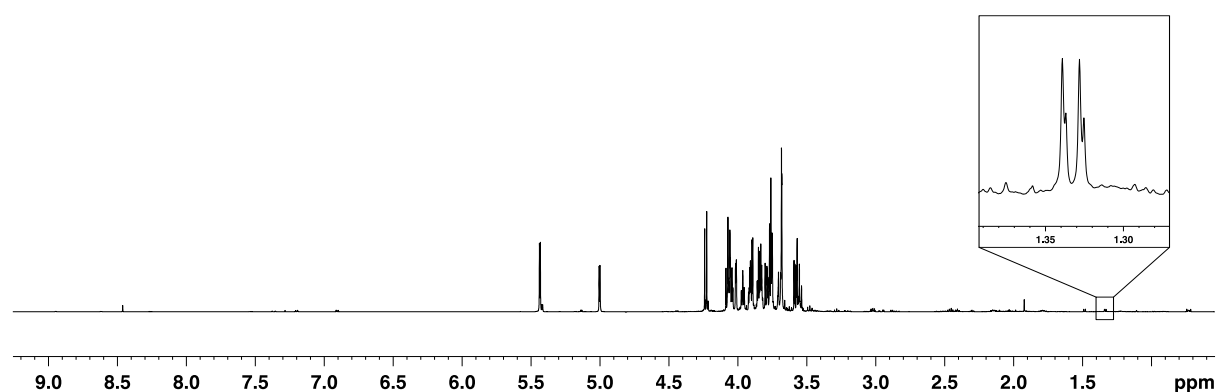


Abbildung 56: ¹H-NMR-Spektrum der Fraktion 43 des polaren Pinienkernextraktes. Die überlagerten Dubletts von Threonin (**28**) und Milchsäure (**55**) bei einer chemischen Verschiebung von 1.33 ppm sind vergrößert dargestellt.

Durch das *Bucketing* mit einer Größe von 0.005 ppm sind die Schultern des leicht verschobenen Dubletts innerhalb der Analyse mit Hilfe der *SCORE-metabolite-ID* App nicht mehr sichtbar. Zusätzlich sind die säulenchromatographischen Verläufe der beiden Metaboliten nicht unterscheidbar. Das selektive TOCSY-Spektrum mit selektiver Einstrahlung auf das Dublett bei 1.34 ppm, das zur Verifizierung der Aminosäure Threonin (**28**) akquiriert wurde, zeigt neben der Kopplung zu Signalen des Threonins (**28**) eine weitere Kopplung zu einem Signal bei 4.11 ppm (vgl. Abbildung 143 im Anhang). Dieses Signal tritt mit hoher Wahrscheinlichkeit somit durch die Anregung der beiden überlagerten Dubletts im selektiven TOCSY-Spektrum auf. Aufgrund der chemischen Verschiebung der NMR-Signale sowie des ähnlichen Elutionsverlaufes wurde ein ähnliches Strukturmotiv dieser Analyten vermutet, sodass Milchsäure als potenzieller Metabolit

in Frage kommt. Über ein *Spike-In* Experiment, das im Anhang in Abbildung 150 gezeigt ist, konnte das leicht ins Hochfeld verschobene Dublett dem Metaboliten Milchsäure (54) zugeordnet werden.

5.2.1.6 Identifizierung von β -D-Glucopyranosyl-1-*N*-indol-3-acetyl-*N*-L-asparaginsäure

In den Fraktionen 51 bis 59 des polaren Pinienkernextraktes treten mehrere Signale im aromatischen Bereich auf, die ein ähnliches Elutionsprofil aufweisen. Die Korrelationskoeffizienten zwischen den EDCs sind in Tabelle 11 angegeben und weisen allesamt Werte > 0.99 auf.

Tabelle 11: Korrelationskoeffizienten berechnet über die Fraktionen 50-65 zwischen EDCs im aromatischen Bereich der NMR-Spektren, die in den Fraktionen 51-59 des polaren Pinienkernextraktes auftreten.

EDC [ppm]	7.672	7.621	7.485	7.355	7.265
7.672	1.0000	0.9982	0.9985	0.9928	0.9997
7.621	0.9982	1.0000	0.9995	0.9969	0.9972
7.485	0.9985	0.9995	1.0000	0.9959	0.9976
7.355	0.9928	0.9969	0.9959	1.0000	0.9921
7.265	0.9997	0.9972	0.9976	0.9921	1.0000

Bei den in Tabelle 11 aufgeführten EDCs handelt es sich um zwei tieffeldverschobene Dubletts bei 7.67 ppm ($J = 8.0$ Hz) und 7.63 ppm ($J = 8.3$ Hz), um zwei Dubletts eines Dubletts bei 7.35 ppm ($J = 7.9, 8.1$ Hz) und 7.27 ppm ($J = 7.9, 7.9$ Hz) sowie um ein Singulett bei 7.49 ppm. Die Kombination der beiden Dubletts und der beiden Dubletts eines Dubletts lässt auf das Strukturmotiv eines *ortho*-substituierten Aromaten schließen. Das zusätzliche Singulett sowie die Signallagen und Integrale der Signale lassen auf ein Spinsystem eines Indol-Strukturmotivs schließen. Ausgehend von dem Singulett bei einer chemischen Verschiebung von 7.485 ppm werden eine Reihe von EMCs mit einem Korrelationskoeffizienten > 0.95 detektiert, die in Abbildung 57 graphisch dargestellt sind.

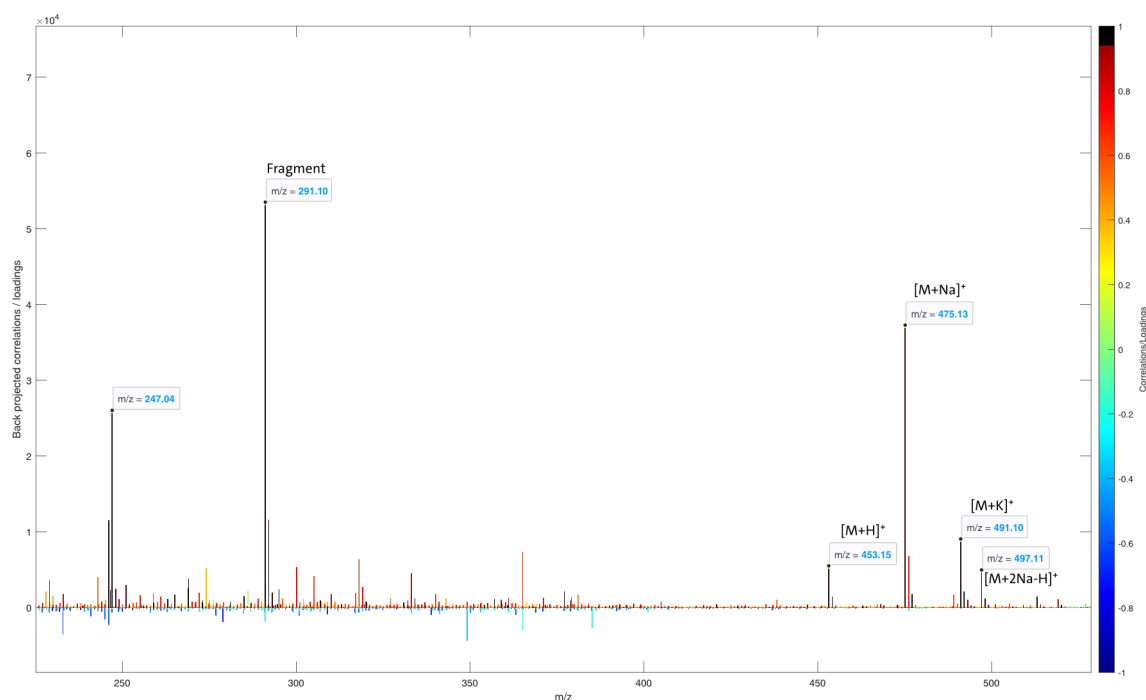


Abbildung 57: Graphische Darstellung der EMCs im positiven Ionisationsmodus, die einen Korrelationskoeffizienten > 0.95 zu dem EDC bei 7.485 ppm aufweisen. Die angegebene Zuordnung der Signale wurde unter Betrachtung der Differenzen der Signale getroffen.

Unter Betrachtung der Differenzen der hoch-korrelierenden EMCs können diese als die verschiedenen Addukt-Ionen identifiziert werden (siehe Zuordnung in Abbildung 57). Durch Subtraktion der Massen der jeweiligen Addukte lässt sich somit eine Annahme über die monoisotopische Masse des Analyten von $452.1435 (\pm 0.0004)$ treffen. In Tabelle 12 sind beispielhaft mögliche Summenformeln zu dieser angenommenen monoisotopischen Masse aufgelistet.

Tabelle 12: Mögliche Summenformeln zur ermittelten monoisotopischen Masse des Analyten.

Summenformel	Monoisotopische Masse	Doppelbindungsäquivalente
$C_5H_{20}N_{14}O_{11}$	452.1436	3
$C_{20}H_{24}N_2O_{10}$	452.1431	10
$C_{18}H_{12}N_{16}$	452.1431	21

Aufgrund der Kenntnis über die Indol-Substruktur, die selbst bereits acht Kohlenstoffatome beinhaltet, lässt sich die Summenformel $C_5H_{20}N_{14}O_{11}$ ausschließen. Die Summenformel $C_{18}H_{12}N_{16}$ erscheint aufgrund der hohen Anzahl der Doppelbindungsäquivalente von 21 als

unwahrscheinlich. Mit hoher Wahrscheinlichkeit lässt sich dadurch die Summenformel zu $C_{20}H_{24}N_2O_{10}$ und damit die monoisotopische Masse des Indol-Derivates zu 452.1431 Da ermitteln.

Die weitere Analyse mittels *SCORE-metabolite-ID* zeigt, dass zusätzlich auch Signale im aliphatischen und im Kohlenhydratbereich des NMR-Spektrums hohe Korrelationskoeffizienten zu den Signalen im aromatischen Bereich sowie zu den EMCs aufweisen. Die Korrelationskoeffizienten sind in Tabelle 13 aufgeführt.

Tabelle 13: Korrelationskoeffizienten zwischen den EDCs und EMCs (im positiven Ionisationsmodus) berechnet über die Fraktionen 51-65.

		EMCs [<i>m/z</i>]				
		491.10 [M+K] ⁺	475.13 [M+Na] ⁺	453.15 [M+H] ⁺	333.11 Fragment	291.10 Fragment
EDCs [ppm]	7.672	0.9802	0.9917	0.9730	0.9762	0.9828
	7.621	0.9677	0.9829	0.9581	0.9638	0.9714
	7.485	0.9701	0.9846	0.9609	0.9647	0.9734
	7.355	0.9532	0.9720	0.9447	0.9563	0.9583
	7.265	0.9817	0.9925	0.9761	0.9789	0.9836
	5.624	0.9477	0.9641	0.9321	0.9474	0.9574
	4.435	0.9631	0.9789	0.9567	0.9694	0.9675
	4.099	0.9664	0.9793	0.9643	0.9699	0.9662
	2.636	0.9662	0.9810	0.9595	0.9703	0.9700
	2.509	0.9705	0.9821	0.9719	0.9822	0.9710

Somit lassen sich über die Korrelation durch die *SCORE-metabolite-ID* Methode weitere NMR-Signale im aliphatischen Bereich zwischen 2.509 und 5.624 ppm identifizieren. Durch die Anwendung von selektiven TOCSY-Experimenten lassen sich dabei individuelle Spinsysteme detektieren. Zum einen lässt sich das Indol-Fragment als ein Strukturmotiv verifizieren (siehe Abbildung 58, zweites Spektrum von unten). Signale im Bereich von 5.624 ppm bis 4.099 ppm deuten auf eine Kohlenhydrat-Struktur hin. Auf Grundlage der Multiplizitäten und Kopplungskonstanten konnte das Kohlenhydrat über ein selektives TOCSY-Experiment als Glucose identifiziert werden (siehe Abbildung 58, zweites Spektrum von oben). Aus der Kopplungskonstante des anomeren Protons bei einer chemischen Verschiebung von 5.624 ppm von 9.2 Hz kann geschlossen werden, dass Glucose in der β -Konfiguration vorliegt. Die selektiven TOCSY-Spektren zeigen außerdem ein weiteres Strukturmotiv (oberes Spektrum in

Abbildung 58). Dabei ergeben die Signale ein ABX-Spinsystem, wie sie beispielsweise in Aminosäuren auftreten.

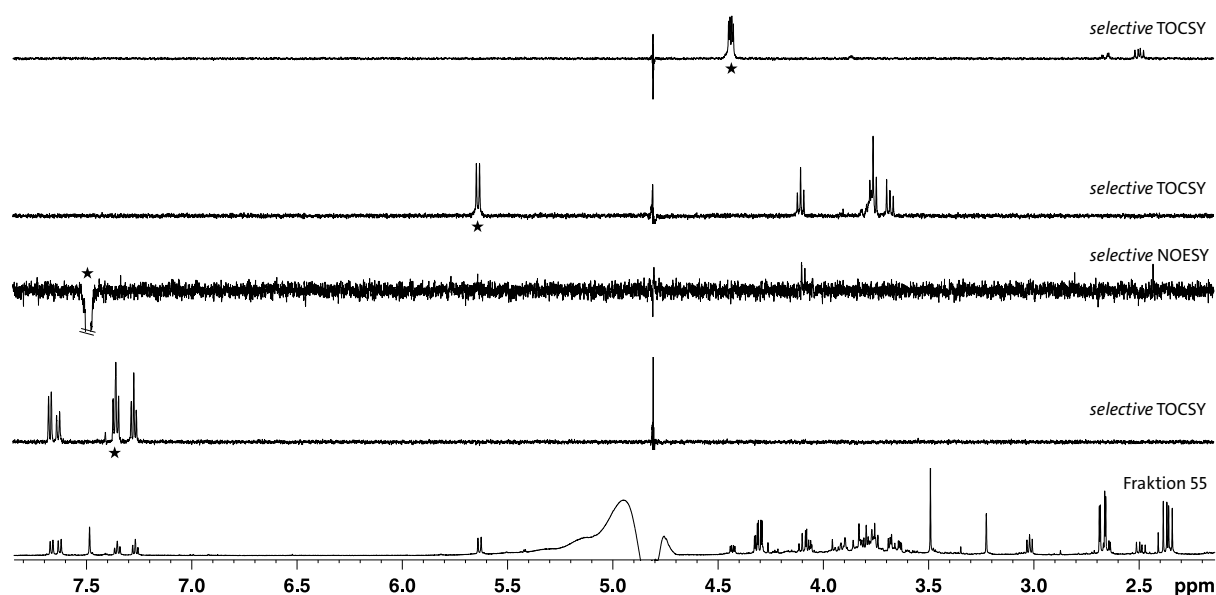


Abbildung 58: Selektive TOCSY und NOESY Experimente der Fraktion 55 des polaren Pinienkernextraktes. Es wurde jeweils das mit einem Sternchen markierte Signal selektiv angeregt.

Durch die *SCORE-metabolite-ID* Methode konnten somit zunächst für alle Substrukturen relevante Signale identifiziert werden, die anschließend durch weitere NMR-Experimente analysiert und somit vollständige Spinsysteme charakterisiert werden konnten.

Das selektive NOESY-Spektrum in Abbildung 58 zeigt eine räumliche Kopplung zwischen dem Singulett des Indols zu einem Signal im Kohlenhydrat-Bereich (drittes Spektrum von unten). Aus der vorangegangenen Analyse wird deshalb geschlossen, dass Glucose über das anomere Kohlenstoffatom mit einem Indol-Substituenten verknüpft ist. Das HMBC-Spektrum zeigt eine Kopplung zwischen dem anomeren Proton H-1' bei 5.63 ppm und den beiden Kohlenstoff-Atomen C2 (127.2 ppm) und C-7a (139.4 ppm) des Indol-Rings und bestätigt somit die Glucose-*N*-Indol-Bindung.

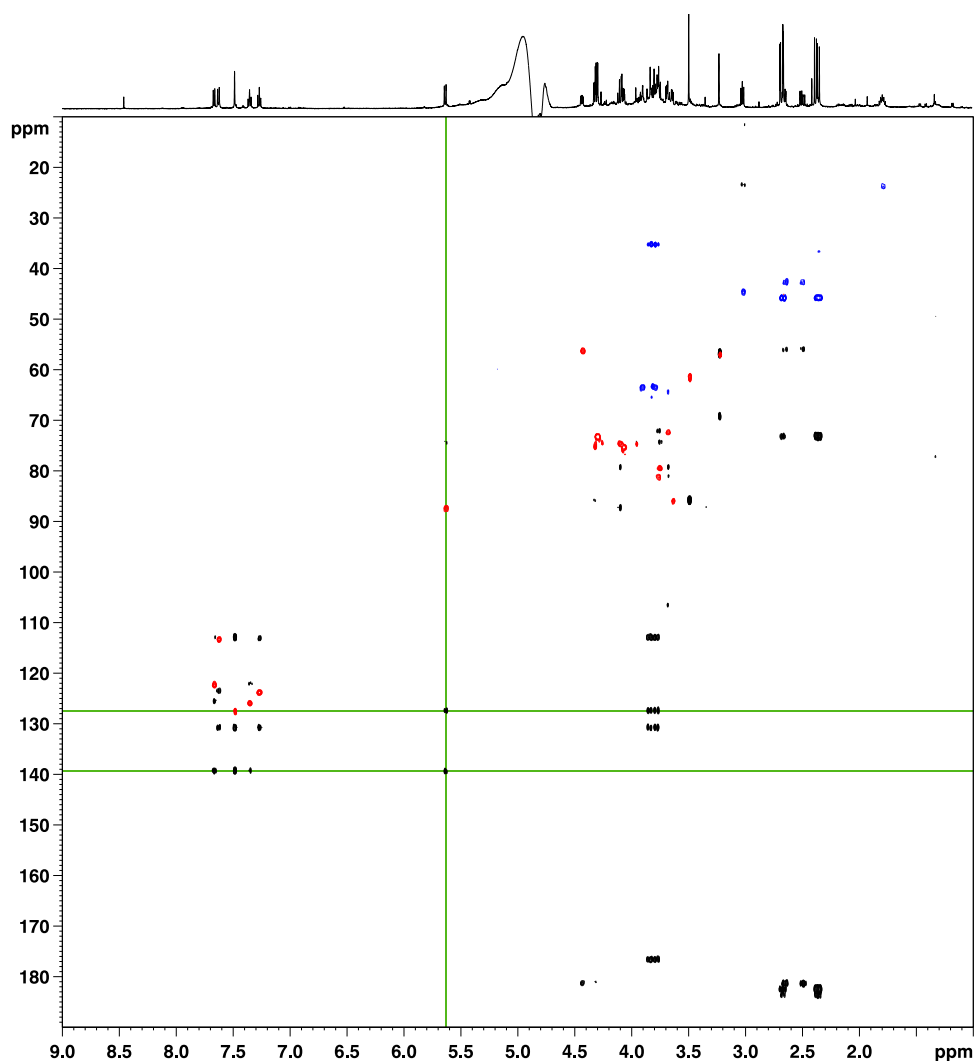


Abbildung 59: Ausschnitt des HSQC-Spektrums (rote und blaue Signale) und HMBC-Spektrums (schwarze Signale) der Fraktion 55 des polaren Pinienkernextraktes. Die grünen Linien zeigen die Kopplung des anomeren Protons der Glucose (H-1') bei 5.63 ppm zu C-2 (127.2 ppm) und C-7a (139.4 ppm) des Indol-Restes. Dies bestätigt die Glucose-*N*-Indol-Bindung.

Im obersten Spektrum in Abbildung 58 ist die Kopplung zwischen den aliphatischen Signalen bei 2.49 und 2.65 ppm zu einem Dublett vom Dublett bei 4.43 ppm entsprechend eines ABX-Spinsystems zu erkennen. Dieses Spinsystem kann als Asparaginsäure-Substruktur identifiziert werden. Unter Betrachtung der bereits identifizierten Teile des Metaboliten und der getroffenen Annahme der Summenformel ($C_{20}H_{24}N_2O_{10}$) wird geschlussfolgert, dass es sich bei der Indol-Substruktur um Indol-3-essigsäure handelt. Das HMBC-Spektrum in Abbildung 59 zeigt eine Kopplung des Indol-Spinsystems zu zwei diastereotopen Protonen bei 3.78 und 3.84 ppm (H-8a/b). Das Auftreten der Diastereotopie kann durch das Stereozentrum der Asparaginsäure erklärt werden, das in der Nähe der Protonen H-8 auftritt, wenn die Asparaginsäure und Indol-3-

essigsäure über eine Amid-Bindung miteinander verknüpft sind. Die Struktur des als β -D-Glucopyranosyl-1-*N*-indol-3-acetyl-*N*-L-asparaginsäure (**56**) (IAA-Asp-*N*-Glc) identifizierten Metaboliten ist in Abbildung 60 gezeigt.

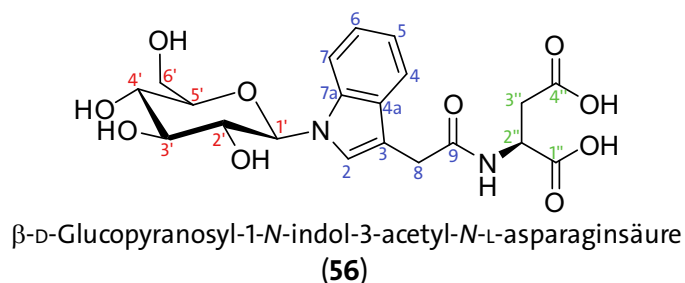


Abbildung 60: Strukturformel des Metaboliten β -D-Glucopyranosyl-1-*N*-indol-3-acetyl-*N*-L-asparaginsäure (**56**) (IAA-Asp-*N*-Glc), der in den Fraktionen 51-59 des polaren Extraktes der Pinienkerne identifiziert werden konnte.

Aufgrund der Tatsache, dass im HMBC-Spektrum keine Kopplung zwischen H-2'' und C-9 beobachtet wird, musste zum Nachweis der Amid-Bindung auf ein alternatives Verfahren zurückgegriffen werden. Zu diesem Zweck wurde die Fraktion 55 lyophilisiert und anschließend in einem Gemisch aus Deuteriumoxid und Wasser (Verhältnis 1:9) aufgenommen. Die NMR-Spektren sind in Abbildung 61 im Vergleich abgebildet.

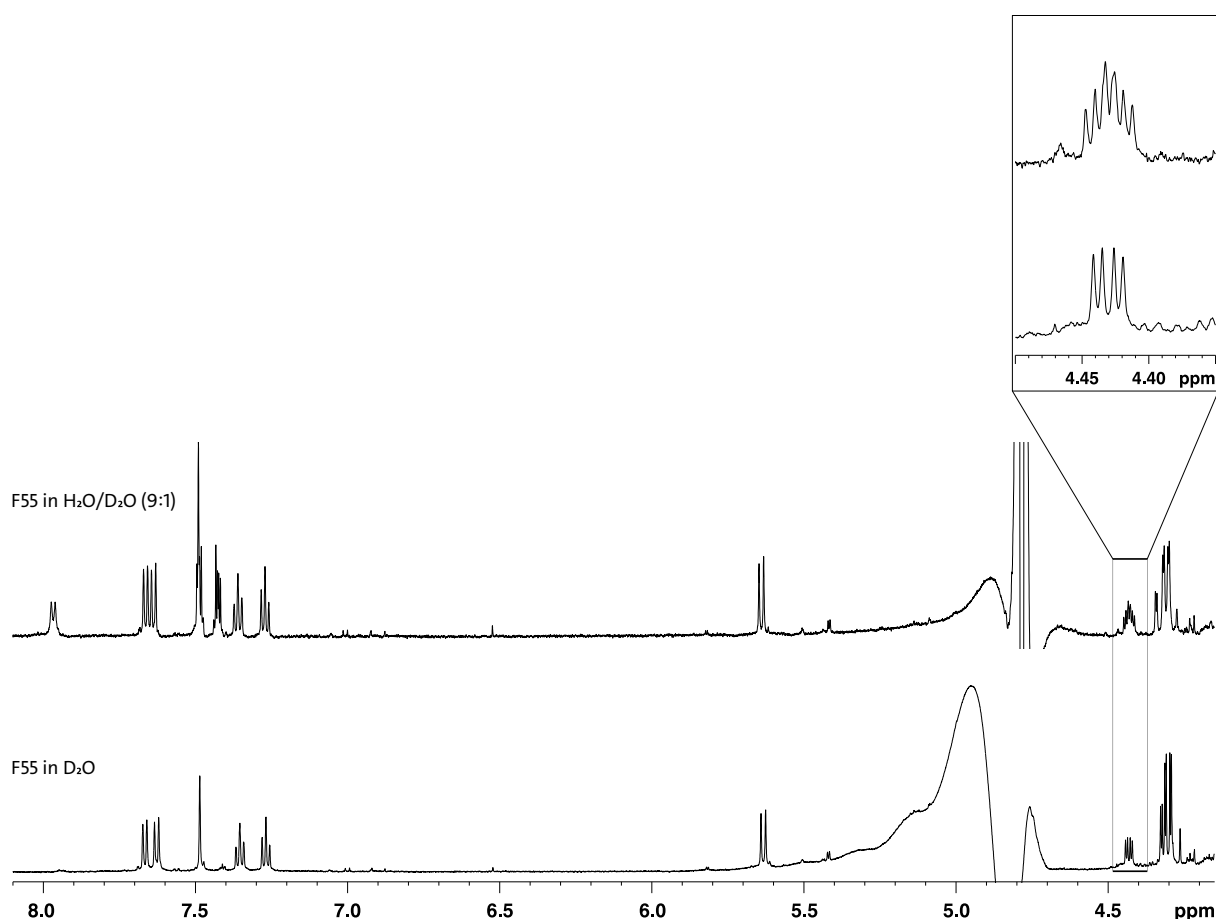


Abbildung 61: ^1H -NMR-Spektrum der Fraktion 55 des polaren Pinienkernextraktes in Deuteriumoxid (unten) sowie in einem Gemisch aus Deuteriumoxid und Wasser im Verhältnis 1:9 (oben). Im Spektrum, das in $\text{H}_2\text{O}/\text{D}_2\text{O}$ aufgenommen wurde, tritt ein neues Signal bei 7.97 ppm in Form eines Dubletts auf. Zusätzlich ändert sich die Signalaufspaltung des Signals bei 4.43 ppm von einem Dublett eines Dubletts in D_2O zu einem Dublett vom Dublett vom Dublett in $\text{H}_2\text{O}/\text{D}_2\text{O}$ durch die zusätzliche Kopplung zu dem Dublett des Amid-Protons bei 7.97 ppm.

Das ^1H -NMR-Spektrum der Fraktion 55, das in $\text{H}_2\text{O}/\text{D}_2\text{O}$ (Verhältnis 9:1) aufgenommen wurde, zeigt ein zusätzliches Dublett bei 7.97 ppm ($J = 7.6$ Hz, 1 H), welches vom Amid-Proton hervorgerufen wird. Diese Annahme wird durch die Änderung der Multiplizität des H-2'' bei einer chemischen Verschiebung von 4.43 ppm von einem Dublett eines Dubletts ($J = 9.3, 4.0$ Hz) in D_2O zu einem Dublett vom Dublett vom Dublett ($J = 4.1, 8.6, 8.7$ Hz) in $\text{H}_2\text{O}/\text{D}_2\text{O}$ (9:1) bestätigt. Die Änderung der Multiplizität des Signals tritt aufgrund der zusätzlichen Kopplung zum Amid-Proton auf. Die vollständige Zuordnung der NMR-Daten in D_2O sind in Tabelle 14 aufgeführt.

Tabelle 14: Zuordnung der ^1H - und ^{13}C -NMR-Daten des Metaboliten β -D-Glucopyranosyl-1-*N*-indol-3-acetyl-*N*-L-asparaginsäure (**56**) (IAA-Asp-*N*-Glc). 600.13 MHz (^1H -NMR) bzw. 150.90 MHz (^{13}C -NMR), D_2O , pH-Wert = 7.0, kalibriert auf TSP- d_4 . Die ^{13}C -NMR-Daten wurden aus den HSQC- und HMBC-Spektren (siehe Abbildung 59) entnommen.

$\delta_{13\text{C}}$ [ppm]	$\delta_{1\text{H}}$ [ppm]	Multiplizität	J [Hz]	Zuordnung
181.5	-	-	-	4'', 1''
176.7	-	-	-	9
139.4	-	-	-	7a
130.8	-	-	-	4a
127.2	7.49	s	-	2
125.8	7.35	dd	7.9, 8.1	6
123.7	7.27	dd	7.9, 7.9	5
122.2	7.67	d	8.0	4
113.2	7.63	d	8.3	7
112.9	-	-	-	3
87.1	5.63	d	9.2	1'
80.9	3.74-3.79	m	-	5'
79.3	3.73-3.78	m	-	3'
74.4	4.10	dd	9.2, 9.2	2'
72.3	3.68	dd	9.4, 9.4	4'
63.2	3.89-3.93	m	-	6'a
	3.83-3.77	m	-	6'b
56.1	4.43	dd	9.3, 4.0	2''
42.6	2.65	dd	15.5, 3.9	3''a
	2.49	dd	15.6, 9.3	3''b
35.0	3.84	d	16.3	8a
	3.78	d	16.0	8b

Der Metabolit IAA-Asp-*N*-Glc (**56**) wurde bereits mittels LC-MS/MS in Reispflanzen (*Oryza sativa*)^[148] und in keimenden Kiefernsemen (*Pinus sylvestris*)^[149] sowie nach Isolierung und Analyse mittels NMR-Spektroskopie und Massenspektrometrie aus Gingko-Samen (*Ginkgo biloba*)^[150] identifiziert. Ein Vergleich der Daten weist hohe Übereinstimmungen auf. Auch die hoch-korrelierenden Fragment-Ionen mit m/z -Werten von 291.10 und 333.11 stimmen mit Ergebnissen aus LC-MS/MS-Experimenten von Kai *et al.* überein.^[148] Die postulierten Fragmentierungsreaktionen zur Entstehung dieser beiden Fragment-Ionen sind in Abbildung 62 dargestellt.

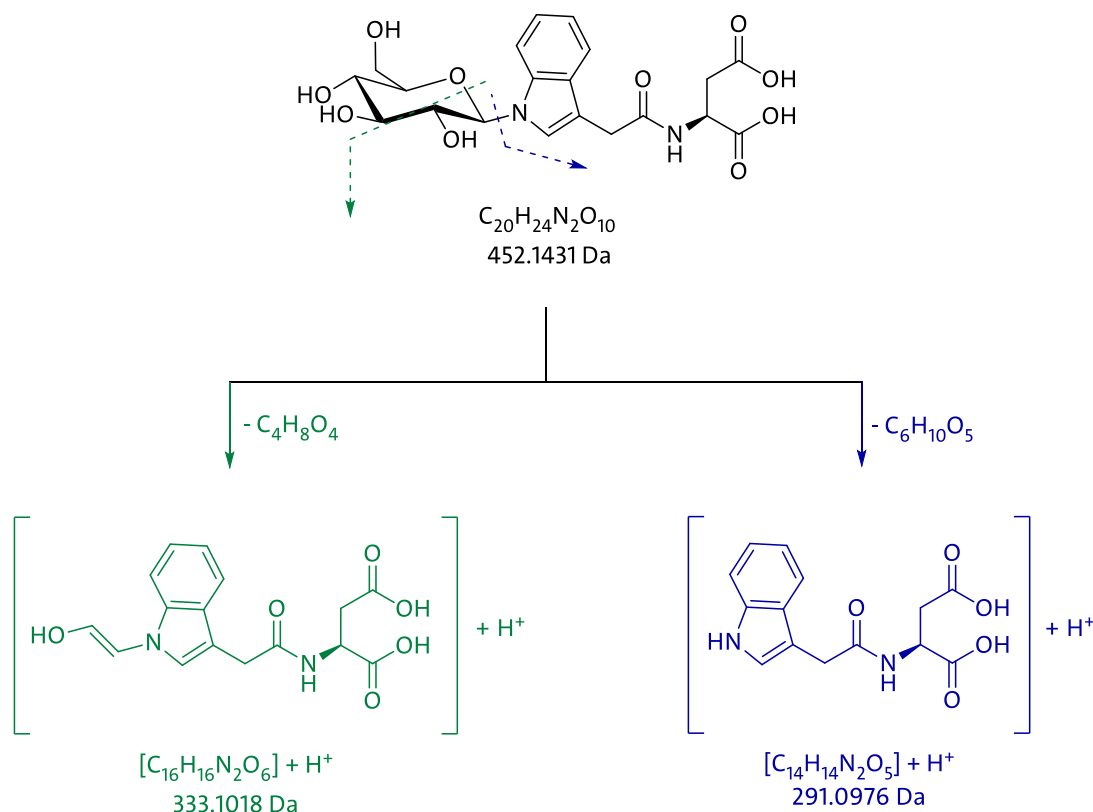


Abbildung 62: Postulierte Fragmentierungsreaktion des in den Pinienkernen identifizierten Metaboliten IAA-Asp-*N*-Glc (**56**) im positiven Ionisationsmodus.

Ebenso stimmen die ermittelten NMR-Daten unter Berücksichtigung der zu erwartenden Abweichungen durch Verwendung unterschiedlicher Lösungsmittel oder pH-Werte weitestgehend mit denen der Literaturdaten überein.^[148,150]

Nach CHENG *et al.* ist der Metabolit als Ginkgoside B^[150] benannt, jedoch ist ein anderer Naturstoff mit gänzlich anderer Struktur ebenfalls unter dem Namen Ginkgoside B^[151] in der Literatur beschrieben, weshalb dieser Name im Rahmen dieser Arbeit nicht verwendet wird. Nach Studien von LUNG *et al.* dient IAA-Asp-*N*-Glc (**56**) nicht als Speicher für freie Indol-3-essigsäure, die als Vertreter eines Auxins bspw. für das Wachstum und die Entwicklung von Pflanzen benötigt wird.^[149,152] Stattdessen tritt das Konjugat als Intermediat im Katabolismus von Indol-3-essigsäure als irreversible Form der Inaktivierung der Indol-3-essigsäure auf.^[149] Während des Katabolismus von Indol-3-essigsäure findet zuerst die Knüpfung der Amid-Bindung zu Asparaginsäure statt, wodurch IAA-Asp erhalten wird, und anschließend die Verknüpfung der Glucose über das Stickstoff-Atom zu IAA-Asp-*N*-Glc (**56**).^[149,152]

Die vorgestellten Ergebnisse zur Identifizierung von IAA-Asp-*N*-Glc (**56**) zeigen, dass die Anwendung von *SCORE-metabolite-ID* zwar keine direkte vollständige Strukturaufklärung ermöglicht, jedoch liefert die Methode die entscheidenden Anhaltspunkte, die in Kombination mit weiteren, gezielt eingesetzten NMR-Experimenten anschließend die vollständige Strukturanalyse ermöglichen. Dieses Beispiel verdeutlicht, dass auch unbekannte und vergleichsweise komplexe Analyten ohne vorherige Isolierung mittels *SCORE-metabolite-ID* zuverlässig und effektiv identifiziert werden können.

5.2.1.7 Identifizierung von *N*^α-(2-Hydroxy-2-carboxymethylsuccinyl)-L-arginin

Im aliphatischen Bereich der NMR-Spektren der Fraktionen 64 bis 70 fallen die AB-Spinsysteme auf (2.47 ppm und 2.66 ppm mit $J = 15.1$ Hz sowie 2.56 ppm und 2.71 ppm mit $J = 15.2$ Hz), deren Signale beispielhaft anhand der Fraktion 67 in Abbildung 63 (unten) gezeigt sind. Die semi-automatische Korrelation des Dubletts bei 2.456 ppm zu EMCs im positiven und negativen Ionisationsmodus ergibt eine Reihe von m/z -Werten, die durch Betrachtung von charakteristischen Differenzen den verschiedenen Quasimolekül- und Addukt-Ionen zugeordnet werden können, die in Tabelle 15 angegeben sind. Darunter wird auch ein EMC mit einem m/z -Verhältnis von 175.12 mit einem hohen Korrelationskoeffizienten detektiert. Der m/z -Wert lässt darauf schließen, dass es sich um ein Fragment des Metaboliten handelt, das zu diesem Zeitpunkt allerdings noch nicht eindeutig zugeordnet werden kann.

Tabelle 15: Intensivste EMCs im positiven und negativen Ionisationsmodus, die einen Korrelationskoeffizient > 0.95 zu dem vorgegebenen EDC bei 2.456 ppm in Fraktion 60-73 aufweisen. Neben dem m/z -Wert und dem Korrelationskoeffizient ist die Zuordnung der EMCs angegeben.

EMC [m/z]	Korrelationskoeffizient	Zuordnung
175.12	0.9737	Fragment
349.13	0.9933	$[M+H]^+$
350.14	0.9936	$[M+H]^+$ Isotopenpeak
351.14	0.9629	$[M+H]^+$ Isotopenpeak
393.10	0.9650	$[M+2Na-H]^+$
329.11	0.9757	$[M-H_2O-H]^-$
347.12	0.9938	$[M-H]^-$

Über die postulierte Zuordnung der EMCs lässt sich die wahrscheinlichste Summenformel des Metaboliten zu $C_{12}H_{20}N_4O_8$ mit einer monoisotopischen Masse von 348.1281 Da bestimmen. Durch die semi-automatische Korrelation der oben angegebenen EMCs werden weitere korrelierende EDCs bei einer chemischen Verschiebung von 4.24 ppm, 3.19 ppm, 1.88 ppm, 1.64 ppm und 1.57 ppm ermittelt. Die Korrelationskoeffizienten zwischen den EDCs und EMCs sind in Tabelle 16 gegeben.

Tabelle 16: Korrelationskoeffizienten zwischen EDCs und EMCs im positiven und negativen Ionisationsmodus berechnet über die Fraktionen 60-73.

		EMCs [m/z]			
		349.13 [M+H] ⁺	175.12 Fragment	347.12 [M-H] ⁻	329.11 [M-H ₂ O-H] ⁻
EDCs [ppm]	4.241	0.9685	0.9643	0.9719	0.9674
	3.196	0.9936	0.9798	0.9946	0.9755
	2.651	0.9713	0.9176	0.9689	0.9662
	2.456	0.9933	0.9737	0.9938	0.9757
	1.876	0.9559	0.9709	0.9602	0.9438
	1.636	0.9700	0.9702	0.9724	0.9517
	1.566	0.9611	0.9642	0.9642	0.9438

Die selektiven TOCSY-Spektren in Abbildung 63 verifizieren die Zusammengehörigkeit einzelner Signale, die durch die semi-automatische Korrelation zuvor detektiert wurden. Es sind drei unabhängige Spinsysteme zu erkennen. Das zweite und dritte Spektrum von oben zeigen die beiden AB-Spinsysteme. Das obere Spektrum zeigt ein selektives TOCSY-Spektrum, das starke Ähnlichkeiten zu dem ¹H-NMR-Spektrum der Aminosäure Arginin (**34**) aufweist (vgl. selektives TOCSY-Spektrum in Abbildung 144 im Anhang). Die Zuordnung des EMCs bei $m/z = 175.12$, der zuvor als Fragment zugeordnet wurde (siehe Tabelle 15 und Tabelle 16), kann nun als protoniertes Molekülion des Arginin-Fragmentes definiert werden.

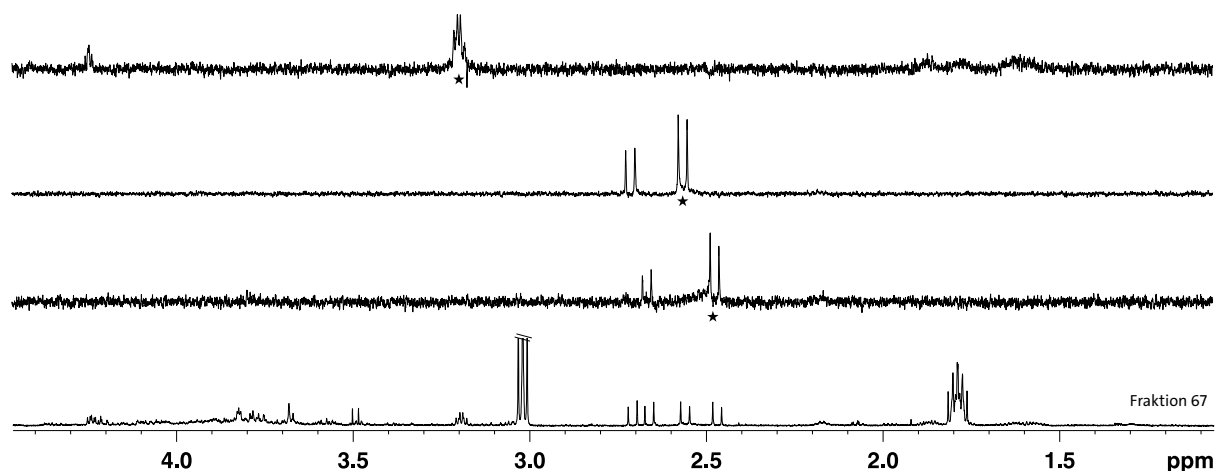


Abbildung 63: ^1H -NMR-Spektrum der Fraktion 67 des polaren Pinienkernextraktes (unten) sowie selektive TOCSY-Spektren. Die Sternchen in den selektiven TOCSY-Spektren markieren das selektiv angeregte Signal.

Aus den gewonnenen Informationen kann geschlossen werden, dass es sich bei dem Metaboliten um das Kondensationsprodukt von Zitronensäure (**52**) und Arginin (**34**) (N^α -(2-Hydroxy-2-carboxymethylsuccinyl)-L-arginin (**57**)) handelt (siehe Abbildung 63).

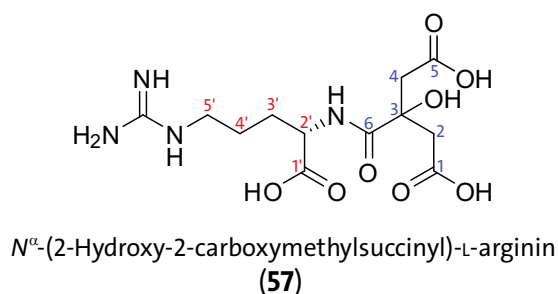


Abbildung 64: Strukturformel des Kondensationsproduktes aus Zitronensäure (**52**) und Arginin (**34**) (N^α -(2-Hydroxy-2-carboxymethylsuccinyl)-L-arginin (**57**)), das im polaren Extrakt der Pinienkerne in den Fraktionen 64-70 identifiziert werden konnte.

Prinzipiell wäre auch das Kondensationsprodukt über die Carboxylgruppe des Arginins (**34**) zur Hydroxylgruppe der Zitronensäure (**52**) in Form eines Esters denkbar. Die Bindung über das Amid, die durch die Kondensation zwischen der Amino-Gruppe des Arginins (**34**) und der Carboxylgruppe der Zitronensäure (**52**) auftritt, kann über das NMR-Spektrum der Fraktion in einem Gemisch aus Wasser und Deuteriumoxid (Verhältnis 9:1) nachgewiesen werden. Im ^1H -NMR-Spektrum der Fraktion 67 in $\text{H}_2\text{O}/\text{D}_2\text{O}$ tritt ein zusätzliches Dublett bei einer chemischen Verschiebung von 8.01 ppm auf, was dem Amid-Proton von (N^α -(2-Hydroxy-2-carboxymethylsuccinyl)-L-arginin (**57**)) zuzuordnen ist. Das 2D-TOCSY-Spektrum dieser Probe zeigt die Kopplung dieses Dubletts zu den Protonen H-2' bei 4.24 ppm und zu den Protonen

H-3'a/b und H-4'a/b und verifiziert damit die Verknüpfung des Arginins (**34**) und der Zitronensäure (**52**) über diese Amidbindung. Das 2D-TOCSY-Spektrum ist im Anhang in Abbildung 151 gezeigt. Die Zuordnung der ^1H - und ^{13}C -NMR-Daten ist in Tabelle 17 aufgeführt.

Tabelle 17: Zuordnung der ^1H - und ^{13}C -NMR-Daten des Metaboliten N^α -(2-Hydroxy-2-carboxymethylsuccinyl)-L-arginin (**57**). 600.13 MHz (^1H -NMR) bzw. 150.90 MHz (^{13}C -NMR), D_2O , pH-Wert = 7.0, kalibriert auf TSP- d_4 . Die ^{13}C -NMR-Daten wurden aus den HSQC- und HMBC-Spektren entnommen, die im Anhang in Abbildung 152 und Abbildung 153 gezeigt sind.

$\delta_{13\text{C}}$ [ppm]	$\delta_{1\text{H}}$ [ppm]	Multiplizität	J [Hz]	Zuordnung
181.2, 181.1	-	-	-	1, 5
181.1	-	-	-	1'
179.4	-	-	-	6
77.6	-	-	-	3
59.6	4.24	dd	6.6, 4.5	2'
47.4	2.71	d	15.2	2a
	2.56	d	15.2	2b
47.5	2.66	d	15.1	4a
	2.47	d	15.1	4b
43.4	3.16-3.22	m	-	5'
31.6	1.86-1.90	m	-	3'a
	1.76-1.81	m	-	3'b
26.2	1.60-1.65	m	-	4'a
	1.54-1.60	m	-	4'b

Zur Zuordnung der quartären Carbonylkohlenstoffatome C-1/C-5 und C-6 der Zitronensäure-Substruktur in N^α -(2-Hydroxy-2-carboxymethylsuccinyl)-L-arginin (**57**) wurde ein HMBC-Vergleichsspektrum von Zitronensäure (**52**) herangezogen. Anhand der Signalintensitäten im ^{13}C -NMR-Spektrum von Zitronensäure (**52**) lässt sich eindeutig zuordnen, dass das Signal bei 176.4 ppm von den Kohlenstoffatomen C-1 und C-5 und das Signal bei 179.8 ppm von dem Carbonylkohlenstoffatom C-6 herrührt. Das HMBC-Spektrum zeigt Kopplungen zwischen den Protonen H-2a/H-4a (3.04 ppm) und H-2b/H-4b (2.86 ppm) zu allen Carbonylkohlenstoffatomen (siehe Abbildung 65 A). In einer Reihe von HMBC-Spektren mit unterschiedlichen Werten der mittleren ^1H - ^{13}C -Kopplungskonstante (CNST13) fällt allerdings auf, dass die Intensität des Signals der Kopplung zwischen H-2a/H-4a zu C-6 (3.04/179.8 ppm) variiert. Dies ist in Abbildung 65 B für die vier Signale des HMBC-Spektrums, die in Abbildung 65 A dargestellt sind, für die

verschiedene Werte der *long range* ^1H - ^{13}C -Kopplungskonstante (CNST13) anhand des absoluten Integrals der Signale gezeigt. Die absoluten Signalintensitäten zeigen die gleiche Abhängigkeit.

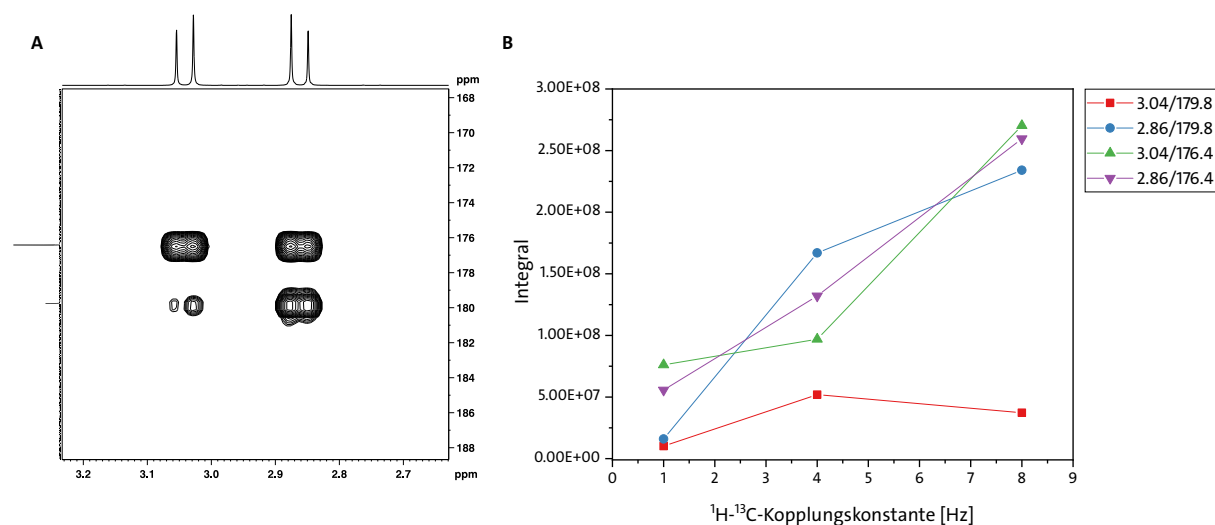


Abbildung 65: A: Ausschnitt des HMBC-Spektrums von Zitronensäure (**52**) in Deuteriumoxid. Die *long range* ^1H - ^{13}C -Kopplungskonstante (CNST13) wurde im Experiment auf einen Wert von 8 Hz eingestellt. Das Signal zwischen 3.04 ppm und 179.8 ppm ist im Vergleich zu den anderen drei Signalen deutlich weniger intensiv. B: Integrale der HMBC-Kopplungen zwischen den in A dargestellten Signalen von Zitronensäure (**52**) in Abhängigkeit von der *long range* ^1H - ^{13}C -Kopplungskonstante im HMBC-Experiment (CNST13).

Bei einer mittleren ^1H - ^{13}C Kopplungskonstante von 8 Hz, was dem Standard-Wert für HMBC-Experimente bei Routine-Messungen entspricht, ist die Intensität bzw. das Integral des Signals der Kopplung der Protonen H-2a/H-4a zu C-6 (3.04/179.8 ppm) im Vergleich mit den anderen Kopplungssignalen deutlich geringer. Die Intensität der Kopplung der Protonen H-2b/H-4b zu C-6 (2.86/179.8) liegt dahingegen im ähnlichen Bereich wie die anderen Kopplungen im HMBC-Spektrum. Die 3J -(^1H - ^{13}C)-Kopplungskonstante ist, wie auch die 3J -(^1H - ^1H)-Kopplungskonstante, entsprechend der KARPLUS-Beziehung abhängig vom Diederwinkel.^[70] Im kristallinen Zustand ist das Kohlenstoffgerüst von Zitronensäure (**52**) (C-1 bis C-5) vollständig gestreckt und die Carboxyl- und Hydroxylgruppe an C-3 stehen senkrecht dazu.^[153] Damit beträgt der Winkel zwischen H-2a/H-4a zu C-6 60° und der Winkel zwischen H-2b/H-4b zu C-6 beträgt 180° . Entsprechend der KARPLUS-Beziehung beträgt die vicinale ^1H - ^{13}C -Kopplungskonstante bei einem Winkel von 60° ca. 2 Hz und bei einem Winkel von 180° ca. 9 Hz.^[70] Für den Fall, dass das HMBC-Experiment mit einer mittleren vicinalen ^1H - ^{13}C -Kopplungskonstante von 8 Hz aufgenommen

wird, ist der Polarisationstransfer für die Protonen H-2a/H-4a zu C-6 dementsprechend verringert, weshalb dieses Signal weniger intensiv erscheint.

Das HMBC-Spektrum des Metaboliten *N^α-(2-Hydroxy-2-carboxymethylsuccinyl)-L-arginin (57)* ist im Anhang in Abbildung 153 gezeigt und weist ein ähnliches Phänomen bzgl. der unterschiedlichen Intensitäten der HMBC-Kopplungen zwischen den Dubletts und den Carbonylkohlenstoffatomen wie das Spektrum von Zitronensäure (52) auf. Das HMBC-Spektrum, das mit einer mittleren ¹H-¹³C-Kopplungskonstante von 8 Hz aufgenommen wurde, zeigt intensive Kopplungen zwischen den Dubletts der Protonen H-2b und H-4b (2.56 ppm und 2.47 ppm) zu einem Kohlenstoffatom mit einer chemischen Verschiebung von 179.4 ppm. Die Kopplung der Protonen H-2a und H-4a (2.71 ppm und 2.66 ppm) zu diesem ¹³C-Signal sind deutlich weniger intensiv. Zusätzlich zeigen alle vier Dubletts intensive Kopplungen zu Signalen von Kohlenstoffatomen mit einer chemischen Verschiebung von 181.1 ppm und 181.2 ppm. Unter der Annahme, dass die Zitronensäure-Substruktur im Metaboliten *N^α-(2-Hydroxy-2-carboxymethylsuccinyl)-L-arginin (57)* eine ähnliche Vorzugskonformation einnimmt wie reine Zitronensäure (52), kann eine Zuordnung der Signale der Carbonylkohlenstoffatome C-1/C-5 und C-6 getroffen werden. Das Carbonylkohlenstoffatom C-6 hat dementsprechend eine chemische Verschiebung von 179.4 ppm und die Kohlenstoffatome C-1 und C-5 von 181.1 ppm und 181.2 ppm. Wie in Abbildung 153 zu sehen ist, tritt im HMBC-Spektrum der Fraktion 68 in diesem Bereich außerdem eine Kopplung zwischen dem Dublett bei einer chemischen Verschiebung von 2.56 ppm zu einem Kohlenstoffatom mit einer chemischen Verschiebung von 184.3 ppm auf. Diese Kopplung tritt im HMBC-Spektrum der Fraktion 67 nicht auf und kann dem Metaboliten Zitronensäure (52) zugeordnet werden. Zitronensäure (52) eluiert ab Fraktion 68 und führt dementsprechend in der ¹H-Dimension ab Fraktion 68 zu Überlagerungen des Dubletts bei einer chemischen Verschiebung von 2.56 ppm.

Der Metabolit *N^α-(2-Hydroxy-2-carboxymethylsuccinyl)-L-arginin (57)* wurde bereits in den Jahren 1983 und 1984 in Blumenzwiebeln von Lilien (*Lilium maximowiczii*) sowie in Jahrestrieben von Birnenbäumen nach Isolierung identifiziert.^[154,155] In der verwendeten Literatur ist bisher kein weiterer Nachweis dieses Metaboliten in einem Lebensmittel bekannt.

5.2.1.8 Identifizierung einer 3-O-Methylaldarsäure

In den Fraktionen 51 bis 60 des polaren Pinienkernextraktes konnte eine 3-O-Methylaldarsäure (**58**) als Metabolit identifiziert werden. Die Identifizierung erfolgte ausgehend von einem Singulett bei einer chemischen Verschiebung von 3.49 ppm. Über die *SCORE-metabolite-ID* App werden im positiven und negativen Ionisationsmodus die in der Tabelle 18 aufgeführten EMCs mit einem hohen Korrelationskoeffizienten detektiert.

Tabelle 18: Intensivste EMCs im positiven und negativen Ionisationsmodus, die einen Korrelationskoeffizient > 0.95 zu dem vorgegebenen EDC bei 3.49 ppm in Fraktion 51-80 aufweisen. Neben dem m/z -Wert und dem Korrelationskoeffizient ist die Zuordnung der EMCs angegeben.

EMC [m/z]	Korrelationskoeffizient	Zuordnung
247.04	0.9709	$[M+Na]^+$
269.02	0.9856	$[M+2Na-H]^+$
223.05	0.9505	$[M-H]^-$
245.03	0.9701	$[M+Na-2H]^-$

Über die postulierte Zuordnung der EMCs auf Basis der beobachteten charakteristischen Massendifferenzen kann die exakte Masse des Metaboliten berechnet werden und damit die Summenformel $C_7H_{12}O_8$ (monoisotopische Masse: 224.0532 Da) als wahrscheinlichste Summenformel bestimmt werden. Das Singulett bei einer chemischen Verschiebung von 3.49 ppm lässt auf einen Methoxy-Substituenten schließen. Das selektive NOESY-Spektrum in Abbildung 66 (Mitte) zeigt eine räumliche Kopplung zu einem Dublett bei 4.33 ppm sowie zu einem Dublett vom Dublett bei 3.63 ppm. Das selektive TOCSY-Spektrum dieses Dublett eines Dubletts in Abbildung 66 (oben) zeigt die weitere Kopplung zu einem Dublett vom Dublett bei 4.07 ppm sowie zu einem Dublett bei 4.08 ppm.

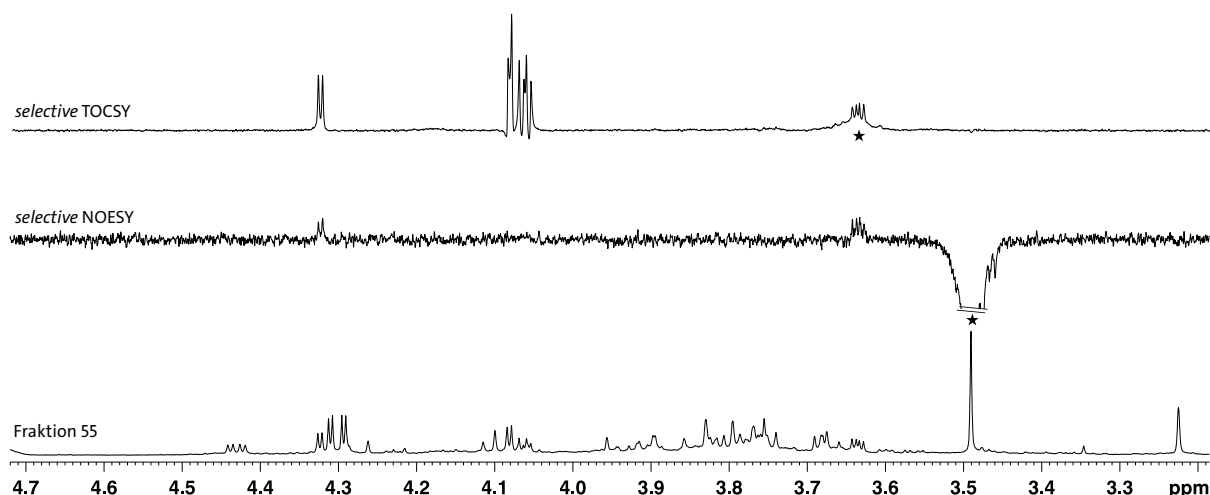
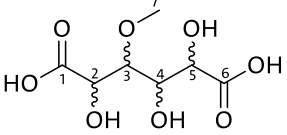


Abbildung 66: ^1H -NMR-Spektrum der Fraktion 55 (unten) sowie selektive NOESY- und TOCSY-Spektren. Es wurde jeweils das mit einem Sternchen markierte Signal angeregt.

Unter Berücksichtigung aller Informationen sowie des HSQC- und HMBC-Spektrums der Fraktion 55 kann darauf geschlossen werden, dass es sich um eine Aldarsäure handelt, die in Position 3 eine *O*-Methylgruppe beinhaltet. Die Zuordnung der ^1H - und ^{13}C -Verschiebungen ist in Tabelle 19 angegeben.

Tabelle 19: Zuordnung der ^1H - und ^{13}C -Verschiebungen der 3-*O*-Methylaldarsäure (**58**). Die ^{13}C -Verschiebungen wurden dem HSQC- und HMBC-Spektrum entnommen. Die HMBC-Kopplungen zu den quartären Kohlenstoffatomen können aufgrund von zu geringer Sensitivität nicht detektiert werden.

Metabolit	$\delta_{^{13}\text{C}}$ [ppm]	$\delta_{^1\text{H}}$ [ppm]	Multiplizität	J [Hz]	Zuordnung
 3- <i>O</i> -Methylaldarsäure (58)	85.7	3.64	dd	5.6, 3.1	3
	75.2	4.07	dd	5.6, 3.6	4
	75.2	4.08	d	2.8	5
	75.0	4.33	d	3.3	2
	61.3	3.49	s	-	7

Aufgrund der flexiblen Struktur des Metaboliten und damit verbunden der Vielzahl an möglichen Konformationen, kann auf Grundlage der ^1H -NMR-Spektren keine Information über die Konfiguration der vier Stereozentren gegeben werden. In einer Studie von CALLAHAN *et al.* konnte ebenfalls eine 3-*O*-Methylaldarsäure (**58**) als Ligand in einem Nickel-(II)-Komplex identifiziert werden.^[156] Der Nickelkomplex stammt aus dem Milchsaft des hyperakkumulierenden Nickelbaums *Sebertia acuminata*.^[156] Die in der Publikation

beschriebenen NMR-Daten stimmen mit den in Tabelle 19 angegebenen Werten weitestgehend überein, jedoch weichen alle ^1H -Verschiebungen um jeweils 0.13 ppm und alle ^{13}C -Verschiebungen um jeweils ca. 3 ppm voneinander ab.^[156] Die Abweichungen treten jedoch bei allen Signalen gleichmäßig auf. Die Diskrepanz kann zum einen möglicherweise durch den in der Literatur isolierten Nickel-(II)-Komplex bedingt sein.^[156] Zum anderen können die Abweichungen möglicherweise auf einen unterschiedlichen pH-Wert der Probe zurückgeführt werden. Es ist anzunehmen, dass die isolierte Substanz aufgrund der beiden Carboxylat-Gruppen sauer reagiert, wohingegen die Proben des Pinienkernextraktes in einem gepufferten System bei einem pH-Wert von 7.0 gemessen wurden.

5.2.1.9 Fazit zur Identifizierung von Metaboliten in Pinienkernen durch Anwendung von SCORE-metabolite-ID

Durch Anwendung der entwickelten Methode *SCORE-metabolite-ID* konnten im polaren Extrakt der Pinienkerne 45 Metaboliten verschiedener Substanzklassen identifiziert werden. Neben Kohlenhydraten, Aminosäuren, Betainen und anderen in biologischen Matrices häufig vorkommenden Metaboliten sind dabei insbesondere die sekundären Metaboliten β -D-Glucopyranosyl-1-*N*-indol-3-acetyl-*N*-L-asparaginsäure (**56**) und *N* ^{α} -(2-Hydroxy-2-carboxymethylsuccinyl)-L-arginin (**57**) sowie die 3-*O*-Methylaldarsäure (**58**) hervorzuheben, wobei im letzten Fall die absolute Konfiguration nicht eindeutig geklärt werden konnte. Die Anwendung von *SCORE-metabolite-ID* stellte dabei die entscheidende Grundlage zur Identifizierung der Metaboliten dar. Die Kombination der beiden fundamentalen Analysemethoden NMR-Spektroskopie und Massenspektrometrie und deren Datenkorrelation ermöglichte z.B. auf Grundlage der MS-Signale eine wahrscheinliche Summenformel zu postulieren oder auf Grundlage der NMR-Signale Strukturfragmente zu identifizieren. Je nach spezifischem Analyten und in Anbetracht der zur Verfügung stehenden Informationen (z.B. Vergleichsspektren aus Datenbanken, kommerzielle Verfügbarkeit für *Spike-In* Experimente, Literaturnachweise etc.) können die weiterführenden Experimente (z.B. selektive oder zweidimensionale NMR-Experimente oder LC-MS-Messungen) wesentlich gezielter geplant und eingesetzt werden, was den Aufwand der individuellen Identifizierung stark reduziert. Die Anwendung von *SCORE-metabolite-ID* eignet sich deshalb sowohl für die effiziente Dereplikation bereits bekannter Metaboliten als auch für die Identifizierung neuer und unbekannter Metaboliten.

5.2.2 Sortendifferenzierung von Apfelproben

Im Rahmen einer *untargeted* Authentizitätsstudie wurden 86 Apfelproben in Hinblick auf die Sortendifferenzierung untersucht. Diese Arbeiten wurden im Zuge einer Bachelorarbeit durchgeführt.^[157] Wie bereits unter Abschnitt 3.1 erläutert, können über *untargeted Metabolomics*-Studien durch die spektroskopischen *Fingerprints* der Proben und der anschließenden multivariaten Datenauswertung Unterschiede zwischen den Probengruppen untersucht und dadurch idealerweise Biomarker detektiert werden. Diese spezifischen Metaboliten können anschließend mit Hilfe der *SCORE-metabolite-ID* Methode identifiziert werden.

Der Apfel, *Malus domestica* BORKH., stellt insbesondere in Deutschland basierend auf dem Pro-Kopf-Konsum die beliebteste Obstsorte dar (siehe Abbildung 67 A).^[158] Nach Bananen und Wassermelonen belegen Äpfel mit einer weltweiten Erntemenge von 90 490 295 Tonnen im Jahr 2020 den dritten Platz bei Vergleich der weltweiten Produktionsmenge verschiedener Obstsorten (siehe Abbildung 67 B).^[159] Fruchtsäfte, insbesondere Orangen- und Apfelsaft, gehören zu den Lebensmitteln, die am häufigsten von Lebensmittelverfälschung und -betrug betroffen sind.^[160,161] Im Fall von Fruchtsäften findet eine Verfälschung meist durch Verdünnen des Saftes mit Wasser, Beimischung von Zucker oder anderen nicht-deklarierten Zusatzstoffen statt.^[160] Ebenso gilt die Beimischung von Apfelsaft einer beliebigen Sorte als Verfälschung, wenn der Apfelsaft bspw. als reiner Saft einer einzigen Sorte deklariert ist.

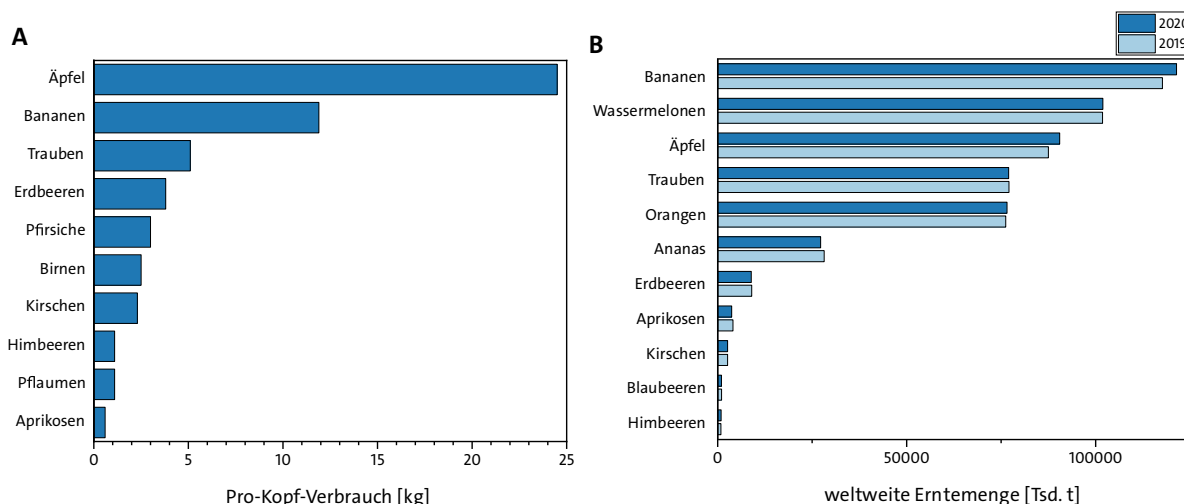


Abbildung 67: A: Pro-Kopf-Konsum von Obst in Deutschland im Jahr 2020/21.^[158] B: Weltweite Erntemenge in Tausend Tonnen verschiedener Obstsorten in den Jahren 2019 und 2020.^[159]

Äpfel bestehen zu einem Großteil aus Wasser (ca. 85 %) und Kohlenhydraten (ca. 14 %) inkl. Cellulose und Pektin als unlösliche und lösliche Ballaststoffe.^[162] Daneben beinhalten Äpfel u.a. Vitamine (z.B. Vitamin C und Vitamin E), organische Säuren, Mineralstoffe (z.B. Kalium) sowie Polyphenole.^[162,163] Insgesamt gelten Äpfel aufgrund der Inhaltsstoffe als gesundheitsfördernde Lebensmittel. So führen beispielsweise hohe Gehalte an Vitamin C, an Ballaststoffen sowie insbesondere ein hoher Polyphenol-Gehalt zu antioxidativen Eigenschaften.^[162] Phenolische Verbindungen werden mit positiven gesundheitlichen Auswirkungen in Verbindung gebracht, wie z.B. der Entzündungshemmung, der Verringerung des Risikos von Herz-Kreislauf-Erkrankungen, Lungenkrebs, Asthma und Diabetes.^[164,165]

5.2.2.1 Identifizierung von Metaboliten durch *Spike-In*-Experimente

Auf Grundlage der bereits zuvor im Extrakt der Pinienkerne identifizierten Metaboliten und durch Datenbankrecherchen wurden Strukturvorschläge von den in Apfelproben potenziell vorhandenen Metaboliten erstellt. Das Vorhandensein der Analyten wurde daraufhin mittels *Spike-In* Experimente anhand der exemplarischen Probe 20-B-SH-9 verifiziert, wodurch die Analyten als identifiziert betrachtet werden können. Zu den über *Spike-In* Experimente identifizierten Metaboliten zählen die in Abbildung 68 dargestellten Analyten Äpfelsäure (**51**), Zitronensäure (**52**), Chinasäure (**3**), Chlorogensäure (**59**), Ethanol (**60**), Essigsäure (**61**) und Ameisensäure (**62**), die Kohlenhydrate Glucose (**12**) und Saccharose (**2**) sowie die Aminosäuren L-Alanin (**10**), L-Valin (**22**), L-Asparagin (**31**), L-Glutaminsäure (**32**) und L-Asparaginsäure (**33**).

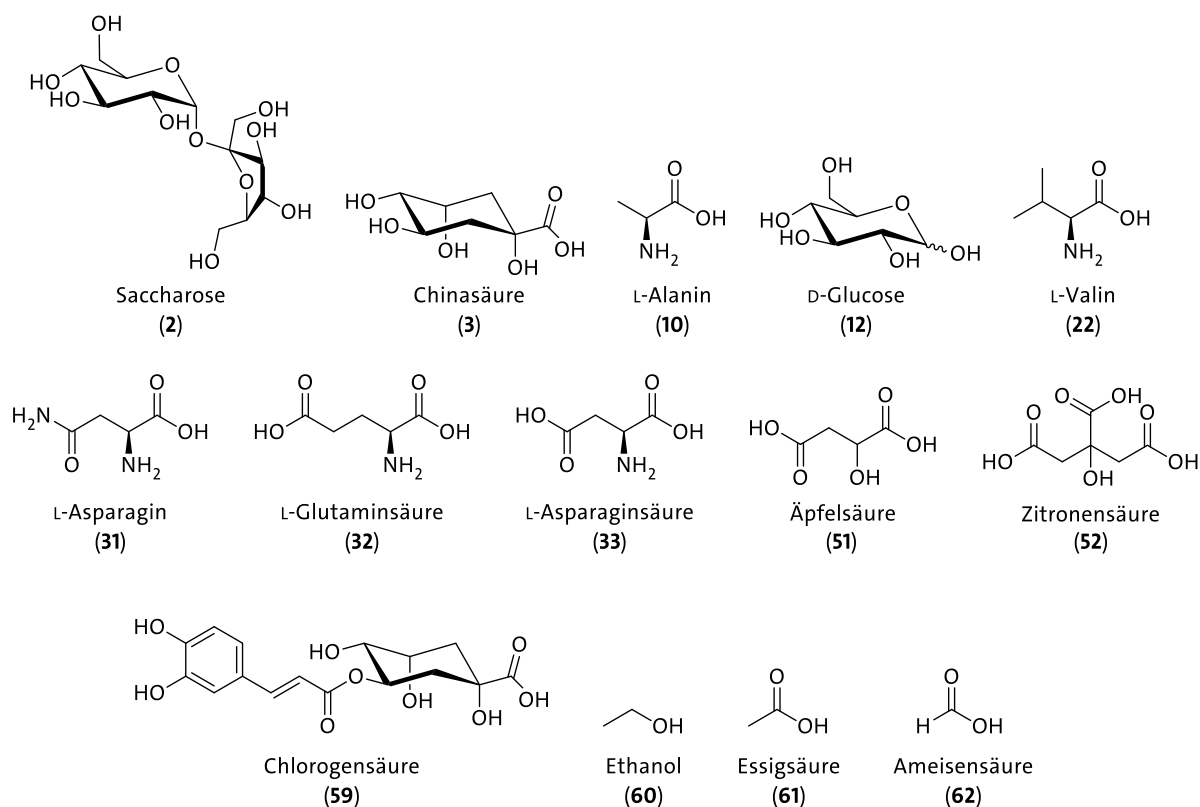


Abbildung 68: Übersicht der Analyten, die per *Spike-In* Experiment in der Apfelprobe identifiziert werden konnten.

In Abbildung 69 ist das ^1H -NMR-Spektrum der Probe 20-B-SH-9 gezeigt. Die durch *Spike-In* Experimente zugeordneten Signale sind entsprechend gekennzeichnet.

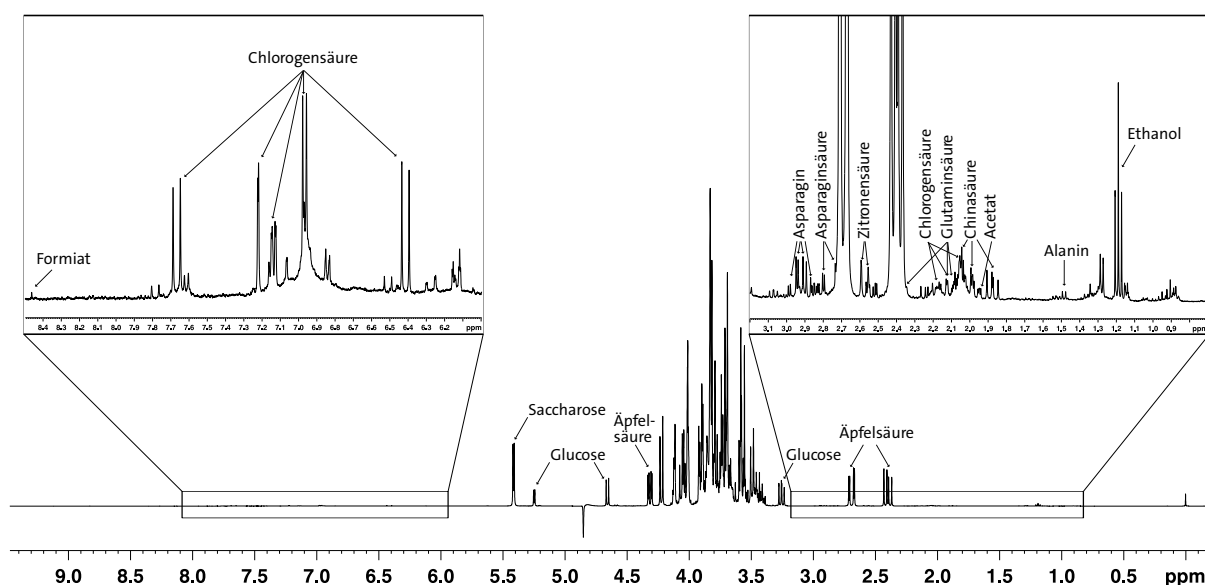


Abbildung 69: ^1H -NMR-Spektrum der Apfelprobe 20-B-SH-9 inkl. Beschriftung der durch *Spike-In* Experimente verifizierten Metaboliten Ethanol (60), Alanin (10), Acetat (61), Chinasäure (3), Glutaminsäure (32), Chlorogensäure (59), Äpfelsäure (51), Zitronensäure (52), Asparaginsäure (33), Asparagin (31), Glucose (12), Saccharose (2) und Formiat (62).

5.2.2.2 Stabilitätsmessung

Zunächst wurde die Stabilität der Proben untersucht. Biologische Proben natürlichen Ursprungs können sich bei längerer Lagerung verändern, was sich in den ^1H -NMR-Spektren, z.B. durch Signalverschiebungen oder Veränderungen in den Signalintensitäten, widerspiegeln könnte. Insbesondere bei *untargeted Metabolomics*-Studien sind derartige zeitliche Veränderungen der Proben problematisch. Zur Untersuchung der Stabilität wurde die exemplarisch ausgewählte Probe 20-H-SH-4 täglich innerhalb von einer Woche sowie erneut nach 14 Tagen gemessen. Die ^1H -NMR-Spektren der Messungen sind in Abbildung 70 gezeigt. Der zeitliche Verlauf von Tag 1 zu Tag 14 ist über den Farbverlauf der Spektren von schwarz zu lila dargestellt.

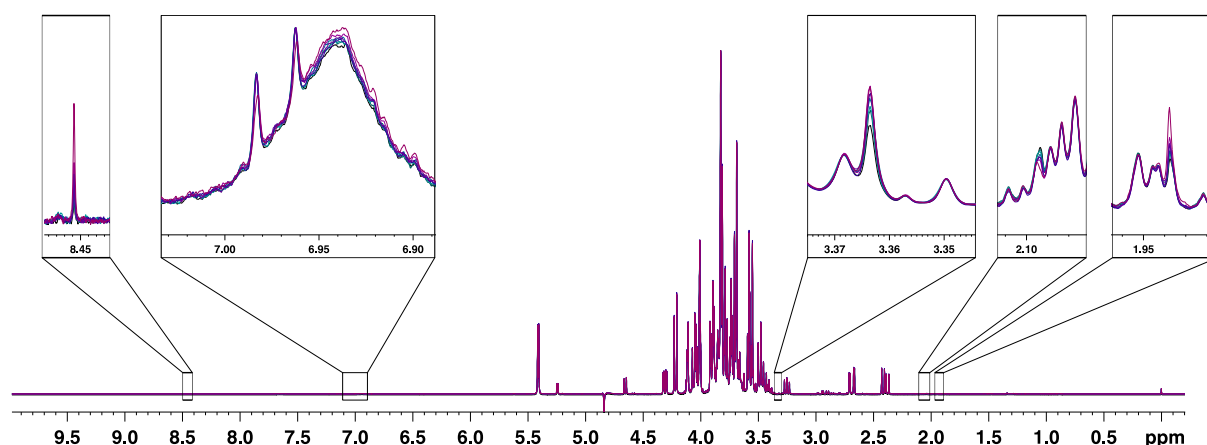


Abbildung 70: ^1H -NMR-Spektren der exemplarisch ausgewählten Apfelprobe 20-H-SH-4, die zur Untersuchung der Stabilität täglich innerhalb von sieben Tagen sowie erneut nach 14 Tagen gemessen wurde. Der zeitliche Verlauf ist über den Farbverlauf der Spektren von schwarz (Tag 1) zu lila (Tag 14) gekennzeichnet. Bereiche, in denen Veränderungen aufgetreten sind, sind vergrößert dargestellt.

Die Signale, die im zeitlichen Verlauf der Lagerung Veränderungen im Spektrum aufweisen, sind in Abbildung 70 vergrößert abgebildet. Insgesamt sind nur geringe Veränderungen detektierbar, wie z.B. die Zunahme des Signals des Acetats (**61**) bei einer chemischen Verschiebung von 1.94 ppm sowie des Formiats (**62**) bei 8.46 ppm. Weitere Veränderungen treten im Bereich von 2.09 ppm, 3.36 ppm und im aromatischen Bereich bei 6.9 bis 7.0 ppm auf. Die ersten geringfügigen Veränderungen, die beispielsweise auf enzymatische oder chemische Reaktionen zurückzuführen sein können, treten bereits nach einem Tag Lagerung bei Raumtemperatur auf. Die Messung der Proben erfolgte daher grundsätzlich innerhalb von 24 Stunden nach der Probenvorbereitung.

5.2.2.3 Differenzierung der fünf Apfelsorten

Alle 86 Proben wurden im Triplikat aufgearbeitet und anschließend jeweils in randomisierter Reihenfolge NMR-spektroskopisch durch Aufnahme eines 1D ^1H -NOESY NMR-Spektrums analysiert. Nach Prozessierung der NMR-Spektren erfolgte ein manuelles *Bucketing*. Dabei wurden 159 *Buckets* in Bereichen unterschiedlicher Größe, individuell je nach Signal, ausgewählt. Die Bereiche der *Buckets* sind im Anhang in Abschnitt 8.3.1 aufgeführt. Mit Hilfe der ausgewählten *Buckets* wurden zunächst Hauptkomponentenanalysen der Triplikat-Messungen jeder einzelnen Sorte durchgeführt, um Ausreißer zu ermitteln. In Abbildung 71 ist exemplarisch

der PCA Score Plot der Triplikatmessungen der Sorte Boskoop gezeigt. Die Score Plots der Triplikatmessungen der anderen vier Sorten sind im Anhang in Abbildung 133-136 aufgeführt.

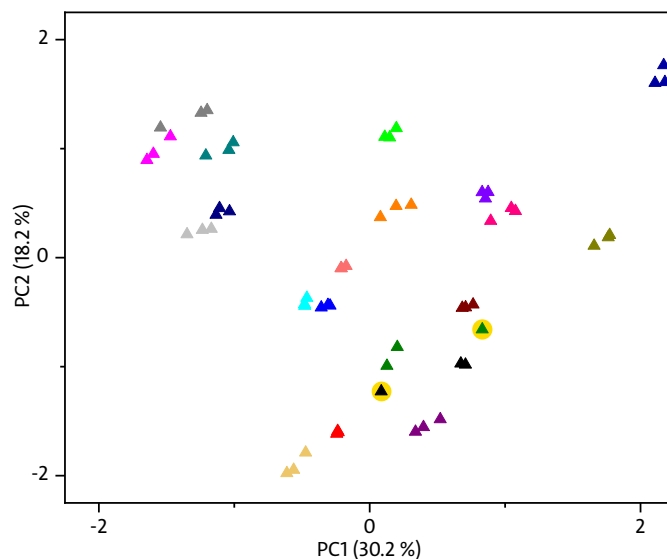


Abbildung 71: PCA Score Plot der im Triplikat aufgearbeiteten Proben der Sorte Boskoop zur Ermittlung von Ausreißern. Messungen derselben Probe sind jeweils in der gleichen Farbe dargestellt. Ausreißer, die aus folgenden Analysen eliminiert wurden, sind gelb markiert. In diesem Fall tritt dies auf jeweils eine Messung der Probe 20-B-BW-1 (schwarz) und 20-B-SH-8 (dunkelgrün) zu.

Im Fall der Boskoop-Proben wurden zwei Messungen als Ausreißer detektiert, die in der obigen Abbildung 71 gelb markiert sind. Diese wurden aus den folgenden Analysen eliminiert. Abweichungen in einzelnen Messungen der Proben können beispielsweise durch präparative Fehler während der Probenvorbereitung oder durch Anomalien, beispielsweise aufgrund von Temperaturschwankungen oder allgemeinen Instabilitäten des Spektrometers, während der Messung auftreten. Nach Exklusion der als Ausreißer detektierten Messungen wurden die Intensitäten der *Buckets* der übrigen Messungen derselben Proben gemittelt. In Abbildung 72 ist die dreidimensionale Darstellung des PCA Score Plots (PC1 gegen PC2 gegen PC3) der gemittelten 159 *Buckets* aller fünf Sorten dargestellt.

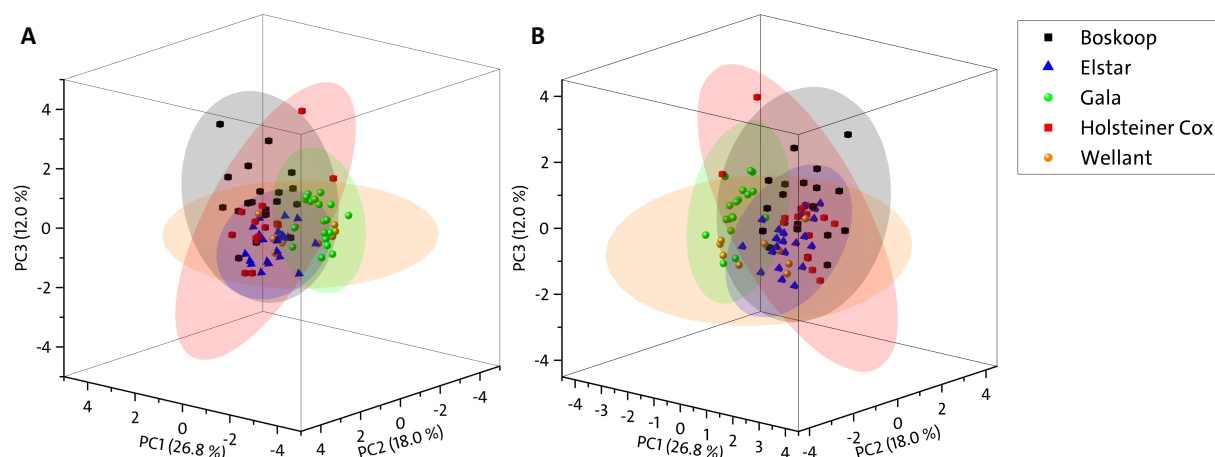


Abbildung 72: Dreidimensionale Darstellung des PCA Score Plot der entsprechend der Dreifachbestimmung gemittelten Variablen. Die Datenpunkte sind entsprechend der Apfelsorte gefärbt. Die Ellipsoiden stellen die 95 %-Konfidenzintervalle dar. Teil **B** zeigt die um 180° gedrehte Ansicht zur Teil **A**. Erklärte Varianz: PC1: 26.8 %, PC2: 18.0 %, PC3: 12.0 %.

Zur Erkennung von Clustern sind die Datenpunkte im *Score Plot* entsprechend der Sortenzugehörigkeit in unterschiedlichen Farben dargestellt. Der dreidimensionale Plot ist beispielhaft aus zwei Perspektiven (Abbildung 72 A und B) gezeigt. Insgesamt liegen die Datenpunkte der verschiedenen Cluster räumlich nur wenig separiert vor. Einzelne Cluster weisen jedoch bei Betrachtung des *Score Plots* aus verschiedenen weiteren Perspektiven eine deutlichere Separation auf. So ist z. B. erkennbar, dass die Cluster der Boskoop- und Gala-Proben entlang der ersten Hauptkomponente nur geringe Überschneidungen aufweisen. Auch die Cluster der Elstar- und Gala-Proben weisen z. B. durch die Kombination der ersten und zweiten Hauptkomponente eine deutlichere Separation auf. Auffällig ist außerdem, dass die Proben der Sorte Boskoop, Holsteiner Cox und Wellant größere Cluster bilden und damit eine größere Intragruppenvarianz aufweisen, wohingegen die Proben der Sorten Elstar und Gala näher beieinander liegen.

In Abbildung 73 ist die zweidimensionale Darstellung der *Score Plots* von PC1 gegen PC2 (A) und PC1 gegen PC3 (C) sowie die zugehörigen *Loading Plots* (B und D) gezeigt. In den *Loading Plots* sind Bereiche spezifischer Signale bzw. Signalgruppen farblich hervorgehoben, wie z. B. Bereiche von Kohlenhydrat-Signalen, aliphatischen und aromatischen Signalen. Außerdem sind die Signale der Äpfelsäure (51) grün hervorgehoben. Im *Loading Plot* der Auftragung PC1 gegen PC3 sind außerdem Signale, die hauptsächlich Glucose (12) zugeordnet werden können, violett schraffiert gekennzeichnet.

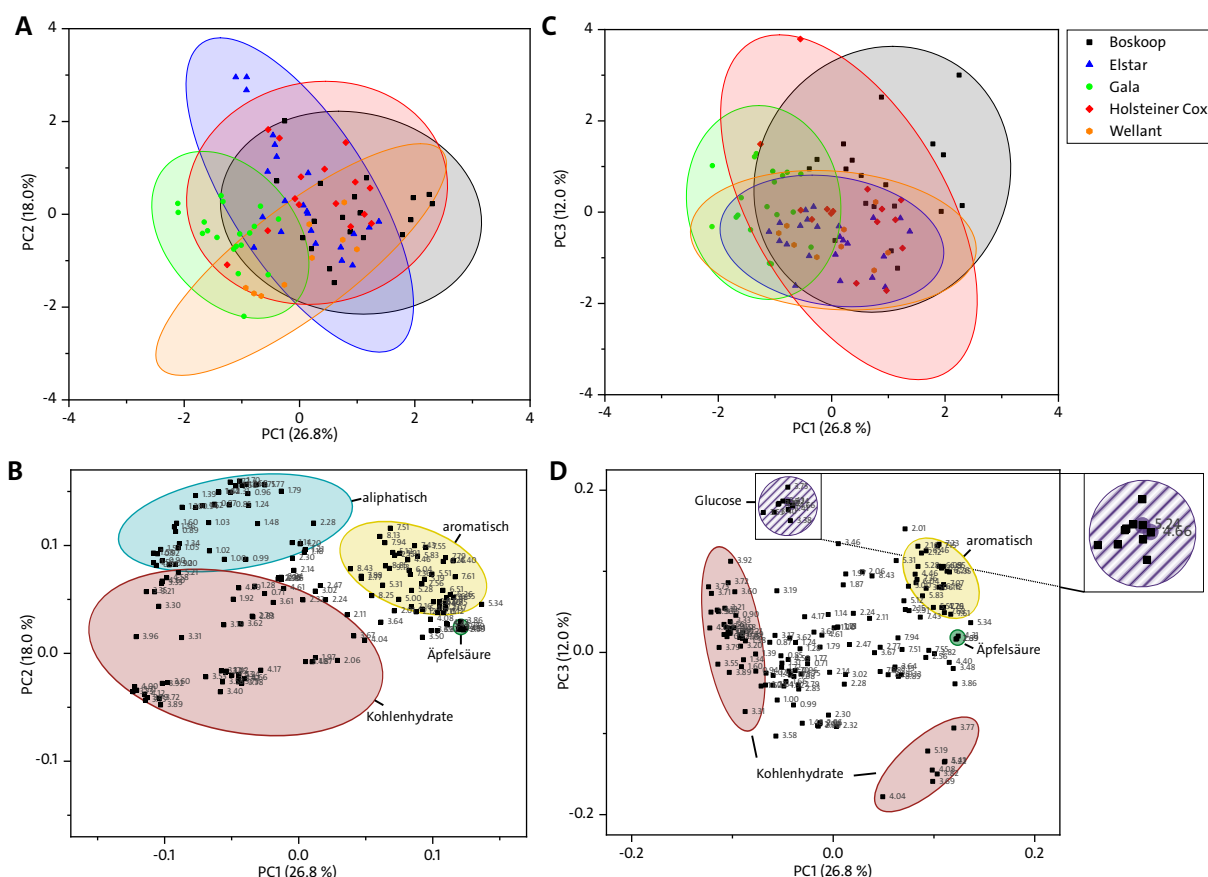


Abbildung 73: PCA Score und Loading Plots der gemittelten Buckets aller untersuchten Apfelproben. **A:** Score Plot mit PC1 gegen PC2. **B:** Zugehöriger Loading Plot mit PC1 gegen PC2. **C:** Score Plot mit PC1 gegen PC3. **D:** Loading Plot mit PC1 gegen PC3. Erklärte Varianz: PC1: 26.8 %, PC2: 18.0 %, PC3: 12.0 %. In den Score Plots sind die 95 %-Konfidenzintervalle jeder Apfelsorte ellipsenförmig gekennzeichnet. In den Loading Plots sind einzelne Signale bzw. Signalgruppen farblich hervorgehoben. Im Loading Plot D sind die Signale der anomeren Protonen von Glucose (12) violett hervorgehoben. Die Signale im violett schraffierten Bereich können zumindest anteilig ebenfalls Glucose (12) zugeordnet werden, jedoch sind die Signale teilweise überlagert.

Durch Vergleich der räumlichen Position der Datenpunkte im Score und Loading Plot können Aussagen über den Einfluss der Variablen im Loading Plot auf die Trennung der Proben im Score Plot getroffen werden. Aufgrund der mittenzentrierten Darstellung der Plots korrelieren positive Werte im Loading Plot bezüglich der gleichen Hauptkomponente mit positiven Werten im Score Plot. Liegen der Variablenwert (Loading Plot) und der Datenpunkt (Score Plot) also in der gleichen Richtung, ist dies gleichbedeutend mit einem überdurchschnittlichen Wert der Variable, d.h. einer höheren Signalintensität innerhalb des Buckets. Analog bedeutet ein Variablenwert im positiven Bereich des Loading Plots in Kombination mit einem Datenpunkt im negativen Bereich des Score Plots, dass eine unterdurchschnittliche Signalintensität innerhalb des Buckets dieses Datenpunktes vorliegt.

Im stark positiven Bereich der ersten Hauptkomponente im *Loading Plot B* sind die Signale von Äpfelsäure (51) zu erkennen, die demnach einen großen Einfluss auf die Trennung der Proben entlang dieser Komponente haben. Somit ist bereits zu erwarten, dass beispielsweise in Proben der Sorte Boskoop der Metabolit Äpfelsäure (51) vergleichsweise hohe Konzentrationen aufweist, wohingegen in den Proben der Sorte Gala geringere Konzentrationen an Äpfelsäure (51) vorliegen.

Nach der ungerichteten Analyse des Datensatzes mittels Hauptkomponentenanalyse wurde im nächsten Schritt ein Klassifizierungsmodell zur Sortenunterscheidung der Apfelproben mittels *Support Vector Machine* Algorithmus erstellt. Die Grundlage bilden die gemittelten Variablen der 159 manuell gewählten *Buckets*. Das Klassifizierungsmodell wurde mittels fünffacher *repeated nested cross-validation* (RNCV) validiert (siehe Abschnitt 3.2.3). Die Konfusionsmatrix des Fünfklassenmodells zur Klassifizierung der Apfelproben entsprechend der Sorten in Abbildung 74 zeigt die gemittelten Werte der RNCV, wobei die Klassenzuordnungen als absolute Werte (in Klammern) und die Genauigkeit in Prozent angegeben ist.

		Zugeordnete Klasse				
		Boskoop	Elstar	Gala	Holsteiner Cox	Wellant
Wahre Klasse	Boskoop	84.0 % (16.8)	7.0 % (1.4)	0.0 % (0)	9.0 % (1.8)	0.0 % (0)
	Elstar	0.9 % (0.2)	91.3 % (21)	0.0 % (0)	7.8 % (1.8)	0.0 % (0)
	Gala	0.0 % (0)	0.0 % (0)	94.7 % (18)	0.0 % (0)	5.3 % (1)
	Holsteiner Cox	4.0 % (0.6)	1.3 % (0.2)	4.0 % (0.6)	90.7 % (13.6)	0.0 % (0)
	Wellant	11.1 % (1)	11.1 % (1)	0.0 % (0)	0.0 % (0)	77.7 % (7)

Abbildung 74: Kombinierte gemittelte Konfusionsmatrix nach fünffacher *repeated nested cross-validation* bei Verwendung aller 159 manuell gewählten *Buckets*. Die Gesamtgenauigkeit des Fünfklassenmodells beträgt $87.7\% \pm 1.6\%$. Unter der relativen Angabe der jeweiligen Zuordnung zu einer Klasse in Prozent sind die absoluten Werte der Klassenzuordnung der Proben in Klammern angegeben.

Die Gesamtgenauigkeit des Modells beträgt $87.7\% \pm 1.6\%$. Die Proben der Sorte Gala werden zu 94.7 % korrekt zugeordnet. In einem ähnlichen Bereich liegen die Proben der Sorten Elstar (91.3 %) und Holsteiner Cox (90.7 %). Die Proben der Sorte Boskoop werden zu 84.0 % und die der Sorte Wellant zu 77.7 % korrekt zugeordnet. Die vergleichsweise geringe Klassifizierungsgenauigkeit der Wellant-Proben kann teilweise durch die geringere Probenzahl begründet werden. Gleichzeitig ist in der Hauptkomponentenanalyse in Abbildung 73 bereits erkennbar, dass die Proben der Sorte Wellant einen vergleichsweise großen Cluster bilden und damit eine größere Intergruppenvarianz aufweisen. Die geringe Klassifizierungsgenauigkeit ist im Speziellen auf die zwei Proben 20-W-SH-7 und 21-W-BW-9 zurückzuführen, die in jeder Iteration der RNCV als Boskoop bzw. Elstar missklassifiziert wurden. In Abbildung 75 sind alle ^1H -NMR-Spektren aller Sorten im Triplikate gezeigt und entsprechend ihrer Sortenzugehörigkeit eingefärbt. Im Vergleich fällt auf, dass das Signal bei einer chemischen Verschiebung von 1.14 ppm bei allen drei Messungen der Probe 20-W-SH-7 deutlich höhere Signalintensitäten aufweist.

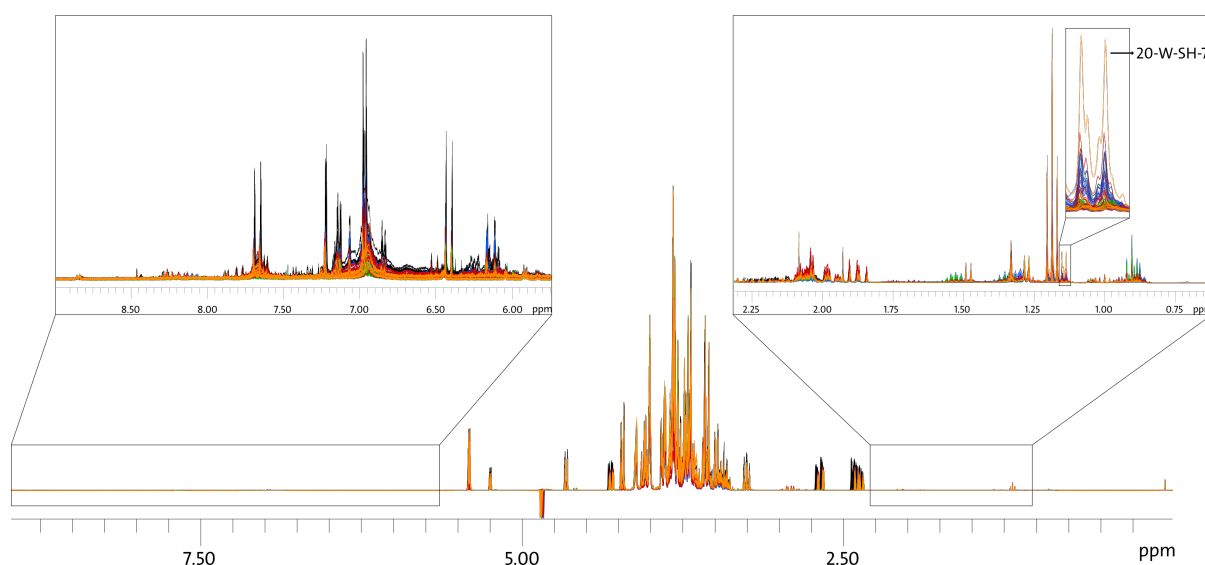


Abbildung 75: ^1H -NMR-Spektren aller 86 Apfel-Proben als Triplikate. Signale im aliphatischen und aromatischen Bereich sind vergrößert dargestellt. Die Spektren sind entsprechend der fünf Sorten eingefärbt, d. h. Boskoop schwarz, Elstar blau, Gala grün, Holsteiner Cox rot und Wellant orange.

Der in Abbildung 76 dargestellte Boxplot des gemittelten *Buckets* bei 1.14 ppm zeigt ebenfalls einen Extremwert (nicht-gefüllte Raute) bei den Proben der Sorte Wellant. Dieser Datenpunkt im Boxplot entspricht dem gemittelten Wert des *Buckets* der Probe 20-W-SH-7 und verdeutlicht den starken Unterschied dieser Einzelprobe.

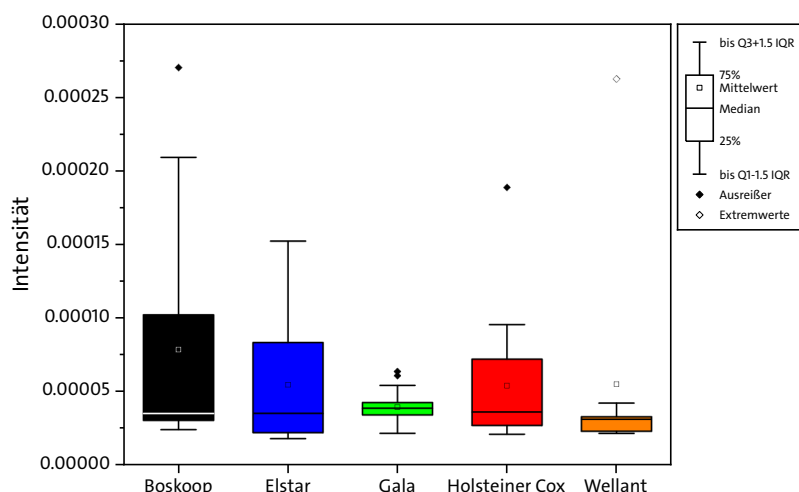


Abbildung 76: Boxplots des *Buckets* bei einer chemischen Verschiebung von 1.14 ppm zur Unterscheidung der Apfelproben unterschiedlicher Sorten. Der Boxplot der Wellant-Proben zeigt einen extremen Ausreißer (nicht gefüllte Raute), der der gemittelten Triplikatmessung der Probe 20-W-SH-7 entspricht.

Solch hohe Intensitäten dieses *Buckets* sind nur bei einer anderen Probe der Sorte Boskoop (21-B-BW-17) zu sehen; die Proben der Sorten Elstar, Gala und Holsteiner Cox weisen deutlich geringere Intensitäten auf. Die Missklassifizierung der Wellant-Probe 20-W-SH-7 als Boskoop-Probe kann deshalb zumindest anteilig auf dieses Signal zurückgeführt werden. Die Identifizierung dieses Signals ist im folgenden Abschnitt 5.2.2.3.1 erläutert.

5.2.2.3.1 Identifizierung von 2,3-Butandiol

Wie in Abbildung 75 zu sehen, zeigt das ^1H -NMR-Spektrum ein Signal höherer Ordnung bei einer geringen chemischen Verschiebung von 1.14 ppm. Die charakteristische Multiplizität dieses Signals lässt zunächst vermuten, dass es sich möglicherweise um ein Molekül mit Symmetriezentrum handelt. Die chemische Verschiebung des Signals deutet außerdem auf ein Molekül mit Hydroxyl-Substituenten hin. Infolgedessen wird zunächst der Strukturvorschlag 2,3-Butandiol (**64**) postuliert. Unter Ausschluss von Sauerstoff kann die 2,3-Butandiol-Gärung stattfinden, bei der unter anaeroben Bedingungen aus Glucose (**12**) das charakteristische Produkt 2,3-Butandiol (**64**) gebildet wird.^[166] Dieser Fermentationsprozess tritt bei den fakultativ anaeroben *Enterobacteriaceae* beispielsweise der Gattung *Enterobacter* oder *Klebsiella* auf.^[166] Dabei wird, wie im Stoffwechselweg in Abbildung 77 gezeigt, ausgehend von zwei Molekülen Pyruvat (**60**) durch die Acetolactat-Synthase (1) unter Abspaltung von Kohlenstoffdioxid (S)-2-Acetolactat (**61**) gebildet. Dieses kann durch die Acetolactat-Decarboxylase (2) zu

(*R*)-Acetoin (*R*-**63**) decarboxyliert werden. Alternativ kann über die Decarboxylierung von (*S*)-2-Acetylactat (**61**) das Zwischenprodukt Diacetyl (**62**) gebildet werden, das wiederum zu (*S*)-Acetoin (*S*-**63**) oder (*R*)-Acetoin (*R*-**63**) reduziert werden kann. Über die Acetoin-Racemase (5) stehen beide Konfigurationen im Gleichgewicht zueinander. Acetoin (**63**) kann daraufhin durch die Butandiol-Dehydrogenase (6) mit NADH zu 2,3-Butandiol (**64**) reduziert werden.^[166–168]

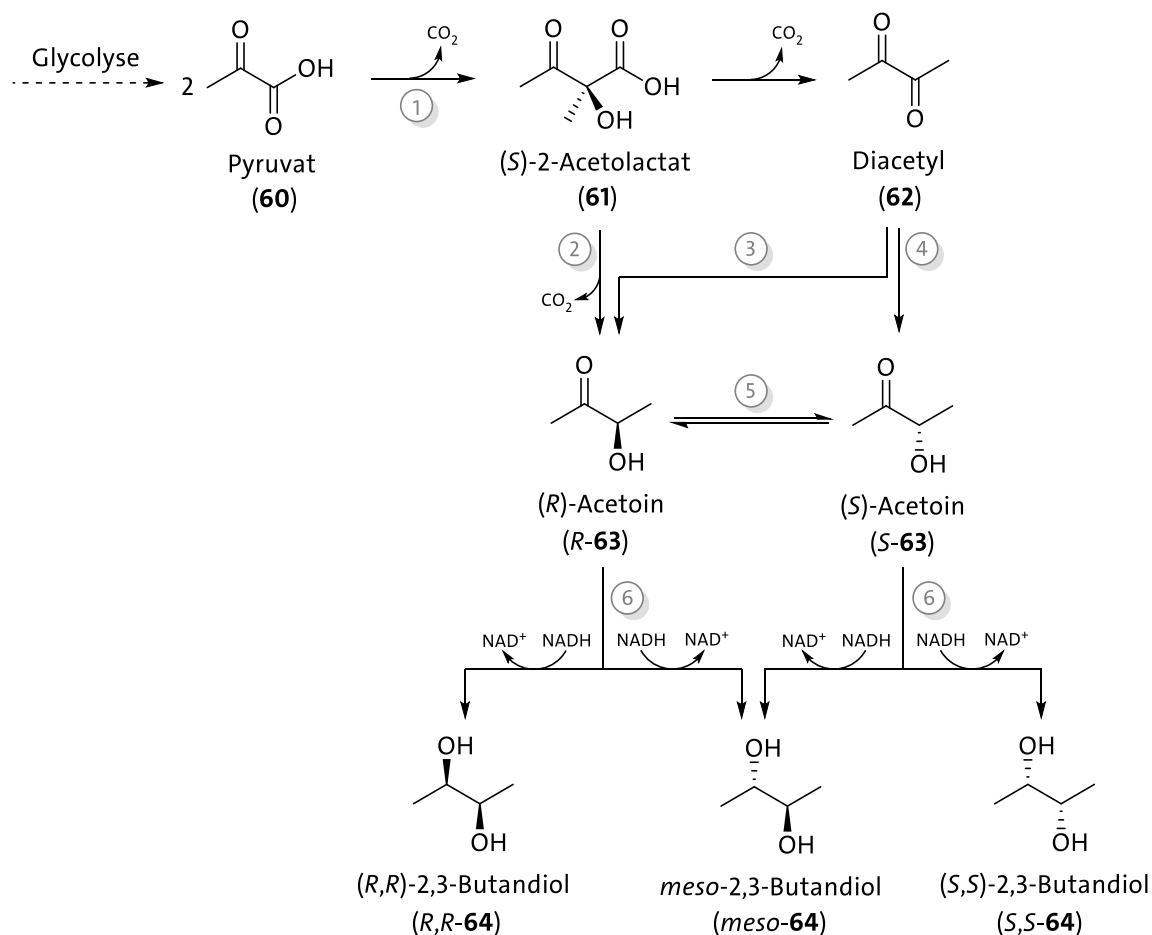


Abbildung 77: Biochemischer Stoffwechselweg der Butandiol-Gärung in Bakterien, bei der aus Pyruvat (**60**) der charakteristische Metabolit 2,3-Butandiol (**64**) gebildet wird. Die an dem Stoffwechselweg beteiligten Enzyme sind durch die grauen Kreise gekennzeichnet: 1: Acetylactat-Synthase, 2: Acetylactat-Decarboxylase, 3: Diacetyl-Reductase [(*R*)-Acetoin-bildend], 4: Diacetyl-Reductase [(*S*)-Acetoin-bildend], 5: Acetoin-Racemase, 6: Butandiol-Dehydrogenase.^[166–168]

Acetoin (**63**) kann aufgrund des Stereozentrums und den verschiedenen Stoffwechselwegen sowohl in der (*S*)- als auch in der (*R*)-Konfiguration vorliegen. Durch die folgende Reduktion der Butandiol-Dehydrogenase (6) können somit drei Stereoisomere des 2,3-Butandiols (**64**) erhalten werden: (*R,R*)-2,3-Butandiol (*R,R*-**64**), (*S,S*)-2,3-Butandiol (*S,S*-**64**) und *meso*-2,3-Butandiol (*meso*-**64**). Butandiol-Dehydrogenasen (6) der unterschiedlichen Gattungen der

Enterobacteriaceae weisen unterschiedliche Selektivitäten bzgl. der Konfiguration auf. So erfolgt die Reduktion der Butandiol-Dehydrogenasen (6) der *Klebsiella pneumoniae*, der *Serratia marcescens* und der *Enterobacter cloacae* spezifisch in der (S)-Konfiguration, sodass selektiv (R)-Acetoin (**R-63**) zu *meso*-2,3-Butandiol (*meso*-**64**) und (S)-Acetoin (**S-63**) zu (S,S)-2,3-Butandiol (**S,S-64**) reduziert wird.^[169,170] Die Butandiol-Dehydrogenasen (6) der *Saccharomyces cerevisiae* und der *Paenibacillus polymyxa* arbeiten dahingegen (R)-spezifisch, wodurch (R)-Acetoin (**R-63**) selektiv zu (R,R)-2,3-Butandiol (**R,R-64**) und (S)-Acetoin (**S-63**) selektiv zu *meso*-2,3-Butandiol (*meso*-**64**) reduziert wird.^[168,171,172] In vielen Fällen wird ein Gemisch der verschiedenen Stereoisomere des 2,3-Butandiols (**64**) produziert.^[168]

Die Verifizierung des postulierten Strukturvorschlags des Metaboliten erfolgt durch ein *Spike-In* Experiment. Zunächst wurde ein Gemisch des (R,R)-, (S,S)- und *meso*-2,3-Butandiols (**64**) zur Probe 20-W-SH-7 zugegeben. Da die Signale im Bereich der Kohlenhydrat-Signale und dementsprechend im erwarteten Bereich des Signals der Methin-Gruppe des 2,3-Butandiols (**64**) stark überlagert sind, wurde ein 2D-TOCSY-Experiment zur Überprüfung eingesetzt (siehe Abbildung 78). In schwarz sind die 2D-TOCSY-Signale der Probe vor der Zugabe und in rot nach der Zugabe der Referenzlösung gezeigt. Hierbei fällt auf, dass ein zusätzliches Korrelationssignal bei 1.14/3.73 ppm nach Zugabe des Gemisches der 2,3-Butandiol-Stereoisomere (**64**) auftritt, welches in der Probe vor Zugabe der Referenzlösung nicht auftritt.

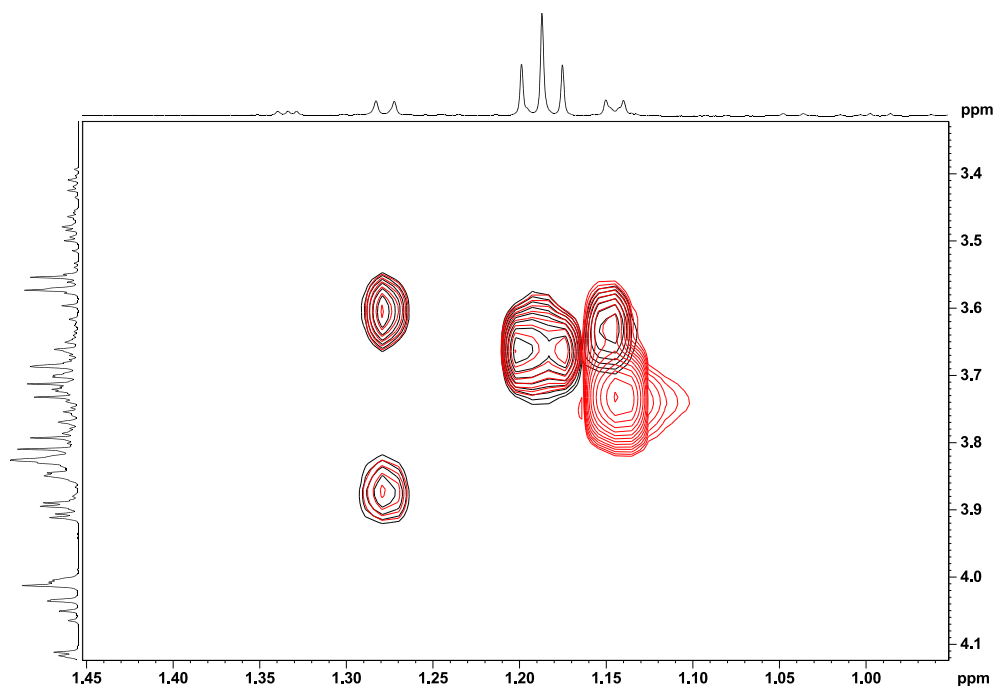


Abbildung 78: Überlagerte Darstellung der 2D-TOCSY-NMR-Spektren der Probe 20-W-SH-7. In schwarz ist ein Ausschnitt der Signale vor der Zugabe eines Gemisches der Stereoisomere des 2,3-Butandiols (**64**) gezeigt. In rot ist das Spektrum nach Zugabe von 0.1 μ L der Referenzlösung gezeigt. Nach Zugabe der Referenzlösung wird ein zusätzliches Korrelationssignal bei 1.14/3.73 ppm detektiert. Es wurde jeweils ein TOCSY-Experiment mit dem Pulsprogramm *dipsi2esgpph* mit 12 Scans und einer Wasserunterdrückung bei 4.707 ppm an einem BRUKER Avance III HD 600 MHz Spektrometer aufgenommen.

Im nächsten Schritt wurde zu der Probe 20-W-SH-7 das reine Stereoisomer (*R,R*)-2,3-Butandiol (*R,R*-**64**) hinzugegeben und es kann, wie in Abbildung 79 gezeigt ist, ein Anstieg des ^1H -NMR-Signals bei 1.14 ppm beobachtet werden. Im 2D-TOCSY-Experiment wird kein zusätzliches Korrelationssignal detektiert, sondern es wird lediglich ein Anstieg des Korrelationssignals bei 1.14/3.63 ppm detektiert.

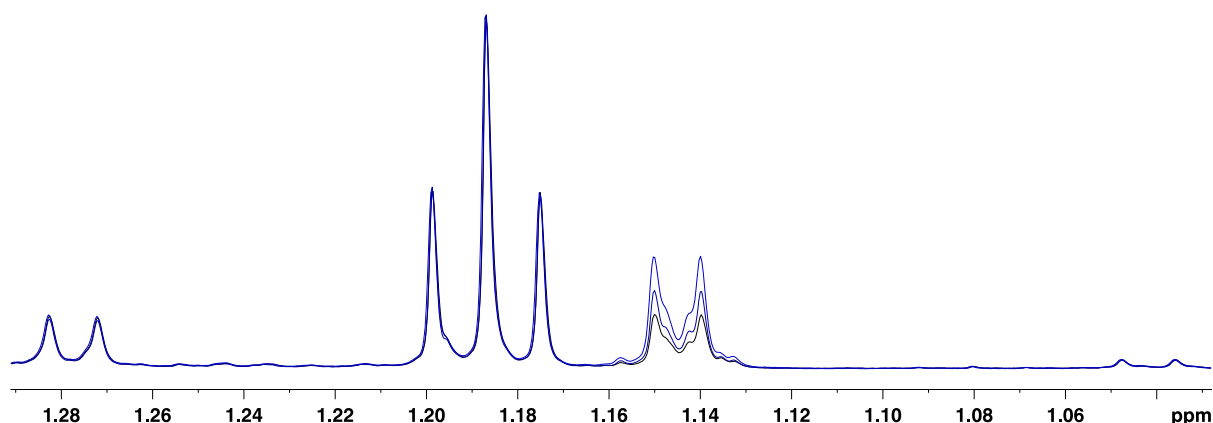
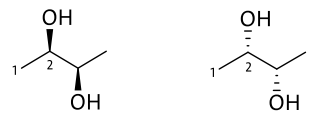
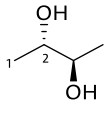


Abbildung 79: ^1H -NMR-Spektrum der Probe 20-W-SH-7 vor und nach Zugabe von (*R,R*)-2,3-Butandiol (*R,R*-**64**). In schwarz ist das Spektrum der ursprünglichen Probe gezeigt. In blau ist das Spektrum derselben Probe nach Zugabe von jeweils 0.5 μL (*R,R*)-2,3-Butandiol (*R,R*-**64**) (0.1 M in D_2O) gezeigt. Es wurde die *noesygppr1d*-Pulssequenz mit einer Wasserunterdrückung bei 4.706 ppm verwendet und jeweils 32 *Scans* an einem BRUKER Avance III HD 600 MHz Spektrometer aufgenommen.

Da die Enantiomere des 2,3-Butandiols (**64**), d.h. (*R,R*)- und (*S,S*)-2,3-Butandiol, NMR-spektroskopisch nicht unterschieden werden können, lässt sich nicht zuordnen, welches der beiden Enantiomere in der Apfelprobe vorliegt. Das Stereoisomer *meso*-2,3-Butandiol (*meso*-**64**) lässt sich allerdings aufgrund der zuvor gezeigten Ergebnisse als Metabolit in den untersuchten Apfelproben ausschließen. Die Probe 20-W-SH-7 enthält dementsprechend überdurchschnittlich hohe Konzentrationen des Metaboliten (*R,R*)- oder (*S,S*)-2,3-Butandiol (*R,R*-**64** oder *S,S*-**64**), der mit hoher Wahrscheinlichkeit für die Missklassifizierung dieser Probe als Boskoop-Probe verantwortlich ist. Der Nachweis dieses Metaboliten könnte möglicherweise ein Hinweis auf mikrobielle Belastung und beginnende Fermentation dieser Probe sein. In Tabelle 20 sind die ^1H - und ^{13}C -NMR-Daten der 2,3-Butandiol-Stereoisomere (**64**) aufgeführt.

Tabelle 20: Zuordnung der NMR-Signale der Enantiomere (*R,R*)- und (*S,S*)-2,3-Butandiol (*R,R*-**64** oder *S,S*-**64**) sowie des Isomers *meso*-2,3-Butandiols (*meso*-**64**).

δ_{H} [ppm]	Multiplizität	J [Hz]	δ_{C} [ppm]	Zuordnung	Molekül
3.63	m	-	74.4	2	 <i>(R,R)</i> -2,3-Butandiol (<i>R,R</i> - 64) <i>(S,S)</i> -2,3-Butandiol (<i>S,S</i> - 64)
1.14	m	-	20.6	1	

3.73	m	-	73.9	2	 <i>meso</i> -2,3-Butandiol (<i>meso</i> - 64)
1.14	m	-	19.6	1	

5.2.2.3.2 Variablenselektion durch *Boruta*

In der bisherigen Betrachtung ist das Variablen-zu-Proben Verhältnis mit 159 *Buckets* zu 86 Proben relativ hoch. Um ein *Overfitting* des Modells zu vermeiden, wurde deshalb eine Variablenselektion mittels *Boruta* durchgeführt. Dies dient außerdem der einfacheren Detektion von Markermetaboliten, die eine Unterscheidung einzelner Apfelsorten zulassen. Durch *Boruta* konnten 82 Variablen selektiert werden, die im Anhang unter Abschnitt 8.3.2 aufgeführt sind. Zur Überprüfung der Selektion wurde zunächst erneut eine Hauptkomponentenanalyse durchgeführt. Trotz Selektion und Reduktion der Variablen, ist zu erwarten, dass der *Score* und *Loading* Plot grundsätzlich ähnliche Ergebnisse liefern. In Abbildung 80 sind die *Score* und *Loading* Plots auf Grundlage der 82 selektierten *Buckets* gezeigt.

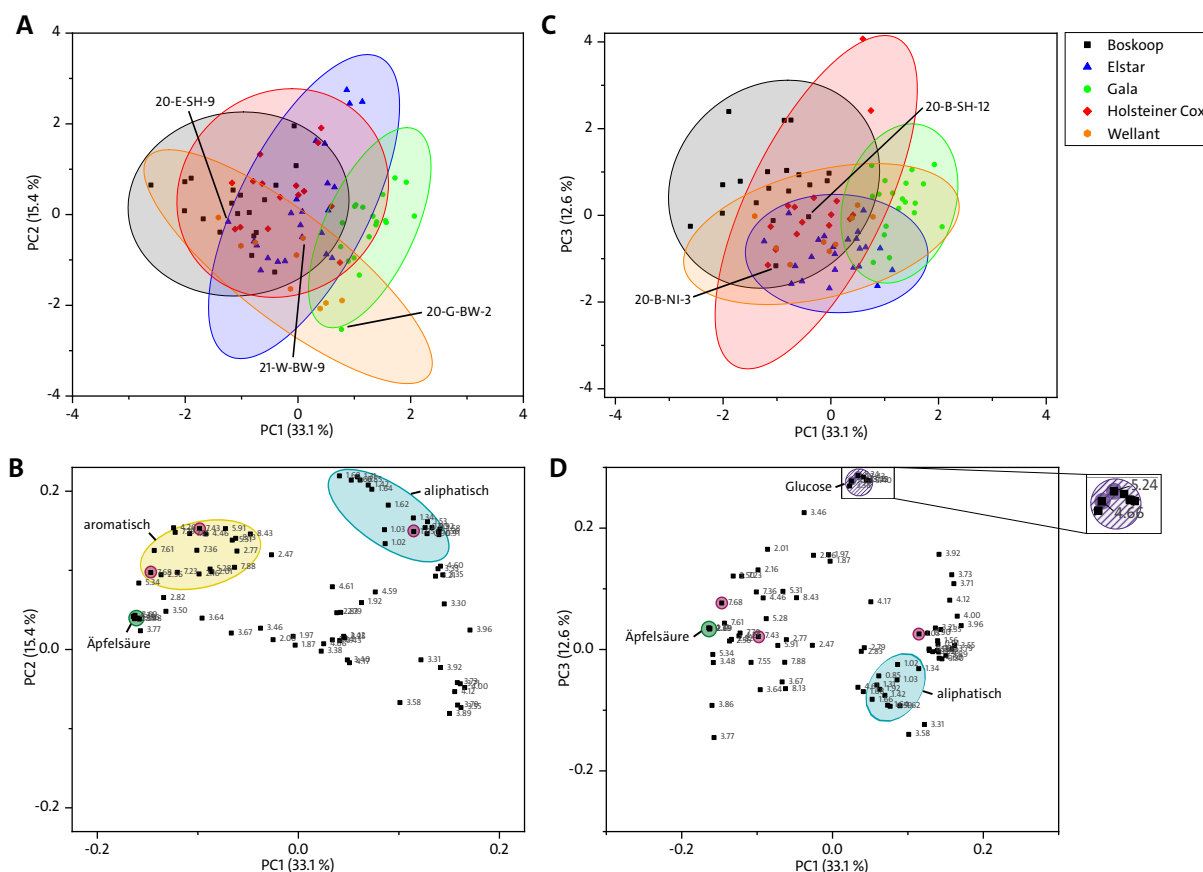


Abbildung 80: PCA Score und Loading Plots der gemittelten Buckets aller untersuchten Apfelproben nach Variablenselektion. Die Liste der selektierten Buckets kann dem Anhang (Abschnitt 8.3.2) entnommen werden. **A:** Score Plot mit PC1 gegen PC2. **B:** Zugehöriger Loading Plot mit PC1 gegen PC2. **C:** Score Plot mit PC1 gegen PC3. **D:** Loading Plot mit PC1 gegen PC3. Erklärte Varianz: PC1: 33.1 %, PC2: 15.4 %, PC3: 12.6 %. In den Score Plots sind die 95 %-Konfidenzintervalle jeder Apfelsorte ellipsenförmig dargestellt und einzelne Datenpunkte gekennzeichnet. In den Loading Plots sind einzelne Signale bzw. Signalgruppen farblich hervorgehoben. Zu den magentafarbenen hervorgehobenen Signalen sind die sortenspezifischen Boxplots im Folgenden gezeigt. Im Loading Plot D sind die Signale der anomeren Protonen von Glucose (**12**) violett hervorgehoben sowie weitere, möglicherweise ebenfalls zu Glucose (**12**) zugehörige Signale violett schraffiert gekennzeichnet. Diese Signale sind teilweise überlagert.

Im Allgemeinen ist ein ähnliches Ergebnis der Hauptkomponentenanalyse basierend auf den 82 selektierten Buckets (Abbildung 80) im Vergleich zur Hauptkomponentenanalyse basierend auf allen 159 Buckets (Abbildung 73) zu erkennen. Dementsprechend kann geschlussfolgert werden, dass die eliminierten Variablen nur einen untergeordneten Einfluss auf die Differenzierung der Apfelproben haben. Zusätzlich konnte durch die Fokussierung auf die wesentlichen Variablen die erklärte Varianz bis zur 3. Hauptkomponente von zuvor 56.8 % auf nun 61.1 % gesteigert werden.

Im Vergleich zu den *Score Plots* der Hauptkomponentenanalyse basierend auf allen 159 *Buckets* (siehe Abbildung 73), ist in den *Score Plots* A und C in Abbildung 80 zudem teilweise eine deutlichere Separation der Cluster erkennbar. Beispielsweise überlagern die Proben der Sorte Gala im *Score Plot* mit PC1 gegen PC2 etwas weniger mit den Proben der übrigen Sorten, v. a. mit den Proben der Sorte Boskoop. Insgesamt liegen die Datenpunkte innerhalb der Cluster der jeweiligen Sorte etwas näher beieinander. Der Einfluss der einzelnen Variablen in den *Loading Plots* ist ebenfalls ähnlich zu den Ergebnissen in Abbildung 73 unter Verwendung aller 159 *Buckets*. Äpfelsäure (**51**) hat einen großen Einfluss auf die Varianz der Hauptkomponente 1 und Glucose (**12**) auf die Varianz der Hauptkomponente 3. Außerdem deutet der *Loading Plot* B darauf hin, dass Proben der Sorte Boskoop höhere Konzentrationen an Metaboliten aufweisen, die Signale im aromatischen Bereich des Spektrums verursachen.

Auf Grundlage der selektierten 82 *Buckets* wurde ebenfalls ein Klassifizierungsmodell erstellt. In Abbildung 81 ist die kombinierte Konfusionsmatrix nach fünffacher RNCV dargestellt.

		Zugeordnete Klasse				
		Boskoop	Elstar	Gala	Holsteiner Cox	Wellant
Wahre Klasse	Boskoop	83.0 % (16.6)	7.0 % (1.4)	0.0 % (0)	10.0 % (2)	0.0 % (0)
	Elstar	0.9 % (0.2)	91.3 % (21)	0.0 % (0)	7.8 % (1.8)	0.0 % (0)
	Gala	0.0 % (0)	0.0 % (0)	94.7 % (18)	0.0 % (0)	5.3 % (1)
	Holsteiner Cox	1.3 % (0.2)	4.0 % (0.6)	1.3 % (0.2)	92.0 % (13.8)	1.3 % (0.2)
	Wellant	0.0 % (0)	8.9 % (0.8)	0.0 % (0)	0.0 % (0)	91.1 % (8.2)

Abbildung 81: Kombinierte gemittelte Konfusionsmatrix nach fünffacher *repeated nested cross-validation* zur Differenzierung der Apfelsorten unter Verwendung der durch *Boruta* selektierten 82 *Buckets*. Die Gesamtgenauigkeit des Modells beträgt 90.4 % $\pm$ 1.8 %. Unter der relativen Angabe der jeweiligen Zuordnung zu einer Klasse in Prozent sind die absoluten Werte der Klassenzuordnung der Proben in Klammern angegeben.

Die Gesamtgenauigkeit des Modells beträgt $90.4 \% \pm 1.8 \%$ und ist somit etwas höher als in dem Klassifizierungsmodell auf Grundlage aller 159 *Buckets*. Die Proben der Sorte Gala werden zu 94.7 % korrekt zugeordnet. In einem ähnlichen Bereich liegen die Proben der Sorten Elstar (91.3 %), Holsteiner Cox (92.0 %) und Wellant (91.1 %). Die Proben der Sorte Boskoop werden zu 83.0 % korrekt zugeordnet. Im Vergleich zum Klassifizierungsmodell basierend auf den unselektierten 159 *Buckets* fällt v.a. die hier deutlich höhere Klassifizierungsgenauigkeit der Wellant-Proben von 91.1 % (im Vergleich zu 77.7 % auf Basis aller 159 *Buckets*) auf. Die zuvor in allen Fällen als Boskoop-Probe missklassifizierte Wellant-Probe 20-W-SH-7 wird in dem Klassifizierungsmodell der selektierten *Buckets* nun immer korrekt zugeordnet. Das *Bucket*, das das Signal von (R,R)- oder (S,S)-2,3-Butandiol (R,R-**64** oder S,S-**64**) bei 1.14 ppm beinhaltet, wurde durch *Boruta* als nicht relevant eingestuft und eliminiert, weshalb diese Missklassifizierung im Klassifizierungsmodell der selektierten *Buckets* nicht mehr auftritt.

In den *Score Plots* in Abbildung 80 sind einzelne Datenpunkte mit der genauen Probenbezeichnung gekennzeichnet. Der Datenpunkt der Elstar-Probe 20-E-SH-9 liegt beispielsweise deutlich am Rand des 95 %-Konfidenzintervalls. Diese Probe wird im Klassifizierungsmodell in allen Fällen der Sorte Holsteiner Cox zugeordnet. Die Wellant-Probe 21-W-BW-9 wird in vier von fünf Fällen als Elstar-Probe zugeordnet. Im *Score Plot A* in Abbildung 80 liegt der Datenpunkt dieser Probe an der gleichen Position bzgl. PC1 und PC2 wie der einer Elstar-Probe. Im *Score Plot C* sind die beiden Proben der Sorte Boskoop 20-B-NI-3 sowie 20-B-SH-12 hervorgehoben, die im Klassifizierungsmodell in allen Fällen als Proben der Sorte Elstar und Holsteiner Cox missklassifiziert werden. Beide Proben liegen zwar deutlich im 95 %-Konfidenzintervall des Clusters der Boskoop-Proben, doch insbesondere diese Proben überlagern stark mit den Clustern der Sorten Elstar und Holsteiner Cox. Die Probe der Sorte Gala 20-G-BW-2 wird im obigen Klassifizierungsmodell in allen Fällen als Wellant-Probe klassifiziert. Im *Score Plot A* der Auftragung PC1 gegen PC2 liegt der Datenpunkt dieser Probe außerhalb des 95 %-Konfidenzintervalls des Clusters der Gala-Proben und überlagert mit dem Cluster der Wellant-Proben. Im *Score Plot* ist ebenfalls eine sorteninterne Auftrennung innerhalb des Clusters der Datenpunkte der Gala-Proben entlang der PC2 zu erkennen. Dies kann unter anderem durch die im *Loading Plot B* hervorgehobenen Signale des aliphatischen und aromatischen Bereichs begründet werden. Die Variablen mit stark positiven Werten bzgl. PC2 stehen in reziproker Korrelation zu der Gala-Probe 20-G-BW-2, die bei stark negativen Werten bzgl. PC2 im *Score Plot* liegt. Ein Beispiel hierfür ist das Signal des *Buckets* bei 7.43 ppm. Die

Signalintensitäten innerhalb dieses *Buckets* der einzelnen Sorten sind in Abbildung 82 in Form eines Boxplots dargestellt.

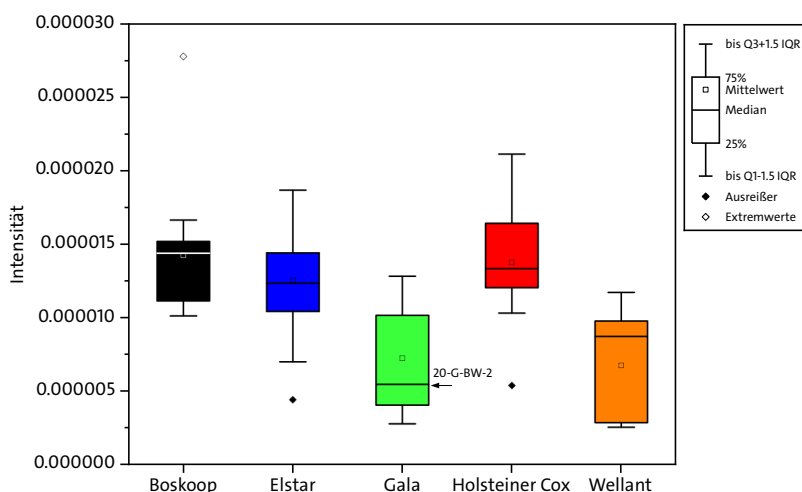


Abbildung 82: Boxplots des *Buckets* bei einer chemischen Verschiebung von 7.43 ppm zur Unterscheidung der Apfelproben unterschiedlicher Sorten. Die Position der Probe 20-G-BW-2 im Boxplot der Gala-Proben ist mit einem Pfeil gekennzeichnet. Die Identifizierung des Metaboliten, der das NMR-Signal bei 7.43 ppm hervorruft, ist unter Abschnitt 5.2.3.1 beschrieben.

Es fällt auf, dass die Proben der Sorten Boskoop, Elstar und Holsteiner Cox in diesem *Bucket* höhere Signalintensitäten aufweisen, als die Proben der Sorten Gala und Wellant. Die Boxplots der Gala- und Wellant-Proben liegen außerdem in einem sehr ähnlichen Bereich. Die Position der Probe 20-G-BW-2 innerhalb des Boxplots der Gala-Proben ist mit einem Pfeil gekennzeichnet und liegt ebenfalls auf einer Höhe mit der Box des Boxplots der Wellant-Proben, als die sie im Klassifizierungsmodell in allen Fällen missklassifiziert wird. In Abschnitt 5.2.3.1 wird die Identifizierung des Metaboliten beschrieben, welcher das betreffende Signal bei 7.43 ppm hervorruft.

Auch das *Bucket* bei einer chemischen Verschiebung von 1.05 ppm bei positiven Werten bzgl. PC2 im *Loading Plot* in Abbildung 80 B steht in Antikorrelation zu der Probe 20-G-BW-2 bei stark negativen Werten bzgl. PC2 im *Score Plot*. Wie in Abbildung 83 anhand von Boxplots gezeigt ist, weist die Probe 20-G-BW-2 in diesem *Bucket* unterdurchschnittliche Signalintensitäten im Vergleich zu den anderen Gala-Proben auf. Im Vergleich zu den anderen Apfelsorten weisen jedoch Gala-Proben sowie Proben der Sorte Holsteiner Cox die höchsten Signalintensitäten innerhalb dieses *Buckets* auf. Die Identifizierung des zugrunde liegenden Metaboliten ist unter Abschnitt 5.2.3.2 beschrieben.

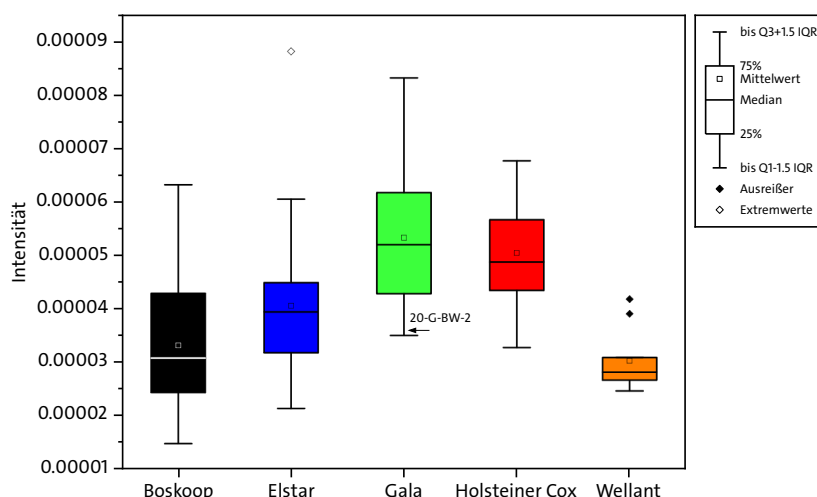


Abbildung 83: Boxplots des *Buckets* bei einer chemischen Verschiebung von 1.05 ppm der fünf untersuchten Apfelsorten. Die Position der Gala-Probe 20-G-BW-2 ist mit einem Pfeil im entsprechenden Boxplot gekennzeichnet. Die Identifizierung des Metaboliten, der das NMR-Signal bei 1.05 ppm in den Apfelproben hervorruft, ist unter Abschnitt 5.2.3.2 beschrieben.

5.2.2.4 Differenzierung der Apfelsorten Boskoop und Gala

Wie bereits anhand des *Score Plots* aller 86 Apfelproben zu sehen ist, bilden die Proben der Apfelsorten Boskoop und Gala weitestgehend separierte Cluster (vgl. Abbildung 80). Die Differenzierung dieser beiden Apfelsorten untereinander wurde deshalb erneut untersucht. In Abbildung 84 ist der *Score* und *Loading Plot* der Hauptkomponentenanalyse bei Verwendung der 20 Proben der Sorte Boskoop (schwarze Datenpunkte) und 19 Proben der Sorte Gala (grüne Datenpunkte).

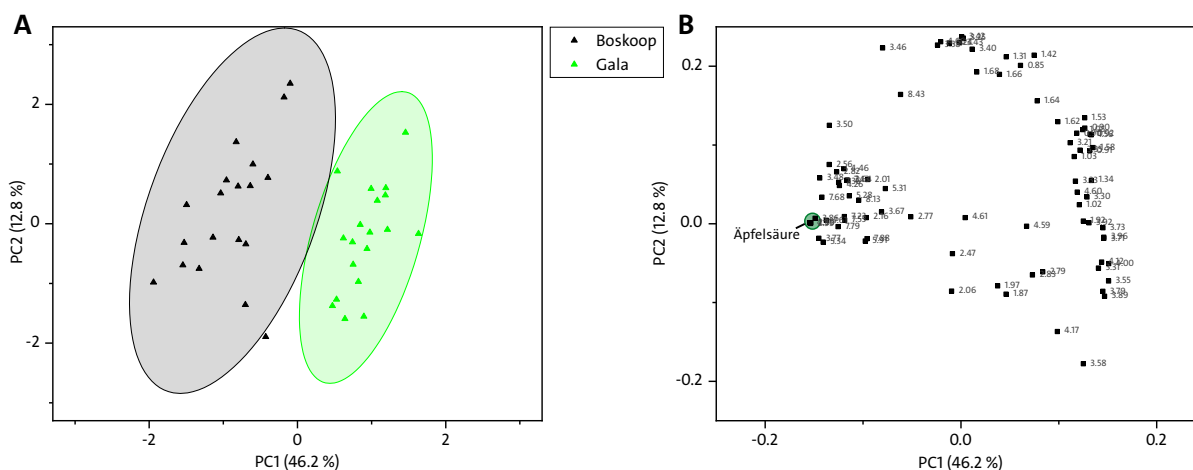


Abbildung 84: PCA Score (A) und Loading (B) Plots der Apfelsorten Boskoop und Gala. Es wurde der gemittelte *Buckettable* nach Variablenselektion verwendet (82 *Buckets*). Die Liste der *Buckets* ist im Anhang unter Abschnitt 8.3.2 aufgeführt. Erklärte Varianz: PC1: 46.2 %, PC2: 12.8 %. Im Score Plot sind die 95 %-Konfidenzintervalle ellipsenförmig gekennzeichnet. Im Loadings Plot sind Signale des Metaboliten Äpfelsäure (51) hervorgehoben.

Die Datenpunkte der verschiedenen Apfelsorten Boskoop und Gala bilden im Score Plot (Abbildung 84 A) deutlich separierte Cluster, die durch das ellipsenförmige 95 %-Konfidenzintervall gekennzeichnet sind. Die Separation der Cluster erfolgt entlang der ersten Hauptkomponente, die eine erklärte Varianz von 46.2 % aufweist. Im Loading Plot in Abbildung 84 B sind die *Buckets*, die Signale des Metaboliten Äpfelsäure (51) beinhalten, hervorgehoben. Diese *Buckets* befinden sich bei stark negativen Werten der ersten Hauptkomponente und haben dementsprechend einen großen Einfluss auf die Separation der Daten entlang dieser Hauptkomponente. Diese Signale bei stark negativen Werten im Loadings Plot korrelieren stark mit den Datenpunkten der Boskoop-Proben, die im Score Plot ebenfalls bei stark negativen Werten bezüglich der ersten Hauptkomponente auftreten. Die Proben der Sorte Boskoop weisen demzufolge höhere Konzentrationen Äpfelsäure (51) auf, als die Proben der Sorte Gala. Dies zeigen auch die in Abbildung 85 dargestellten Boxplots eines Signals der Äpfelsäure (51) bei einer chemischen Verschiebung von 4.32 ppm.

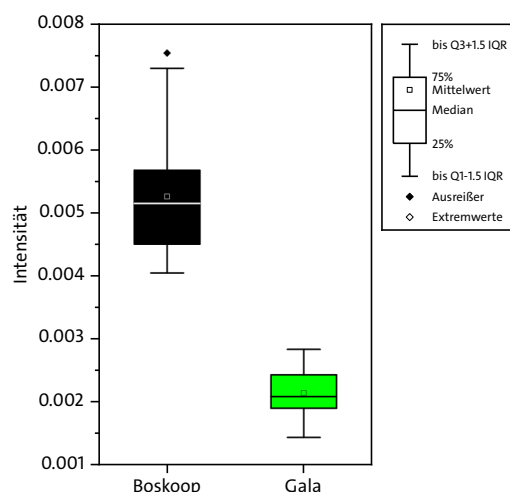


Abbildung 85: Boxplots des *Buckets* bei einer chemischen Verschiebung von 4.32 ppm, dessen Signal dem Metaboliten Äpfelsäure (51) zuzuordnen ist.

Bei Vergleich der Boxplots dieses *Buckets* der Proben Gruppen Boskoop und Gala fällt auf, dass diese keinerlei Überschneidung aufweisen. Deshalb liegt die Vermutung nahe, dass sich diese beiden Proben Gruppen in einem Zweiklassenmodell einzig auf Basis des Gehaltes von Äpfelsäure (51) unterscheiden lassen. Es wurde deshalb ein Klassifizierungsmodell zur Differenzierung der beiden Sorten in Form einer univariaten Analyse lediglich auf Grundlage des *Buckets* bei einer chemischen Verschiebung von 4.32 ppm erstellt. Die gemittelte Konfusionsmatrix nach fünffacher RNCV ist in Abbildung 86 dargestellt.

		Zugeordnete Klasse	
		Boskoop	Gala
Wahre Klasse	Boskoop	100.0 % (20)	0.0 % (0)
	Gala	0.0 % (0)	100.0 % (19)

Abbildung 86: Kombinierte gemittelte Konfusionsmatrix nach fünffacher *repeated nested cross-validation* zur Unterscheidung der Sorten Boskoop und Gala. Zur Klassifizierung der Proben wurde lediglich das *Bucket* bei einer chemischen Verschiebung von 4.32 ppm, das das Signal des Metaboliten Äpfelsäure (51) beinhaltet, verwendet. Die Genauigkeit des Modells beträgt 100.0 % $\pm$ 0.0 %. Unter der relativen Angabe der jeweiligen Zuordnung zu einer Klasse in Prozent sind die absoluten Werte der Klassenzuordnung der Proben in Klammern angegeben.

Erwartungsgemäß hat das Klassifizierungsmodell eine Gesamtgenauigkeit von $100.0 \% \pm 0.0 \%$. Durch die univariate Analyse eines Signals des Metaboliten Äpfelsäure (51) ist für diesen Datensatz demnach eine eindeutige Unterscheidung der Sorten Boskoop und Gala möglich.

5.2.3 Identifizierung von potenziellen Markermetaboliten in Apfelproben

Zur Identifizierung von möglichen Markermetaboliten in Äpfeln, die im Rahmen der multivariaten Datenanalyse als relevant zur Unterscheidung bestimmter Sorten erachtet werden, wurde exemplarisch eine Probe der Sorte Elstar (Probe 20-E-SH-9) ausgewählt. Nach Fraktionierung unter HILIC-Bedingungen wurden 80 Fraktionen erhalten, von denen jeweils ein ^1H -NMR-Spektrum und ein DI-MS-Spektrum aufgenommen wurden. Zur Identifizierung von Metaboliten wurden spezifische Signale innerhalb dieses Datensatzes daraufhin mit Hilfe der *SCORE-metabolite-ID* Methode analysiert. Zur Analyse der NMR-Daten wurde innerhalb der App ein *Bucketing* mit einer Größe der *Buckets* von 0.005 ppm verwendet.

5.2.3.1 Identifizierung von Triandrin

Wie bereits unter Abschnitt 5.2.2.3.2 erläutert und in Abbildung 82 anhand der Boxplots gezeigt, weisen die Proben der Sorten Gala und Wellant deutlich geringere Intensitäten innerhalb des *Buckets* bei einer chemischen Verschiebung von 7.43 ppm auf als die Proben der Sorten Boskoop, Elstar und Holsteiner Cox. Durch die semi-automatische Korrelation des EDCs bei einer chemischen Verschiebung von 7.439 ppm über die Fraktionen 1 bis 30 werden einige Signale im aromatischen Bereich sowie Signale im Bereich von 4.4 bis 4.6 ppm als hoch-korrelierende EDCs detektiert. In Abbildung 87 sind diese Signale in der Fraktion 28, in der der EDC bei 7.439 ppm das Maximum aufweist, farblich hinterlegt abgebildet.

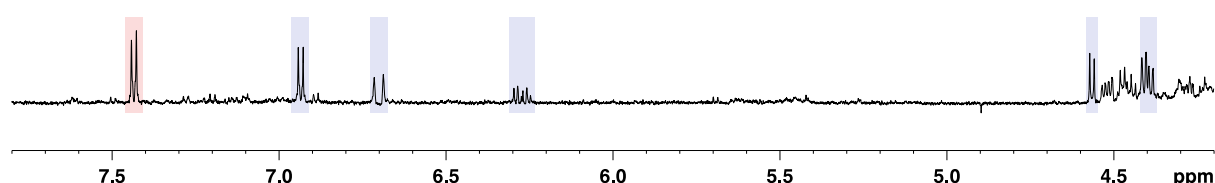


Abbildung 87: Ausschnitt des ^1H -NMR-Spektrums der Fraktion 28 der Apfelprobe 20-E-SH-9. Durch die semi-automatische Korrelation des EDCs bei 7.493 ppm (rot markiertes Signal) innerhalb der *SCORE-metabolite-ID* App wurden die blau markierten NMR-Signale als EDCs mit einem Korrelationskoeffizienten > 0.95 gefunden.

Die Analyse der ^1H -NMR-Signale zwischen 6.2 und 7.5 ppm lässt einen ersten Rückschluss auf ein Strukturfragment des Metaboliten zu. So lassen die beiden charakteristischen Multipletts bei 7.42-7.45 ppm und 6.92-6.95 ppm zunächst auf einen *para*-substituierten Aromaten schließen. Das Dublett bei 6.70 ppm mit einer großen Kopplungskonstante von 16.4 Hz deutet auf eine *trans*-konfigurierte Doppelbindung hin. Die starke Tieffeldverschiebung dieses Signals lässt vermuten, dass die Doppelbindung in Konjugation zum Aromaten steht. Das vermeintliche Dublett eines Tripletts bei 6.27 ppm (später identifiziert als Dublett eines Dubletts eines Dubletts) weist ebenfalls eine Kopplungskonstante von 16.4 Hz auf. Die weitere Kopplung in Form eines Tripletts (später identifiziert als Dublett eines Dubletts) lässt auf eine benachbarte Methylen-Gruppe schließen. Das postulierte Strukturfragment, wie es beispielsweise in (*E*)-Cumarylalkohol (**65**) vorkommt, ist in Abbildung 88 dargestellt.

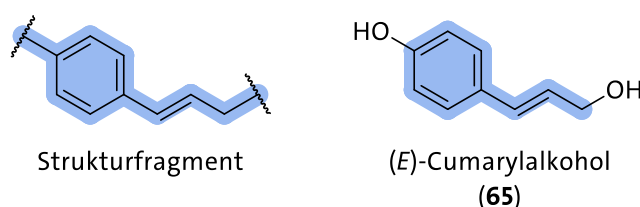


Abbildung 88: Postuliertes Strukturfragment (links) aus den hoch-korrelierenden ^1H -NMR-Signalen der Fraktion 28 (vgl. Abbildung 87). Dieses Strukturfragment kommt beispielsweise natürlich in (*E*)-Cumarylalkohol (**65**) (rechts) vor.

Das Dublett bei einer chemischen Verschiebung von 4.56 ppm ($J = 8.0$ Hz) könnte vom anomeren Proton eines Kohlenhydrates herrühren. Die Korrelation von EMCs zu dem EDC bei 7.439 ppm ergibt eine Reihe von m/z -Werten mit einem Korrelationskoeffizienten > 0.95 . Die zehn intensivsten EMCs sind im Anhang in Tabelle 43 aufgeführt. Zusätzlich sind die Signale graphisch in Abbildung 89 abgebildet.

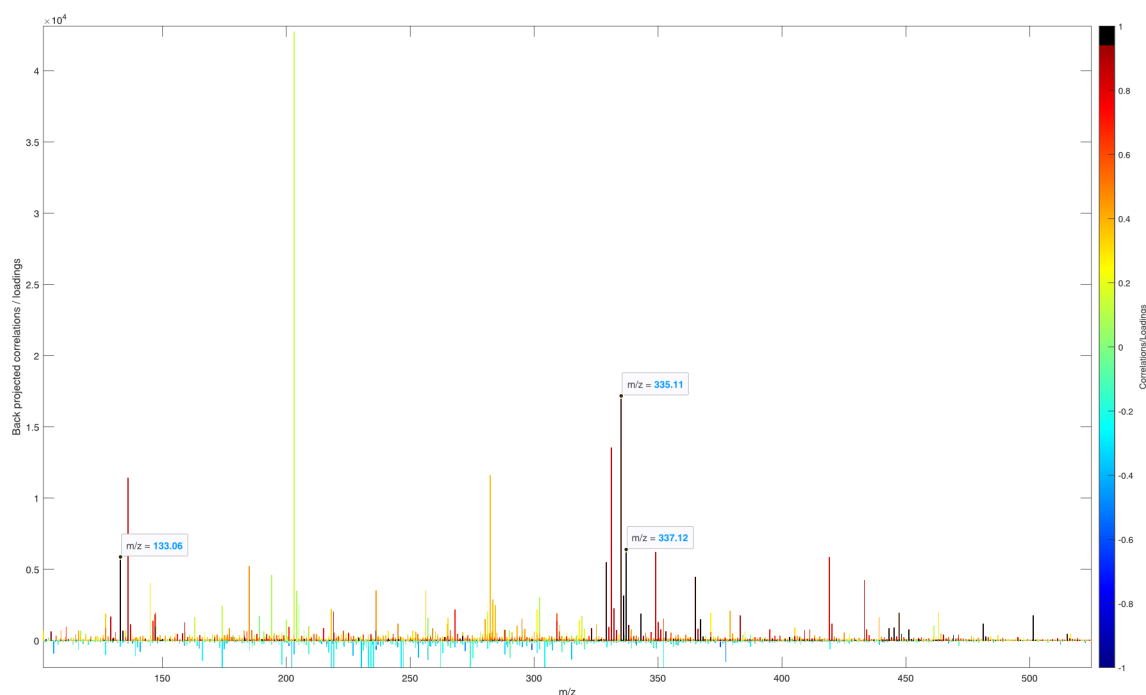


Abbildung 89: Graphische Darstellung der EMCs, die einen Korrelationskoeffizienten > 0.95 zu dem EDC bei einer chemischen Verschiebung von 7.439 ppm, berechnet über die Fraktionen 1-30, aufweisen.

Das intensivste Signal mit einem Korrelationskoeffizienten > 0.95 liegt bei einem m/z -Wert von 335.11. Auffällig ist dabei allerdings das vermeintliche Isotopenmuster dieses Signals. Unter der Annahme, dass der m/z -Wert von 335.11 dem Quasimolekülion M^+ entspricht, entspräche demnach das Signal bei $m/z = 337.12$ dem $[M+2]^+$ -Signal. Die Intensität dieses Signals entspricht etwa einem Drittel des Quasimolekülion-Signals, wohingegen die Intensität des $[M+1]^+$ -Signals nur ca. 18 % der Intensität des Signals des Quasimolekülions aufweist. Diese Isotopenverteilung ist für natürlich-vorkommende Metaboliten terrestrischen Ursprungs untypisch. Im Gegensatz zu Marineorganismen enthalten natürlich vorkommende Metaboliten der terrestrischen Pflanzen im Regelfall keine Halogene wie Chlor, das solch ein Isotopenmuster hervorrufen würde. Aus diesem Grund wurde die Fraktion 28 in einem weiteren Schritt mittels LC-MS analysiert. Die relevanten EICs sind in Abbildung 90 dargestellt.

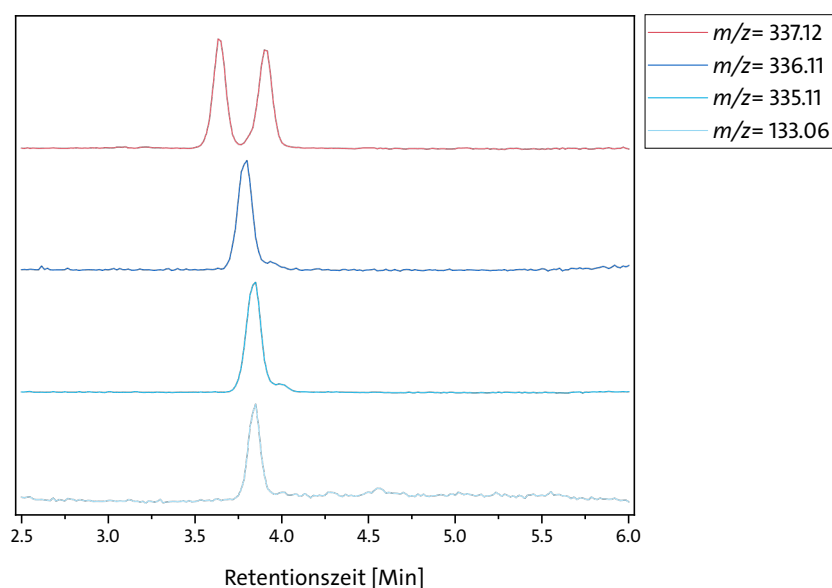


Abbildung 90: EICs der LC-MS-Messung der in der Legende angegebenen Masse-zu-Ladungs-Verhältnisse von der Fraktion 28 der Apfelprobe. Es ist der Ausschnitt im Retentionszeitfenster zwischen 2.5 und 6.0 Minuten gezeigt.

Anhand der EICs der m/z -Werte von 133.06, 335.11, 336.11 und 337.12 in Abbildung 90 zeigt sich, dass lediglich die Signale mit den Masse-zu-Ladungs-Verhältnissen von 133.06, 335.11 und 336.11 aufgrund desselben Verlaufes zu einem Metaboliten gehören. Der EIC von $m/z = 337.12$ zeigt zwei Maxima, die zu leicht anderen Retentionszeiten auftreten. Trotz des hohen Korrelationskoeffizienten dieses Signals, der innerhalb der *SCORE-metabolite-ID* App auf Grundlage der DI-Massenspektren ermittelt wird, lässt sich somit über das LC-MS-Chromatogramm der spezifischen EICs die Zusammengehörigkeit dieses Signals bei $m/z = 337.12$ zu den übrigen Signalen bei $m/z = 133.06$, 335.11 und 336.11 ausschließen. Somit wird postuliert, dass das Signal bei $m/z = 335.11$ dem Natrium-Addukt des Metaboliten entspricht und das Signal bei 336.11 dem ^{13}C -Isotopenpeak des Natrium-Adduktes. Über die Intensitäten der beiden Signale kann in erster Näherung eine grobe Abschätzung bzgl. der Anzahl der Kohlenstoffatome entsprechend Gleichung (9) getroffen werden.^[105,173] Zur Berechnung wurden die Intensitätswerte der beiden Signale innerhalb eines Massenspektrums der Fraktion 28 herangezogen.

$$N_{\text{C-Atome}} = \frac{98.93 * (\text{Int. } [M+1])}{1.07 * (\text{Int. } [M])} = \frac{98.93 * 13692}{1.07 * 77796} = 16.27 \quad (9)$$

Die ungefähre Anzahl der Kohlenstoffatome beträgt demnach 16. Dies ist jedoch lediglich als erste Näherung zu betrachten, da die Berechnung aufgrund von geringen Signal-zu-Rausch-Verhältnissen oder Interferenzen mit anderen Signalen mit Fehlern behaftet sein kann.^[105]

Sofern es sich bei dem Masse-zu-Ladungs-Verhältnis von 335.11 (bzw. 335.1097) tatsächlich um das Natrium-Addukt des Metaboliten handelt, kann die monoisotopische Masse zu 312.1204 Da bestimmt werden. Die wahrscheinlichste Summenformel mit einer Abweichung von 1.61 ppm ist deshalb $C_{15}H_{20}O_7$ (monoisotopische Masse: 312.1209 Da).

Wie der Abbildung 90 zusätzlich zu entnehmen ist, weist der EIC des m/z -Wertes von 133.06, der innerhalb der *SCORE-metabolite-ID* App basierend auf den DI-Massenspektren ebenfalls einen hohen Korrelationskoeffizienten aufweist, den gleichen Verlauf wie die in blau dargestellten EICs ($m/z = 335.11$ und 336.11) auf und kann dementsprechend ebenfalls diesem Metaboliten zugeordnet werden. Die Differenz zwischen dem vermeintlichen Natrium-Addukt des Metaboliten ($m/z = 335.11$) und dem Fragment ($m/z = 133.06$) beträgt 202.05. Diese Differenz lässt auf das Natrium-Addukt einer Hexose ($[C_6H_{12}O_6+Na]^+$) schließen, das ein Masse-zu-Ladungs-Verhältnis von 203.05 aufweist. Das Fragment mit $m/z = 133.06$ würde demnach dem Quasimolekülion des Aglycon-Teils des Metaboliten ($[Aglycon+H]^+$) entsprechen. Dies stützt die zu Beginn getroffene Annahme, dass es sich im NMR-Spektrum bei dem Dublett bei einer chemischen Verschiebung von 4.56 ppm um ein Signal eines Kohlenhydrates handelt. In Abbildung 91 sind selektive TOCSY-Spektren der Fraktion 28 gezeigt. Die selektiven TOCSY-Spektren des Kohlenhydrat-Teils des Metaboliten (obere beide Spektren) ermöglichen über die Analyse der Kopplungskonstanten die Identifizierung des Kohlenhydrat-Teils als Glucose (**12**).

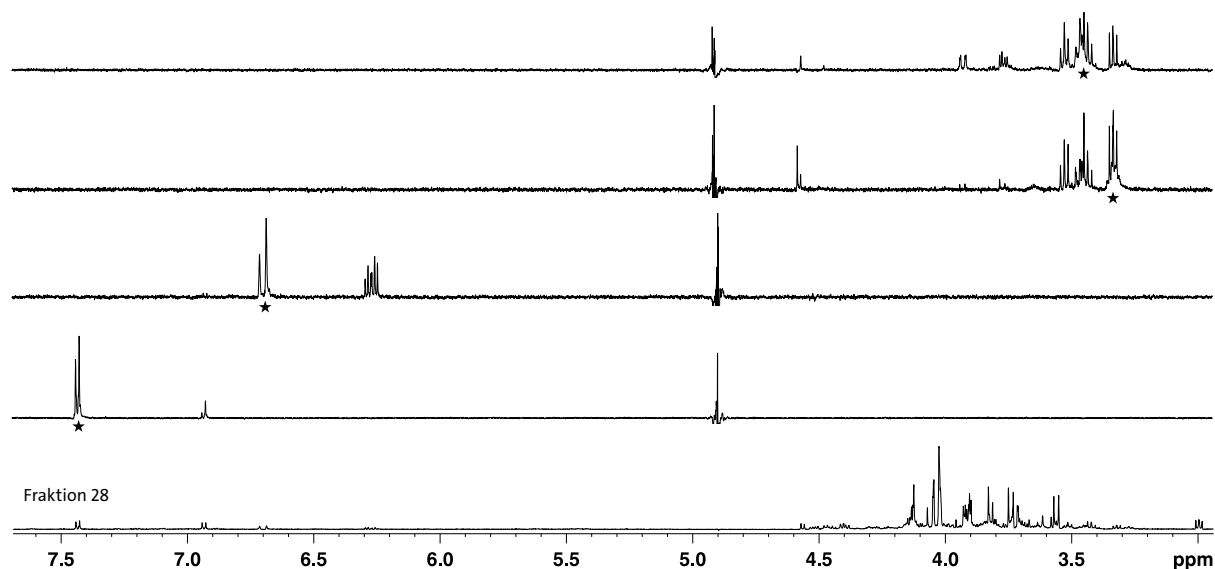


Abbildung 91: Selektive TOCSY-Spektren der Fraktion 28 der Apfelprobe nach Fraktionierung. Die Signale, die selektiv angeregt wurden, sind mit einem Sternchen gekennzeichnet. Die Kopplung von H-8 zu H-9a und H-9b ist im selektiven TOCSY nicht zu sehen. Das 2D-COSY-Spektrum, das im Anhang in Abbildung 154 gezeigt ist, zeigt diese Kopplung allerdings eindeutig. Das unterste Spektrum zeigt das ^1H -NMR-Spektrum der Fraktion 28.

Die Verknüpfung der beiden Substrukturen Glucose (**12**) und (*E*)-Cumarylalkohol (**65**) kann prinzipiell sowohl über die phenolische Hydroxylgruppe als auch über die alkoholische Hydroxylgruppe des (*E*)-Cumarylalkohols (**65**) erfolgen. Über ein selektives NOESY-Spektrum mit selektiver Anregung des anomeren Protons (H-1') kann die Verknüpfung über die alkoholische Hydroxylgruppe nachgewiesen werden, da eine räumliche Kopplung zu einem der diastereotopen Protonen an C-9 des (*E*)-Cumarylalkohols (**65**) bei 4.41 ppm detektiert werden kann.

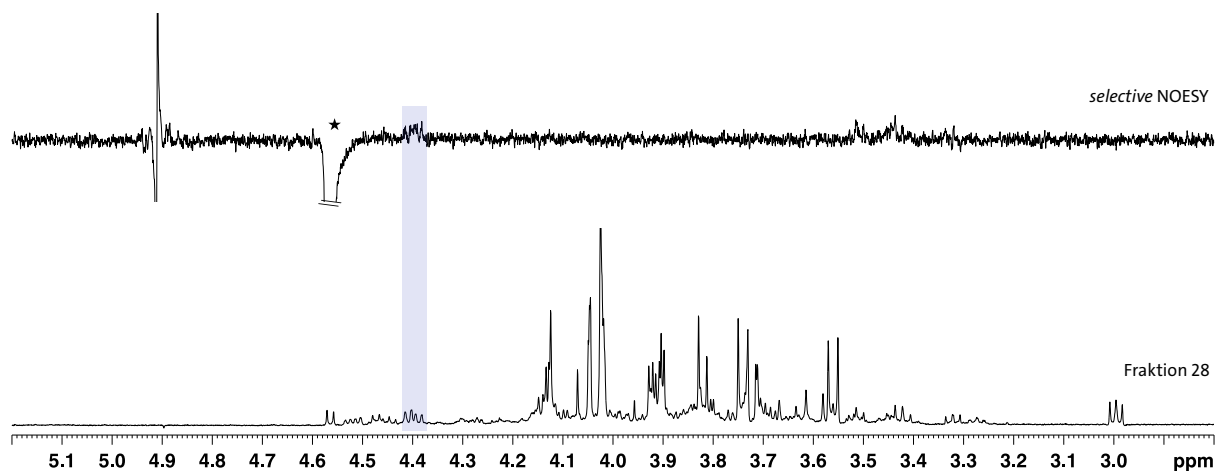
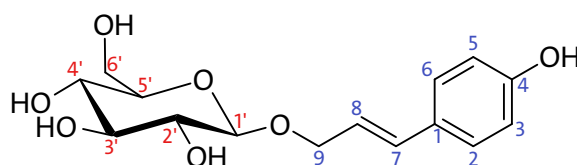


Abbildung 92: Selektives NOESY-Spektrum der Fraktion 28 (oben) mit selektiver Anregung des anomeren Protons der Glucose-Einheit bei 4.56 ppm. Es ist eine räumliche Kopplung zu H-9b bei 4.41 ppm zu sehen (blau markiertes Signal).

Insgesamt kann dieser Metabolit demnach als Triandrin (**66**) (*para*-Hydroxy-*trans*-cinnamyl-*O*- β -D-glucopyranosid) mit einer monoisotopischen Masse von 312.1209 Da identifiziert werden. Dieser Metabolit ist ebenfalls unter dem Namen Sachaliside 1 in der Literatur beschrieben.^[174] Der Metabolit wurde bereits nach Isolierung in der Apfelsorte Annurca^[175] sowie auch in anderen Matrices wie der Weiderinde^[176], den Blüten der Rübsen (*Brassica rapa*)^[177] und der Rosenwurz (*Rhodiola rosea*)^[178] identifiziert.



Triandrin (*para*-Hydroxy-*trans*-cinnamyl-*O*- β -D-glucopyranosid)
(**66**)

Abbildung 93: Strukturformel des Metaboliten Triandrin (**66**) (*para*-Hydroxy-*trans*-cinnamyl-*O*- β -D-glucopyranosid), der in Fraktion 28 der Apfelprobe identifiziert werden konnte.

Als Derivat eines Phenylpropanoids erfolgt die Biosynthese vermutlich über die Shikimat-Biosynthese.^[178] Nach einer Studie zu Äpfeln der Sorte Annurca weist Triandrin als Radikalfänger antioxidative Eigenschaften auf.^[175] Die ¹H-NMR-Daten des Metaboliten Triandrin (**66**) sind in Tabelle 21 aufgeführt.

Tabelle 21: ¹H-NMR-Daten des Metaboliten Triandrin (**66**) (*para*-Hydroxy-*trans*-cinnamyl-*O*- β -D-glucopyranosid), der in Fraktion 28 der Apfelprobe identifiziert werden konnte. 600.13 MHz, D₂O, pH-Wert = 7.0, kalibriert auf TSP-*d*₄.

δ_{H} [ppm]	Multiplizität	<i>J</i> [Hz]	Zuordnung
7.45-7.42	m	-	2 & 6
6.95-6.92	m	-	3 & 5
6.70	d	16.4	7
6.27	ddd	16.4, 6.7, 6.7	8
4.56	d	8.0	1'
4.52	dd	11.6, 5.7	9a
4.41	dd	12.5, 7.3	9b
3.93	dd	12.5, 2.0	6'a
3.77	dd	12.4, 5.6	6'b
3.52	dd	9.0, 9.0	3'
3.40-3.48	m	-	4', 5'
3.32	dd	8.6, 8.1	2'

Wie anhand der Boxplots zuvor in Abbildung 82 zu sehen ist, ist der Metabolit Triandrin (**66**) in höheren Konzentrationen in den untersuchten Proben der Sorten Boskoop, Elstar und Holsteiner Cox enthalten. Die Proben der Sorten Gala und Wellant weisen im Mittel geringere Konzentrationen Triandrin (**66**) auf.

5.2.3.2 Identifizierung eines Vomifoliol-Disaccharid Glycosids

Die zuvor in Abbildung 83 gezeigten Boxplots des *Buckets* bei einer chemischen Verschiebung von 1.05 ppm zeigen, dass der diesem Signal zugrundeliegende Metabolit im Mittel höhere Konzentrationen in den Apfelsorten Gala und Holsteiner Cox aufweist, als in den Sorten Boskoop, Elstar und Wellant. Nach Fraktionierung der Apfelprobe der Sorte Elstar (Probe 20-E-SH-9) weist die Fraktion 30 das Maximum dieses Signals auf. Zur Identifizierung des Metaboliten wurden zunächst durch die semi-automatische Korrelation diejenigen EMCs detektiert, die innerhalb der 80 Fraktionen einen Korrelationskoeffizienten > 0.95 zu dem EDC bei einer chemischen Verschiebung von 1.05 ppm aufweisen. Die zehn intensivsten EMCs sind unter Angabe des Korrelationskoeffizienten, der maximalen Intensität und der potenziellen Zuordnung in Tabelle 22 angegeben.

Tabelle 22: Zehn intensivste EMCs im positiven Ionisationsmodus mit einem Korrelationskoeffizienten > 0.95 zu dem EDC bei einer chemischen Verschiebung von 1.05 ppm in den Fraktionen 1-80 nach Fraktionierung der Apfelprobe 20-E-SH-9.

EMC [m/z]	Korrelationskoeffizient	max. Intensität [abs.]	Zuordnung
207.14	0.9627	57928	$[\text{Vomifoliol}-\text{H}_2\text{O}+\text{H}]^+$
225.15	0.9862	40975	$[\text{Vomifoliol}+\text{H}]^+$
288.10	0.9549	6871	-
369.19	0.9779	10029	$[\text{M}+\text{H}-\text{C}_5\text{H}_{10}\text{O}_5]^+$
387.20	0.9897	78484	$[\text{M}+\text{H}-\text{C}_5\text{H}_8\text{O}_4]^+$
389.21	0.9862	6485	$[\text{M}+\text{H}-\text{C}_5\text{H}_8\text{O}_4]^+$ Isotopenpeak
519.24	0.9769	6303	$[\text{M}+\text{H}]^+$
541.23	0.9787	71005	$[\text{M}+\text{Na}]^+$
542.23	0.9913	19253	$[\text{M}+\text{Na}]^+$ Isotopenpeak
543.24	0.9946	8687	$[\text{M}+\text{Na}]^+$ Isotopenpeak

Durch Betrachtung der Differenzen zwischen den Masse-zu-Ladungs-Verhältnissen der EMCs kann der EMC bei $m/z = 519.24$ dem protonierten Molekülion und der EMC bei $m/z = 541.23$ dem Natrium-Addukt des Metaboliten zugeordnet werden. Die exakte Masse des Metaboliten kann demnach zu 518.24 Da (bzw. $518.2364 \pm 0.0002 \text{ Da}$) bestimmt werden. Entsprechend der Gleichung (10) kann über das Verhältnis des Signals des Natrium-Adduktes zu dem Isotopensignal des Natrium-Adduktes die maximale Anzahl der Kohlenstoffatome des Metaboliten abgeschätzt werden.^[105,173] Zur Berechnung wurden die Intensitätswerte innerhalb eines Massenspektrums der Fraktion 30 herangezogen.

$$N_{\text{C-Atome}} = \frac{98.93 * (\text{Int. [M+1]})}{1.07 * (\text{Int. [M]})} = \frac{98.93 * 11292}{1.07 * 43464} = 24.02 \quad (10)$$

Die maximale Anzahl der Kohlenstoffatome beträgt demnach 24. Somit kann die wahrscheinlichste Summenformel des Metaboliten zu $\text{C}_{24}\text{H}_{38}\text{O}_{12}$ bestimmt werden.

Über die semi-automatische Korrelation des EDCs bei 1.05 ppm wird eine Reihe von weiteren EDCs detektiert, die einen Korrelationskoeffizienten > 0.95 zueinander aufweisen. Die Signale der EDCs unter der zusätzlichen Angabe der Multiplizitäten und Kopplungskonstanten sind in Tabelle 23 aufgeführt.

Tabelle 23: EDCs mit einem Korrelationskoeffizienten > 0.95 zu dem EDC bei einer chemischen Verschiebung von 1.05 ppm, berechnet über die Fraktionen 1-31 unter zusätzlicher Angabe der Multiplizitäten und Kopplungskonstanten der Signale.

EDC [ppm]	Multiplizität	Korrelationskoeffizient
1.038	s	0.9990
1.048	s	1.0000
1.323	d ($J = 6.6$ Hz)	0.9661
1.959	d ($J = 1.2$ Hz)	0.9996
1.964		0.9934
2.244	d ($J = 18.0$ Hz)	0.9817
2.2735		0.9674
2.609	d ($J = 18.0$ Hz)	0.9618
4.434	d ($J = 7.9$ Hz)	0.9520
4.539	d ($J = 7.9$ Hz)	0.9717
5.794	dd ($J = 15.7, 6.8$ Hz)	0.9862
5.805		0.9858
5.819		0.9948
5.829		0.9963
5.909	d ($J = 15.6$ Hz)	0.9851
5.935		0.9541
6.049	s (br)	0.9917
6.054		0.9828

Die NMR-Signale in der obigen Tabelle erlauben die Identifizierung einzelner Strukturmotive. Auffällig ist die *trans*-konfigurierte Doppelbindung, die sich aufgrund der Kopplungen und der chemischen Verschiebungen der Signale bei 5.81 ppm ($J = 15.7$ und 6.8 Hz) und 5.92 ppm ($J = 15.6$ Hz) ergibt. Das breite Singulett bei einer chemischen Verschiebung von 6.05 ppm sowie das Dublett mit einer kleinen Kopplungskonstante von 1.2 Hz bei einer chemischen Verschiebung von 1.96 ppm lässt vermuten, dass diese Signale skalar über vier Bindungen miteinander koppeln. Das Signal bei 1.96 ppm mit einem Integral von drei Protonen ist darüber hinaus charakteristisch für eine Methylgruppe in Nachbarschaft zu einer Doppelbindung. Die beiden Singuletts bei 1.04 und 1.05 ppm weisen im HSQC-Spektrum eine Kopplung zu zwei Kohlenstoffatomen bei 25.6 und 26.5 ppm auf. Im HMBC-Spektrum der Fraktion 30 zeigen beide

Signale eine Kopplung zu dem gleichen Kohlenstoffatom bei 44.0 ppm auf. Dementsprechend kann geschlussfolgert werden, dass es sich um zwei Methylgruppen an einem quartären Kohlenstoffatom handelt. Über das 2D-COSY-Spektrum werden außerdem weitere NMR-Signale detektiert, die ebenfalls von diesem Metaboliten stammen. Unter Zuhilfenahme aller Informationen kann darauf geschlossen werden, dass der Metabolit die Substruktur des Vomifoliols (**67**) (6-Hydroxy-3-oxo- α -ionol) beinhaltet. Dies deckt sich mit den in Tabelle 22 aufgeführten hoch-korrelierenden EMCs. Der EMC bei $m/z = 225.15$ entspricht demnach dem protonierten Molekülion des Vomifoliol-Fragmentes (**67**). Der EMC bei $m/z = 207.14$ entsteht durch die formale Eliminierung eines Hydroxid-Ions dieses Fragmentes.

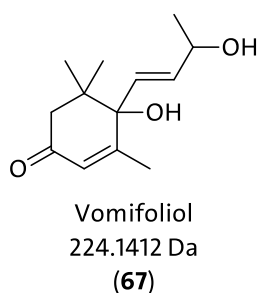


Abbildung 94: Strukturformel von Vomifoliol (**67**), das als Substruktur des Metaboliten in Fraktion 30 der Apfelprobe identifiziert werden konnte.

Die Differenz zwischen den Signalen des Vomifoliol-Fragmentes (**67**) ($m/z = 225.15$) und dem Quasimolekülion des Metaboliten ($m/z = 519.24$) beträgt 294.09 Da. Die Summenformel des übrigen Fragments beläuft sich auf $C_{11}H_{18}O_9$. Zusätzlich werden, wie zuvor in Tabelle 22 gezeigt, EMCs bei $m/z = 369.19$ und $m/z = 387.20$ mit einem hohen Korrelationskoeffizienten zu dem EDC eines Metaboliten-Signals detektiert. Das Signal bei $m/z = 369.19$ weist eine Massendifferenz von 150.05 Da zu dem Quasimolekülion des Metaboliten auf und kann demzufolge dem formalen Verlust von $C_5H_{10}O_5$ zugeordnet werden. Diese Summenformel lässt auf eine Pentose schließen. Das Signal bei $m/z = 387.20$ entspricht dem protonierten Fragment nach Verlust von $C_5H_8O_4$ des Quasimolekülions, d.h. das verbleibende Fragment besteht aus dem Kondensationsprodukt von Vomifoliol (**67**) und einer Hexose. Bei dem Metaboliten handelt es sich demnach um ein Vomifoliol-Disaccharid (**68**), wobei das Disaccharid aus einer Hexose und einer Pentose besteht und die Pentose die terminale Einheit darstellt. Die beschriebenen Fragmentierungsreaktionen sind in Abbildung 95 dargestellt.

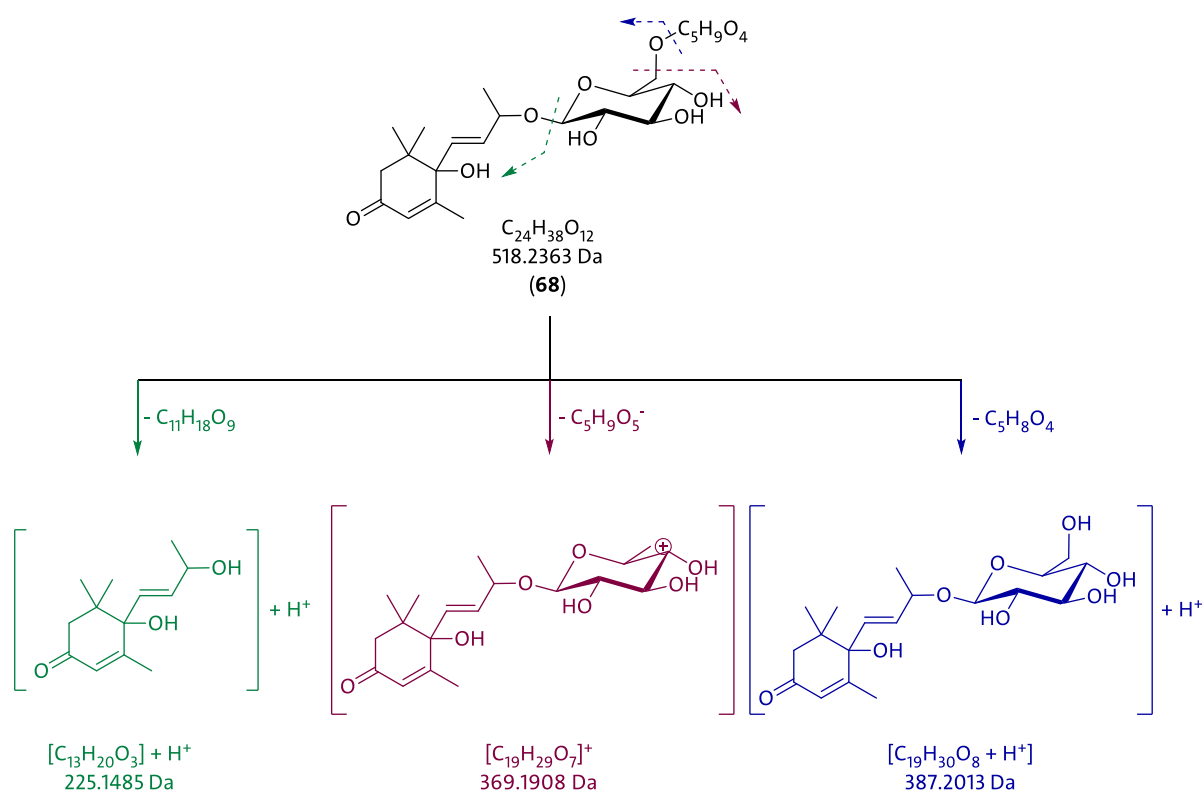


Abbildung 95: Postulierte Fragmentierungsreaktionen des in der Apfelprobe identifizierten Vomifoliol-Disaccharid Glycosids (68). Durch die Signale der hoch-korrelierenden EMCs bei $m/z = 369.19$ und $m/z = 387.20$ kann geschlossen werden, dass das Aglycon Vomifoliol (67) mit einem Disaccharid verknüpft ist, welches wiederum aus einer Hexose und einer Pentose besteht. Die Verknüpfung findet über die Hexose statt, d. h. die Pentose stellt die terminale Einheit des Glycosids dar.

Die anomeren Protonen der beiden Kohlenhydrat-Einheiten können den Dubletts bei 4.43 ppm ($J = 7.9$ Hz; ^{13}C -Signal aus HSQC: 106.4 ppm) und 4.53 ppm ($J = 7.9$ Hz; ^{13}C -Signal aus HSQC: 103.6 ppm) zugeordnet werden. Über das HMBC-Spektrum der Fraktion 30 ist die Kopplung zwischen dem Dublett bei 4.53 ppm zu dem Signal bei 80.3 ppm des Kohlenstoffatoms C-9 des Vomifoliol-Fragmentes zu detektieren. Demzufolge kann das Dublett bei 4.53 ppm dem anomeren Proton der Hexose zugeordnet werden. Das selektive TOCSY-Spektrum der Fraktion 30 mit selektiver Einstrahlung auf das anomere Proton der Hexose sowie das COSY-Spektrum zeigen die Kopplung zu zwei Dubletts eines Dubletts bei 3.27 ppm ($J = 8.2$ Hz, 8.7 Hz) und 3.52 ppm ($J = 8.2$ Hz, 6.1 Hz), die dementsprechend den Protonen H-2' (3.27 ppm) und H-3' (3.52 ppm) zugeordnet werden können (siehe Abbildung 96). In dem selektiven TOCSY-Spektrum ist neben der zuvor beschriebenen Kopplung auch die Kopplung zu den zwei Signalen der Doppelbindungs-Protonen bei 5.81 ppm und 5.92 ppm des Vomifoliol-Fragmentes des Metaboliten detektierbar, die ausgehend von dem Signal des zur Doppelbindung benachbarten

Methin-Protons auftritt. Das Signal des Methin-Protons bei 4.51 ppm des Vomifoliol-Fragments überlagert teilweise mit dem Signal des anomeren Protons der Hexose-Einheit bei 4.53 ppm. Die Signale der Protonen H-4', H-5' und H-6' können aufgrund von Überlagerung in den 2D-Spektren bzw. unvollständigen Magnetisierungstransfer im selektiven TOCSY-Spektrum nicht detektiert werden. Nichtsdestotrotz kann aufgrund der aus dem selektiven TOCSY-Spektrum ermittelten Kopplungskonstanten der Protonen H-1', H-2' und H-3' darauf geschlossen werden, dass es sich bei der Hexose um β -Glucose handelt.

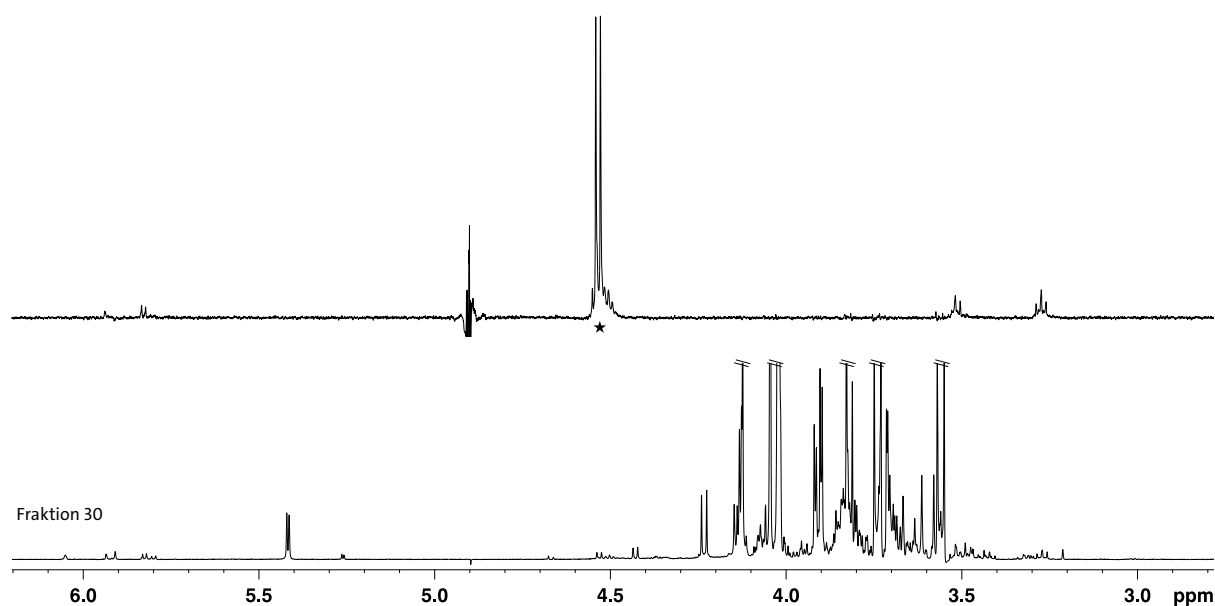


Abbildung 96: Selektives TOCSY-Spektrum der Fraktion 30 (oben) mit selektiver Anregung des anomeren Protons der Glucose-Einheit bei 4.53 ppm (gekennzeichnet durch Sternchen) sowie das ¹H-NMR-Spektrum der Fraktion 30 (unten). Das selektive TOCSY-Spektrum zeigt eine Kopplung zu zwei Dubletts eines Dubletts bei 3.27 ppm und 3.52 ppm, die unter Zuhilfenahme des COSY-Spektrums den Protonen H-2' und H-3' der Hexose zugeordnet werden können. Zusätzlich zeigt das Spektrum eine Kopplung zu den Protonen der Doppelbindung des Vomifoliol-Fragmentes des Metaboliten bei 5.81 und 5.92 ppm. Diese Kopplung tritt aufgrund der Überlagerung des Signals des anomeren Protons mit dem Signal eines Methin-Protons des Vomifoliol-Fragmentes bei 4.51 ppm auf.

Das Dublett bei einer chemischen Verschiebung von 4.43 ppm kann demzufolge dem anomeren Proton der Pentose zugeordnet werden. Im selektiven und im zweidimensionalen TOCSY-Spektrum kann eine Kopplung zu einem Dublett vom Dublett bei 3.31 ppm ($J = 8.0$ Hz, 9.0 Hz) und 3.47 ppm ($J = 9.0$ Hz, 9.1 Hz) detektiert werden (siehe Abbildung 97). Außerdem ist die Kopplung zu zwei weiteren Signalen bei 3.63 ppm sowie 3.97 ppm zu erkennen, dessen Multiplizitäten und Kopplungskonstanten jedoch aufgrund des geringen Signal-zu-Rausch-Verhältnisses von 4.26 bzw. 4.77 nicht eindeutig bestimmbar sind.

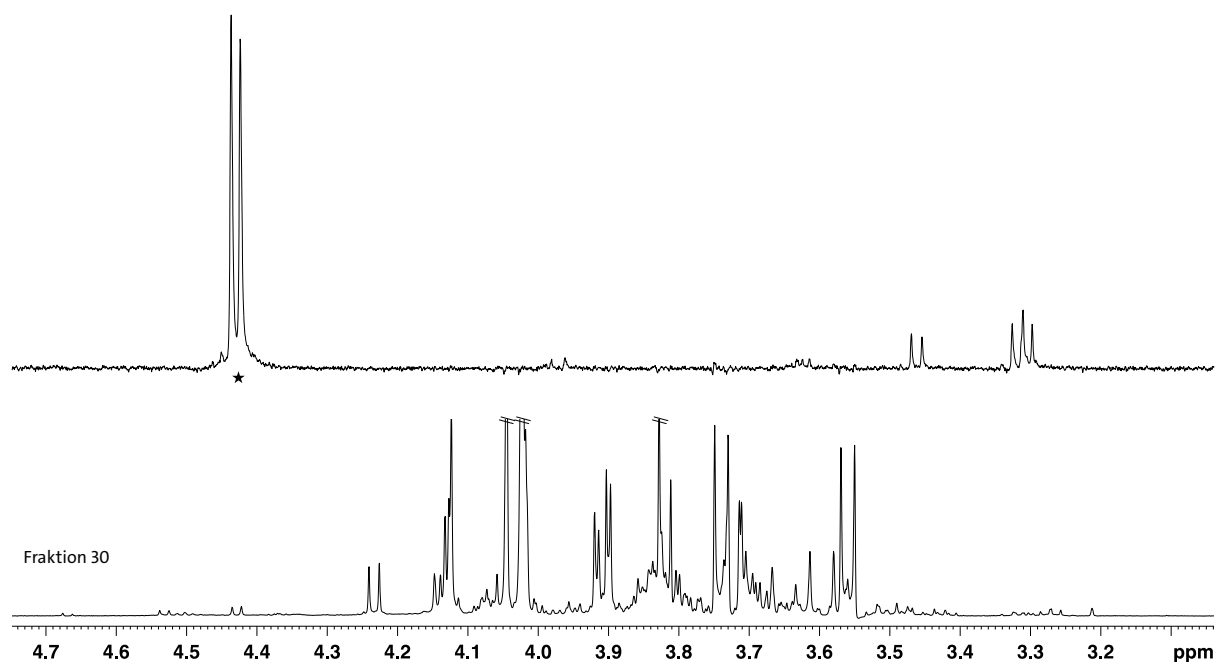


Abbildung 97: ^1H -NMR-Spektrum der Fraktion 30 der Apfelprobe (unten) sowie das selektive TOCSY-Spektrum der Fraktion mit selektiver Anregung des anomeren Protons der Pentose bei 4.43 ppm (oben).

Aufgrund der geringen Signalintensitäten in den NMR-Spektren sowie der Überlagerung der Signale im Bereich der Kohlenhydrat-Signale ist an dieser Stelle keine eindeutige Strukturaufklärung der Pentose-Substruktur des Metaboliten möglich. In der verwendeten Literatur sind die Vomifoliol-Disaccharid Glycoside bestehend aus den Disacchariden β -D-Xylopyranosyl-(1 $\prime\prime$ $\rightarrow$ 6 $\prime$)-O- β -D-glucopyranosid (Primverose) (**69**)^[179–181], α -L-Arabinopyranosyl-(1 $\prime\prime$ $\rightarrow$ 6 $\prime$)-O- β -D-glucopyranosid (**70**)^[182], α -L-Arabinofuranosyl-(1 $\prime\prime$ $\rightarrow$ 6 $\prime$)-O- β -D-glucopyranosid (**71**)^[183] und β -D-Apiofuranosyl-(1 $\prime\prime$ $\rightarrow$ 6 $\prime$)-O- β -D-glucopyranosid (**72**)^[179,184–186] beschrieben. Die Strukturformeln sind in Abbildung 98 dargestellt.

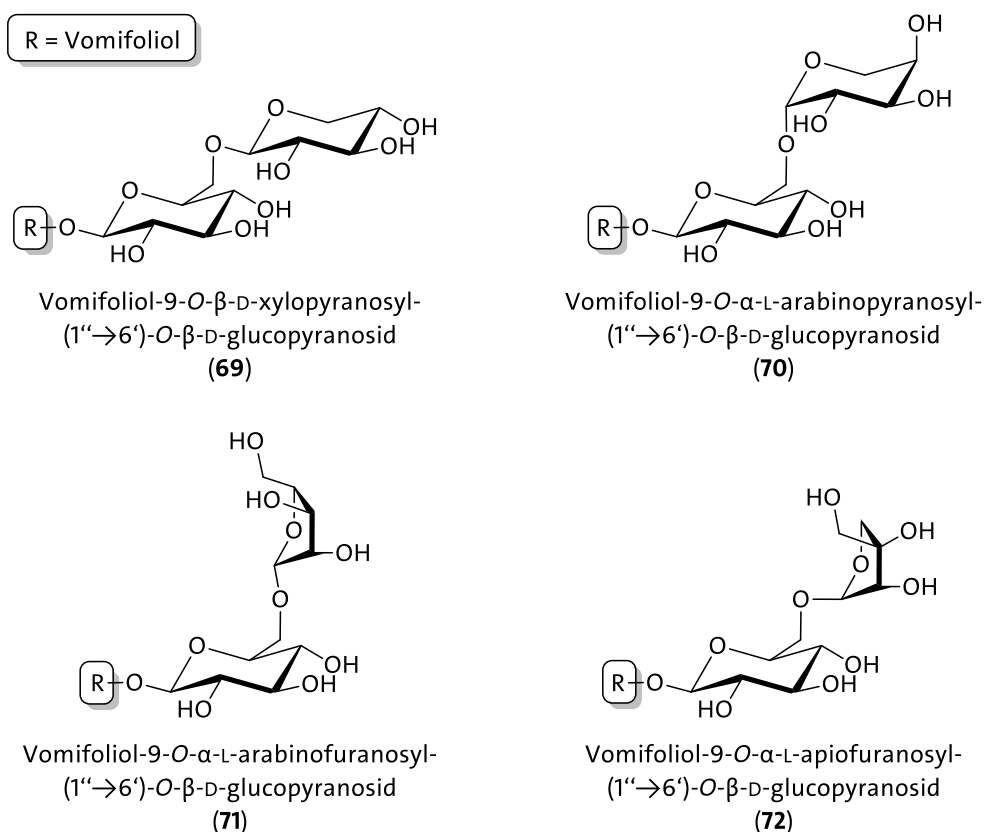


Abbildung 98: Strukturformeln der in der verwendeten Literatur beschriebenen Vomifoliol-Disaccharide (69)^[179–181], (70)^[182], (71)^[183] und (72)^[179,184–186].

Das Vorhandensein des Disaccharids bestehend aus der Pentose Apiofuranose (72) kann ausgeschlossen werden, da in dem Fall ausgehend vom anomeren Proton lediglich die Kopplung zu einem weiteren Dublett des H-2 Protons zu erwarten ist. Wie dem selektiven TOCSY-Spektrum in Abbildung 97 zu entnehmen ist, wird bei selektiver Anregung des anomeren Protons jedoch die Kopplung zu einem Dublett vom Dublett sowie weiterreichende Kopplungssignale detektiert. Weitere eindeutigen Aussagen zur exakten Struktur des Disaccharids können jedoch nicht getroffen werden. In Abbildung 99 ist die Strukturformel des im Rahmen dieser Arbeit identifizierten Vomifoliol-Disaccharid Glycosids (68) dargestellt, wobei die Pentose möglicherweise β-D-Xylopyranose, α-L-Arabinofuranose oder α-L-Arabinopyranose entspricht.

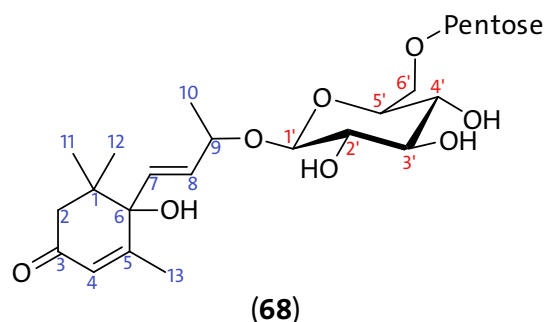


Abbildung 99: Strukturformel des in der Apfelprobe 20-E-SH-9 identifizierten Vomifoliol-Disaccharid Glycosids (68). Das Disaccharid besteht dabei aus β -D-Glucopyranose sowie einer Pentose, dessen Struktur nicht eindeutig aufgeklärt werden konnte.

Die Zuordnung der ^1H -NMR-Daten des Vomifoliol-Disaccharid Glycosids (68) ist in Tabelle 24 aufgeführt. Dabei ist die Zuordnung der Kohlenhydrat-Signale von β -D-Glucose und der Pentose aufgrund von Signalüberlagerungen und geringen Signalintensitäten unvollständig.

Tabelle 24: Zuordnung der ^1H -NMR-Daten des Vomifoliol-Disaccharid Glycosids (68). Das Disaccharid besteht aus β -D-Glucose sowie einer Pentose, dessen Struktur jedoch nicht eindeutig bestimmt werden konnte. Für Glucose und die Pentose können aufgrund von Signalüberlagerungen nicht alle Signale der Protonen detektiert und zugeordnet werden.

δ_{H} [ppm]	Multiplizität	J [Hz]	Zuordnung
6.05	s (br)	-	4
5.92	d	15.6	7
5.81	dd	15.7, 6.8	8
4.53	d	7.9	1'
4.47-4.52	m	-	9
4.43	d	7.9	1'' (Pentose)
3.52	dd	8.2, 6.1	3'
3.47	dd	9.1, 9.0	Pentose
3.31	dd	9.0, 8.0	Pentose
3.27	dd	8.7, 8.7	2'
2.59	d	18.0	2a
2.26	d	18.0	2b
1.96	d	1.2	13
1.32	d	6.6	10
1.05	s	-	11
1.04	s	-	12

Das Vomifoliol-Disaccharid bestehend aus dem Disaccharid β -D-Xylopyranosyl-(1''→6')-O- β -D-glucopyranosid (**69**) sowie das strukturell verwandte Vomifoliol-Trisaccharid bestehend aus dem Trisaccharid β -D-Glucopyranosyl-(1'''→4'')-O- β -D-xylopyranosyl-(1''→6')-O- β -D-glucopyranosid (**73**) (siehe Abbildung 100) wurden bereits in Apfelproben identifiziert.^[181] Das enthaltene Disaccharid Primverose kommt in einer Vielzahl von Pflanzen in Form von verschiedenen Glycosiden (Primverosiden) vor.^[187] Das Aglycon Vomifoliol ist strukturell verwandt mit Abscisinsäure (**74**) (siehe Abbildung 100) und beinhaltet als Phytohormon biologische Aktivität.^[188,189]

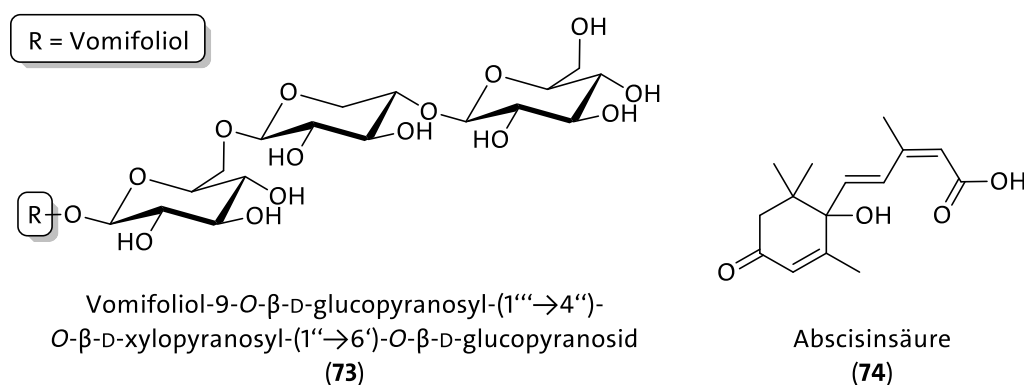


Abbildung 100: Strukturformeln des Vomifoliol-Trisaccharids (**73**), das bereits in Apfelproben identifiziert werden konnte, sowie der Abscisinsäure (**74**), das strukturell verwandt mit Vomifoliol (**67**) ist.^[181]

5.2.3.3 Fazit zur Identifizierung von Markermetaboliten in Apfelproben durch Anwendung von SCORE-metabolite-ID

Neben dem polaren Extrakt der Pinienkerne wurde die Methode *SCORE-metabolite-ID* zusätzlich zur Identifizierung von potenziellen Markermetaboliten in Äpfeln verwendet. Durch die multivariate Datenanalyse der Apfelproben der fünf Sorten Boskoop, Elstar, Gala, Holsteiner Cox und Wellant konnten zunächst spezifische Signale ermittelt werden, die einen signifikanten Einfluss auf die Differenzierung der Proben z. B. in Bezug auf die Sortenunterscheidung haben. Diese Signale wurden anschließend spezifisch mittels *SCORE-metabolite-ID* analysiert, um die zugrundeliegenden Metaboliten zu identifizieren. Dabei konnte zum einen der Metabolit Triandrin (**66**) identifiziert werden, der jeweils höhere Konzentrationen in den Proben der Sorten Boskoop, Elstar und Holsteiner Cox aufweist. Zum anderen konnte ein Vomifoliol-Disaccharid

Glycosid (**68**) identifiziert werden, das in höheren Konzentrationen in den Apfelsorten Gala und Holsteiner Cox auftritt.

In beiden Fällen konnten durch die Korrelation der NMR- und MS-Daten zunächst ein Großteil der zugehörigen Signale der jeweiligen Metaboliten aus dem Gemisch heraus ermittelt werden. Die Analyse der NMR-Signale ermöglichte daraufhin die Herleitung von Strukturfragmenten. Zusätzlich konnten durch die MS-Signale die Anzahl der Kohlenstoffatome sowie die Summenformel der Metaboliten postuliert werden. Die Analyse der Fragment-Ionen aus den Massenspektren lieferte weitere relevante Informationen zur Identifizierung der Analyten. Durch die Ergänzung mit weiteren (selektiven und zweidimensionalen) NMR-Experimenten konnten somit die beiden komplexen Metaboliten in vergleichsweise kurzer Zeit und ohne aufwendige Isolierung der Substanzen aus dem Gemisch heraus identifiziert werden. Diese Beispiele stellen erneut die Vorteile der Anwendung von *SCORE-metabolite-ID* heraus und zeigen, dass diese Methodik eine adäquate Ergänzung zur multivariaten Datenanalyse im Rahmen von *Metabolomics* Studien darstellt, um daran anschließend spezifische Metaboliten zu identifizieren.

5.3 Anwendung von SCORE-metabolite-ID zur Reaktionsverfolgung

Das bisherige Anwendungsgebiet der *SCORE-metabolite-ID* Methode liegt in der Identifizierung von Metaboliten im Rahmen von *Metabolomics* Studien. Dabei liegt der dreidimensionalen Korrelationsstrategie, auf der die *SCORE-metabolite-ID* Methode basiert, ein zeitlicher Verlauf von analytischen Daten mit komplementären Informationen (z.B. NMR- und MS-Daten) zu Grunde. Die zeitliche Domäne wird durch die chromatographische Fraktionierung intentionell erstellt. Bei der Durchführung chemischer Reaktionen ist die Dimension der zeitlichen Abfolge jedoch eine inhärente Eigenschaft, wodurch sich die Möglichkeit der Nutzung von *SCORE-metabolite-ID* zur Reaktionsverfolgung ergibt.

Die Verfolgung chemischer Reaktionen dient dem tiefergehenden Verständnis der Reaktionsmechanismen, um beispielsweise Intermediate und Nebenprodukte zu identifizieren. Damit einhergehend ermöglicht die detaillierte Kenntnis des Reaktionsmechanismus die Optimierung der Reaktionsbedingungen, um dadurch die Effizienz der Synthese zu steigern.^[190] Im Folgenden wird die chemische Reaktion zur Darstellung von 2,4,6-Tri-(6-aminocapronsäure)-1,3,5-triazin (**78**) (TACT) durch die Korrelation von NMR- und MS-Daten untersucht. Die Experimente zur Synthese von TACT (**78**) wurden in Zusammenarbeit mit dem Institut für Technische und Makromolekulare Chemie mit CHRISTIAN DECKERT durchgeführt. In Folge dessen wurden die gleichen Rohdaten sowohl in dieser, als auch in der Dissertation von CHRISTIAN DECKERT verwendet.

5.3.1 Darstellung von 2,4,6-Tri-(6-aminocapronsäure)-1,3,5-triazin (TACT)

Die Verbindung 2,4,6-Tri-(6-aminocapronsäure)-1,3,5-triazin (**78**) (TACT) wirkt als Korrosionsinhibitor, der als aktive Komponente in Kühlschmierstoffen bei der Metallverarbeitung eingesetzt wird. Die antikorrosive Wirkung ergibt sich durch die Adsorption des Korrosionsinhibitors an die metallische Oberfläche, wodurch diese vor dem direkten Kontakt mit dem korrosiv wirkenden Agens geschützt wird.^[191]

TACT (**78**) und strukturell verwandte symmetrische oder asymmetrische Melamin-Derivate sind synthetisch ausgehend von Cyanurchlorid (**75**) (2,4,6-Trichlor-1,3,5-triazin) durch dreifache nucleophile aromatische Substitution der Chlorid-Substituenten mit den entsprechenden primären oder sekundären Aminen zugänglich. Die Tendenz des 1,3,5-Triazin-Rings eine

nucleophile aromatische Substitution einzugehen sinkt im Laufe der Reaktion mit abnehmender Anzahl der Chlorid-Substituenten. In vereinfachter Weise kann die Darstellung derartiger Melamin-Derivate durch die sequentielle und temperaturabhängige Substitution der Chlorid-Substituenten gesehen werden, die schematisch in Abbildung 101 gezeigt ist.^[192,193]

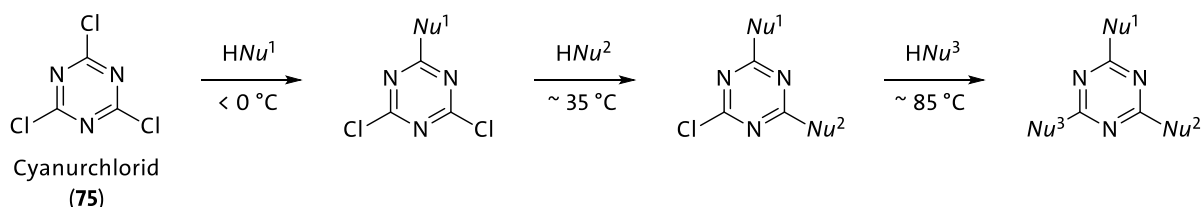


Abbildung 101: Vereinfachte Darstellung der sequentiellen und temperaturabhängigen Synthese von 1,3,5-Triazin-Derivaten ausgehend von Cyanurchlorid (75). Die Reagenzien $\text{H}Nu^1$, $\text{H}Nu^2$ und $\text{H}Nu^3$ sind beispielsweise Amine, wodurch Melamin-Derivate erhalten werden können. Aufgrund der abnehmenden Reaktivität des aromatischen Systems in Folge der Substitution der Chlorid-Substituenten ist mit zunehmendem Substitutionsgrad der Nucleophile eine Temperaturerhöhung vonnöten. Die Abbildung wurde verändert nach BRETTERBAUER *et al.* bzw. SMOLIN *et al.*^[192,193]

Aufgrund der abnehmenden Reaktivität des aromatischen 1,3,5-Triazin-Systems ist im Laufe der Reaktion für eine vollständige Substitution aller Chlorid-Substituenten je nach eingesetztem Amin eine Temperaturerhöhung auf bis zu 100°C notwendig.^[194] Die zu wählende Temperatur steht in starker Abhängigkeit der Nucleophilie des eingesetzten Amins.^[192] Beispielsweise resultiert die Umsetzung von Cyanurchlorid (75) mit Ammoniak bei 0°C zunächst in 2-Amino-4,6-dichloro-1,3,5-triazin. Bei einer Temperatur von $50\text{--}60^\circ\text{C}$ erfolgt die Substitution von zwei Chlorid-Substituenten und zur Darstellung von Melamin (2,4,6-Triamino-1,3,5-triazin) wird eine Temperatur von 100°C benötigt.^[193] SMOLIN *et al.* zu Folge stellt der in Abbildung 101 dargestellte Reaktionsverlauf im Allgemeinen jedoch lediglich eine grobe Näherung dar.^[193] Die genauen Reaktionsbedingungen müssen demnach individuell ermittelt und zur Maximierung der Ausbeute optimiert werden.

Zur Darstellung von TACT (78) ist das erforderliche Amin das Natriumsalz der 6-Aminocapronsäure (77), das durch basische Hydrolyse aus kommerziell erhältlichem ϵ -Caprolactam (76) darstellbar ist. Die Reaktion zu TACT (78) erfolgt NESTLER *et al.* zu Folge im wässrigen Milieu unter stark basischen Bedingungen in Ausbeuten von ca. 76%.^[194] Die Reaktionsgleichung ist in Abbildung 102 gezeigt.

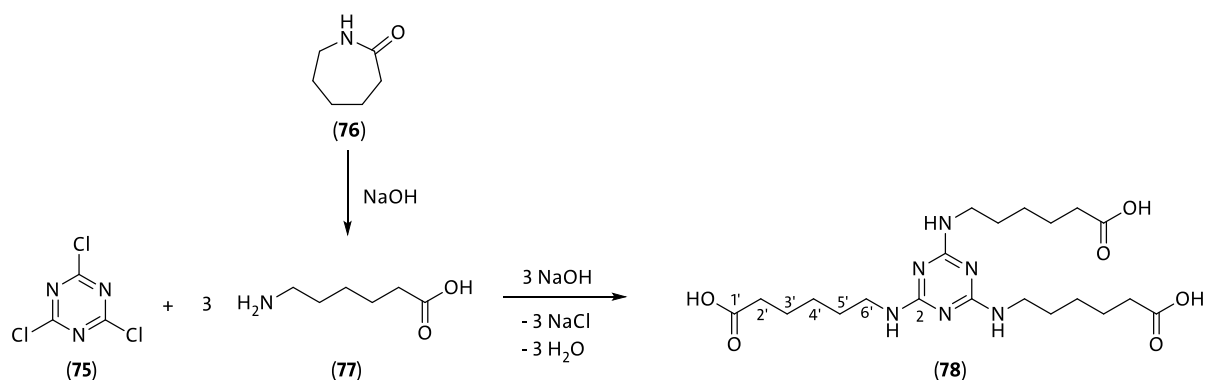


Abbildung 102: Reaktionsgleichung der Darstellung von TACT (**78**) (2,4,6-Tri-(6-aminocapronsäure)-1,3,5-triazin) ausgehend von Cyanurchlorid (**75**) und 6-Aminocapronsäure (**77**), das wiederum aus ϵ -Caprolactam (**76**) dargestellt werden kann. Die Reaktion findet unter stark basischen Bedingungen statt.

In wässriger Umgebung findet bei Temperaturen $> 10\text{ }^{\circ}\text{C}$ die Hydrolyse von Cyanurchlorid (**75**) zu Cyanursäure (**84**) (2,4,6-Trihydroxy-1,3,5-triazin) statt, die insbesondere unter den basischen Bedingungen in der oben gezeigten Reaktion beschleunigt wird.^[193] Die Hydrolyse eines oder mehrerer Chlorid-Substituenten stellt deshalb eine entscheidende Konkurrenzreaktion bei der Synthese von TACT (**78**) dar. Der gebildete Hydroxyl-Substituent kann anschließend nicht mehr im Rahmen einer nucleophilen aromatischen Substitution mit 6-Aminocapronsäure (**77**) reagieren, da das Hydroxid-Anion keine geeignete Abgangsgruppe ist. Die Hydrolyse eines Chlorid-Substituenten ist demzufolge zu vermeiden, um die Ausbeute der Synthese von TACT (**78**) zu steigern.

Da es sich bei der Synthese von TACT (**78**) um eine stufenweise ablaufende Reaktion handelt (vgl. Abbildung 101), bei der fortlaufend ein weiteres Chlor-Atom substituiert werden soll, ist die Hydrolyse eines Chlorid-Substituenten nicht nur zu Beginn der Reaktion bei Cyanurchlorid (**75**) möglich. Auch im weiteren Verlauf der Reaktion, d.h. nach bspw. der bereits stattgefundenen ein- oder zweifachen Substitution durch 6-Aminocapronsäure (**77**), ist die Hydrolyse eines Chlor-Atoms weiterhin möglich. In Abbildung 103 sind die parallel ablaufenden Substitutionsreaktionen im Überblick dargestellt. Ausgehend von Cyanurchlorid (**75**) können dementsprechend eine Vielzahl von Produkten über verschiedenste Reaktionswege und unterschiedliche Zwischenprodukte gebildet werden.

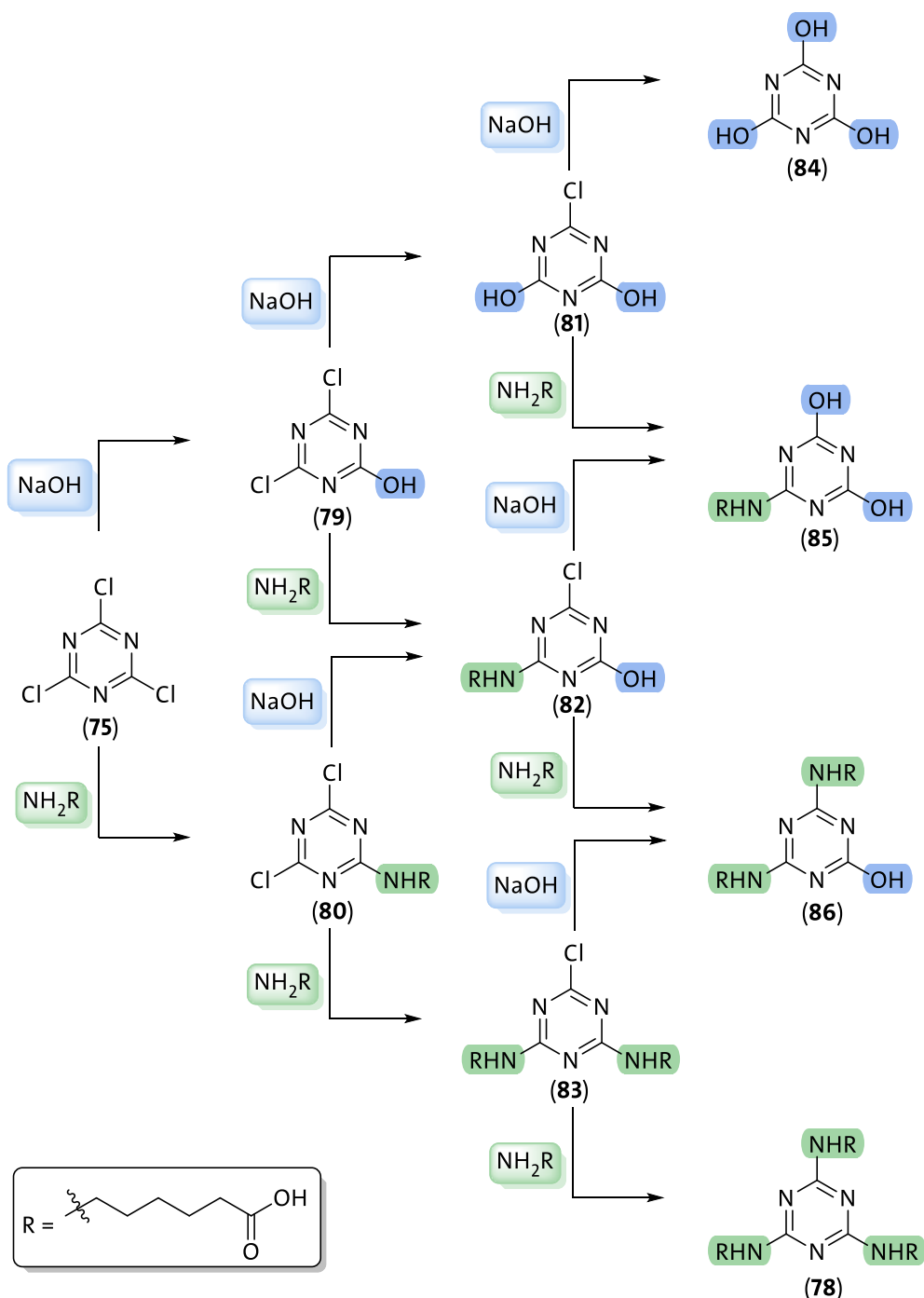


Abbildung 103: Parallel- und Folgereaktionen, die bei der Synthese von TACT (78) ausgehend von Cyanurchlorid (75) ablaufen können. Jeder Chlorid-Substituent kann sowohl durch die 6-Aminocapronsäure (77) substituiert werden, als auch in Folge der basischen Bedingungen hydrolysiert werden, wodurch die weitere Umsetzung mit 6-Aminocapronsäure (77) verhindert wird.

Um die Ausbeute und Selektivität der Reaktion in Hinblick auf die Synthese von TACT (78) zu maximieren, ist eine genaue Kenntnis des Reaktionsmechanismus der Synthese erforderlich. Für die Optimierung der technischen Prozessführung ist insbesondere die Reihenfolge der

ablaufenden Parallel- und Folgereaktionen sowie die Beeinflussung dieser ein wichtiges Kriterium. Die besondere Komplexität liegt an dieser Stelle daran, dass das gleiche Molekül über verschiedene Reaktionswege gebildet werden kann. Die stattgefundenen Reaktionswege lassen sich dabei in der finalen Reaktionslösung bzw. dem Produkt nicht mehr unterscheiden.

5.3.2 NMR-spektroskopische Untersuchung von TACT

Aufgrund des aromatischen Ringsystems und der Vielzahl quartärer Kohlenstoffatome ist die ^1H -NMR-Spektroskopie nicht die geeignete Methode zur Untersuchung von TACT (**78**). Stattdessen werden deshalb die ^{13}C -Kerne von TACT (**78**) NMR-spektroskopisch mittels eindimensionaler ^{13}C -NMR-Spektren untersucht. Wie bereits unter Abschnitt 3.3.2.1 beschrieben, ist die Sensitivität von ^{13}C -NMR-Spektren infolge der geringen natürlichen Häufigkeit von ^{13}C -Kernen im Vergleich zu bspw. ^1H -NMR-Spektren deutlich geringer, sodass insgesamt längere Messzeiten notwendig sind. Zur Signalverstärkung werden sich bei der ^{13}C -NMR-Spektroskopie deshalb zwei Effekte zunutze gemacht. Zum einen werden die ^{13}C -NMR-Spektren in der Regel als ^1H -entkoppelte Spektren aufgenommen, wodurch die Signalintensität durch den Zusammenfall der einzelnen Signalspitzen eines Multipletts zu einem Singulett erhöht wird.^[195] Daneben findet eine Erhöhung der Signalintensität durch den *Nuclear Overhauser Effect* (NOE) statt, der insbesondere zur Signalerhöhung von denjenigen Kohlenstoffatomen führt, die direkt an Wasserstoffatome verknüpft sind. Quartäre Kohlenstoffatome erfahren demnach kaum Signalerhöhung, da diese nur einen schwachen NOE erfahren und auch die Entkopplung nicht zu einer Signalerhöhung führt.^[195] Zusätzlich ist die Spin-Gitter-Relaxationszeit T_1 (longitudinale Relaxationszeit) von quartären ^{13}C -Kernen im Allgemeinen länger, als von primären, sekundären oder tertiären Kohlenstoffatomen. Um dennoch ausreichend hohe Signalintensitäten für die quartären Kohlenstoffatome von TACT (**78**) in den ^{13}C -NMR-Spektren zu erreichen, sollte der Relaxationsdelay zwischen den einzelnen *Scans* deshalb ausreichend lang gewählt werden. Der Relaxationsdelay sollte das ein- bis fünffache der T_1 -Relaxationszeit betragen. Die Ermittlung der T_1 -Relaxationszeiten der Kohlenstoffatome von TACT (**78**) erfolgt mit Hilfe des *Inversion-Recovery* Experimentes (Pulsprogramm *t1irpg*). Das Experiment besteht zunächst aus einem 180° -Puls zur Inversion der z-Magnetisierung. Nach dem Delay τ erfolgt ein 90° -Puls, um die Magnetisierung zur Detektion in die xy-Ebene zu überführen. Der Delay τ wird während des

Experiments seriell variiert. Über einen *Fit* nach Gleichung (11) kann anschließend die T_1 -Relaxationszeit bestimmt werden.^[59]

$$I(\tau) = I_0 (1 - 2 A e^{-\tau/T_1}) \quad (11)$$

Der Faktor A in Gleichung (11) ist aufgrund der unvollständigen Inversion der Magnetisierung auf die $-z$ -Achse zu Beginn des Experimentes inkludiert.^[59] In Abbildung 104 sind die Ergebnisse von vier beispielhaften Signalen des Analyten TACT (**78**) dargestellt. Die Auswertung erfolgt dabei auf Basis der ermittelten Signalintensitäten.

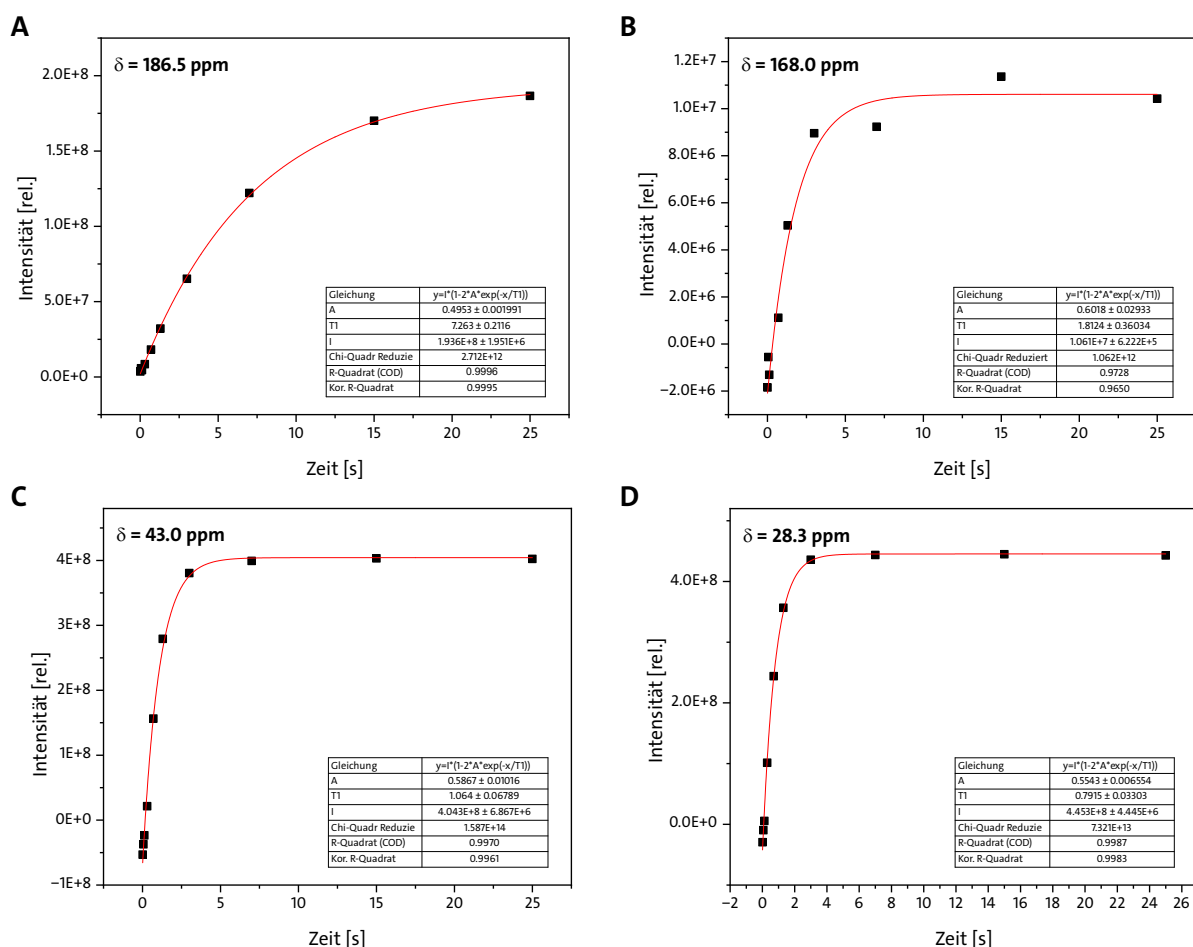


Abbildung 104: Bestimmung der T_1 -Relaxationszeiten von vier beispielhaften ^{13}C -NMR-Signalen des Analyten TACT (**78**) bei 186.5 ppm (A), 168.0 ppm (B), 43.0 ppm (C) und 28.3 ppm (D).

Erwartungsgemäß ergeben sich für die Relaxationszeiten der quartären Kohlenstoffatome C-1' (186.5 ppm) und C-2 (168.0 ppm) mit 7.26 Sekunden und 1.81 Sekunden im Vergleich deutlich längere Relaxationszeiten als für die übrigen Kohlenstoffatome C-2' bis C-6'. In Tabelle 25 sind die Relaxationszeiten aller Kohlenstoffatome des Analyten zusammengefasst.

Tabelle 25: Ermittelte T_1 -Relaxationszeiten in Sekunden aller ^{13}C -NMR-Signale von TACT (**78**) durch das *Inversion Recovery* Experiment und Auswertung der Daten mittels *Fit* nach Gleichung (11).

$\delta_{13\text{C}}$ [ppm]	Zuordnung	T_1 [s]
186.5	C-1'	7.26 ± 0.21
168.0	C-2	1.81 ± 0.36
43.0	C-6'	0.42 ± 0.01
40.2	C-2'	1.06 ± 0.07
31.3	C-5'	0.49 ± 0.01
28.8	C-4'	0.73 ± 0.02
28.3	C-3'	0.79 ± 0.03

Auf Grundlage der Ergebnisse wurde für die NMR-spektroskopische Reaktionsverfolgung ein Relaxationsdelay von 15 Sekunden für die ^{13}C -NMR-Spektren festgelegt. Dies entspricht in etwa dem Doppelten der ermittelten T_1 -Relaxationszeit für C-1' bei einer chemischen Verschiebung von 186.5 ppm, was bedeutet, dass ca. 86.5 % der Magnetisierung bereits relaxiert sind. Dadurch wird einerseits eine hinreichende Relaxation der Kerne gewährleistet, andererseits aber die Dauer der Messungen dennoch in einem vertretbaren Rahmen gehalten.

5.3.3 Reaktionsverfolgung durch Korrelation der Daten

Im Folgenden soll die Reaktion von TACT (**78**) ausgehend von Cyanurchlorid (**75**) durch Umsetzung mit 6-Aminocaprinsäure (**77**) genauer untersucht werden. Das Ziel ist dabei die Untersuchung der zeitlichen Abfolge der Parallel- und Folgereaktionen, wie zuvor in Abbildung 103 gezeigt wurde. Da sich das Edukt, die Zwischenprodukte und die Produkte in den NMR-Spektren nur minimal unterscheiden, ist die Reaktionsverfolgung lediglich auf Grundlage von NMR-Spektren nicht möglich. Die gezielte Synthese der Nebenprodukte, um die Zuordnung der NMR-Signale auf Grundlage der Reinsubstanzen zu ermöglichen, ist ebenfalls nur schwer möglich. Deshalb bietet sich die Korrelation der komplementären NMR-Daten mit MS-Daten an, anhand dessen die Analyten eindeutig unterschieden werden können. Durch die Korrelation der Daten soll auch die Zuordnung der NMR-Signale zu den Analyten aus dem Gemisch heraus ermöglicht werden. Für die Optimierung des Reaktionsverlaufs soll die Temperaturabhängigkeit der dreifachen nucleophilen Substitution überprüft werden. Weiterhin ist insbesondere die

Frage nach dem tatsächlichen Zeitpunkt der Hydrolyse, die zur Bildung eines der Hydroxy-substituierten Nebenprodukte führt, von besonderem Interesse.

Für die NMR-spektroskopische und massenspektrometrische Verfolgung der Reaktion wurde die Synthese von TACT (**78**) in einem Reaktionskalorimeter mit aktiver Manteltemperierung und isothermer Betriebsweise in Form einer dreistufigen Temperaturrampe durchgeführt (siehe Abschnitt 6.10). Der Temperaturverlauf während der Reaktion ist in Abbildung 105 dargestellt.

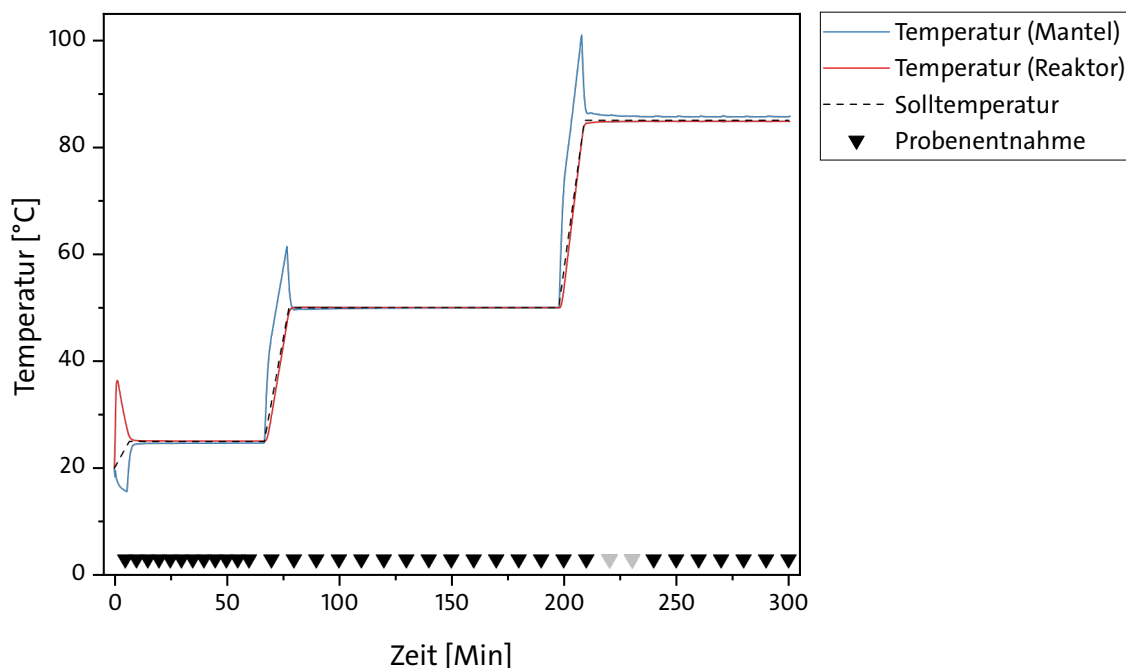


Abbildung 105: Temperaturverlauf der im Reaktionskalorimeter durchgeführten Synthese von TACT (**78**). Die Solltemperatur wurde aktiv über den Reaktormantel gesteuert und ist durch die gestrichelte schwarze Linie gekennzeichnet. Die schwarzen Dreiecke kennzeichnen die Zeitpunkte der Probenentnahme. Die beiden grau markierten Dreiecke entsprechen den Proben 28 und 29, die aufgrund von Signalverschiebungen innerhalb der ^{13}C -NMR-Spektren von der weiteren Auswertung exkludiert wurden.

Zu Beginn der Reaktion steigt die Temperatur durch die Zugabe des festen Cyanurchlorids (**75**) zur Reaktionslösung bestehend aus 6-Aminocapronsäure (**77**) in Natriumhydroxid so schnell an, dass die konstante Temperierung der Reaktionslösung nicht gelingt und die Temperatur um 16 °C ansteigt. Daraufhin erfolgt die Temperierung der Reaktionslösung auf 25 °C für einen Zeitraum von 60 Minuten. Die nächste Temperaturstufe von 50 °C wird für 120 Minuten gehalten, bis anschließend für weitere 90 Minuten auf 85 °C temperiert wird. Während der fünfstündigen Reaktion wurden in regelmäßigem zeitlichen Abstand Proben entnommen.

Innerhalb der ersten Stunde erfolgte die Probenentnahme jeweils alle fünf Minuten und anschließend jeweils alle zehn Minuten, sodass insgesamt 36 Proben entnommen wurden.

Nach entsprechender Probenvorbereitung wurden DI-Massenspektren sowie ^{13}C -NMR-Spektren aufgenommen. Die erhaltenen Daten wurden daraufhin mit Hilfe der *SCORE-metabolite-ID* App ausgewertet. Während der Analyse der Daten fiel auf, dass einzelne Signale in den ^{13}C -NMR-Spektren der Proben 28 und 29 leicht verschoben im Vergleich zu den Signalen in den übrigen Proben waren. Auch eine erneute Messung der Spektren sowie eine erneute Probenvorbereitung der Rückstellproben aus der Reaktionsverfolgung führte nicht zu einer Verbesserung. Aus diesem Grund wurden die Daten dieser beiden Messungen exkludiert, sodass insgesamt 34 Proben zur Analyse zur Verfügung standen. Die 34 Proben werden im Folgenden als Fraktionen benannt.

Da die exakten Signallagen innerhalb der NMR-Spektren bzgl. der chemischen Verschiebung im Verlauf der Fraktionen leicht variieren, wurde ein *Alignment* der Spektren in MestReNova durchgeführt, um diese minimalen Unterschiede in den Signallagen im Verlauf der Proben anzugleichen. Innerhalb der *SCORE-metabolite-ID* App wurde anschließend für die NMR-Spektren ein *Bucketing* mit einer *Bucket*-Größe von 0.02 ppm angewendet. Außerdem wurde zur Minimierung von einzelnen Ausreißern innerhalb der Konzentrationsverläufe eine Datenglättung über vier Datenpunkte, d. h. über vier aufeinanderfolgende Fraktionen (Proben), angewendet. Eine Glättung der Daten erscheint im Fall einer Reaktionsverfolgung besonders sinnvoll, da angenommen werden kann, dass die Konzentrationsverläufe stetigen Entwicklungen folgen und spontane Schwankungen eher in Folge von minimalen Unterschieden bei der Probennahme und Probenvorbereitung sowie in Folge von Messungenauigkeiten auftreten, als dass sie den tatsächlichen Konzentrationsverlauf der beteiligten Stoffe abbilden. In Abbildung 106 sind die ^{13}C -NMR-Spektren in gestaffelter Darstellung über die 34 Fraktionen dargestellt.

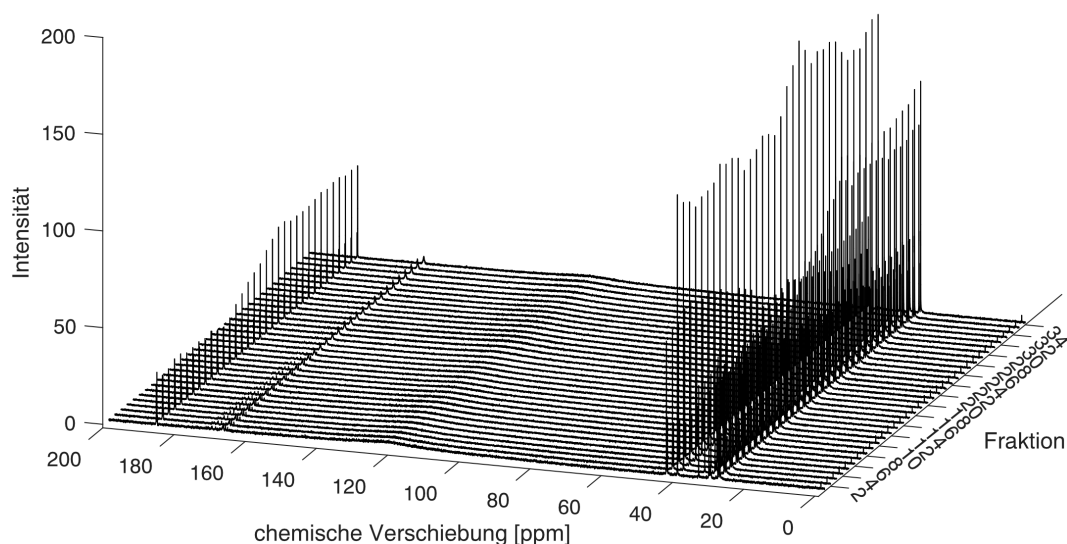


Abbildung 106: Gestaffelte Darstellung der ^{13}C -NMR-Spektren der Proben 1-34 der Synthese von TACT (**78**) im Bereich von 0-200 ppm. Für die NMR-Daten wurde zunächst ein *Alignment* in MestReNova durchgeführt und anschließend innerhalb der *SCORE-metabolite-ID* App ein *Bucketing* mit einer *Bucket*-Größe von 0.02 ppm sowie eine Datenglättung über vier fortfolgende Fraktionen durchgeführt.

Die in Abbildung 106 dargestellten NMR-Spektren zeigen die Intensitätsänderungen der NMR-Signale im Verlauf der Fraktionen durch die Konzentrationszu- oder -abnahme der Produkte, Zwischenprodukte und Edukte. Trotz der auf den ersten Blick übersichtlich zu scheinenden Signale im Verlauf der Reaktion ist eine eindeutige Zuordnung einzelner Signale zu den jeweiligen Analyten nicht ohne weiteres möglich, da die Signale der Analyten teilweise nur minimale Unterschiede in den chemischen Verschiebungen zeigen und teilweise nur als Schultern zu erkennen sind. Dies entspricht den Erwartungen, da zwar einerseits Annahmen über die zu erwartenden Substanzen getroffen werden können, sich die Zwischen- und Nebenprodukte sowie das Produkt strukturell aber nur sehr gering unterscheiden. Durch die Vielzahl der verschiedenen Signale und dem Fehlen von Referenzspektren der Reinsubstanzen ist die Zuordnung ohne weitere analytische Daten so nicht möglich. In Abbildung 107 sind in Analogie die zugehörigen Massenspektren als Verlauf über die Fraktionen dargestellt.

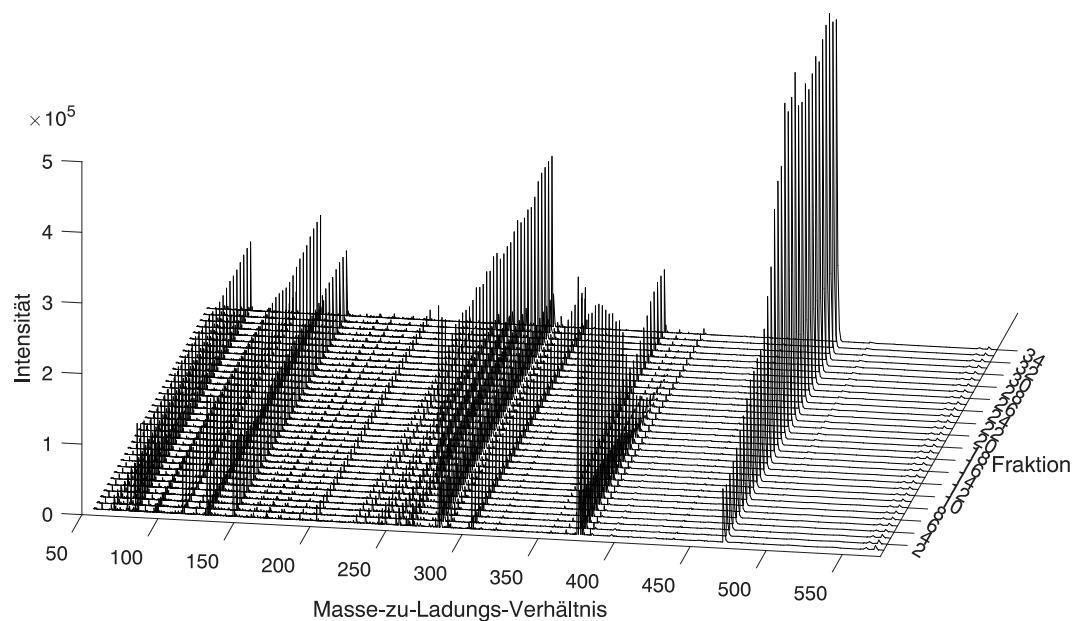


Abbildung 107: Gestaffelte Darstellung der DI-Massenspektren im positiven Ionisationsmodus über die 34 Fraktionen im m/z -Bereich zwischen 50 und 575. Die Daten wurden analog zu den NMR-Spektren über vier Datenpunkte geglättet.

Innerhalb der Massenspektren sind ebenfalls eindeutige Signalzunahmen und Signalabnahmen im Verlauf der Fraktionen ersichtlich. Aufgrund der bereits vorhandenen Kenntnis über die möglichen Reaktionspfade entsprechend Abbildung 103 sind die Analyten von Interesse bereits bekannt, sodass die MS-Signale derjenigen Analyten gezielt untersucht werden können. In Abbildung 108 sind die relevanten Analyten, die in den obigen Massenspektren detektiert werden können, aufgeführt.

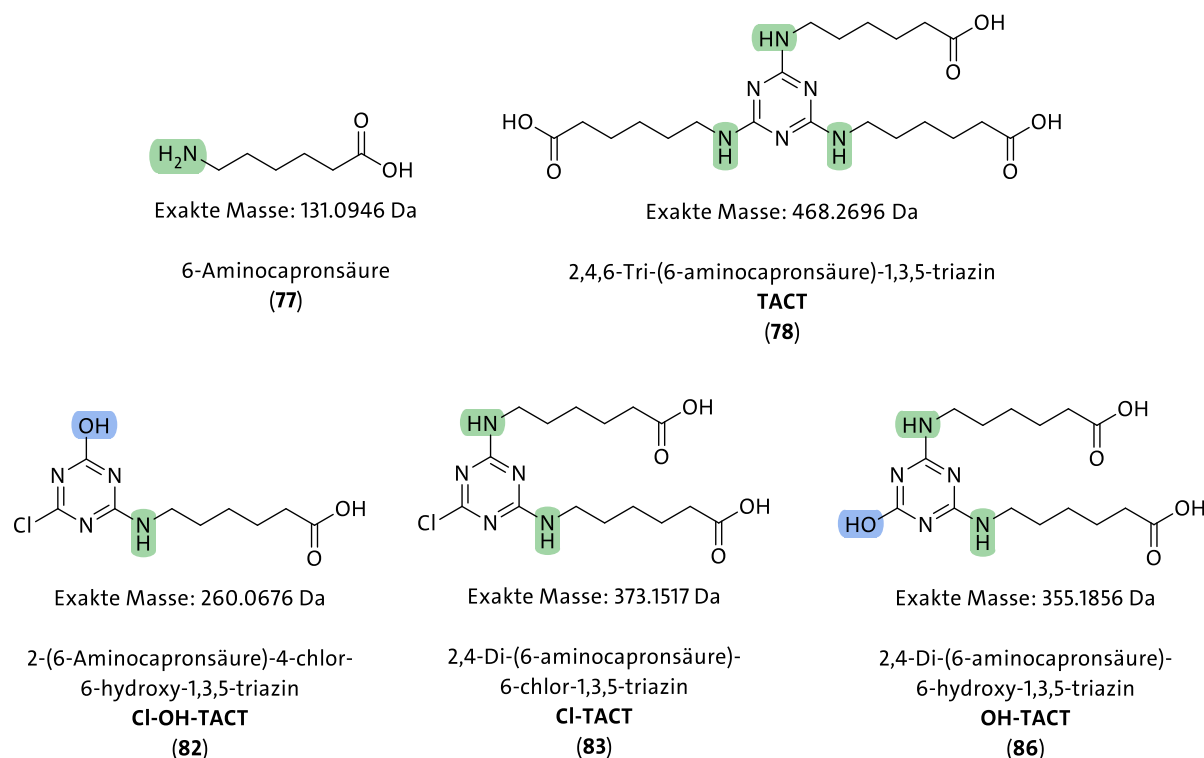


Abbildung 108: Übersicht der in den Massenspektren detektierten Analyten während der Reaktionsverfolgung der Synthese von TACT (78) mit der Angabe der jeweiligen monoisotopischen Masse sowie der im Rahmen dieser Arbeit verwendeten Kurzbenennung der Analyten.

Dazu zählen das Edukt 6-Aminocapronsäure (77) und das dreifach-substituierte Produkt TACT (78) sowie daneben das einfach-substituierte und einfach-hydroxylierte Zwischenprodukt 2-(6-Aminocapronsäure)-4-chlor-6-hydroxy-1,3,5-triazin (Cl-OH-TACT) (82), das zweifach-substituierte Zwischenprodukt 2,4-Di-(6-aminocapronsäure)-6-chlor-1,3,5-triazin (Cl-TACT) (83) und das zweifach-substituierte und einfach-hydroxylierte Produkt 2,4-Di-(6-aminocapronsäure)-6-hydroxy-1,3,5-triazin (OH-TACT) (86). Das zweifach-hydrolysierte und einfach-substituierte Nebenprodukt (85) ist in den Massenspektren nicht zu detektieren, sodass an dieser Stelle davon ausgegangen wird, dass dieses Nebenprodukt nicht oder nur in sehr geringem Maße gebildet wird.

Aufgrund der strukturellen Ähnlichkeit der TACT-Derivate ist die eindeutige Zuordnung einzelner NMR-Signale aus dem Gemisch heraus ohne Weiteres kaum möglich. Die Korrelation der NMR- und MS-Daten würde aber möglicherweise die Detektion von Marker-Signalen in den NMR-Spektren erlauben, anhand derer z.B. in folgenden Untersuchungen der Umsatz zu TACT (78) und die Menge der gebildeten Nebenprodukte untersucht werden kann. Es wäre zu erwarten, dass die TACT-Derivate sich anhand der EDC-Verläufe im aromatischen Bereich der

^{13}C -NMR-Spektren unterscheiden, da diese ^{13}C -Kerne durch die unterschiedlichen Substituenten eine stark unterschiedliche chemische Umgebung aufweisen. Die Intensität der Signale der aromatischen ^{13}C -Kerne ist allerdings sehr gering, sodass die Signale teilweise nicht vom Rauschen unterschieden werden können und sich deshalb nicht in allen Fällen für die Korrelationsanalyse eignen. Stattdessen ist womöglich über die ^{13}C -NMR-Signale der Carbonylkohlenstoffatome im Bereich von ca. 186 ppm eine deutlichere Unterscheidung der TACT-Derivate möglich. Die Signale der Carbonylkohlenstoffatome liegen zwar nahe beieinander, sind aber dennoch voneinander separiert. In Tabelle 26 sind die Korrelationskoeffizienten zwischen den EMCs und den EDCs in diesem Bereich der chemischen Verschiebung angegeben.

Tabelle 26: Korrelationskoeffizienten zwischen den EMCs der protonierten Molekülonen von TACT (**78**), Cl-TACT (**83**), OH-TACT (**86**), Cl-OH-TACT (**82**) und 6-Aminocaprinsäure (**77**) zu den EDCs der Carbonylsignale im Bereich von ca. 186 ppm berechnet über alle Fraktionen der Reaktionsverfolgung.

		EMCs [m/z]				
		469.28 TACT (78)	374.16 Cl-TACT (83)	356.19 OH-TACT (86)	261.07 Cl-OH-TACT (82)	132.10 6-Amino- caprinsäure (77)
EDCs [ppm]	186.37	-0.9733	0.9602	-0.7797	0.6989	0.9418
	186.45	-0.9737	0.9664	-0.7856	0.7107	0.9511
	186.51	0.9695	-0.9511	0.8206	-0.7383	-0.9337
	186.59	0.7932	-0.7840	0.9829	-0.9633	-0.8448
	186.67	-0.9564	0.9482	-0.7714	0.6962	0.9355

Wie der Tabelle 26 zu entnehmen ist, ist für das Nebenprodukt OH-TACT (**86**) die eindeutige Zuordnung des Signals bei 186.59 ppm möglich. Ebenso ist das Signal bei 186.51 ppm eindeutig dem Produkt TACT (**78**) zuzuordnen. Zur visuellen Darstellung der Korrelationen aller EDCs mit einem Korrelationskoeffizienten > 0.95 zu dem EMC bei $m/z = 469.28$ ($[\text{M}+\text{H}]^+$ von TACT (**78**)) ist in Abbildung 109 A das auf dieser Grundlage erhaltene berechnete ^{13}C -NMR-Spektrum von TACT (**78**) gezeigt. Abbildung 109 B zeigt zum Vergleich das Referenzspektrum von TACT (**78**).

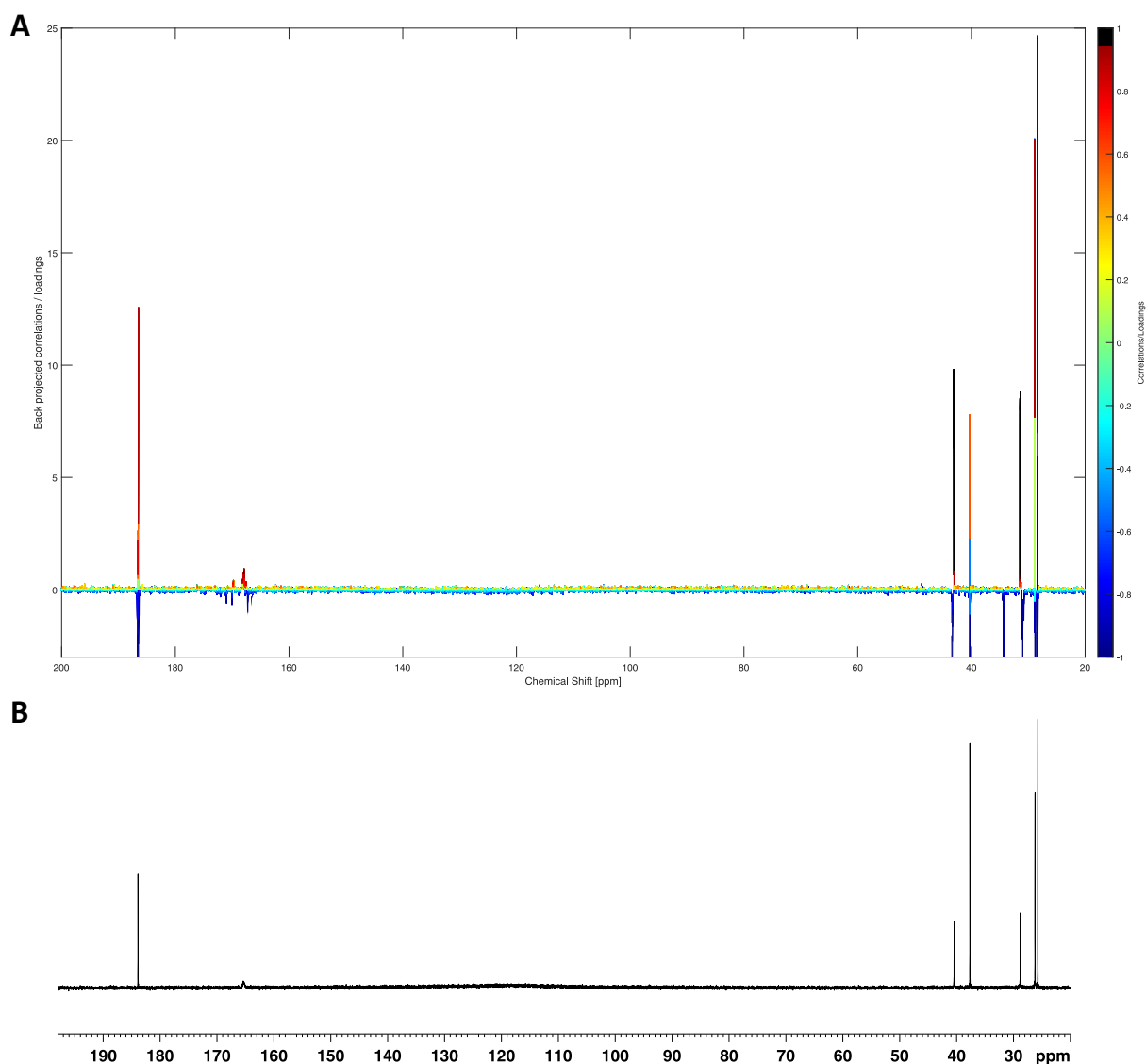


Abbildung 109: A: Berechnetes ^{13}C -NMR-Spektrum durch die semi-automatische Korrelation von EDCs > 0.95 zu dem EMC von TACT (**78**) bei $m/z = 469.28$ über alle 34 Fraktionen der Reaktionsverfolgung. B: Referenzspektrum von TACT (**78**) im Bereich von 20-200 ppm.

Ein Vergleich des Referenzspektrums von TACT (**78**) zu dem berechneten Korrelationsspektrum in Abbildung 109 zeigt, dass alle Signale des Analyten über die Korrelation detektierbar sind. Die Signale des Analyten TACT (**78**) weisen einen hohen Korrelationskoeffizienten zu dem EMC des protonierten Moleküls auf.

Für Cl-OH-TACT (**82**) ist auf diese Weise allerdings keine eindeutige Zuordnung eines ^{13}C -NMR-Signals möglich. Ein möglicher Grund ist die im Vergleich zu TACT (**78**), Cl-TACT (**83**) und 6-Aminocaprinsäure (**77**) geringere Konzentration des Analyten im Verlauf der Reaktion. Im Allgemeinen sind die NMR-Signale der strukturell ähnlichen Moleküle teilweise überlagert und

insbesondere die gering konzentrierten Analyten teilweise nur als Schultern der intensiveren Signale zu erkennen. Das *Bucketing* mit einer *Bucket*-Größe von 0.02 ppm führt zusätzlich dazu, dass die Schultern nicht mehr als solche detektiert werden, sondern vollständig mit dem intensiveren Signal überlagern. Aufgrund von Unterschieden in der genauen chemischen Verschiebung der Signalspitzen im Verlauf der Proben kann auf das *Bucketing* jedoch nicht verzichtet werden. Der Tabelle 26 ist außerdem zu entnehmen, dass sowohl CI-TACT (**83**) als auch 6-Aminocaprinsäure (**77**) hohe Korrelationskoeffizienten zu den Carbonylsignalen bei 186.37 ppm, 186.45 ppm und 186.67 ppm aufweisen. Allein auf der Basis der Korrelationskoeffizienten ist hier also keine eindeutige Zuordnung dieser Signale zu einem der Moleküle möglich. Stattdessen ist an dieser Stelle die Zuordnung lediglich durch Zuhilfenahme der Spektren und Vergleich der relativen Intensitäten möglich. In Abbildung 110 A sind zu diesem Zweck die NMR-Spektren im Verlauf der Reaktion im Bereich von 186 bis 187 ppm gezeigt. Die Verläufe der EDCs sind farblich hervorgehoben. Zusätzlich sind die EDC-Verläufe in Abbildung 110 B in zweidimensionaler Darstellung gezeigt.

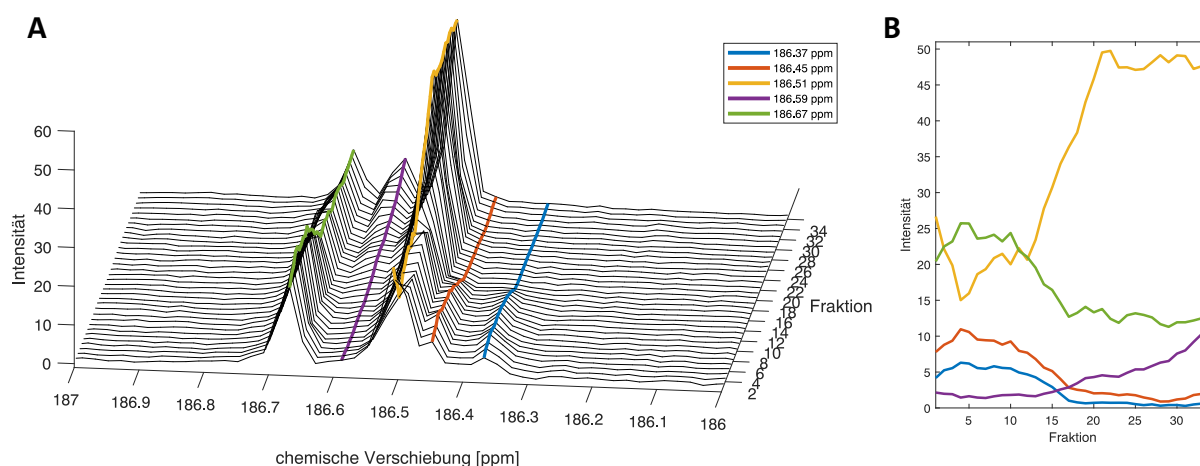


Abbildung 110: A: Gestaffelte Darstellung der ¹³C-NMR-Spektren zur Reaktionsverfolgung der Synthese von TACT im Bereich von 186 bis 187 ppm. Die EDCs sind farblich entsprechend der Legende hervorgehoben. B: Zweidimensionale Darstellung der gleichen EDCs.

Es fällt auf, dass der EDC bei einer chemischen Verschiebung von 186.67 ppm (grüner EDC in Abbildung 110) einen abnehmenden Verlauf aufweist, jedoch zum Ende der Reaktionszeit nicht vollständig auf null sinkt. Dahingegen fallen die EDCs bei den chemischen Verschiebungen 186.37 ppm und 186.45 ppm (blauer und orangefarbener EDC in Abbildung 110) bereits ab Fraktion 17 auf null ab. Zusätzlich ist die absolute Intensität der letzteren beiden EDCs im Vergleich deutlich geringer. Die 6-Aminocaprinsäure (**77**) wurde während der Reaktion im

Überschuss (3.2 Äquivalente) eingesetzt. Somit ist zu erwarten, dass das Edukt nicht vollständig verbraucht wird, sondern zum Ende der Reaktion ein Plateau erreicht. Gleichzeitig ist zu erwarten, dass Cl-TACT (**83**) lediglich als Zwischenprodukt auftritt und im Fortschritt der Reaktion weiter umgesetzt wird. Somit kann an dieser Stelle festgehalten werden, dass alleinig auf Basis der Korrelationskoeffizienten keine eindeutige Zuordnung der Signale getroffen werden kann. Unter Zuhilfenahme aller vorliegenden Informationen kann jedoch geschlussfolgert werden, dass das Signal bei 186.67 ppm der 6-Aminocapronsäure (**77**) und die Signale bei 186.37 ppm und 186.45 ppm dem Cl-TACT (**83**) zugeordnet werden können. Die semi-automatische Berechnung aller EDCs, die einen Korrelationskoeffizienten > 0.95 zu dem EMC von Cl-TACT (**83**) bei $m/z = 374.16$ aufweisen, ist in der folgenden Abbildung 111 gezeigt.

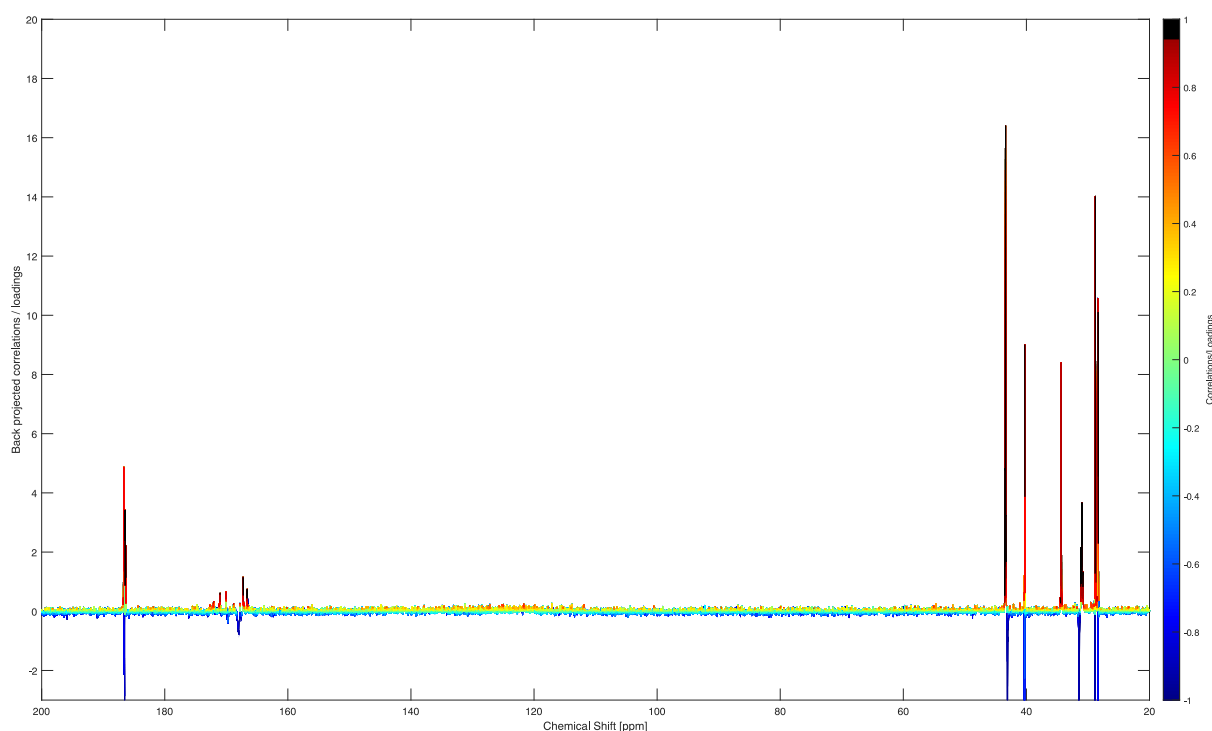


Abbildung 111: Berechnetes ^{13}C -NMR-Spektrum von Cl-TACT (**61**) durch die graphische Darstellung nach semi-automatischer Korrelation aller EDCs mit einem Korrelationskoeffizienten > 0.95 zu dem EMC des protonierten Moleküls bei $m/z = 374.16$.

Wie bereits durch die in Tabelle 26 gezeigten Korrelationskoeffizienten zwischen den EMCs der Analyten und den EDCs im Carbonylbereich ersichtlich wurde, werden auch in der visuellen Darstellung aller korrelierender EDCs in Abbildung 111 teilweise falsch-positive Korrelationen erhalten, die auf den qualitativ nicht zu unterscheidenden Verlauf von Cl-TACT (**83**) und 6-Aminocapronsäure (**77**) zurückzuführen sind. Nichtsdestotrotz zeigt das berechnete ^{13}C -NMR-

Spektrum in Abbildung 111 beispielsweise im aromatischen Bereich eindeutig die Signale von Cl-TACT (**83**), da das Edukt 6-Aminocapronsäure aufgrund der fehlenden Aromatizität in diesem Bereich des Spektrums keine Signale aufweisen kann.

Zur Untersuchung der ablaufenden Reaktionen und deren zeitlicher Verläufe sind in Abbildung 112 die normierten EMC-Verläufe der vier TACT-Derivate anhand der protonierten Molekulationen-Signale in zweidimensionaler Darstellung gezeigt. Zusätzlich sind die Temperaturstufen während der Reaktion dargestellt.

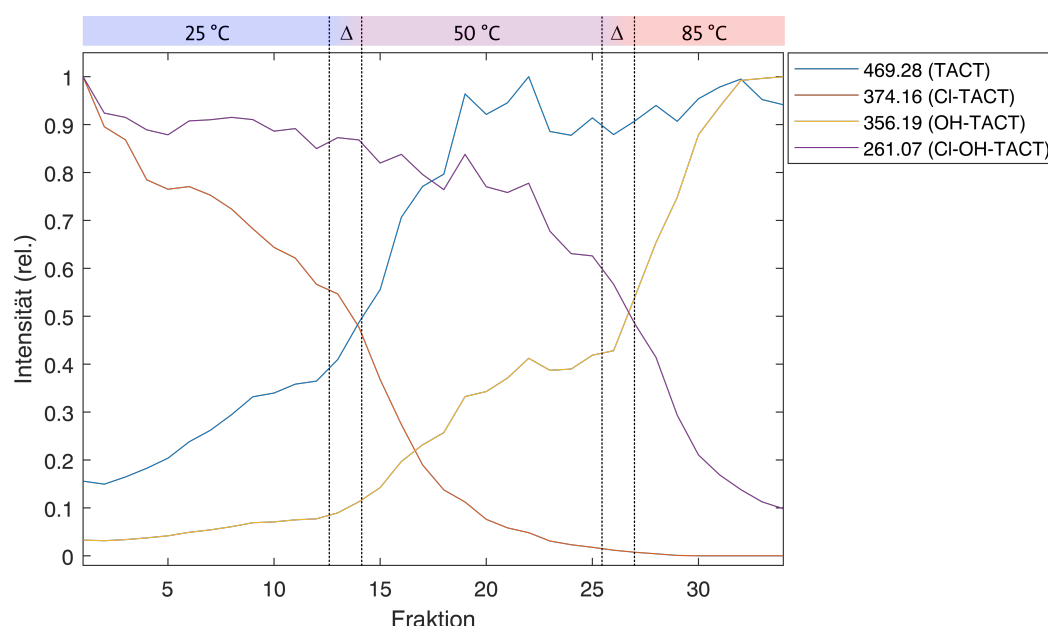


Abbildung 112: Normierte Verläufe der EMCs der vier Analyten TACT (**78**) ($m/z = 469.28$ ($[M+H]^+$)) in blau, Cl-TACT (**83**) ($m/z = 374.16$ ($[M+H]^+$)) in orange, OH-TACT (**86**) ($m/z = 356.19$ ($[M+H]^+$)) in gelb und Cl-OH-TACT (**82**) ($m/z = 261.07$ ($[M+H]^+$)) in violett in den Proben 1-34 der Reaktionsverfolgung. Die gestrichelten senkrechten Linien verweisen auf den über den EMCs angegebenen Temperaturverlauf während der Reaktion.

Anhand der visuellen Auswertung der EMC-Verläufe können bereits vielerlei Rückschlüsse auf die zeitliche Abfolge der ablaufenden Reaktionen gezogen werden. So zeigt sich anhand des EMCs des protonierten Molekulations von TACT (**78**) bei $m/z = 469.28$, dass bereits in der ersten Probe, die nach einer Reaktionszeit von 5 Minuten entnommen wurde, das dreifach-substituierte Produkt vorliegt. Würde die Reaktion nach dem zuvor in Abbildung 101 postulierten temperaturabhängigen Dreischritt ablaufen, so sollte das Produkt erst bei einer erhöhten Temperatur von 85 °C, d. h. in diesem Fall ab Fraktion 27, gebildet werden. Stattdessen wird bereits bei einer Temperatur von 25 °C direkt zu Beginn der Reaktion dreifach-substituiertes Produkt (TACT) (**78**) erhalten. Im weiteren Verlauf der Reaktion nimmt die

gebildete Menge des TACT (**78**) im Temperaturbereich von 50 °C in höherem Maße zu. Bereits in Fraktion 19, d.h. nach einer Reaktionszeit von 2 Stunden und 10 Minuten, ist allerdings die maximal gebildete Menge von TACT (**78**) erreicht. Die Temperaturerhöhung auf 85 °C im weiteren Verlauf der Reaktion hat demzufolge keinen Einfluss mehr auf die (weitere) Bildung von TACT (**78**).

In Ergänzung zu den in Abbildung 112 gezeigten EMC-Verläufen sind in der Tabelle 27 die Korrelationskoeffizienten zwischen den jeweiligen EMCs angegeben.

Tabelle 27: Korrelationskoeffizienten zwischen den EMCs von TACT (**78**), Cl-TACT (**83**), OH-TACT (**86**) und Cl-OH-TACT (**82**) berechnet über alle 34 Fraktionen mittels der *SCORE-metabolite-ID* App.

EMC [<i>m/z</i>]	469.28 TACT (78)	374.16 Cl-TACT (83)	356.19 OH-TACT (86)	261.07 Cl-OH-TACT (82)
469.28 TACT (78)	1.000	−0.9894	0.8106	−0.7266
374.16 Cl-TACT (83)	−0.9894	1.000	−0.8037	0.7411
356.19 OH-TACT (86)	0.8106	−0.8037	1.000	−0.9822
261.07 Cl-OH-TACT (82)	−0.7266	0.7411	−0.9822	1.000

Der hohe Antikorrelationskoeffizient von −0.9894 zwischen dem EMC von TACT (**78**) und dem EMC von Cl-TACT (**83**) (siehe Tabelle 27) zeigt die direkte Abhängigkeit der beiden Substanzen voneinander in dieser Reaktion. Die Bildung von TACT (**78**) geht direkt mit dem Verbrauch von Cl-TACT (**83**) einher, wie es entsprechend den dargestellten Folge- und Parallelreaktionen in Abbildung 103 zu erwarten ist. Weiterhin zeigt diese starke Antikorrelation zwischen dem EMC von TACT (**78**) und dem EMC von Cl-TACT (**83**) aber auch, dass Cl-TACT (**83**) im Reaktionsverlauf nicht oder nur in sehr geringem Maße zu dem Nebenprodukt OH-TACT (**86**) hydrolysiert wird. Der EMC von OH-TACT (**86**) weist gleichermaßen einen hohen Antikorrelationskoeffizienten in Höhe von −0.9822 zu dem EMC von Cl-OH-TACT (**82**) auf. Dies lässt also darauf schließen, dass OH-TACT (**86**) hauptsächlich ausgehend von Cl-OH-TACT (**82**) gebildet wird, d.h. zu diesem Zeitpunkt wird der Chlorid-Substituent des Cl-OH-TACT (**82**) ein weiteres Mal durch 6-Aminocaprinsäure (**77**) substituiert. Ein Blick in die zeitlichen Verläufe in Abbildung 112 zeigt, dass ein starker Anstieg der Steigung des EMCs von OH-TACT (**86**) zwischen der Temperaturstufe bei 50 °C zu der Temperaturstufe bei 85 °C erkennbar ist. Das Nebenprodukt

OH-TACT (**86**) wird demnach hauptsächlich zum Ende der Reaktion bei einer Reaktionstemperatur von 85 °C gebildet. Die Konzentration von Cl-OH-TACT (**82**) ist dahingegen zu Beginn der Reaktion im Zeitrahmen dieser Reaktionsverfolgung maximal. Die Konzentration von Cl-OH-TACT (**82**) nimmt zu Beginn nur leicht ab. Eine Steigerung der Konzentrationsabnahme findet erst im Verlauf der Reaktion statt. Dabei nimmt die negative Steigung der Kurve in gleicher Weise wie die Bildungskurve von OH-TACT (**86**) im Temperaturbereich von 50 °C leicht zu und im Temperaturbereich von 85 °C weitaus stärker zu. Zu diesem Zeitpunkt wird demnach die im Überschuss eingesetzte 6-Aminocapronsäure (**77**) durch die Substitution des Chlorid-Substituenten in Cl-OH-TACT (**82**) zur Bildung von OH-TACT (**86**) weiter verbraucht. Aus den oben aufgeführten Beobachtungen kann insgesamt geschlossen werden, dass die Parallelreaktion der Hydrolyse bereits zu Beginn der gesamten Reaktion stattfindet. Die Frage, ob die Hydrolyse direkt aus dem Edukt Cyanurchlorid (**75**) oder nach bereits einfacher Substitution eines Chlorid-Substituenten durch 6-Aminocapronsäure (**77**) erfolgt, kann an dieser Stelle nicht eindeutig geklärt werden. Aufgrund der Tatsache, dass Cyanurchlorid (**75**) als Feststoff in die Reaktionslösung eingebracht wird und sich erst in Folge von mindestens einer Substitutionsreaktion löst, kann die Erstsitution analytisch auf diese Weise nicht erfasst werden. Der Zeitpunkt der ersten Probenahme wird durch den Zeitpunkt der vollständig homogenen Lösung bestimmt. Insgesamt kann jedoch festgehalten werden, dass weder ein 6-Aminocapronsäure-Substituent noch die Chlorid-Substituenten des Cl-TACT (**83**) oder Cl-OH-TACT (**82**) im Laufe der Reaktion wieder hydrolysiert werden.

Die vorgestellten Ergebnisse zeigen, dass die Korrelation von komplementären NMR- und MS-Daten mittels *SCORE-metabolite-ID* bei der Aufklärung von Reaktionsverläufen einen wertvollen Beitrag leisten kann. Bei der Untersuchung der Reaktion von Cyanurchlorid (**75**) zu TACT (**78**) konnte der zeitliche Konzentrationsverlauf der bedeutenden Zwischen- und Nebenprodukte nachvollzogen werden und damit das System aus Parallel- und Folgereaktionen detailliert analysiert werden. Die Ergebnisse zeigen, dass sich die Selektivität der Reaktion bereits in den ersten Minuten der Reaktionszeit ergibt und die vereinfachte Annahme der thermischen Kontrolle des Substitutionsgrades des Cyanurchlorids unter den gewählten Reaktionsbedingungen nicht oder nur bedingt dem realen Reaktionsverlauf entspricht. Dabei ist zu beachten, dass das gleiche Haupt- oder Nebenprodukt über unterschiedliche Reaktionswege zugänglich ist und sich daher diese Ergebnisse nicht aus der finalen Reaktionslösung ableiten lassen. Aufgrund der strukturellen Ähnlichkeit der Haupt-, Zwischen- und Nebenprodukte fällt

eine manuelle Zuordnung von charakteristischen Signalen zunächst schwer und konnte durch den Einsatz von *SCORE-metabolite-ID* deutlich vereinfacht werden. Dabei muss allerdings berücksichtigt werden, dass nicht alle Signale der NMR-Spektren mit ausreichender Genauigkeit eindeutig einer Substanz zugeordnet werden konnten.

In Hinblick auf die allgemeine Anwendbarkeit von *SCORE-metabolite-ID* zur Reaktionsverfolgung müssen verschiedene Kriterien beachtet werden, welche den Anwendungsbereich von *SCORE-metabolite-ID* teilweise einschränken. Aufgrund der Tatsache, dass die Unterschiede in den Konzentrationsverläufen die Basis der Korrelation bilden, ist die Anwendung von *SCORE-metabolite-ID* nicht bei allen Reaktionssystemen sinnvoll. Eine Einschränkung besteht z. B. in Fällen, in denen die Konzentration von mehreren Stoffen zeitgleich einen quasistationären Zustand erreichen und sich die Konzentrationsverläufe in Folge dessen nicht unterscheiden lassen, wie es u. a. bei Katalysezyklen der Fall sein kann. Außerdem folgen die Konzentrationsprofile in einer Reaktionsverfolgung häufig einem stetigen Verlauf. Wenn die Konzentration von zwei unterschiedlichen Substanzen im Verlauf der Reaktion zunehmen, besteht das Risiko, dass sich die Korrelationskoeffizienten einzelner Signale nicht eindeutig zuordnen lassen, während dies z. B. bei zeitlich verschobenen GAUß-förmigen Konzentrationsverläufen leichter fällt. In Folge der genannten und möglicherweise weiterer Einschränkungen muss die Anwendbarkeit von *SCORE-metabolite-ID* im Einzelfall geprüft werden.

5.4 Entwicklung von SCORE-metabolite-ID² basierend auf 2D HSQC-NMR-Spektren

Im Rahmen von *Metabolomics*-Studien werden in den meisten Fällen eindimensionale ¹H-NMR-Daten verwendet, die sich insbesondere in Bezug auf die vergleichsweise kurze Messdauer als vorteilhaft erweisen.^[196] Bei der Analyse komplexer Gemische treten in eindimensionalen ¹H-NMR-Spektren jedoch erhebliche Signalüberlagerungen auf, die die Analyse und insbesondere die Quantifizierung und Identifizierung von einzelnen Metaboliten erschweren. Bei der Verwendung zweidimensionaler NMR-Spektren werden Signalüberlagerungen durch die zusätzliche Frequenzdomäne minimiert. Im Fall von 2D HSQC-Daten werden außerdem zusätzliche Informationen über die direkte Konnektivität zwischen ¹H- und ¹³C-Kernen über eine Bindung erhalten. Dieser Informationsgehalt birgt einen großen Vorteil im Zuge der Identifizierung von Metaboliten. Die größere Dispersion der Signale bzgl. der ¹³C-Domäne führt nicht nur zu geringerer Signalüberlagerung, sondern bietet auch weitere Vorteile. Beispielsweise gestaltet sich der Vergleich von gemessenen NMR-Spektren mit Datenbank-Spektren bei Verwendung der ¹³C-Verschiebungen als vorteilhaft, da diese eine höhere Genauigkeit sowie Reproduzierbarkeit aufweisen und somit ein effizienterer Datenbankvergleich möglich ist.^[197,198] Auch die auf artifiziellen neuronalen Netzwerken basierte Methode SMART 2.0 (b), die bereits in Abschnitt 3.3.2.4 erläutert wurde, beruht auf der Analyse von 2D HSQC-Daten.^[90,91]

Aufbauend auf das generelle Konzept von *SCORE-metabolite-ID* wurde deshalb die Methode auf die Verwendung von 2D HSQC-NMR-Daten erweitert. In Analogie zu dem EDC wird der Begriff *Extracted Double Delta Chromatogram* (EDDC) eingeführt. Ein EDDC wird durch die Auftragung der Signalintensität eines HSQC-Korrelationssignals, d. h. bei jeweils einer spezifischen ¹H- und ¹³C-Verschiebung, gegen die Zeitdomäne der chromatographischen Separation, d. h. die konsekutiven Fraktionen, generiert. Schematisch ist dies in Abbildung 113 gezeigt.

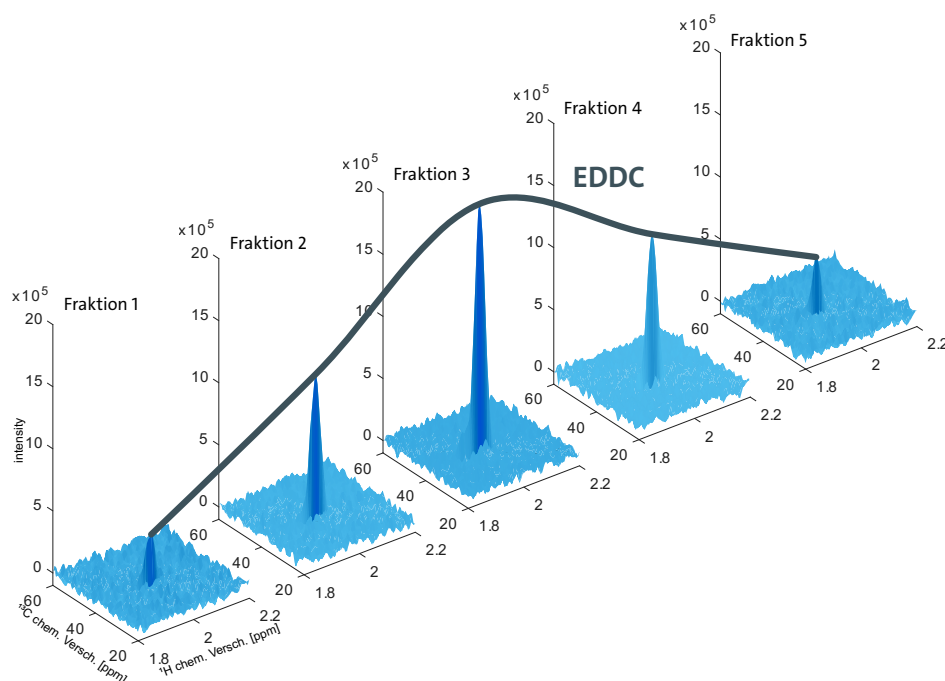


Abbildung 113: Schematische Darstellung der Bildung eines EDDCs (*Extracted Double Delta Chromatogram*). Ein EDDC entspricht dem chromatographischen Signalamplitudenprofil eines spezifischen HSQC-Korrelationssignals über die gewählte Anzahl an Fraktionen.

Durch die Berechnung von Korrelationskoeffizienten zwischen EDDCs untereinander, als auch zwischen EDDCs und EMCs aus den Massenspektren können somit HSQC-Korrelationssignale bestimmten Masse-zu-Ladungs-Verhältnissen zugeordnet werden, was die Identifizierung von Metaboliten vereinfacht. Die Analyse der Daten und die Berechnung von Korrelationskoeffizienten kann über die MATLAB-Applikation *SCORE-metabolite-ID*² erfolgen. Die Entwicklung der Applikation ist unter folgendem Abschnitt 5.4.1 erläutert.

5.4.1 Entwicklung der graphischen Benutzeroberfläche

Die MATLAB-Applikation *SCORE-metabolite-ID*² wurde zur Analyse von HSQC- und MS-Daten entwickelt. Das Grundkonzept basiert auf der bereits zuvor erläuterten Methode *SCORE-metabolite-ID*. In Abbildung 114 ist eine Übersicht der graphischen Benutzeroberfläche der App gezeigt.

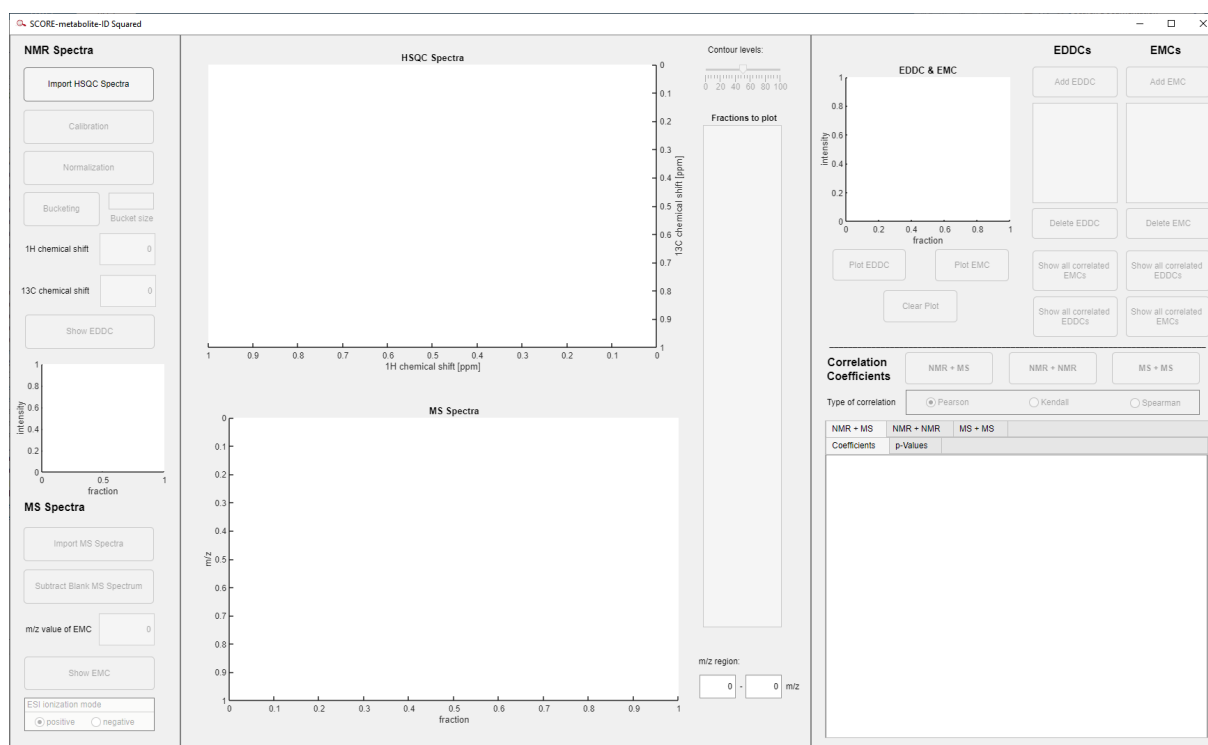


Abbildung 114: Übersicht der graphischen Benutzeroberfläche der MATLAB App *SCORE-metabolite-ID²* zur Korrelation von 2D-HSQC-NMR-Daten und MS-Daten über die dritte Dimension der chromatographischen Separation.

Der grundsätzliche Aufbau der App ist analog zu dem der *SCORE-metabolite-ID* App, d.h. im linken Teil findet der Import und die Prozessierung der Daten statt, im mittleren Teil werden die Daten visualisiert und können graphisch analysiert werden und im rechten Teil findet die individuelle sowie semi-automatische Korrelation der Daten statt. In den folgenden Unterkapiteln sind die Besonderheiten der *SCORE-metabolite-ID²* App erläutert. Der Import der MS-Daten, die grundsätzliche Berechnung der Korrelationskoeffizienten sowie die Fraktionsauswahl unterscheiden sich nicht zur bereits zuvor vorgestellten *SCORE-metabolite-ID* App und werden deshalb mit Verweis auf die Kapitel 5.1.1.2 bis 5.1.1.4 an dieser Stelle nicht erneut erläutert.

5.4.1.1 Import und Prozessierung der HSQC-Daten

Nach Anklicken der Schaltfläche „*Import HSQC Spectra*“ kann der Ordner, in dem sich die HSQC-Daten aller Fraktionen im BRUKER Format befinden, ausgewählt werden. Die Daten werden unter Zuhilfenahme der Funktion „*rbnmr*“ importiert.^[127] Im Gegensatz zu eindimensionalen ¹H-NMR-Daten, die als Vektor dargestellt werden können, liegen die zweidimensionalen HSQC-Daten

entsprechend als Matrix vor. Ein HSQC-Spektrum wird im Regelfall als Konturplot dargestellt, wobei die ^1H - und ^{13}C -Verschiebung im zweidimensionalen Plot aufgetragen und die Intensitäten als dritte Dimension durch Konturlinien dargestellt werden. Diese Art der dreidimensionalen Datenmenge wird bei der *SCORE-metabolite-ID²* Methode über die chromatographische Fraktionierung durch eine vierte Dimension erweitert. Zur visuellen Darstellung der konsekutiven HSQC-Spektren wäre deshalb prinzipiell eine gestaffelte Darstellung von Konturplots denkbar, jedoch ist diese Art der Auftragung rechenintensiv und dementsprechend vergleichsweise langsam. Die zur Analyse benötigten Amplitudenprofile der Signale sind außerdem nicht direkt visuell erkennbar. Deshalb wurde die graphische Darstellung innerhalb der *SCORE-metabolite-ID²* App insofern vereinfacht, als dass jeweils die Summe der HSQC-Korrelationssignale der Fraktionen gebildet und dargestellt wird. Dies ist schematisch in Abbildung 115 gezeigt. Zur visuellen Darstellung werden die Intensitätswerte der HSQC-Spektren der Fraktionen über die jeweiligen Fraktionen hinweg summiert und kombiniert dargestellt. In Abbildung 115 sind die summierten Intensitätswerte zur Veranschaulichung in den jeweiligen Farben der ursprünglichen Fraktionen gekennzeichnet. Ausgehend von dem summierten HSQC-Spektrum können innerhalb der *SCORE-metabolite-ID²* App die EDDC-Verläufe generiert werden. Dies ist in Abbildung 115 schematisch anhand des violett hinterlegten HSQC-Korrelationssignals bei 2.18 ppm/40.2 ppm gezeigt. Der EDDC-Verlauf in Abbildung 115 unten rechts entspricht also dem Signalamplitudenprofil dieses ^1H - ^{13}C -Korrelationssignals über die Fraktionen. Wie bereits in der *SCORE-metabolite-ID* App können auch in der *SCORE-metabolite-ID²* App die relevanten Fraktionen für die jeweilige Analyse individuell ausgewählt werden. Die dargestellte Summe der HSQC-Spektren entspricht den zum jeweiligen Zeitpunkt ausgewählten Fraktionen.

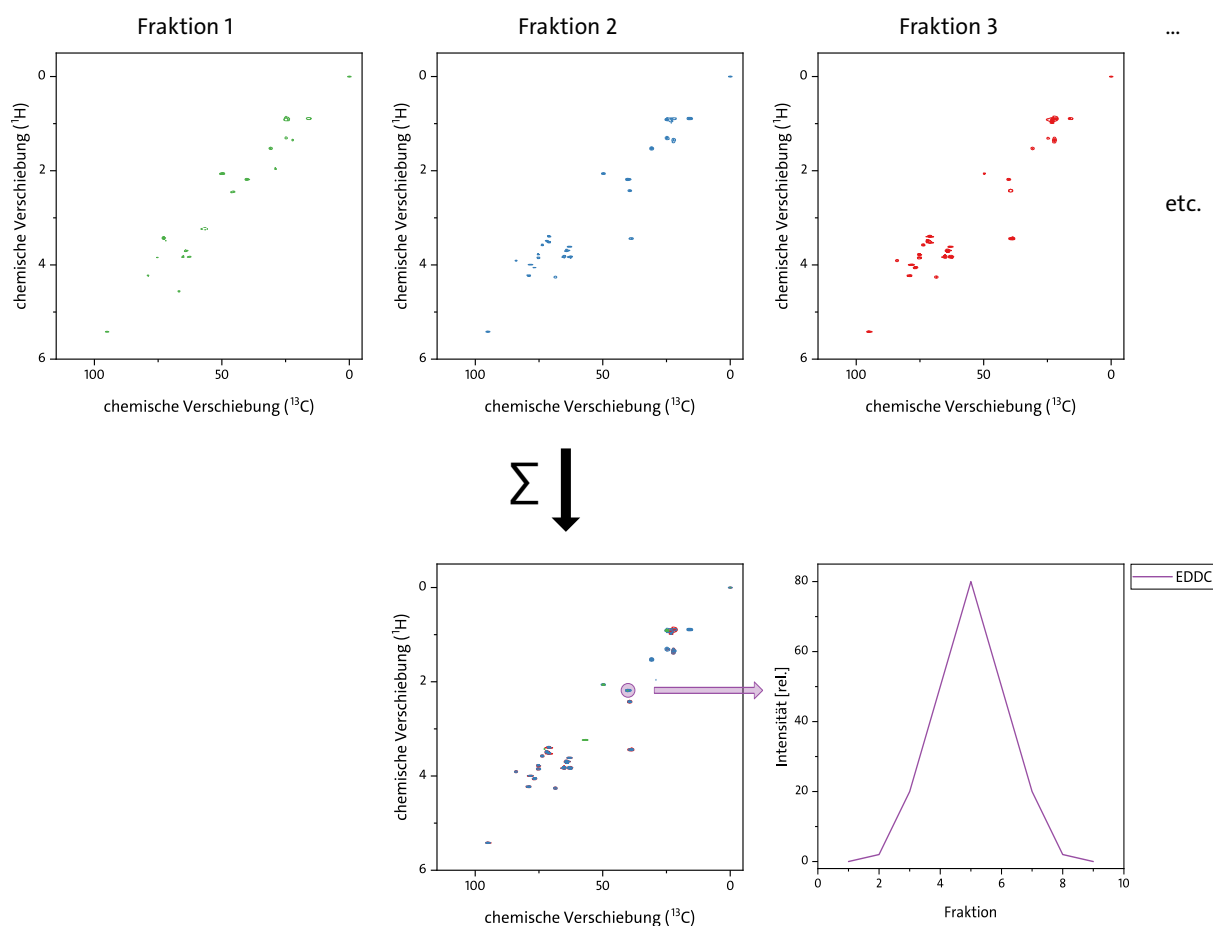


Abbildung 115: Nach Aufnahme von HSQC-Spektren der einzelnen Fraktionen ergibt sich prinzipiell eine vierdimensionale Datenmenge. Zur graphischen Veranschaulichung werden die HSQC-Korrelationspeaks zunächst als Konturplots dargestellt (siehe grüne, blaue und rote Signale der Fraktionen 1, 2 und 3 oben). Um eine gestaffelte Darstellung der Konturplots zu vermeiden, wird stattdessen die Summe der HSQC-Korrelationspeaks der einzelnen Fraktionen gebildet. In dieser schematischen Abbildung der drei Fraktionen sind die Signale im Summenspektrum (Mitte unten) zur Veranschaulichung entsprechend der ursprünglichen Farben gekennzeichnet. In der *SCORE-metabolite-ID²*-App werden jedoch alle Signale in der gleichen Farbe dargestellt. Anhand des violett hinterlegten Signals des Summenspektrums ist exemplarisch die Erstellung eines zweidimensionalen EDDCs (*Extracted Double Delta Chromatogram*) gezeigt. Der Verlauf eines EDDCs entspricht demnach dem Intensitätsverlauf eines HSQC-Korrelationssignals konsekutiver Fraktionen.

Nach dem Import der Daten erfolgt eine Kalibrierung der Spektren auf das Signal des internen Standards (z.B. TSP-*d*₄) über die Schaltfläche „*Calibration*“. Diese Kalibrierung der Spektren innerhalb der App ist erforderlich, um die Datenpunkte bzgl. der x- und y-Achse für alle Fraktionen gleich anzupassen. Es erfolgt zunächst die Abfrage der chemischen Verschiebung bzgl. der ¹H-Domäne und der ¹³C-Domäne, auf das die Kalibrierung erfolgen soll (z.B. 0.00 ppm für die ¹H- und ¹³C-Domäne für TSP-*d*₄) und anschließend die Abfrage der Bereiche, in dem sich

das Signal des internen Standards bzgl. der ^1H -Domäne (z. B. -0.50 ppm bis 0.50 ppm) und bzgl. der ^{13}C -Domäne (z. B. -1.50 ppm bis 1.50 ppm) aktuell befindet. Über die Schaltfläche „*Normalization*“ wird eine Normalisierung aller importierten HSQC-Spektren auf die Signalintensität des internen Standards durchgeführt.

Nach Anklicken der Schaltfläche „*Bucketing*“ ist optional ein *Bucketing* der HSQC-Spektren innerhalb der *SCORE-metabolite-ID²* App möglich. Die Größe der *Buckets* kann individuell festgelegt werden. Das *Bucketing* wird ausschließlich in der ^1H -Dimension durchgeführt. Die Größe der *Buckets* muss dabei größer als die spektrale Auflösung des Spektrums sein (vgl. Abschnitt 5.1.1.1). Ein *Bucketing* innerhalb der ^{13}C -Dimension ist nicht implementiert, da die Anzahl der Datenpunkte in dieser Domäne aufgrund der zweidimensionalen Aufnahmetechnik deutlich geringer ist und ein *Bucketing* deshalb redundant wäre.

Nach erfolgreichem Import und Prozessierung der Spektren können nun über die Eingabefelder „ ^1H chemical shift“ und „ ^{13}C chemical shift“ einzelne EDDC-Verläufe über die Schaltfläche „*Show EDDC*“ zunächst zweidimensional visualisiert werden. Die für die Analyse relevanten EDDCs können für die weitere (semi-automatische) Analyse und zur Berechnung der Korrelationskoeffizienten über die Schaltfläche „*Add EDDC*“ im rechten Teil der graphischen Benutzeroberfläche in die Liste übernommen werden. In Abbildung 116 ist die Ansicht der Benutzeroberfläche nach Import und Prozessierung der HSQC-Spektren dargestellt.

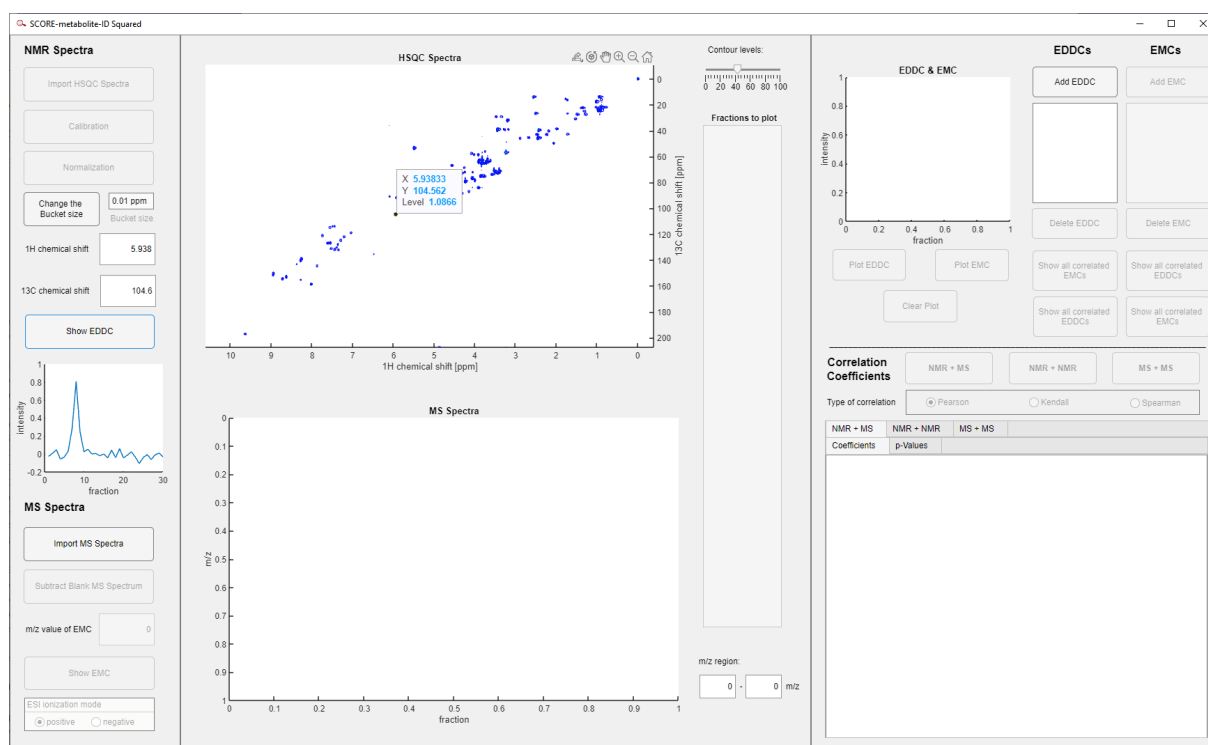


Abbildung 116: Graphische Benutzeroberfläche der *SCORE-metabolite-ID²* App nach Import und Prozessierung der NMR-Daten. Da die HSQC-Spektren als Summe über alle Fraktionen dargestellt werden und somit die Verläufe der Signale über die Fraktionen nicht direkt erkennbar sind, können die zweidimensionalen Verläufe über die Schaltfläche „Show EDDC“ im linken Teil der Benutzeroberfläche visualisiert werden.

Über den Schieberegler unter „*Contour Levels*“ kann die Anzahl der darzustellenden Konturlinien in dem Plot der HSQC-Spektren angepasst werden. Über die Schaltflächen oberhalb der HSQC-Spektren kann bspw. ein Ausschnitt der Spektren vergrößert dargestellt werden.

5.4.1.2 Semi-automatische Nutzung

Nachdem relevante EDDCs und EMCs in die EDDC- und EMC-Listen im rechten Teil der App hinzugefügt wurden, können die Korrelationskoeffizienten sowie die zugehörigen *p*-Werte berechnet werden. Zusätzlich können auch in der *SCORE-metabolite-ID²* App korrelierende Signale über die semi-automatische Korrelation detektiert werden. Sofern ein Eintrag eines bestimmten EDDCs oder EMCs innerhalb der Listen ausgewählt wird, können über die Schaltflächen „*Show all correlated EMCs*“ bzw. „*Show all correlated EDDCs*“ die Signale, die einen höheren Korrelationskoeffizienten als der vorgegebene Schwellenwert aufweisen, angezeigt

werden. Der Schwellenwert beträgt standardmäßig einen Grenzwert von 0.95, jedoch kann dieser Wert über das sich öffnende Eingabefenster angepasst werden.

Bei der semi-automatischen Detektion von hoch-korrelierenden EDDCs werden dabei sowohl Signale mit einem Korrelationskoeffizienten > 0.95 , als auch Signale mit einem Korrelationskoeffizienten < -0.95 aufgeführt. Diese Einstellung wurde gewählt, da im Fall von multiplizitäts-editierten HSQC-Spektren die Signale von CH_2 -Gruppen negative Signalintensitäten aufweisen. Die daraus resultierenden EDDCs haben demnach ebenfalls ein negatives Vorzeichen. Zu demselben Analyten gehörige EMCs oder EDDCs von CH_3 - oder CH -Gruppen weisen jedoch positive Signalintensitäten auf. Die Korrelation dieser Verläufe resultiert demnach in einer starken Antikorrelation mit Korrelationskoeffizienten von bspw. < -0.95 . Neben der tabellarischen Auflistung der korrelierenden EDDCs werden diese auch graphisch in Form eines HSQC-Spektrums als Konturplot dargestellt. Hoch-korrelierende EMCs können wie bereits in der *SCORE-metabolite-ID* App über die Schaltfläche „*Create plot of correlations*“ im Kontextmenü visualisiert werden.

5.4.1.3 Kombination mit SMART 2.0

Die Methode SMART 2.0 (b) (*Small Molecule Accurate Recognition Technology*) dient der schnellen und einfachen Identifizierung von Naturstoffen basierend auf HSQC-Spektren.^[91] In einem chemoinformatischen Ansatz basierend auf einem *Convolutional Neural Network* wurden insgesamt 53 076 gemessene und simulierte HSQC-Spektren von Naturstoffen und Marineorganismen in einen 180-dimensionalen Raum eingeordnet. Dies entspricht ca. 15 % der bisher bekannten Naturstoffe. Durch die Netzwerk-basierte Analyse der Muster können somit auch ähnliche Struktur motive detektiert werden.^[91] Daneben ist die Methodik außerdem zur Dereplikation bereits bekannter Metaboliten hilfreich.

Um eine einfache Verknüpfung zwischen der *SCORE-metabolite-ID*² Methode und dem SMART 2.0 Ansatz zu ermöglichen, wurde ein MATLAB-Skript („*DataInput_SMARTNMR*“) implementiert, durch das die Daten für die direkte Eingabe für SMART 2.0 präpariert werden. Nach der zuvor beschriebenen semi-automatischen Korrelation von bspw. einem EMC zu hoch-korrelierenden EDDCs wird durch das Skript ein *Clustering* durchgeführt, um die durch die Korrelation ermittelten Datenpunkte zu jeweils einem Datenpunkt pro Signal zu reduzieren. Für

das *Clustering* wird die MATLAB-Funktion „*findClusters*“ aus der „2D NMR Spectra Fitter“ Toolbox verwendet.^[199] Gleichzeitig werden die Daten in das entsprechende Format überführt, sodass diese direkt auf der Website des SMART 2.0 Webtools eingegeben werden können.^[200] In Abbildung 117 ist das beschriebene Vorgehen beispielhaft anhand des Metaboliten Tryptophan (**25**) gezeigt.

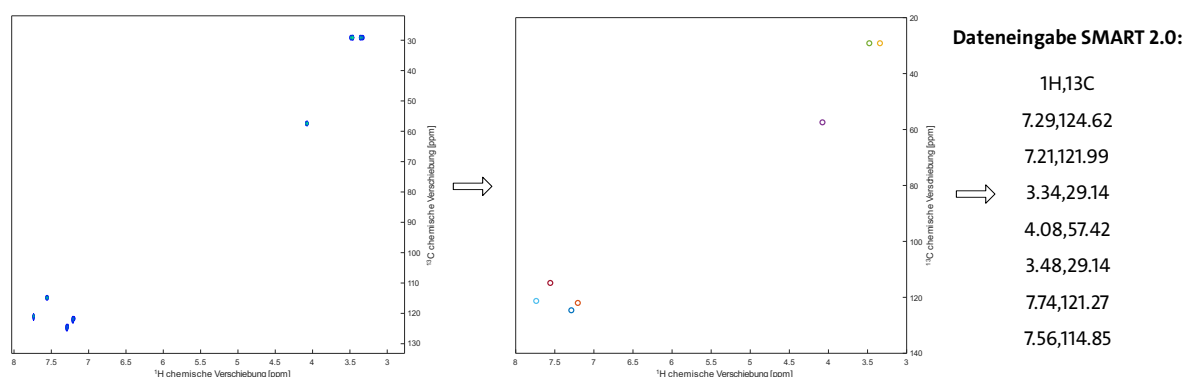


Abbildung 117: Beispielhafte Darstellung des *Clustering* zur Datenvorbereitung für die Kombination mit SMART 2.0. Der linke Plot zeigt ein berechnetes HSQC-Spektrum durch die semi-automatische Korrelation, d. h. es werden die Signale mit einem hohen Korrelationskoeffizienten zu dem vorgegebenen EMC bei $m/z = 205.10$ ($[M+H]^+$ von Tryptophan (**25**)) angezeigt. Das MATLAB-Skript „*DataInput_SMARTNMR*“ führt ein *Clustering* durch, was prinzipiell einem *Peak Picking* entspricht, sodass pro Signal nur ein einzelner Datenpunkt erhalten wird (siehe mittlerer Plot des HSQC-Spektrums). Die Datenpunkte der ermittelten *Cluster* werden außerdem schriftlich ausgegeben, um die Daten direkt in das Webtool von SMART 2.0 eingeben zu können.

5.4.2 Vergleich verschiedener HSQC-Experimente

Ein genereller Nachteil an der Verwendung von 2D NMR-Daten ist die, im Vergleich zu ¹H-NMR-Experimenten, lange Messdauer der 2D NMR-Experimente. Dies ist eine intrinsische Eigenschaft von 2D Experimenten, die sich durch die Inkrementierung von t_1 zur Bildung der zusätzlichen Dimension ergibt. Insbesondere bei großen Probenmengen ist dementsprechend eine Verkürzung der Messdauer wünschenswert. Im Fall von HSQC-Experimenten stellt ein möglicher Ansatz zur Verkürzung der Messdauer das in Abschnitt 3.3.2.3 erläuterte ASAP-HSQC-Experiment dar, bei welchem die Verkürzung der Messdauer hauptsächlich auf der Verkürzung des Relaxationsdelays beruht. Der damit prinzipiell verbundene Sensitivitätsverlust soll bei dem ASAP-HSQC-Experiment durch ein Polarisationsreservoir und einen Polarisationstransfer

ungenutzter Magnetisierung kompensiert werden. Eine weitere Möglichkeit zur Reduktion der Messdauer ist die Verwendung von *Nonuniform Sampling* (NUS). Dabei wird lediglich eine reduzierte Anzahl der Inkremente aufgenommen. Die fehlenden Inkremente werden vor der Fourier-Transformation berechnet.

Im Allgemeinen findet bei HSQC-Experimenten während der Aufnahme des FIDs eine kontinuierliche Entkopplung statt. Durch diesen hohen Energieeintrag kann sich die Probe möglicherweise erwärmen. Eine leichte Erwärmung der Probe ist bereits bei konventionellen HSQC-Experimenten zu beobachten. Die schnellere Abfolge der ASAP-HSQC-Pulssequenz aufgrund des reduzierten Relaxationsdelays führt zur stärkeren Erhitzung der Probe und möglicherweise der Spektrometerelektronik, was schlimmstenfalls zu Defekten führen kann. Um eine geeignete Länge des Relaxationsdelays zu ermitteln, wurde der Relaxationsdelay deshalb schrittweise verkürzt und der jeweilige Temperaturanstieg beobachtet. Bei dem ASAP-HSQC-Experiment resultiert am verwendeten 600 MHz-NMR-Spektrometer ein Relaxationsdelay von 0.3 Sekunden in einem Temperaturanstieg um ca. 1.0 K. Während der Messung sollte eine konstante Temperatur herrschen, da die chemische Verschiebung unter anderem von der Temperatur abhängig ist und Temperaturunterschiede demnach zu Verschiebungen der Signale im Verlauf der Messdauer führen würden. Um Temperaturschwankungen während der Messung zu vermeiden, wurde deshalb die Anzahl der *Dummy Scans* erhöht.

Zunächst wurde das konventionelle multiplizitäts-editierte HSQC-Experiment (Pulssequenz: *hsqcedetgpcp.3*) mit einem standardmäßig verwendeten Relaxationsdelay von 1.0 Sekunden mit dem ASAP-HSQC-Experiment (Pulssequenz: *ASAP-hsqcetgpcp*) mit einem Relaxationsdelay von 0.3 Sekunden verglichen. Zusätzlich wurde zum Vergleich das ASAP-HSQC-Experiment unter Anwendung von *Nonuniform Sampling* getestet. Für den Vergleich wurde eine 67.2 mM Lösung von Saccharose (**2**) in Deuteriumoxid (+ 2.5 mM TSP-*d*₄) verwendet. Der Vergleich der Experimente erfolgte anhand des Signal-zu-Rausch-Verhältnisses des Korrelationssignals des Protons H-1' der Glucose-Einheit (5.34 ppm/92.0 ppm). Das Signal-zu-Rausch-Verhältnis wurde in TopSpin mittels des Python-Skriptes (*sino2d_ok.py*) ermittelt. Die Ergebnisse sind in Abbildung 118 dargestellt.

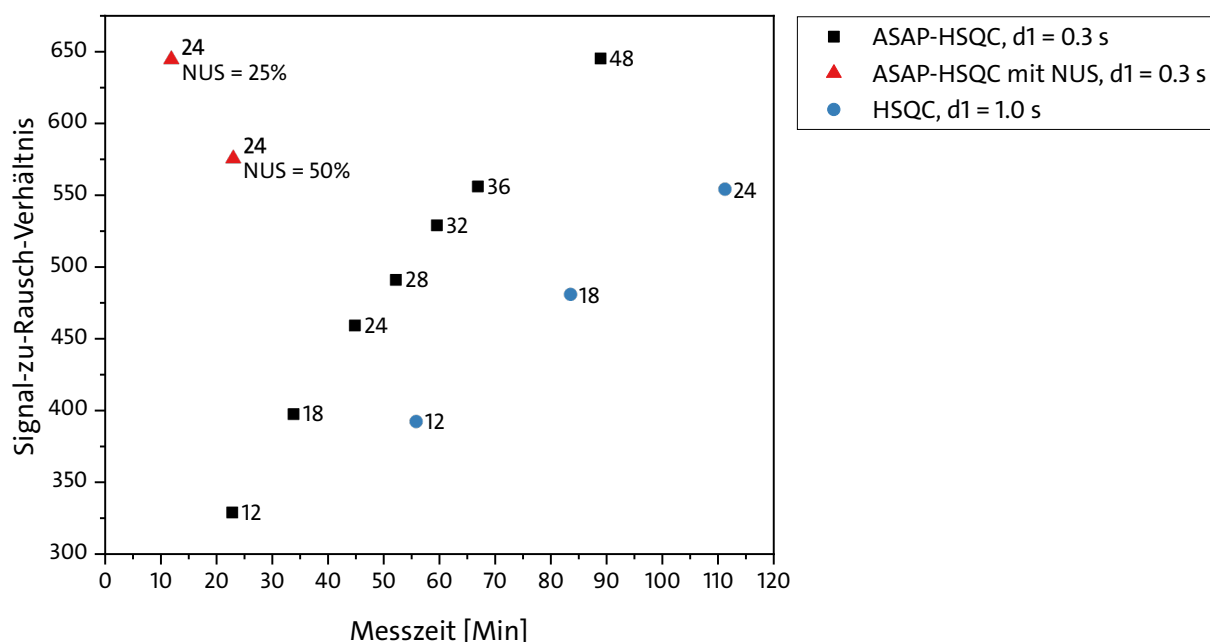


Abbildung 118: Vergleich des ASAP-HSQC-Experimentes (*ASAP-hsqcetgppsp*) mit verkürztem Relaxationsdelay von 0.3 Sekunden (schwarze Quadrate) mit dem konventionellen HSQC-Experiment (*hsqcedetgppsp.3*) mit einem Relaxationsdelay von 1.0 Sekunden (blaue Punkte) anhand des Signal-zu-Rausch-Verhältnisses des Protons H-1' der Saccharose (**2**) (67.2 mM Lösung in Deuteriumoxid + 2.5 mM TSP-*d*₄) bezüglich der Gesamtmessdauer. Die roten Dreiecke zeigen die Ergebnisse des ASAP-HSQC-Experimentes mit verkürztem Relaxationsdelay und der gleichzeitigen Anwendung von *Nonuniform Sampling*. Die Ziffern neben den Datenpunkten geben die Anzahl der *Scans* an sowie für die Messungen mit NUS zusätzlich die prozentuale Anzahl an Inkrementen, die tatsächlich aufgenommen wurde (*NUSAmount*). Die anderen Parameter wurden wie folgt eingestellt: TD2 = 1024, TD1 = 256, DS = 32 (*hsqcedetgppsp.3*) und DS = 128 (*ASAP-hsqcetgppsp*), SW1 = 11 ppm, SW2 = 180 ppm.

Wie der Abbildung 118 zu entnehmen ist, wird bei der Aufnahme der gleichen Anzahl von *Scans* ein höheres Signal-zu-Rausch-Verhältnis bei Verwendung des HSQC-Experimentes im Vergleich zum ASAP-HSQC-Experiment mit einem Relaxationsdelay von 0.3 Sekunden erhalten. Dabei ist jedoch die Messdauer des HSQC-Experimentes mehr als doppelt so lang. Ein in etwa gleich hohes Signal-zu-Rausch-Verhältnis kann somit durch das ASAP-HSQC-Experiment mit verringertem Relaxationsdelay in deutlich kürzerer Zeit erhalten werden. Durch die Verwendung von *Nonuniform Sampling* (NUS) kann das Signal-zu-Rausch-Verhältnis bei gleicher Messdauer (für *NUSAmount* = 50 %) bzw. bei verringerter Messdauer (für *NUSAmount* = 25 %) im Vergleich zur Aufnahme des ASAP-HSQC-Spektrums mit 12 *Scans* zusätzlich deutlich erhöht werden.

Ähnliche Ergebnisse erzielten auch die Vergleiche des ASAP-HSQC-Experimentes mit einem Relaxationsdelay von 0.4 Sekunden und 0.5 Sekunden. Bei dem Vergleich aller Messungen fiel

jedoch auf, dass das Signal-zu-Rausch-Verhältnis scheinbar unabhängig des gewählten Relaxationsdelays ist, sondern direkt von der Messdauer abhängt. Obgleich das HSQC-Pulsprogramm kein Polarisationsreservoir zur Verwendung ungenutzter Magnetisierung nutzt, ist prinzipiell auch bei diesem Pulsprogramm eine Reduktion des Relaxationsdelays möglich. Die Reduktion des Relaxationsdelays auf 0.3 Sekunden führte zu einer Erwärmung der Probe um ca. 0.4 K. In Abbildung 119 ist ein Vergleich der ASAP-HSQC-Experimente und der konventionellen HSQC-Experimente mit variierenden Relaxationsdelays von 0.3, 0.4 und 0.5 Sekunden durch Auswertung der ermittelten Signal-zu-Rausch-Verhältnisse gezeigt.

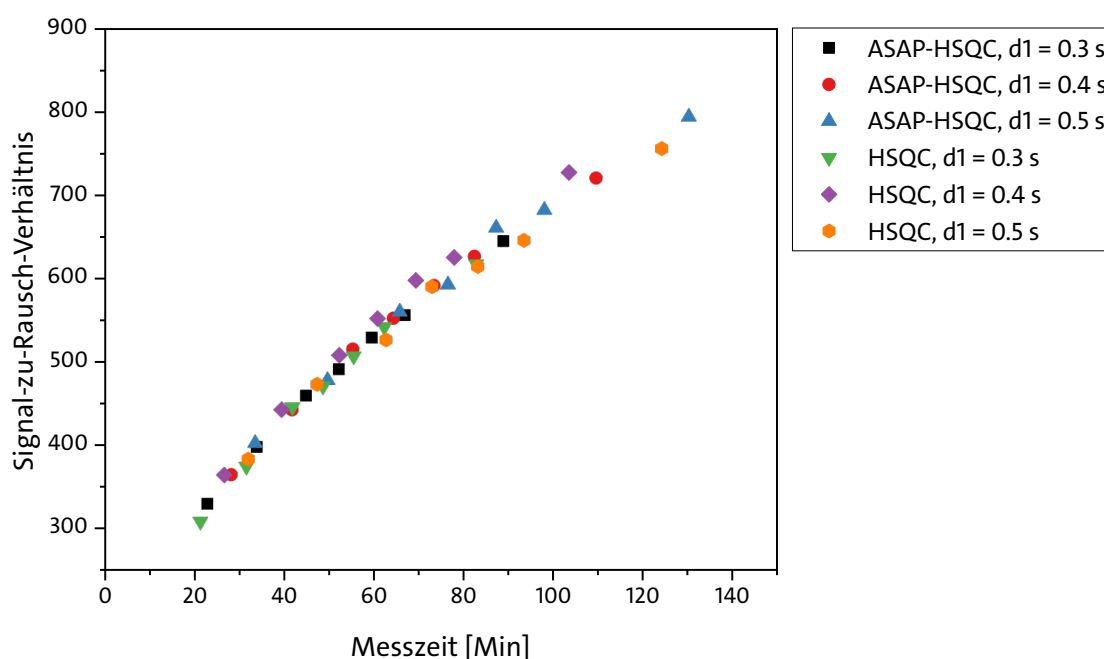


Abbildung 119: Vergleich der ASAP-HSQC-Experimente (*ASAP-hsqcetgpsp*) und der konventionellen HSQC-Experimente (*hsqcedetgpsp.3*) mit variierenden Relaxationsdelays von 0.3, 0.4 und 0.5 Sekunden für beide Experimente. Der Vergleich erfolgt anhand des Signal-zu-Rausch-Verhältnisses des ^1H - ^{13}C -Korrelationssignals des Protons H-1' von Saccharose (**2**) (5.34 ppm/92.0 ppm), das in TopSpin mit Hilfe des Python-Skriptes *sino2d_ok.py* ermittelt wurde. Die verschiedenen Datenpunkte einer Farbe kennzeichnen Experimente mit gleichen Parametern (TD2 = 1024, TD1 = 256, DS = 128, SW1 = 11 ppm, SW2 = 180 ppm), wobei die Anzahl der aufgenommenen *Scans* von 12, 18, 24, 28, 32, 36 bis 48 variiert wurde. Aus Gründen der Übersichtlichkeit wurde auf die Kennzeichnung der Anzahl der *Scans* in der Abbildung verzichtet.

Wie die Abbildung 119 zeigt, implizieren die Signal-zu-Rausch-Verhältnisse des untersuchten Signals eine nahezu direkte Abhängigkeit von der Messdauer. Dabei spielt das gewählte Pulsprogramm eine untergeordnete Rolle. Sowohl für das ASAP-HSQC- als auch für das HSQC-Experiment werden bei gleichem Relaxationsdelay ähnliche Signal-zu-Rausch-Verhältnisse

erhalten. Nach SCHULZE-SÜNNINGHAUSEN *et al.* minimiert sich bei dem ASAP-HSQC-Experiment der Effekt der Ernst-Winkel Anregung und der isotropen Mischperiode zum Polarisationstransfer bei längeren Experimenten.^[83] Durch die fortfolgende Aufnahme mehrerer *Scans* findet keine vollständige Relaxation statt und nach einer gewissen Zeit wird somit ein Gleichgewichtszustand erhalten. Insofern ist im Rahmen dieser Arbeit bei vergleichsweise langen Messzeiten kein eindeutiger Vorteil in einem der beiden Pulsexperimente zu verzeichnen.

Wie bereits zuvor erwähnt, kann neben dem ASAP-HSQC-Experiment auch *Nonuniform Sampling* zur Verkürzung der Messdauer angewendet werden. In den hier verwendeten Spektren der Reinsubstanz Saccharose führt die Verwendung von *Nonuniform Sampling* zu einer deutlichen Steigerung des Signal-zu-Rausch-Verhältnisses bei vergleichbarer Messzeit (vgl. Abbildung 118). Bei Spektren, die aufgrund der Vielzahl von unterschiedlich konzentrierten Metaboliten eine Vielzahl von Signalen beinhalten und dadurch eine höhere Komplexität aufweisen, wie sie im Rahmen dieser Arbeit Anwendung finden, kann die Verwendung von *Nonuniform Sampling* jedoch auch zu nachteiligen Effekten führen. Beispielsweise werden Signale gering konzentrierter Metaboliten in den NMR-Spektren, die mit NUS aufgenommen werden, nicht oder mit deutlich geringerer Signalintensität detektiert. In Abbildung 120 ist dies beispielhaft für zwei ASAP-HSQC-Spektren eines Gemisches aus Phenylalanin (**4**) in einer Konzentration von 0.47 mg/mL, Vanillin (**87**) in einer Konzentration von 0.65 mg/mL und Isoleucin (**6**) in einer Konzentration von 3.00 mg/mL gezeigt. Das in schwarz abgebildete HSQC-Spektrum wurde ohne NUS aufgenommen, bei dem in rot dargestellten HSQC-Spektrum wurden lediglich 50 % der Inkremente tatsächlich aufgenommen.

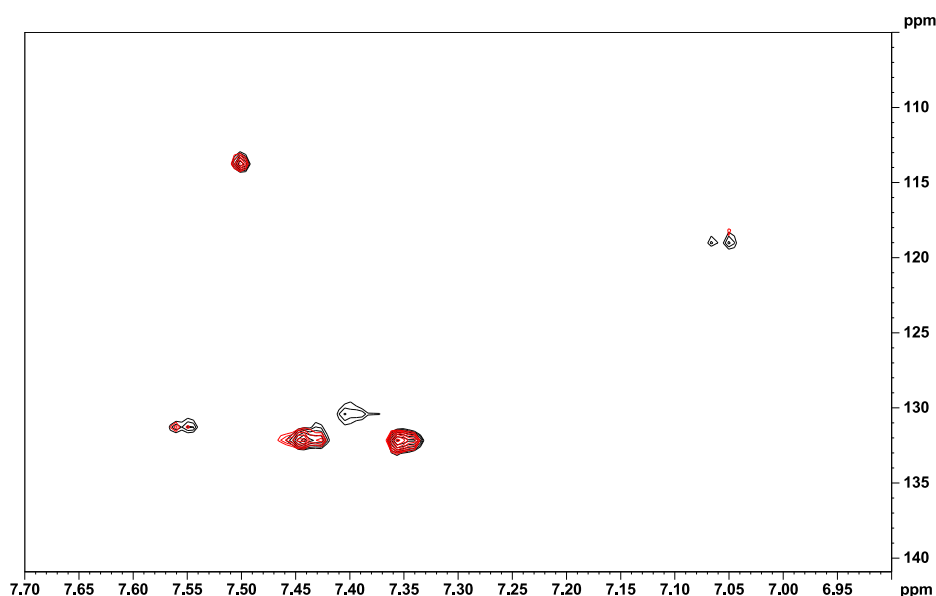


Abbildung 120: Überlagerte Darstellung des aromatischen Bereichs zweier ASAP-HSQC-Spektren. Das schwarze Spektrum wurde ohne NUS aufgenommen, das rote Spektrum mit NUS, wobei lediglich 50 % der Inkremente aufgenommen wurden (*NUSAmount* = 50 %). Die anderen Parameter wurden in beiden Fällen gleich gewählt (*TD2* = 2048, *TD1* = 256, *DS* = 256, *NS* = 16, *d1* = 0.3 s, *SW1* = 11 ppm, *SW2* = 220 ppm). Daraus resultiert eine Gesamtmessdauer von 36:46 Minuten für das Experiment ohne NUS und von 19:27 Minuten für das Experiment mit NUS. Die Probe enthält ein Gemisch aus Phenylalanin (**4**) (*c* = 0.47 mg/mL), Vanillin (**87**) (*c* = 0.65 mg/mL) und Isoleucin (**6**) (*c* = 3.00 mg/mL).

Abbildung 120 zeigt am Beispiel ausgewählte Signale im aromatischen Bereich, dass bei der Verwendung von NUS zur Verkürzung der Messdauer Signale teilweise nicht detektiert werden. In diesem Fall betrifft es ein Signal des Analyten Phenylalanin (**4**) (7.40 ppm/130.3 ppm). Zusätzlich sind zwei Signale des Analyten Vanillin (**87**) (7.05 ppm/119.0 ppm und 7.55 ppm/131.2 ppm) im direkten Vergleich deutlich weniger intensiv. Prinzipiell kann NUS statt zur Verkürzung der Messzeit auch zur Steigerung der Sensitivität genutzt werden, indem alle Parameter beibehalten werden und lediglich die Anzahl der *Scans* erhöht wird. Beispielsweise können bei einem *NUSAmount* von 50 % bezogen auf die gleiche Messdauer doppelt so viele *Scans* aufgenommen werden. In vielen Fällen führt die Berechnung der fehlenden Inkremente jedoch häufig zu vermehrtem Auftreten von Artefakten in den Spektren. Aus diesem Grund wurde im Rahmen dieser Arbeit für die *SCORE-metabolite-ID²* Methode das ASAP-HSQC-Experiment ohne NUS verwendet.

5.4.3 Validierung durch artifizielles Gemisch B

Zur Validierung der *SCORE-metabolite-ID²* Methode wurde das artifizielle Gemisch B bestehend aus 20 Analyten hergestellt. Die enthaltenen Analyten sind in Abbildung 121 aufgeführt. Um die individuellen Elutionsprofile und damit die EDDC- und EMC-Verläufe zu simulieren, wurden variierende Volumina der Stammlösungen der Analyten in insgesamt 30 Proben gegeben. Somit entsprechen die 30 Proben künstlich hergestellten Fraktionen, in denen die verschiedenen Analyten individuelle Elutionsprofile aufweisen.

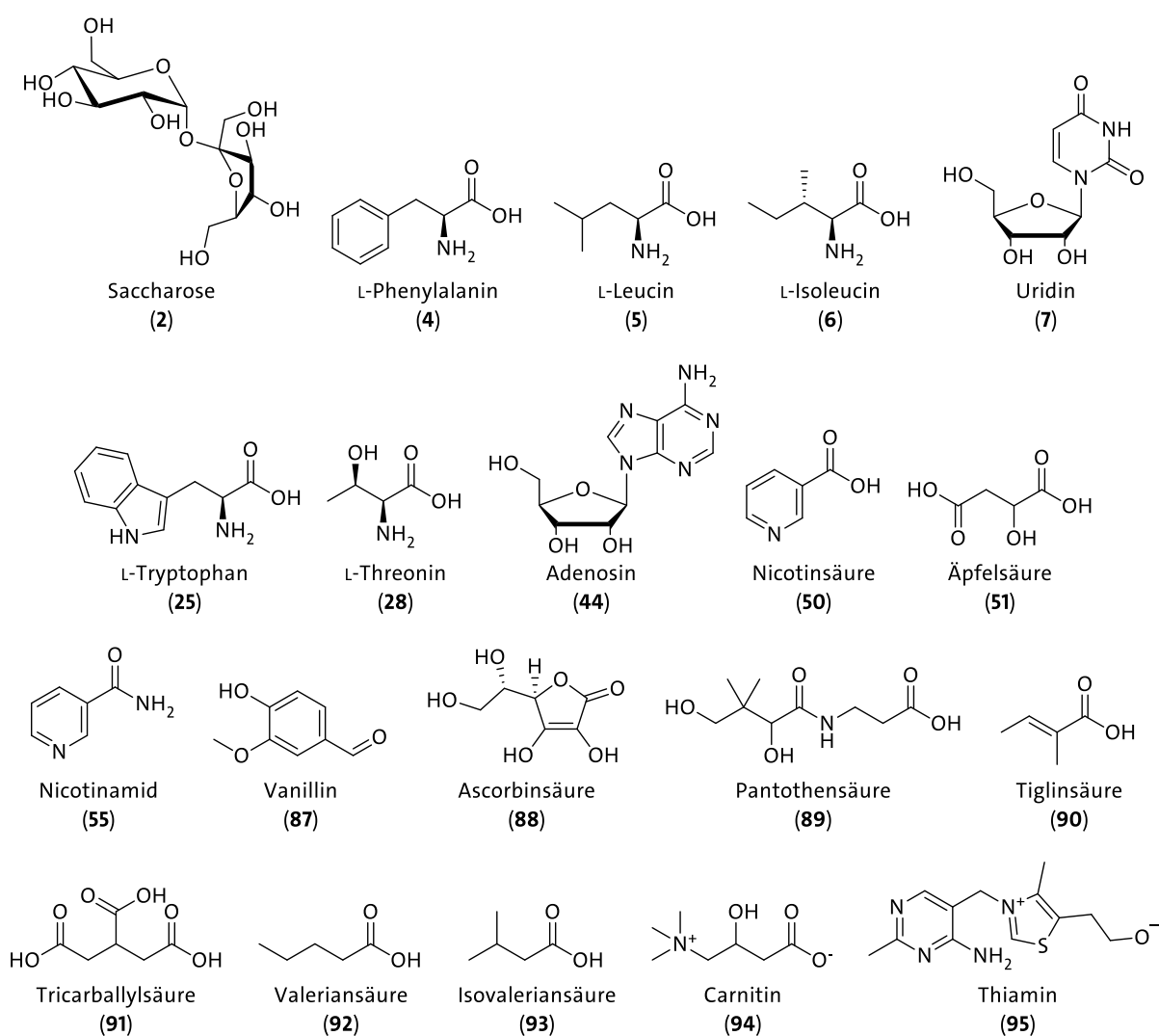


Abbildung 121: Übersicht der 20 Analyten, die zur Erstellung des artifiziiellen Gemisches B eingesetzt wurden.

Um erneut ein möglichst reales Bild einer chromatographischen Separation eines komplexen Gemisches zu simulieren, wurden absichtlich starke Konzentrationsunterschiede zwischen einzelnen Analyten eingesetzt (siehe Tabelle 31 unter Abschnitt 6.5.3). Die maximale

Konzentration liegt z.B. bei 15.3 mg/mL für Adenosin in Fraktion 19. In derselben Fraktion liegt bspw. Äpfelsäure in einer Konzentration von 1.53 mg/mL vor. Die minimale Konzentration von 0.29 mg/mL weist Tiglinsäure in den Fraktionen 21 und 25 auf. Außerdem beinhaltet das artifizielle Gemisch B insgesamt drei Paare isobarer Verbindungen (Valeriansäure (**92**) und Isovaleriansäure (**93**), Leucin (**5**) und Isoleucin (**6**) sowie Tricarballylsäure (**91**) und Ascorbinsäure (**88**)). Um einen Vergleich zwischen den Ansätzen *SCORE-metabolite-ID* und *SCORE-metabolite-ID*² zu ermöglichen, wurde von jeder Fraktion sowohl ein ¹H-NMR-Spektrum als auch ein 2D HSQC-Spektrum mittels ASAP-HSQC-Pulssequenz aufgenommen sowie jeweils ein Massenspektrum.

In Abbildung 122 A sind zunächst die ¹H-NMR-Daten des artifiziellen Gemisches B dargestellt. Insbesondere im Bereich der chemischen Verschiebung zwischen 0.8 und 4.5 ppm sind teilweise deutlich überlagerte Signale zu erkennen, die die Auswertung mittels *SCORE-metabolite-ID* erschweren. In Abbildung 122 B ist die summierte Darstellung der HSQC-Konturplots über alle 30 Fraktionen gezeigt, entsprechend des zuvor unter Abschnitt 5.4.1.1 erläuterten Prinzips. Es wird ersichtlich, dass im Bereich des HSQC-Spektrums zwischen 0.8/13 ppm bis 4.5/90 ppm eine Vielzahl von Signalen auftreten. Im Gegensatz zu den ¹H-NMR-Spektren sind die Signale im HSQC-Spektrum aufgrund der zusätzlichen Dimension jedoch in deutlich geringerem Maße überlagert.

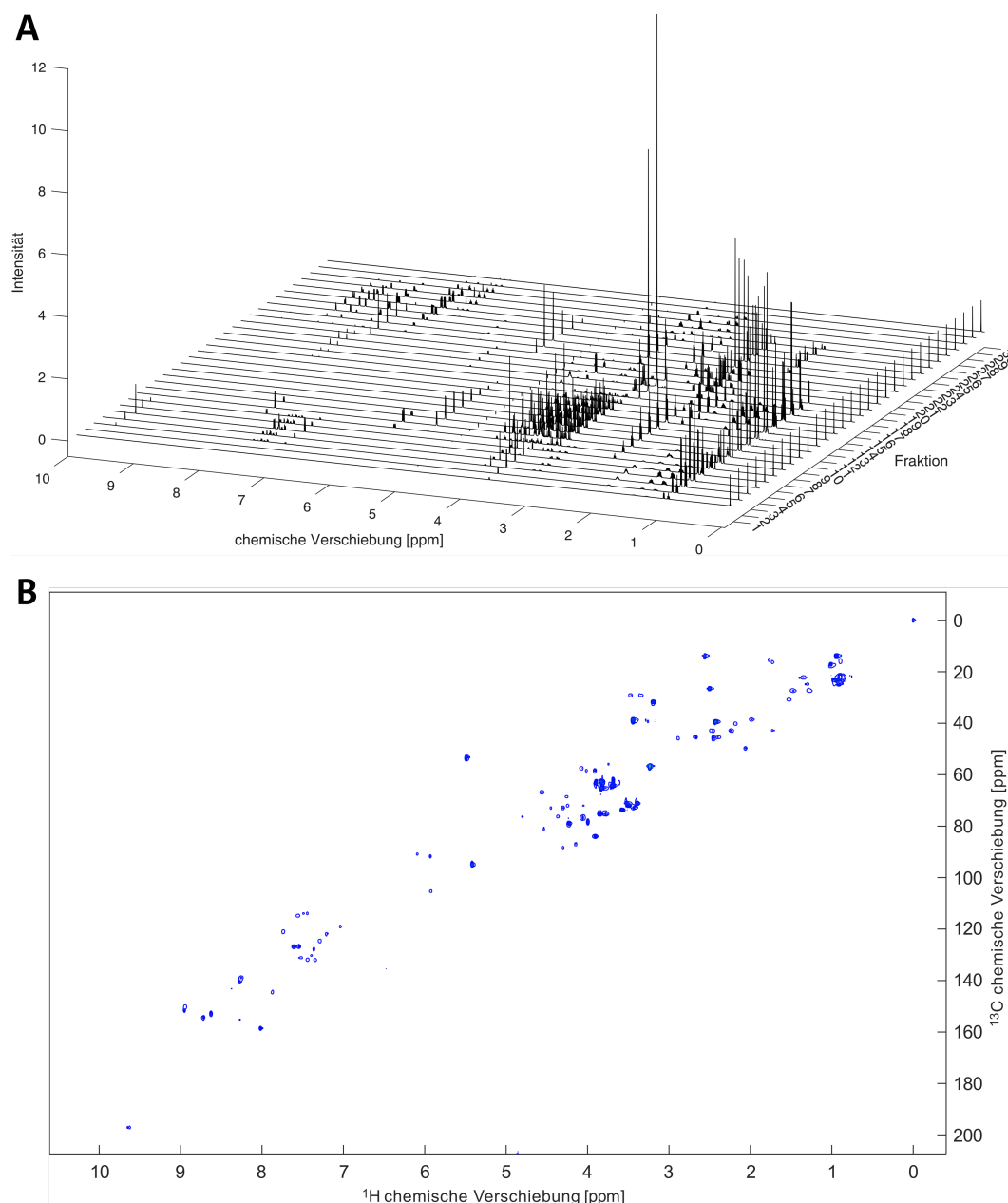


Abbildung 122: A: ^1H -NMR-Spektren des artifiziellen Gemisches B in gestaffelter Darstellung über alle 30 Fraktionen. B: Summierte Darstellung der HSQC-Konturplots aller 30 Fraktionen entsprechend des unter Abschnitt 5.4.1.1 erläuterten Prinzips.

Beispielhaft ist zunächst die Auswertung des Analyten Uridins (**7**) im artifiziellen Gemisch B gezeigt. Abbildung 123 A zeigt zwei EMC-Verläufe des Analyten anhand des Natrium-Adduktes sowie dem protonierten Molekülion im positiven Ionisationsmodus. Durch die semi-automatische Korrelation des EMCs des Natrium-Adduktes ($m/z = 267.06$) erfolgte die Zuordnung der EDDCs, die einen hohen Korrelationskoeffizienten dazu aufweisen. Die graphische Darstellung der hoch-korrelierenden EDDCs ist in Abbildung 123 B in Form eines

berechneten HSQC-Spektrums gezeigt. Demnach korreliert der EMC bei $m/z = 267.06$ mit jeweils hohen Korrelationskoeffizienten > 0.95 zu einem Signal im aromatischen Bereich (7.89/144.8 ppm), zu zwei Signalen bei einer chemischen Verschiebung von 5.93/91.7 ppm und 5.93/104.5 ppm sowie zu mehreren Signalen im Bereich der Kohlenhydrate. Abbildung 123 C zeigt ein HSQC-Referenzspektrum des Analyten Uridin (**7**).

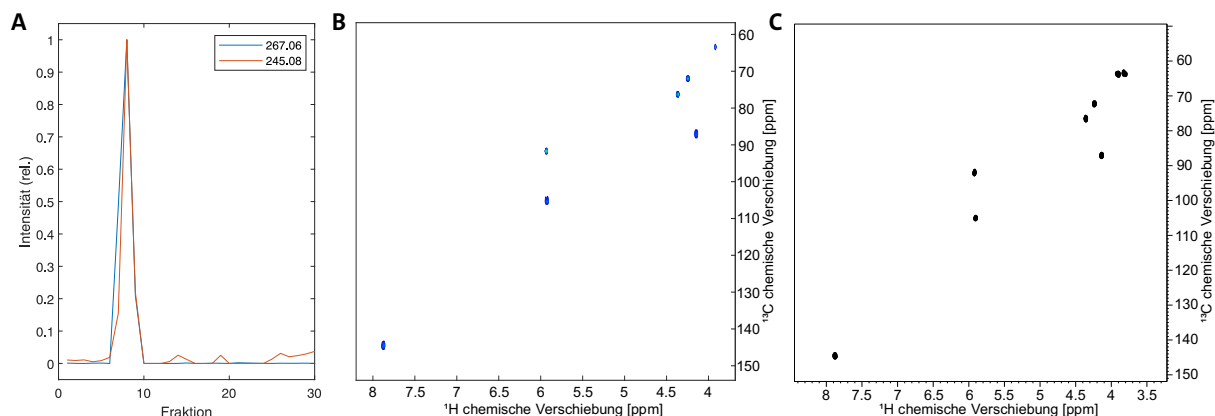


Abbildung 123: A: EMCs des Analyten Uridin (**7**) bei $m/z = 267.06$ ($[M+Na]^+$) und $m/z = 245.08$ ($[M+H]^+$) des artifiziiellen Gemisches B. B: Graphische Darstellung der EDDCs, die einen Korrelationskoeffizienten > 0.95 berechnet über alle 30 Fraktionen zu dem EMC des Natrium-Adduktes von Uridin (**7**) aufweisen. C: HSQC-Vergleichsspektrum von Uridin (**7**) als Reinsubstanz.

Ein Vergleich des berechneten Spektrums (Abbildung 123 B) zu dem Referenzspektrum (Abbildung 123 C) zeigt, dass beide Spektren hohe Ähnlichkeiten zueinander aufweisen. Lediglich ein Signal bei einer chemischen Verschiebung 3.92/63.5 ppm, das einem der diastereotopen Protonen H-5 der Ribose-Substruktur zuzuordnen ist, ist in dem berechneten HSQC-Spektrum nicht detektierbar.

Unter der Annahme, dass die Identität des Analyten nicht bekannt wäre, wurden die über die semi-automatische Korrelation ermittelten Daten zur weiteren Analyse mit Hilfe des Skriptes „DataInput_SMARTNMR“ verarbeitet und anschließend mittels SMART 2.0 analysiert. Die Ergebnisliste von SMART 2.0 ist entsprechend des ermittelten *Cosine Scores* (Kosinus-Ähnlichkeit) geordnet. Hohe *Cosine Scores* werden in diesem Fall für die Nucleotide und Nucleoside Uridin-5'-diphosphat (**98**), Uridin-5'-triphosphat (**99**), Uridin-5'-monophosphat (**48**), Cytidin-5'-monophosphat (**101**), Cytidin (**100**) und Uridin (**7**) erhalten (siehe Tabelle 28).

Tabelle 28: Zusammengefasste Ergebnisliste des Webtools SMART 2.0 nach Eingabe der durch die semi-automatische Korrelation detektierten hoch-korrelierenden EDDCs zu dem EMC des Natrium-Adduktes des Analyten Uridins (**7**). Die Ergebnisliste wurde insofern gekürzt, als dass mehrmalige Einträge des gleichen Analyten eliminiert wurden.

Substanz	Cosine Score	monoisotopische Masse
Uridin-5'-diphosphat	0.9845	404.0022 Da
Uridin-5'-triphosphat	0.9845	483.9685 Da
Uridin-5'-monophosphat	0.9829	324.0359 Da
Cytidin-5'-monophosphat	0.9798	323.0519 Da
Cytidin	0.9776	243.0855 Da
Uridin	0.9752	244.0695 Da

Alle in der obigen Tabelle aufgeführten Nucleotide und Nucleoside weisen hohe strukturelle Ähnlichkeiten zueinander auf und unterscheiden sich teilweise nur in der Anzahl bzw. der generellen Anwesenheit von Phosphatsubstituenten an der Kohlenhydrat-Grundstruktur (siehe Abbildung 124).

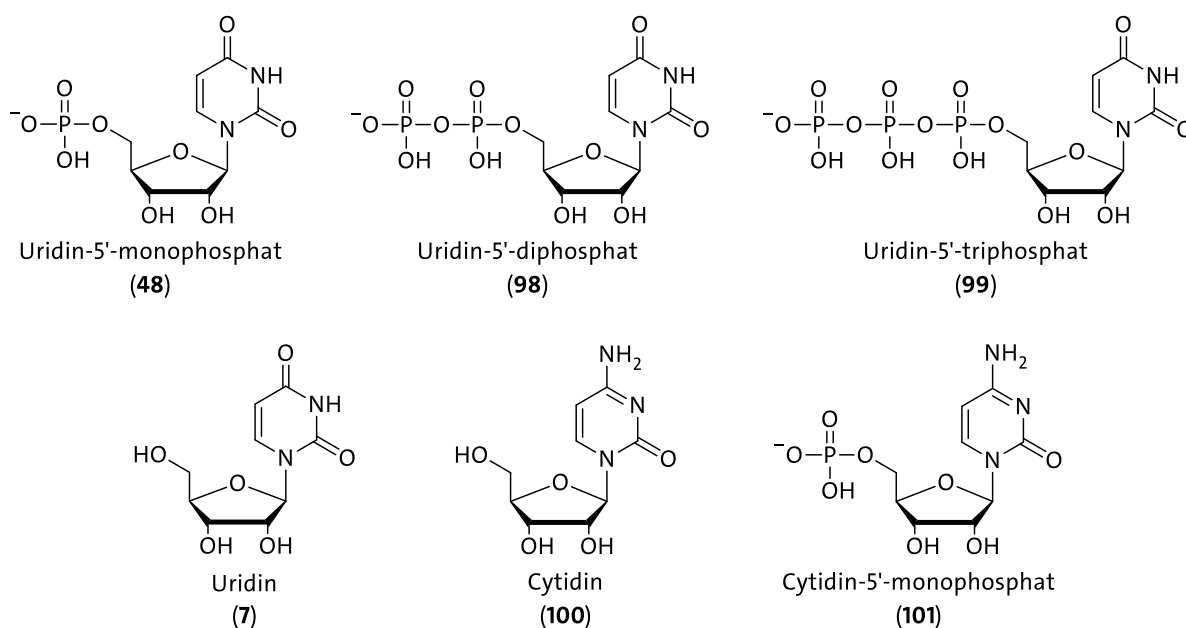


Abbildung 124: Übersicht der über SMART 2.0 vorgeschlagenen Strukturformeln nach Eingabe der durch die semi-automatische Korrelation detektierten hoch-korrelierenden EDDCs zu dem EMC des Natrium-Adduktes des Analyten Uridins (**7**) (siehe Tabelle 28).

Anhand eines $^1\text{H},^{13}\text{C}$ -HSQC-Spektrums sind diese Substanzen dementsprechend kaum zu unterscheiden. Unter Zuhilfenahme der Information über die monoisotopische Masse des

Analyten von 244.07 Da, auf die über die Korrelation zwischen den EDDCs und den EMCs des protonierten Moleküliions und des Natrium-Adduktes geschlossen werden kann, könnte jedoch auch ohne die vorherige Kenntnis eindeutig geschlossen werden, dass es sich bei dem Analyten um Uridin (**7**) mit einer monoisotopischen Masse von 244.0695 Da handelt. Anhand dieses Beispiels kann dementsprechend gezeigt werden, dass analog zu dem Ansatz von *SCORE-metabolite-ID* eine Korrelation zwischen HSQC- und MS-Daten möglich ist. Daneben vereinfacht die graphische Auswertung der semi-automatischen Korrelation eine Strukturzuordnung, bspw. durch einen Datenbankvergleich mit z. B. SMART 2.0.

Für einen Vergleich der Ansätze von *SCORE-metabolite-ID* basierend auf ^1H -NMR-Spektren und von *SCORE-metabolite-ID*² basierend auf 2D-HSQC-NMR-Spektren wurden jeweils beide Arten von Spektren für alle Fraktionen des artifiziellen Gemisches B akquiriert. Anhand der Analyten Äpfelsäure (**51**) und Adenosin (**44**) werden die Ansätze im Folgenden miteinander verglichen. Beide Substanzen befinden sich in den Fraktionen 16-21 des artifiziellen Gemisches B und weisen ein ähnliches artifizielles Elutionsprofil auf, wobei das Konzentrationsmaximum um eine Fraktion zueinander verschoben ist. In der folgenden Tabelle 29 sind die ermittelten Korrelationskoeffizienten nach PEARSON zwischen jeweils einem EMC jedes Analyten zu allen EDCs bzw. EDDCs der Analyten aufgeführt.

Tabelle 29: Berechnete Korrelationskoeffizienten nach PEARSON zwischen den EDDCs und jeweils einem EMC von Adenosin (**44**) und Äpfelsäure (**51**) im linken Teil der Tabelle für die Methode *SCORE-metabolite-ID*². Die Korrelationskoeffizienten wurden über alle 30 Fraktionen berechnet. Im rechten Teil der Tabelle sind die analogen Korrelationskoeffizienten, berechnet lediglich über die Fraktionen 15-22, zwischen den EDCs und den respektiven EMCs in der *SCORE-metabolite-ID* Methode angegeben. Sowohl für die HSQC-Spektren als auch für die ¹H-NMR-Spektren wurde eine einheitliche Größe der *Buckets* von 0.01 ppm gewählt.

<i>SCORE-metabolite-ID</i> ²				<i>SCORE-metabolite-ID</i>			
	<i>m/z</i>	157.01 [M+Na] ⁺	268.10 [M+H] ⁺		<i>m/z</i>	157.01 [M+Na] ⁺	268.10 [M+H] ⁺
EDDC [ppm]	Analyt	Äpfel- säure	Adenosin	EDC [ppm]	Analyt	Äpfel- säure	Adenosin
(2.40; 45.42)	Äpfel- säure	0.9912	0.5979	2.40	Äpfel- säure	0.9986	0.5083
(2.69; 45.42)		0.9959	0.6383	2.69		0.9766	0.4760
(4.30; 72.85)		0.9960	0.6255	4.30		0.9841	0.6440
(3.91; 64.28)	Adenosin	0.6715	0.9686	3.91	Adenosin	-0.1764	-0.1396
(4.29; 88.28)		0.5576	0.9798	4.29		0.9841	0.6440
(4.45; 72.85)		0.7268	0.9595	4.45		0.5560	0.9689
(4.79; 76.28)		0.5596	0.9797	4.80		0.5285	0.9820
(6.09; 90.85)		0.7234	0.9717	6.09		0.5802	0.9681
(8.27; 155.13)		0.7513	0.9505	8.27		-0.2562	-0.4325
(8.37; 143.13)		0.3819	0.8753	8.37		0.6389	0.9497

Für den Analyten Äpfelsäure (**51**) ergeben sich sowohl in der *SCORE-metabolite-ID*² Methode als auch in der *SCORE-metabolite-ID* Methode sehr hohe Korrelationskoeffizienten zwischen allen EDDCs bzw. EDCs, die jeweils ein NMR-Signal der Äpfelsäure (**51**) beinhalten, zu dem EMC des Natrium-Adduktes bei *m/z* = 157.01. Dies wird auch durch die graphische Auftragung der jeweils hoch-korrelierenden EDCs bzw. EDDCs in Abbildung 125 A bzw. B deutlich. Sowohl in den ¹H-NMR-Spektren als auch in den HSQC-Spektren sind alle drei NMR-Signale der Äpfelsäure (**51**) eindeutig zu detektieren. Dabei ist anzumerken, dass in den ¹H-NMR-Spektren die Feinaufspaltung der Signale in diesem Fall nicht erkennbar ist. Dies liegt an der vergleichsweise großen *Bucket*-Größe von 0.01 ppm, die für einen direkteren Vergleich zwischen den ¹H-NMR-Spektren und den HSQC-Spektren einheitlich gewählt wurde. In diesem Fall beinhalten die HSQC-Spektren durch die zusätzliche ¹³C-Dimension demnach einen größeren Informationsgehalt für eine theoretische Identifizierung des Analyten.

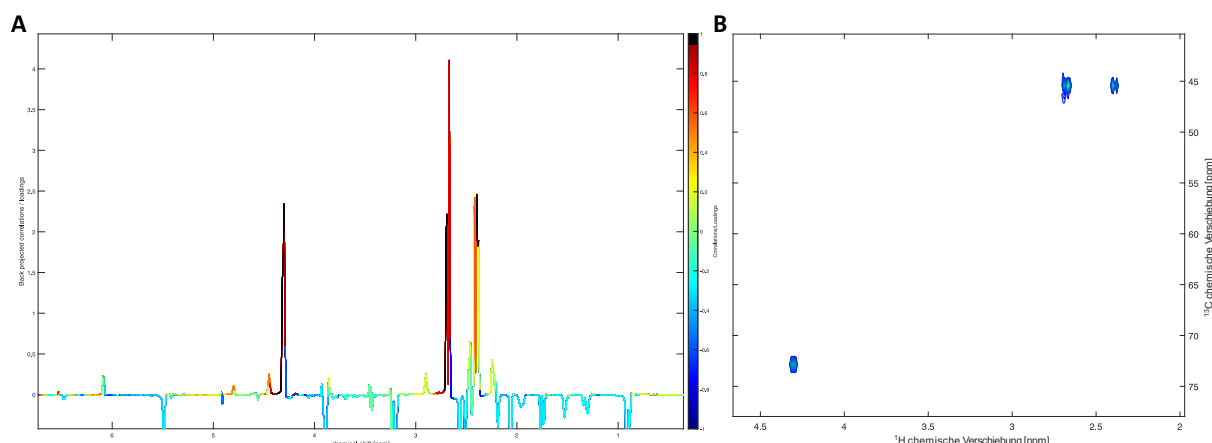


Abbildung 125: A: Graphische Darstellung der hoch-korrelierenden EDCs zu dem EMC des Natrium-Adduktes der Äpfelsäure (**51**) bei $m/z = 157.01$, berechnet über die Fraktionen 15-22, in der *SCORE-metabolite-ID* App ausgehend von den ¹H-NMR-Spektren. B: Graphische Darstellung der EDDCs, die einen Korrelationskoeffizienten > 0.95 in den Fraktionen 1-30 zu dem gleichen EMC der Äpfelsäure (**51**) ($m/z = 157.01$ ($[M+Na]^+$)) in der *SCORE-metabolite-ID*² App basierend auf den HSQC-Spektren aufweisen. Die Größe der *Buckets* beträgt für beide Arten von Spektren jeweils 0.01 ppm.

Im Fall von Adenosin (**44**) zeigen die ermittelten Korrelationskoeffizienten in Tabelle 29, dass einige NMR-Signale des Analyten lediglich in der *SCORE-metabolite-ID*² Methode, nicht aber in der *SCORE-metabolite-ID* Methode, dem EMC des protonierten Moleküls bei $m/z = 268.10$ zugeordnet werden können. Dies betrifft insbesondere die ¹H-NMR-Signale bei einer chemischen Verschiebung von 3.91, 4.29 ppm und 8.27 ppm. Der Grund dafür sind Überlagerungen der NMR-Signale mit Signalen, die entweder bei der gleichen chemischen Verschiebung in den gleichen Fraktionen auftreten (im Fall des EDCs bei 4.29 ppm) oder die bei der gleichen chemischen Verschiebung in vorherigen oder späteren Fraktionen auftreten (im Fall der EDCs bei 3.91 ppm und 8.27 ppm) und damit das Profil des EDCs ebenfalls beeinflussen. Der EDC bei einer chemischen Verschiebung von 4.29 ppm beinhaltet beispielsweise sowohl ein Signal von Adenosin (**44**) (H-4 der Ribose-Substruktur) als auch von Äpfelsäure (**51**) (H-2). Aufgrund der höheren Konzentration der Äpfelsäure (**51**) in diesen Fraktionen, dominiert auch das Signal der Äpfelsäure (**51**) bei der chemischen Verschiebung von 4.29 ppm im EDC. Dies äußert sich auch in dem hohen Korrelationskoeffizienten von 0.9841 zwischen dem EDC bei 4.29 ppm und dem EMC von Äpfelsäure (**51**) in Tabelle 29. Das Signal des H-4 von Adenosin (**44**) kann demnach nur bei Abwesenheit der Äpfelsäure (**51**), d.h. in der Fraktion 15 und in den Fraktionen 21 und 22, detektiert werden und kann demnach über die *SCORE-metabolite-ID* Methode aufgrund von Signalüberlagerungen nicht detektiert werden.

Die ^1H -NMR-Signale bei 3.91 ppm und 8.27 ppm weisen ebenfalls keinen hohen Korrelationskoeffizienten zu dem EMC von Adenosin (**44**) auf, da diese EDCs jeweils weitere Maxima in anderen Fraktionen, d.h. ausgehend von der Anwesenheit anderer Analyten, beinhalten. Zusätzlich weist das Signal bei 8.27 ppm leichte Unterschiede bzgl. der exakten chemischen Verschiebung im Verlauf der Fraktionen auf, was aufgrund von minimalen pH-Wert-Veränderungen oder Unterschieden der Ionenstärke auftreten kann. Aufgrund der zweidimensionalen Aufnahmetechnik bei HSQC-Spektren erscheinen die Signale im Regelfall generell verbreitert, sodass dieses Phänomen teilweise in den Hintergrund tritt. Bei dem Auftreten mehrerer Maxima in einem EDC oder EMCs ist es prinzipiell möglich, die Anzahl der zu analysierenden Fraktionen zu reduzieren, sodass lediglich der relevante Bereich betrachtet wird. Trotz Reduktion der in der *SCORE-metabolite-ID* App zu analysierenden Fraktionen auf die Fraktionen 15-22 konnte jedoch für Adenosin (**44**) keine eindeutige Korrelation zu allen relevanten NMR-Signalen erzielt werden, wie Tabelle 29 zeigt. In solchen Fällen birgt die Nutzung der *SCORE-metabolite-ID*² Methode ein erheblich größeres Potenzial, da die EDDCs sich aufgrund der zusätzlichen ^{13}C -Domäne deutlicher voneinander unterscheiden. Beispielhaft ist dies in Abbildung 126 für vier verschiedene EDDCs bei der gleichen ^1H -Verschiebung von 3.91 ppm (siehe Abbildung 126 A) gezeigt.

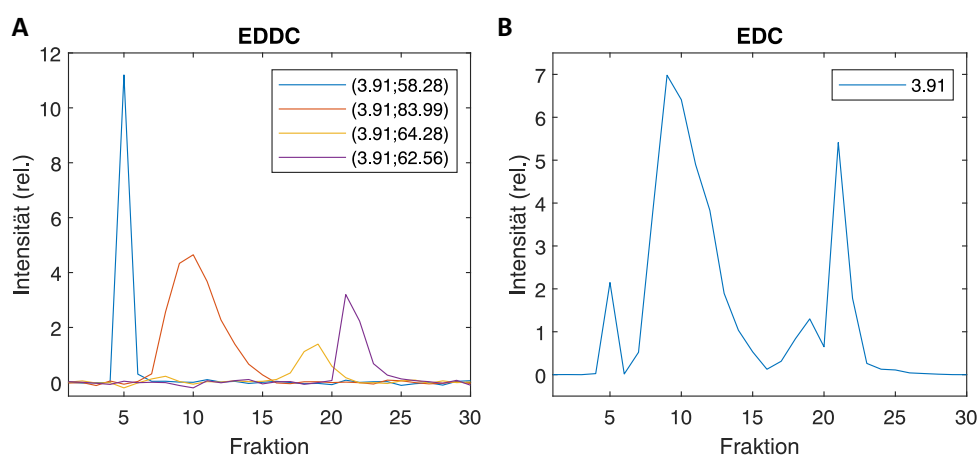


Abbildung 126: Gegenüberstellung der Verläufe von EDDCs und EDCs im künstlichen Gemisch B bei der chemischen Verschiebung in der ^1H -Dimension von 3.91 ppm. **A** zeigt die Verläufe vier verschiedener EDDCs, die allesamt bei der gleichen ^1H -Verschiebung von 3.91 ppm auftreten, sich aber durch die ^{13}C -Verschiebung unterscheiden. In **B** ist die äquivalente Darstellung des EDCs bei der chemischen Verschiebung von 3.91 ppm zu sehen.

Die vier EDDC-Verläufe in Abbildung 126 A sind eindeutig voneinander zu unterscheiden, da die vier Signale bei verschiedenen ^{13}C -Verschiebungen auftreten. Im Gegensatz dazu weist der EDC

der gleichen ^1H -Verschiebung (3.91 ppm) in Abbildung 126 B vier Maxima auf, die demnach auf vier verschiedene Analyten zurückzuführen sind. Für eine Korrelation der ^1H -NMR-Daten ist demnach die Auswahl der individuellen Fraktionen unerlässlich. In den Fraktionen 15-22 ist die Trennung der Verläufe jedoch nicht möglich, da sich die Signalverläufe innerhalb der Fraktionen überlagern (siehe zum Vergleich EDDC-Verläufe bei 3.91/64.28 ppm und 3.91/62.56 ppm in Abbildung 126 A). Hier liegt demzufolge ein erheblicher Vorteil in der Nutzung der zweidimensionalen NMR-Spektren und der *SCORE-metabolite-ID*² Methode. Aufgrund der geringeren Sensitivität des HSQC-Experimentes im Vergleich zu einem regulären ^1H -NMR-Experiment sind allerdings höhere Konzentrationen im Allgemeinen bzw. längere Experimentdauern vonnöten.

Die isobaren Verbindungen Leucin (**5**) und Isoleucin (**6**) (monoisotopische Masse: 131.0946 Da) sowie Valeriansäure (**92**) und Isovaleriansäure (**93**) (monoisotopische Masse: 102.0681 Da) kommen in überschneidenden Fraktionen vor und können demnach nicht über die EMCs voneinander unterschieden werden. Wie auch schon bei der *SCORE-metabolite-ID* Methode können jedoch auch bei der *SCORE-metabolite-ID*² Methode die HSQC-Signale untereinander korreliert werden. Darüber lassen sich jeweils alle NMR-Signale der einzelnen Analyten einander zuordnen. Die aus der semi-automatischen Korrelation berechneten HSQC-Spektren der vier Verbindungen sind in Abbildung 127 gezeigt.

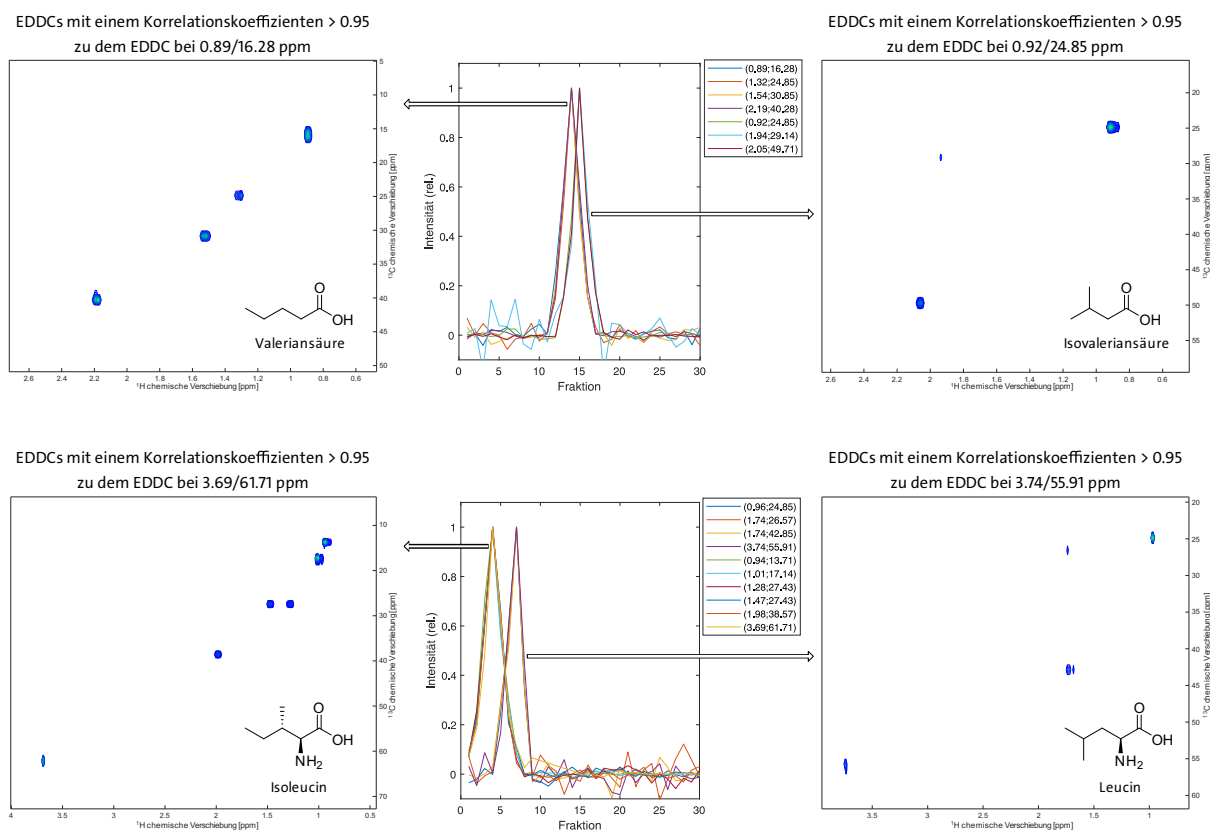


Abbildung 127: Berechnete HSQC-Spektren der isobaren Analyten Valeriansäure (**92**) und Isovaleriansäure (**93**) (oben) sowie von Isoleucin (**6**) und Leucin (**5**) (unten). Die berechneten HSQC-Spektren stellen jeweils diejenigen NMR-Signale dar, die einen Korrelationskoeffizienten > 0.95 zu dem vorgegebenen EDDC (z. B. bei 0,89/16,28 ppm im Fall der Valeriansäure (**92**)) aufweisen. In der Mitte sind jeweils alle EDDC-Verläufe der Substanzen dargestellt.

Die isobaren Substanzen Ascorbinsäure (**88**) und Tricarballoylsäure (**91**) mit einer monoisotopischen Masse von 176,0321 Da treten in verschiedenen Fraktionen auf, sodass in diesem Fall die Anzahl der Fraktionen, über die die Korrelationskoeffizienten berechnet werden sollen, prinzipiell eingeschränkt werden kann. Dadurch ist es möglich, individuelle Korrelationsanalysen für beide Analyten durchzuführen. Dies kann im positiven Ionisationsmodus beispielsweise anhand des Signals des Natrium-Adduktes der Substanzen bei $m/z = 199.02$ durch Einschränkung auf die Fraktionen 1-21 für Tricarballoylsäure (**91**) und auf die Fraktionen 21-30 für Ascorbinsäure (**88**) realisiert werden. Dies ist in der folgenden Abbildung 128 anhand der farblichen Hinterlegung des EMC-Verlaufes dargestellt.

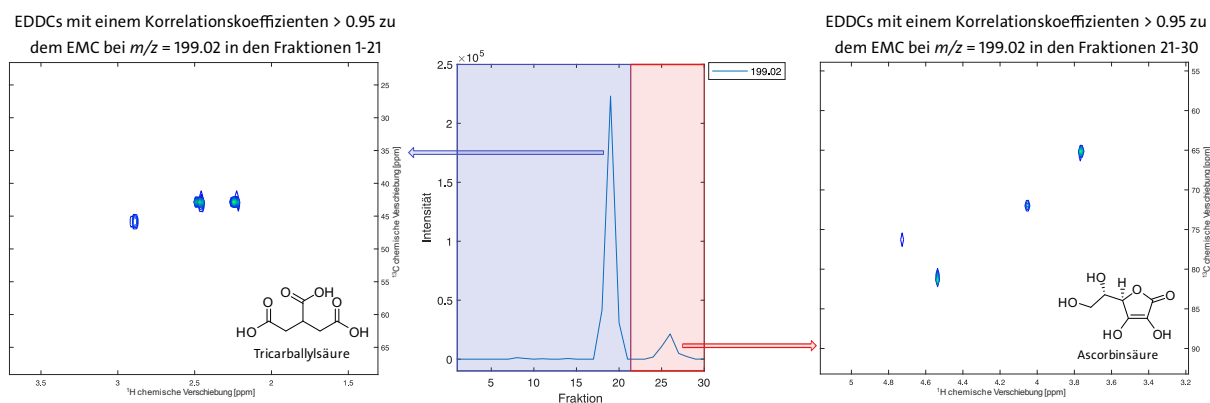


Abbildung 128: Berechnete HSQC-Spektren der isobaren Analyten Tricarballysäure (91) (links) und Ascorbinsäure (88) (rechts) anhand der EDDCs, die einen Korrelationskoeffizienten $> 0,95$ zu dem EMC bei $m/z = 199,02$ in den Fraktionen 1-21 für Tricarballysäure (91) (blau hinterlegter Bereich) bzw. in den Fraktionen 21-30 für Ascorbinsäure (88) (rot hinterlegter Bereich) aufweisen.

Durch die Einschränkung der Fraktionen können die HSQC-Spektren der beiden Verbindungen über die hoch-korrelierenden EDDCs berechnet werden. Es ist zu erwarten, dass in den Massenspektren, die im negativen Ionisationsmodus aufgenommen wurden, der EMC des deprotonierten Moleküls bei $m/z = 175,03$ ebenfalls zwei Maxima in den Fraktionen 19 und 26 aufweist. Entgegen der Erwartung zeigt der EMC jedoch nur noch ein Maximum in Fraktion 19. Die Ursache für diesen Effekt liegt in der Reaktivität des Analyten Ascorbinsäure (88), der ein natürlich vorkommendes Antioxidans ist. Unter den gewählten Bedingungen wird Ascorbinsäure (88) zunächst reversibel zu Dehydroascorbinsäure (96) (monoisotopische Masse: 174,0164 Da) oxidiert. Nach Studien von Szultka *et al.* kann Dehydroascorbinsäure daraufhin durch Hydrolyse des Lactons irreversibel zu 2,3-Diketogulonsäure (97) (monoisotopische Masse: 192,0270 Da) abgebaut werden.^[201]

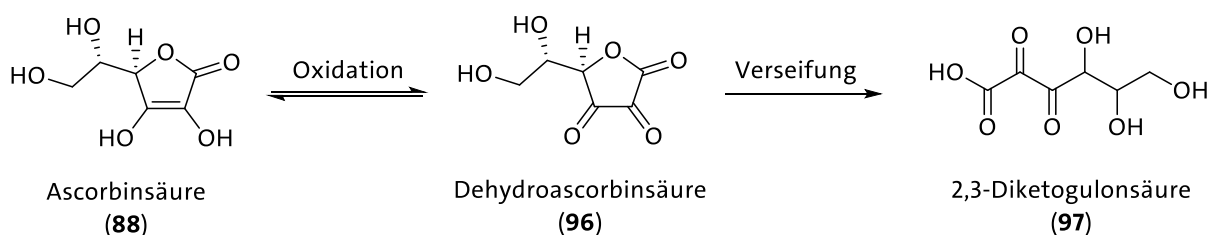


Abbildung 129: Abbaureaktion der Ascorbinsäure (88). Durch Oxidation entsteht zunächst Dehydroascorbinsäure (96), die daraufhin irreversibel zu 2,3-Diketogulonsäure (97) abgebaut wird.^[201]

Im negativen Ionisationsmodus wird das Signal des deprotonierten Moleküls der Ascorbinsäure (88) aus diesem Grund nicht detektiert, sondern stattdessen lediglich die Signale

der deprotonierten Dehydroascorbinsäure (**96**) ($m/z = 173.03$) sowie der deprotonierten 2,3-Diketogulonsäure (**97**) ($m/z = 191.02$) und dessen Hydrat ($m/z = 209.03$). Die isobaren Analyten Ascorbinsäure (**88**) und Tricarballoylsäure (**91**) weisen demzufolge im negativen Ionisationsmodus unterschiedliche Masse-zu-Ladungs-Verhältnisse auf und können anhand dessen theoretisch unterschieden werden.

Die Analyten Nicotinsäure (**50**) und Nicotinamid (**55**) weisen aufgrund des gleichen Strukturmotivs in Form eines Pyridinrings, der in Position 3 entsprechend substituiert ist, sehr ähnliche NMR-Spektren auf. Beide Analyten treten in überschneidenden Fraktionen auf, wobei lediglich das Konzentrationsmaximum um eine Fraktion zueinander verschoben ist (siehe Abbildung 130 A). Die HSQC-Signale der beiden Verbindungen liegen deshalb in der summierten Darstellung der HSQC-Spektren aller Fraktionen teilweise nicht isoliert vor, sondern sind als verbreitertes Signal erkennbar. Zusätzlich erschweren *Shifts* der Signale im Verlauf der Fraktionen die Auswertung an dieser Stelle. Dies betrifft die grau hinterlegten Signale in Abbildung 130 B.

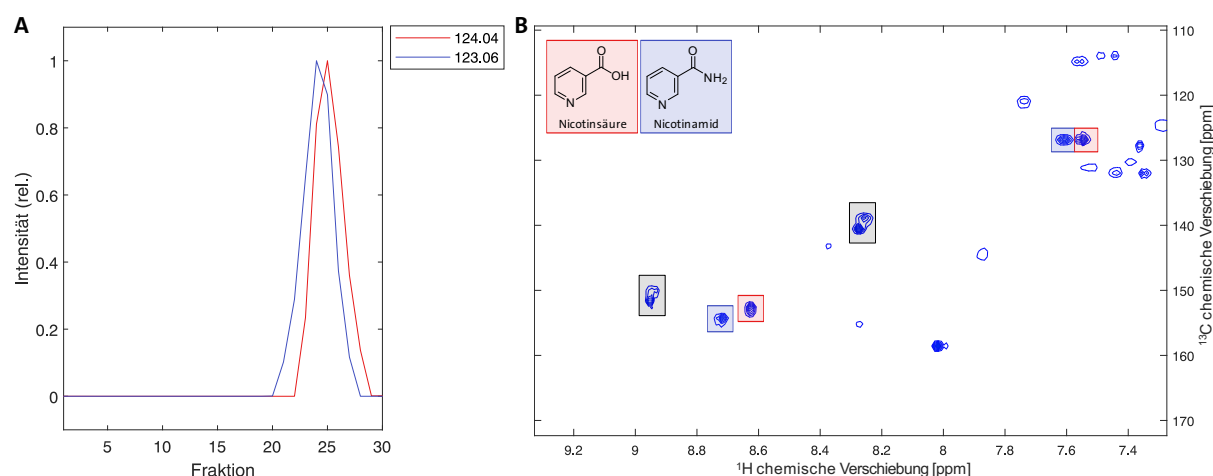


Abbildung 130: A: Verläufe der EMCs im positiven Ionisationsmodus von Nicotinamid (**55**) bei $m/z = 123.06$ ($[\text{M}+\text{H}]^+$) und von Nicotinsäure (**50**) bei $m/z = 124.04$ ($[\text{M}+\text{H}]^+$). B: Ausschnitt aus der summierten Darstellung der HSQC-Spektren der Fraktionen 1-30. Durch die Korrelationsanalyse zu den EMCs können lediglich die rot und blau markierten Signale eindeutig der Nicotinsäure (**50**) und dem Nicotinamid (**55**) zugeordnet werden. Die grau hinterlegten Signale treten als verbreiterte Signale beider Analyten auf. Neben der strukturellen Ähnlichkeit führen auch pH-Wert induzierte *Shifts* der Signale im Verlauf der Fraktionen zu der Verbreiterung der Signale und damit zu uneindeutigen EDDCs.

Im Fall von Nicotinsäure (**50**) und Nicotinamid (**55**) ist demzufolge lediglich die Zuordnung von jeweils zwei HSQC-Signalen durch die Korrelation derjenigen EDDCs zu den in Abbildung 130 A

gezeigten EMCs möglich. Diese Signale sind in der summierten Darstellung der HSQC-Spektren farblich hervorgehoben. Eine eindeutige Zuordnung der übrigen Signale zu den EMCs der beiden Analyten ist aus dem oben genannten Grund für diese beiden Analyten im artifiziellen Gemisch B jedoch nicht eindeutig möglich.

Für die weiteren Analyten des artifiziellen Gemisches B ist analog zu der zu Beginn des Kapitels gezeigten Auswertung von Uridin (**7**) eine eindeutige Korrelation zwischen EDDCs und EMCs sowie die eindeutige semi-automatische Korrelation und graphische Auswertung möglich. Dies ist für die Analyten Tryptophan (**25**), Threonin (**28**) und Pantothersäure (**89**) in der folgenden Abbildung 131 gezeigt. Zum direkten Vergleich sind zusätzlich die Referenzspektren der Reinsubstanzen gezeigt.

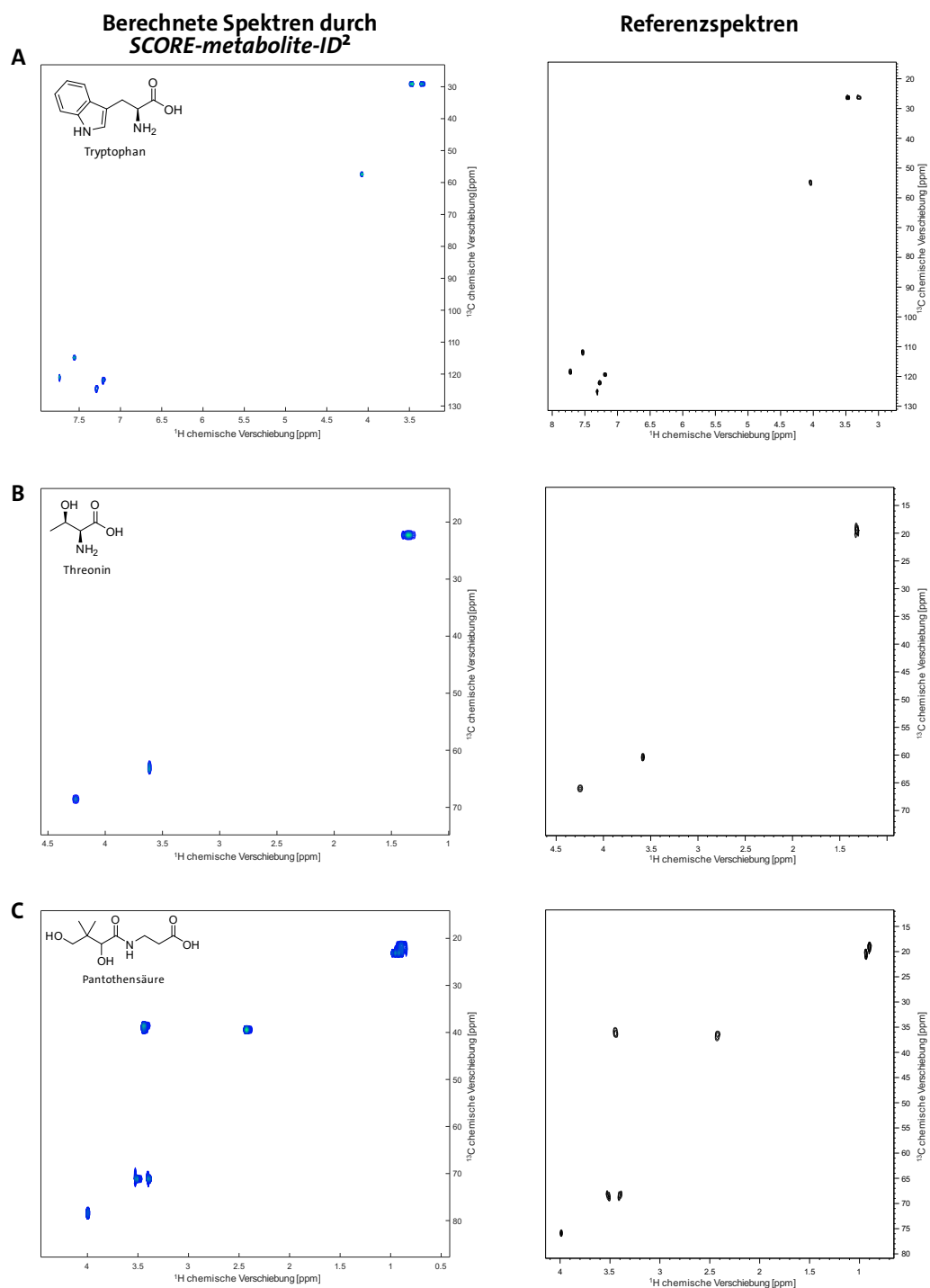


Abbildung 131: Über die semi-automatische Korrelationsanalyse berechnete HSQC-Spektren der Analyten Tryptophan (**25**), Threonin (**28**) und Pantothensäure (**89**). Die berechneten HSQC-Spektren der weiteren Analyten sind im Anhang in Abbildung 155 gezeigt. Für alle Analyten wurde die Korrelation über alle 30 Fraktionen durchgeführt. Die durch *SCORE-metabolite-ID²* berechneten Spektren entsprechen den EDDCs, die einen Korrelationskoeffizienten > 0.95 zu einem EMC des Analyten im positiven Ionisationsmodus aufweisen ($m/z = 205.10$ für Tryptophan (**25**) ($[M+H]^+$), 74.06 für Threonin (**28**) ($[M+H-H_2O-CO]^+$), 220.12 für Pantothensäure (**89**) ($[M+H]^+$)). Rechts sind zum Vergleich die Referenzspektren der Analyten gezeigt.

Der Vergleich der durch *SCORE-metabolite-ID*² berechneten HSQC-Spektren (links) mit den Referenzspektren der Reinsubstanzen (rechts) zeigt, dass in den meisten Fällen alle relevanten Signale der Analyten ermittelt werden. Dies trifft auch auf die weiteren Analyten zu, die im Anhang in Abbildung 155 gezeigt sind. Lediglich im Spektrum von Phenylalanin (**4**) (Abbildung 155 E) tritt ein zusätzliches Signal bei einer chemischen Verschiebung von 7.05/119.13 ppm auf, das dem Analyten Vanillin (**87**) zuzuordnen ist. Die fehlerhafte Zuordnung im berechneten HSQC-Spektrum von Phenylalanin (**4**) tritt auf, da dieses Signal im Verlauf der Fraktionen einem *Shift* der chemischen Verschiebung unterliegt. Der Analyt Tiglinsäure (**90**) konnte in den Massenspektren weder im positiven noch im negativen Ionisationsmodus detektiert werden, sodass in diesem Fall lediglich die Korrelation zwischen den EDDCs durchgeführt wurde. Im Allgemeinen wird jedoch durch diese Berechnung der HSQC-Spektren somit neben der visuellen Analyse ein einfacher Datenbankvergleich ermöglicht, ohne dass die Analyten isoliert werden müssen.

Die vorliegenden Ergebnisse zur Validierung der *SCORE-metabolite-ID*² Methode anhand des artifiziellen Gemisches B zeigen, dass die Korrelation der zweidimensionalen HSQC-Daten fortlaufender Fraktionen mit Daten aus Massenspektren prinzipiell möglich ist und vielversprechende Analyseergebnisse erzielt. Insbesondere die zusätzliche Frequenzdomäne, die einerseits einen höheren Informationsgehalt erzielt und zum anderen die Überlagerung der Signale reduziert, ist für die Analyse von komplexen Gemischen von enormem Vorteil. Bei ausreichender Konzentration der Analyten ist durch die Anwendung von *SCORE-metabolite-ID*² somit ein zusätzlicher Informationsgewinn im Vergleich zur Anwendung von *SCORE-metabolite-ID* gegeben.

6 Experimenteller Teil

6.1 Chemikalien

Die im Rahmen dieser Arbeit verwendeten Chemikalien und Lösungsmittel wurden von den im Folgenden angegebenen Firmen bezogen und ohne weitere Reinigung verwendet. Deuteriumoxid (99.9 %), Acetonitril- d_3 (99.8 %), 3-(Trimethylsilyl)-propionsäure-2,2,3,3- d_4 -Natriumsalz (TSP- d_4 , 99.0 %) wurden von der Firma DEUTERO GmbH bezogen. Methanol- d_4 (99.8 %) und Chloroform- d (99.8 %) wurden von der Firma EURISOTOP erworben. Natriumazid (99.5 %), Kaliumdihydrogenphosphat (≥ 99.0 %), Dikaliumhydrogenphosphat (≥ 98 %), Ameisensäure (≥ 98 %), Äpfelsäure (≥ 99.0 %), Histamin (≥ 97 %), β -Alanin (99 %), Creatinin (≥ 98 %), Vanillin (99 %), Nicotinamid (≥ 99.5 %), Nicotinsäure (≥ 98 %), ($\pm$)-Carnitin-Hydrochlorid (≥ 98 %), Thiamin-Hydrochlorid, Chlorogensäure (≥ 95 %), Chinasäure (98 %), D-Pinitol (95 %), D-Raffinose-Pentahydrat (98 %), Fumarsäure (≥ 99.0 %) und Cyanurchlorid (99 %) wurden von SIGMA ALDRICH bezogen. Acetonitril (≥ 99.9 %) für die Flash-Chromatographie und die Extraktion sowie Cholinchlorid (99 %) und Tiglinsäure (*trans*-2,3-Dimethylacrylsäure; 98 %) wurden von THERMO FISHER SCIENTIFIC erworben. Für die Aufnahme der Massenspektren wurden Wasser (LiChrosolv®, Supelco®) und Acetonitril (LiChrosolv®, Supelco®) von der Firma MERCK KGaA bezogen. Ebenso wurden D-Glucose Monohydrat, Valeriansäure (≥ 98 %), Ascorbinsäure (≥ 99.7 %), Adenosin (98 %), L-Asparaginsäure (99 %), (*R,R*)-2,3-Butandiol (97 %) und 2,3-Butandiol (Gemisch der D- und L-Form; 98 %) von MERCK KGaA erworben. Methanol (≥ 99.8 %), 2-Propanol (≥ 99.8 %) und Stachyose (≥ 97 %) wurden von VWR INTERNATIONAL GmbH erworben. Milchsäure (85 %) wurde von TH. GEYER GmbH & Co. KG bezogen. Natriumhydroxid (99.0 %) wurde bei der Firma GRÜSSING GmbH erworben. L-Asparagin-Monohydrat (≥ 99.5 %), L-Glutaminsäure (≥ 99.5 %), L-Isoleucin (> 99.5 %), L-Leucin (> 99.5 %), L-Phenylalanin (> 99.0 %), L-Valin (≥ 99.5 %), L-Threonin (> 99.5 %), L-Histidin (≥ 99.5 %), L-Tryptophan (> 99.5 %) und L-Alanin wurden von der Firma FLUKA bezogen. Uridin (≥ 99 %) und Pantothensäure (≥ 98 %) wurden bei CARL ROTH GmbH + Co. KG erworben. Tricarallylsäure (98 %) wurde von ALFA AESAR bezogen. Zur Herstellung von 6-Aminocapronsäure wurde ϵ -Caprolactam von der BASF SE und Natronlauge (33 %) von STOCKMEIER CHEMIE GmbH & Co. KG erworben.

6.2 Materialien

Es wurden Reference® 2 Pipetten der Firma EPPENDORF mit den entsprechenden epT.I.P.S® Spitzen verwendet. Als Probengefäße wurden 2 mL Safe-Lock Tubes der Firma EPPENDORF sowie 15 mL und 50 mL Falcon-Tubes der Firma SARSTEDT AG & Co. KG verwendet. Es wurden 5 mm NMR-Röhrchen *boroeco-5-7* der Firma DEUTERO GMBH sowie 3 mm NMR-Röhrchen der Firma HILGENBERG GmbH genutzt. Außerdem wurden Amicon® *Ultra centrifugal Filter (cut-off: 3 kDa)* der Firma MERCK verwendet. Für die Flash-Chromatographie wurde eine CHROMABOND DL 40 sowie eine DL 4 Leerkartusche der Firma MACHEREY-NAGEL verwendet. Als Sorbens wurde das 3-Aminopropyl-funktionalisierte Kieselgel POLYGOPREP 60-30 NH₂ der Firma MACHEREY-NAGEL genutzt. Für LC-MS-Messungen wurde eine NUCLEOSHELL HILIC Säule mit einer Partikelgröße von 2.7 µm, einem Innendurchmesser von 2 mm und einer Länge von 150 mm verwendet.

6.3 Geräte

Zur Probenvorbereitung der Walnuss-Proben wurde eine WAL MAN Handknackmaschine der Firma FEUCHT-OBSTTECHNIK GMBH verwendet. Die Vermahlung der Nussproben fand in einer Grindomix GM 300 Messermühle der Firma RETSCH statt. Zur Gefriertrocknung des Nussmehls und von allen wässrigen Proben wurde eine ALPHA 1-4 LD_{plus} sowie eine ALPHA 1-2 Gefriertrocknungsanlage der Firma CHRIST verwendet. Zur Lösungsmittelentfernung wurde außerdem teilweise der *Speed Vacuum Concentrator* Savant SPD121P der Firma THERMO FISHER SCIENTIFIC verwendet. Des Weiteren wurde der Bead Ruptor 24 der Firma OMNI INTERNATIONAL zur Extraktion von Proben im kleinen Maßstab verwendet. Die Durchmischung von Proben erfolgte mittels des Digital Vortexmischers von VQR INTERNATIONAL. Bidestilliertes Wasser wurde über das arium® pro UV-Reinstwassersystem der Firma SARTORIUS erhalten. Es wurde das Ultraschallbad Sonorex RK 1028 BH der Firma BANDELIN verwendet, um Lösungsmittel für LC-MS Messungen zu entgasen. Des Weiteren wurde die Waage A200S der Firma SARTORIUS sowie die Zentrifugen 5804R, 5417R und MiniSpin® von EPPENDORF verwendet. Zur Einstellung der pH-Werte wurde das Orion Star A111 pH-Meter der Firma THERMO FISHER SCIENTIFIC verwendet. Zusätzlich wurde die Messelektrode *pH Electrode InLab NMR* der Firma METTLER TOLEDO zur Überprüfung der pH-Werte in kleinen Volumina verwendet.

6.4 Software

Für die Entwicklung der *SCORE-metabolite-ID* und *SCORE-metabolite-ID*² Applikation sowie zur Erstellung von Abbildungen wurde MATLAB R2020B sowie MATLAB R2022B (THE MATHWORKS, INC.) verwendet. Zur weiteren Erstellung von Grafiken sowie zur Erstellung der Hauptkomponentenanalyse wurde OriginPro 2022 und OriginPro 2023 (ORIGINLAB CORPORATION) verwendet. Das *Bucketing* der NMR-Spektren wurde in Amix 3.9.14 von BRUKER durchgeführt. Weiterhin wurden RStudio (Version 2022.07.1) von RSTUDIO, PBC, in Verknüpfung mit R 4.2.1 von R CORE TEAM und RTools 4.2 von THE R FOUNDATION, sowie Python 3.2.2 von PYTHON SOFTWARE FOUNDATION verwendet. Die Auswertung von NMR-Spektren erfolgte mit TopSpin 4.1.4 (BRUKER BIOSPIN). Die Auswertung der Massenspektren wurde mit DataAnalysis 4.2 (BRUKER DALTONICS) durchgeführt. Chemische Strukturformeln wurden mit Hilfe von ChemDraw Professional (PERKINELMER INFORMATICS, INC.) erstellt. Ein *Alignment* der Spektren wurde in MestReNova 14.1.1 (MESTRELAB RESEARCH S. L.) durchgeführt.

6.5 Artificielle Gemische A und B

Zur Validierung von *SCORE-metabolite-ID* und *SCORE-metabolite-ID*² wurden die beiden artifiziellen Gemische A und B hergestellt, indem bekannte Substanzen in jeweils unterschiedlichen Konzentrationen so zusammengegeben wurden, dass ein künstlicher Konzentrationsverlauf der einzelnen Substanzen in Analogie zu einem tatsächlichen Elutionsprofil entsteht.

6.5.1 Phosphatpuffer (200 mM, pH 7.0, 1 mM TSP-*d*₄, 3 mM NaN₃)

Für die Proben des artifiziellen Gemisches A wurde ein 200 mM Phosphatpuffer verwendet. Es wurden 6.80 g (50.0 mmol) Kaliumdihydrogenphosphat, 8.71 g (50.0 mmol) Dikaliumhydrogenphosphat, 86.1 mg (0.500 mmol) TSP-*d*₄ und 97.5 mg (1.50 mmol) Natriumazid in 500 mL Deuteriumoxid gelöst. Der pH-Wert betrug 7.0.

6.5.2 Phosphatpuffer (1 M, pH 7.0, 2 mM TSP- d_4 , 3 mM NaN₃)

Für die Proben des artifiziellen Gemisches B wurde ein 1 M Phosphatpuffer verwendet. Zur Herstellung wurden 17.01 g (125.0 mmol) Kaliumdihydrogenphosphat, 21.77 g (125.0 mmol) Dikaliumhydrogenphosphat, 86.2 mg (0.500 mmol) TSP- d_4 und 48.3 mg (0.743 mmol) Natriumazid in 250 mL Deuteriumoxid gelöst und der pH-Wert mittels deuterierter Salzsäure (4 M) und deuterierter Natronlauge (4 M) auf 7.0 eingestellt. Derselbe Puffer wurde auch für die Probenvorbereitung der Fraktionierung der Apfelproben (siehe Abschnitt 6.9.2) verwendet.

6.5.3 Probenvorbereitung

Artifiziellles Gemisch A

Für die Erstellung des artifiziellen Gemisches A, das aus zehn Substanzen in verschiedenen Konzentrationsprofilen besteht, wurden Stammlösungen von L-Isoleucin, L-Leucin, L-Phenylalanin, L-Alanin, Cholinchlorid, Uridin, Histamin, β -Alanin, Creatinin und D-Glucose mit einer Konzentration von 0.1 M in Wasser (LC-MS *grade*) hergestellt. Für jede Substanz wurde ein artifizieller Konzentrationsverlauf erstellt, indem bestimmte Volumina der jeweiligen Stammlösungen in die verschiedenen Proben, d.h. Fraktionen, gegeben wurden. Die verwendeten Mengen sind in Tabelle 30 aufgeführt. Jede Fraktion wurde mit Wasser (LC-MS *grade*) anschließend auf 1000 μ L aufgefüllt.

Tabelle 30: Menge aus den Stammlösungen [μ L], die zu den jeweiligen Proben, d.h. Fraktionen, für die Herstellung des artifiziellen Gemisches A zugegeben wurden. Die Stammlösungen enthalten jeweils die in Wasser (LC-MS *grade*) gelöste Substanz mit einer Konzentration von 0.1 M. Ile: L-Isoleucin, Leu: L-Leucin, Cho.: Cholin, Uri.: Uridin, Creat.: Creatinin, Ala: L-Alanin, Phe: L-Phenylalanin, Glc: D-Glucose, β -Ala: β -Alanin, Hist.: Histamin.

Frakt.	Ile	Leu	Cho.	Uri.	Creat.	Ala	Phe	Glc	β -Ala	Hist.
1	10	0	0	0	0	0	0	0	0	0
2	30	10	0	0	0	0	0	0	0	0
3	80	30	0	0	0	0	0	0	0	0
4	130	80	0	20	0	0	0	0	0	0
5	130	130	0	40	0	0	0	0	0	0
6	80	130	0	80	0	0	0	0	0	0
7	30	80	5	120	0	0	0	0	0	0
8	10	30	10	160	0	0	0	0	0	0

Frakt.	Ile	Leu	Cho.	Uri.	Creat.	Ala	Phe	Glc	β -Ala	Hist.
9	0	10	20	200	0	0	0	0	0	0
10	0	0	25	220	0	0	0	0	0	0
11	0	0	20	210	0	0	0	0	0	0
12	0	0	10	160	20	20	0	0	0	0
13	0	0	5	115	80	80	0	0	0	0
14	0	0	0	70	180	200	10	0	0	0
15	0	0	0	5	160	80	50	0	0	0
16	0	0	0	0	130	20	130	0	0	0
17	0	0	0	0	100	0	180	10	0	0
18	0	0	0	0	70	0	130	100	0	0
19	0	0	0	0	20	0	50	10	0	0
20	0	0	0	0	0	0	10	0	20	5
21	0	0	0	0	0	0	0	0	70	50
22	0	0	0	0	0	0	0	0	95	150
23	0	0	0	0	0	0	0	0	100	350
24	0	0	0	0	0	0	0	0	95	150
25	0	0	0	0	0	0	0	0	70	50
26	0	0	0	0	0	0	0	0	20	5

Von jeder Fraktion wurden 50 μ L abgenommen und für die Aufnahme der Massenspektren 500-fach mit Wasser (LC-MS *grade*) verdünnt. Das Injektionsvolumen zur Aufnahme der Massenspektren betrug 10 μ L. Das Lösungsmittel der verbleibenden 950 μ L jeder Fraktion wurde mit Hilfe des *Speed Vacuum Concentrators* entfernt, der Rückstand in 700 μ L Phosphatpuffer (200 mM) gelöst und 10-fach mit Phosphatpuffer (200 mM) verdünnt. Es wurden 600 μ L der Lösung in ein 5 mm NMR-Röhrchen überführt.

Artifizielles Gemisch B

Das artifizielle Gemisch B besteht aus 20 Substanzen, die über 30 Fraktionen in unterschiedlichen Konzentrationsprofilen zusammengegeben wurden, um ein artifizielles Elutionsprofil jeder Substanz zu erstellen. Das artifizielle Gemisch B wurde zur Validierung der

Methode *SCORE-metabolite-ID*² verwendet. Es wurden Stammlösungen in einer Konzentration von 0.1 mol/L von L-Phenylalanin, Vanillin, Valeriansäure, Isovaleriansäure, L-Threonin, Äpfelsäure, Tricarballoylsäure, L-Leucin, Adenosin, Nicotinamid, Nicotinsäure, Tiglinsäure, L-Isoleucin, L-Tryptophan, (±)-Carnitin-Hydrochlorid, Pantothersäure, Uridin, Thiamin-Hydrochlorid, Ascorbinsäure und Saccharose zum einen in Wasser (LC-MS *grade*) zur späteren Untersuchung mittels Massenspektrometrie und zum anderen in Phosphatpuffer (1 M) für die NMR-spektroskopische Untersuchung hergestellt. Entsprechend der Tabelle 31 wurden die angegebenen Volumina der Stammlösungen in die jeweiligen Proben, d.h. Fraktionen, zusammengegeben, um ein artifizielles Elutionsprofil für jede Substanz zu erstellen. Jede Fraktion wurde im Anschluss mit Wasser (LC-MS *grade*) bzw. mit Phosphatpuffer (1 M) auf 1000 µL aufgefüllt. Zur weiteren Vorbereitung der Proben für die massenspektrometrischen Messungen wurden 50 µL der in Wasser verdünnten Fraktionen entnommen und 1000-fach mit Wasser (LC-MS *grade*) verdünnt. Das Injektionsvolumen für die Aufnahme der Massenspektren betrug 5 µL für Messungen im positiven Ionisationsmodus und 20 µL für Messungen im negativen Ionisationsmodus.

Tabelle 31: Menge aus den Stammlösungen [μL], die zu den jeweiligen Proben, d.h. Fraktionen, für die Herstellung des artifiziellen Gemisches B zugegeben wurden. Die Stammlösungen enthalten jeweils die in Wasser (LC-MS *grade*) gelöste Substanz mit einer Konzentration von 0.1 M. Phe: L-Phenylalanin, Van.: Vanillin, Ile: L-Isoleucin, Leu: L-Leucin, Val.: Valeriansäure, Iso.: Isovaleriansäure, Thr: L-Threonin, Äpf.: Äpfelsäure, Tric.: Tricarallylsäure, Ade.: Adenosin, NiA.: Nicotinamid, NiS.: Nicotinsäure, Tig.: Tiglinsäure, Trp: L-Tryptophan, Car.: Carnitin, Pan.: Pantothersäure, Uri.: Uridin, Thi.: Thiamin, Asc.: Ascorbinsäure, Sac.: Saccharose.

Frakt.	Phe	Van.	Ile	Leu	Val.	Iso.	Thr	Äpf.	Tric.	Ade.	NiA.	NiS.	Tig.	Trp	Car.	Pan.	Uri.	Thi.	Asc.	Sac.
1	0	0	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	20	30	160	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	40	100	260	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	50	360	160	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
6	40	100	60	40	0	0	0	0	0	0	0	0	0	0	0	40	0	0	0	0
7	20	30	20	80	0	0	0	0	0	0	0	0	0	0	0	80	40	0	0	20
8	0	0	0	30	0	0	0	0	0	0	0	0	0	0	0	140	140	0	0	160
9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	220	40	0	0	260
10	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	300	0	0	0	260
11	0	0	0	0	0	0	20	0	0	0	0	0	0	0	0	360	0	0	0	200
12	0	0	0	0	30	0	60	0	0	0	0	0	0	0	0	360	0	0	0	140
13	0	0	0	0	80	30	100	0	0	0	0	0	0	0	0	160	0	0	0	80
14	0	0	0	0	160	80	60	0	0	0	0	0	0	0	0	40	0	0	0	40
15	0	0	0	0	80	160	20	0	0	0	0	0	0	0	40	0	0	0	0	20
16	0	0	0	0	30	80	0	0	0	60	0	0	0	0	160	0	0	0	0	0
17	0	0	0	0	0	30	0	80	0	140	0	0	0	0	260	0	0	0	0	0

Experimenteller Teil

Frakt.	Phe	Van.	Ile	Leu	Val.	Iso.	Thr	Äpf.	Tric.	Ade.	NiA.	NiS.	Tig.	Trp	Car.	Pan.	Uri.	Thi.	Asc.	Sac.
18	0	0	0	0	0	0	0	420	40	300	0	0	0	0	40	0	0	0	0	0
19	0	0	0	0	0	0	0	80	200	400	0	0	0	0	0	0	0	0	0	0
20	0	0	0	0	0	0	0	0	40	180	0	0	0	0	0	0	0	60	0	0
21	0	0	0	0	0	0	0	0	0	60	30	0	20	0	0	0	0	300	0	0
22	0	0	0	0	0	0	0	0	0	0	80	0	50	0	0	0	0	240	0	0
23	0	0	0	0	0	0	0	0	0	0	160	60	60	0	0	0	0	120	0	0
24	0	0	0	0	0	0	0	0	0	0	240	200	50	20	0	0	0	60	20	0
25	0	0	0	0	0	0	0	0	0	0	240	300	20	60	0	0	0	20	40	0
26	0	0	0	0	0	0	0	0	0	0	120	240	0	160	0	0	0	0	100	0
27	0	0	0	0	0	0	0	0	0	0	40	140	0	260	0	0	0	0	40	0
28	0	0	0	0	0	0	0	0	0	0	0	50	0	160	0	0	0	0	20	0
29	0	0	0	0	0	0	0	0	0	0	0	0	0	60	0	0	0	0	0	0
30	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	0	0	0	0	0

6.5.4 NMR-Spektroskopie

Artifiziellles Gemisch A

Die NMR-Spektren des artifiziellen Gemisches A wurden an einem BRUKER Avance III HD 600 MHz NMR Spektrometer, das mit einem 5 mm BBFO Probenkopf ausgestattet ist, bei einer Feldstärke von 600.13 MHz und unter Verwendung der Software TopSpin 3.6.2 aufgenommen. Die ^1H -NMR-Spektren wurden mit Hilfe der *noesygppr1d* Pulssequenz und unter Verwendung des *baseopt* Filters bei einer Temperatur von 298 K akquiriert. Es wurden jeweils 32 *Scans* sowie 4 *Dummy Scans* aufgenommen. Es wurden 65 536 Datenpunkte aufgenommen und die spektrale Weite betrug 20.5 ppm. Der *transmitter frequency offset* wurde auf die Resonanz des Wassersignals von 2824 Hz eingestellt. Der *Receiver gain* wurde auf 64 eingestellt. Es ergab sich eine Gesamtmessdauer von ca. 4 Minuten pro Probe. Die erhaltenen FIDs wurden mit einer Exponentialfunktion mit einem *line broadening* Faktor von 0.3 Hz multipliziert und anschließend einer Fourier-Transformation unterzogen. Es erfolgte eine Kalibrierung gegen TSP- d_4 ($\delta = 0.00$ ppm), eine automatische Phasenkorrektur 0. Ordnung (Befehl *apk0*) sowie eine automatische Basislinienkorrektur (Befehl *absn*).

Artifiziellles Gemisch B

Zur Validierung von *SCORE-metabolite-ID²* wurden 2D HSQC-NMR Spektren an einem BRUKER Avance III HD 600 MHz NMR-Spektrometer unter der Verwendung der Software TopSpin 3.6.4, ausgestattet mit einem 5 mm BBFO-Probenkopf, bei einer Basisfrequenz von 600.13 MHz und einer Temperatur von 298 K akquiriert. Es wurde die ASAP-HSQC-Pulssequenz *ASAP-hsqcetgpsp* verwendet. Es wurden jeweils 16 *Scans* sowie 256 *Dummy Scans* aufgenommen und ein Relaxationsdelay von 0.3 Sekunden gewählt. Die spektrale Weite betrug 220 ppm in der indirekten und 11 ppm in der direkten Dimension. Es wurden 2048 Datenpunkte in der direkten Domäne und 256 Inkremente aufgenommen. Die mittlere Kopplungskonstante $^1J_{\text{CH}}$ für den INEPT-Transfer wurde auf 145 Hz (CNST4) festgelegt. Insgesamt ergab sich eine Experimentdauer von 36 Minuten und 46 Sekunden. Nach zweifacher Fourier-Transformation des FIDs erfolgte die Kalibrierung gegen TSP- d_4 ($\delta_{\text{H}} = 0.00$ ppm, $\delta_{\text{13C}} = 0.00$ ppm), eine automatische Phasenkorrektur (Befehl *apk2d*) sowie eine automatische Basislinienkorrektur (Befehl *abs2D*).

6.6 Pinienkern-Proben

Die Pinienkern-Matrix wurde als komplexe biologische Matrix pflanzlichen Ursprungs zur Testung der Identifizierung von Metaboliten mittels *SCORE-metabolite-ID* ausgewählt. Die Pinienkern-Probe wurde im lokalen Einzelhandel in Hamburg gekauft und stammt von der Firma HERBERT KLUTH GmbH & Co. KG. Zur Probenvorbereitung wurden ca. 200 g Pinienkerne in flüssigem Stickstoff schockgefroren. Unter Zugabe von ca. 300 g Trockeneis wurden die Pinienkerne mit Hilfe der Messermühle gemahlen. Die vermahlene Probe wurde anschließend für 48 Stunden gefriergetrocknet und daraufhin bei $-20\text{ }^{\circ}\text{C}$ gelagert. Andere Nüsse und Kerne (z. B. Haselnüsse, Cashewkerne, Walnüsse etc.) wurden analog aufgearbeitet.

6.6.1 Phosphatpuffer (100 mM, pH 7.0, 1 mM TSP- d_4 , 3 mM NaN_3)

Für die Vorbereitung der NMR-Proben nach der Fraktionierung des Pinienkernextraktes wurde ein 100 mM Phosphatpuffer verwendet. Dazu wurden 1.70 g (125 mmol) Kaliumdihydrogenphosphat, 2.18 g (125 mmol) Dikaliumhydrogenphosphat, 43.3 mg (0.251 mmol) TSP- d_4 sowie 48.8 mg (75.1 mmol) Natriumazid in 250 mL Deuteriumoxid gelöst. Der pH-Wert betrug 7.0.

6.6.2 Polare Extraktionsmethode

Für Testmessungen des polaren Extraktes verschiedener Nüsse und Kerne wurden je $300 \pm 2\text{ mg}$ des jeweiligen Lyophilisats eingewogen und zwei entfettete Stahlkugeln ($\varnothing = 2\text{ mm}$) zugegeben. Anschließend wurden 400 μL Methanol- d_4 , 500 μL Chloroform- d und 600 μL Phosphatpuffer zugegeben. Die Suspension wurde für drei Minuten im Bead Ruptor 24 bei 3.1 m/s extrahiert. Der Ansatz wurde daraufhin für 10 Minuten bei 14 000 rcf zentrifugiert. Es wurden 100 μL des Überstandes abgenommen und mit 600 μL Phosphatpuffer versetzt. Nach Durchmischung wurden 600 μL der Lösung in ein 5 mm NMR-Röhrchen überführt.

Zur Extraktion der Metaboliten in Pinienkernen im größeren Maßstab wurden 18.3 g des Pinienkern-Lyophilisats in 180 mL Methanol suspendiert und für 2 Stunden bei Raumtemperatur gerührt. Nachdem das Methanol unter vermindertem Druck entfernt wurde, wurde das Pinienkern-Lyophilisat in 150 mL Chloroform, 120 mL Methanol und 180 mL Reinstwasser suspendiert und für 24 Stunden bei Raumtemperatur gerührt. Die Suspension wurde

anschließend bei 4 °C und 8000 rpm für 30 Minuten zentrifugiert. Nachdem der Überstand abgenommen und das Methanol unter vermindertem Druck entfernt wurde, wurde der Extrakt gefriergetrocknet. Der Extrakt wurde in 23 mL Reinstwasser gelöst und mit Hilfe von Zentrifugenfiltern (*cut-off*: 3 kDa) zur Entfernung von Makromolekülen gereinigt. Die Zentrifugenfilter wurden vor Benutzung 20 Mal mit je 480 µL einer 0.1 M Natriumhydroxid-Lösung und dann einmal mit 480 µL Phosphatpuffer (100 mM) gewaschen, um das Feuchthaltemittel Glycerin zu entfernen. Die Filtration erfolgte bei 14 000 rcf und Raumtemperatur für jeweils 20 Minuten. Nach der Filtration durch die Zentrifugenfilter wurde der Extrakt erneut gefriergetrocknet. Es wurden 0.86 g Extrakt erhalten, welcher anschließend bei -20 °C gelagert wurde.

6.6.3 Chromatographische Fraktionierung

Zur Entwicklung der chromatographischen Fraktionierungsmethode wurde aufgrund der einfacheren Verfügbarkeit ein polares Extrakt von lyophilisierter Rinderleber verwendet. Die polare Extraktion wurde analog zur Extraktion der Pinienkerne unter Abschnitt 6.6.2 durchgeführt. Die Fraktionierung der Extrakte erfolgte an einem Sepacore® Flash System mit Control Unit C-620 in Kombination mit dem Fraction Collector C-660 und UV-Photometer C-640 der Firma BÜCHI.

6.6.3.1 Entwicklung der chromatographischen Fraktionierungsmethode

Für die Entwicklung und Optimierung der Fraktionierungsmethode wurden verschiedene Flash-Säulen und verschiedene Gradienten verwendet. In Tabelle 32 sind die Gradienten der verschiedenen Methoden aufgelistet. Für die Methoden A, B und C, die einer RP-Chromatographie entsprechen, wurde die RS 40 C₁₈ ec Säule der Firma MACHEREY-NAGEL verwendet. Für die Methoden D und E wurde die Säule EcoFlex Amino 20 g der Firma BÜCHI eingesetzt. Dabei wurde die Säule in Methode D unter RP-Bedingungen und in Methode E unter HILIC-Bedingungen eingesetzt. Für Methode F wurde eine Flash DL 40 Leersäule mit dem 3-Aminopropyl-funktionalisierten Kieselgel POLYGOPREP 60-30 NH₂ der Firma MACHEREY-NAGEL gefüllt.

Tabelle 32: Übersicht der verwendeten Methoden A-F zur Entwicklung und Optimierung der Fraktionierungsmethode. Für die Methoden A-C wurde die RS 40 C₁₈ ec Säule der Firma MACHEREY-NAGEL verwendet, für die Methoden D und E wurde die Säule EcoFlex Amino 20 g der Firma BÜCHI und für die Methode F wurde die 3-Aminopropylfunktionalisierte Kieselgel Säule der Firma MACHEREY-NAGEL verwendet. In der zweiten Zeile der Tabelle sind die jeweils verwendeten Mengen der Rinderleberextrakte angegeben. B: Acetonitril, FA: Ameisensäure.

Methode A (RP)		Methode B (RP)		Methode C (RP)		Methode D (RP)		Methode E (HILIC)		Methode F (HILIC)	
610.8 mg		663.6 mg		705.4 mg		712.9 mg		623.8 mg		486.4 mg (dry-load)	
Zeit [min]	B [%]	Zeit [min]	B +0.1 % FA [%]	Zeit [min]	B +0.1 % FA [%]	Zeit [min]	B [%]	Zeit [min]	B [%]	Zeit [min]	B [%]
0	0	0	0	0	50	0	5	0	95	0	90
15	0	10	100	15	100	2	5	2	95	5	90
25	100	15	100	17	100	12	100	27	5	12	65
30	100	20	0	25	0	15	100	30	5	15	65
		25	0			20	5			27	0
						22	5			30	0

Als Eluenten wurden Wasser (A) und Acetonitril (B) eingesetzt. Bei Methoden B und C wurde zu Wasser und Acetonitril jeweils 0.1 % Ameisensäure (FA) zugesetzt. Die Flussrate betrug für alle Methoden 20 mL/min und es wurden Fraktionen mit einem Volumen von je 10 mL aufgefangen. Die jeweils eingesetzte Menge des Rinderleberextraktes ist in der zweiten Zeile der Tabelle 32 angegeben. Für die Methoden A-E wurde das Rinderleberextrakt im jeweiligen Lösungsmittel, das zu Beginn der Fraktionierungsmethode eingesetzt wird, gelöst. Für die Methode F wurde ein *dry-load* durch Adsorption des Extraktes (486.4 mg) an 1.9 g des 3-Aminopropylfunktionalisierten Kieselgels durchgeführt.

6.6.3.2 Chromatographische Fraktionierung des Pinienkernextraktes

Zur Fraktionierung des Pinienkernextraktes wurde eine Flash DL 40 Leersäule mit dem 3-Aminopropyl-funktionalisierten Kieselgel POLYGOPREP 60-30 NH₂ der Firma MACHEREY-NAGEL gefüllt. Als Eluenten dienten Wasser (A) und Acetonitril (B). Zur Fraktionierung wurden 514 mg des getrockneten polaren Pinienkernextraktes an 2.1 g des 3-Aminofunktionalisierten Kieselgels adsorbiert. Es wurde der in Tabelle 33 gezeigte Gradient unter HILIC-Bedingungen verwendet. Im

Vergleich zu den in Tabelle 32 gezeigten Methoden entspricht dies der Methode F, wobei die Gesamtzeit der Methode um 10 Minuten verlängert wurde. Die Flussrate betrug 20 mL/min und es wurde eine Fraktionsgröße von 10 mL gewählt. Insgesamt wurden somit 80 Fraktionen erhalten.

Tabelle 33: Gradient zur Fraktionierung des polaren Pinienkernextraktes unter HILIC-Bedingungen auf einer Flash-Säule mit 3-Aminopropylfunktionalisiertem Kieselgel. B: Acetonitril.

Zeit [min]	B [%]
0	90
5	90
12	65
15	65
27	0
40	0

6.6.4 Probenvorbereitung

Nach der Fraktionierung des Pinienkernextraktes entsprechend der Fraktionierungsmethode unter Abschnitt 6.6.3.2 wurde das Lösungsmittel jeder Fraktion unter vermindertem Druck entfernt und die Fraktionen gefriergetrocknet. Der Rückstand wurde in jeweils 1000 µL Wasser (LC-MS *grade*) gelöst. Für die Aufnahme der Massenspektren wurden 50 µL abgenommen. Die Probe jeder Fraktion wurde im Verhältnis 1:500 mit Wasser (LC-MS *grade*) verdünnt. Für die Aufnahme der Massenspektren wurden 20 µL der Probe injiziert. Für die Aufnahme der NMR-Spektren wurden die restlichen 950 µL erneut gefriergetrocknet und der Rückstand daraufhin in 700 µL Phosphatpuffer (100 mM, pH 7.0, 1 mM TSP-*d*₄, 3 mM NaN₃) gelöst. Für die Messung wurden 600 µL dieser Lösung in ein 5 mm NMR-Röhrchen überführt.

6.6.5 NMR-Spektroskopie

Die NMR-Spektren der 80 Fraktionen des Pinienkernextraktes wurden an einem Avance III HD 600 MHz NMR-Spektrometer der Firma BRUKER, ausgestattet mit einem 5 mm BBFO-Probenkopf, bei einer Messfrequenz von 600.13 MHz und einer Temperatur von 298 K unter Verwendung der Software TopSpin 3.6.2 aufgenommen. Es wurden ¹H-NMR-Spektren durch Verwendung der *noesygppr1d* Pulssequenz akquiriert, wobei der *transmitter frequency offset* auf 2820 Hz

eingestellt wurde. Die Spektren wurden jeweils mit 65 536 Datenpunkten durch Aufnahme von 128 *Scans* und 4 *Dummy Scans* bei einem festen *Receiver gain* von 32 und einem Relaxationsdelay von 4 Sekunden aufgenommen. Die erhaltenen FIDs wurden mit einer Exponentialfunktion mit einem *line broadening* Faktor von 0.3 Hz multipliziert. Nach Fourier-Transformation erfolgte eine Kalibrierung gegen TSP- d_4 ($\delta = 0.00$ ppm), eine automatische Phasenkorrektur 0. Ordnung (Befehl *apk0*) und eine automatische Basislinienkorrektur (Befehl *absn*).

6.7 Apfel-Proben für multivariate Datenanalyse

Insgesamt standen 86 Apfelproben unterschiedlicher Sorten zur Verfügung, die in Hinblick auf Unterscheidungen auf Basis der verschiedenen Sorten untersucht werden sollten. Die Proben der Sorten Boskoop (20), Holsteiner Cox (15), Wellant (9), Elstar (23) und Gala (19) stammen aus den Erntejahren 2020 und 2021 (siehe Abbildung 132) und wurden vom Landeslabor Schleswig-Holstein zur Verfügung gestellt. Informationen zur Herkunft etc. können der Tabelle 37 im Anhang entnommen werden.

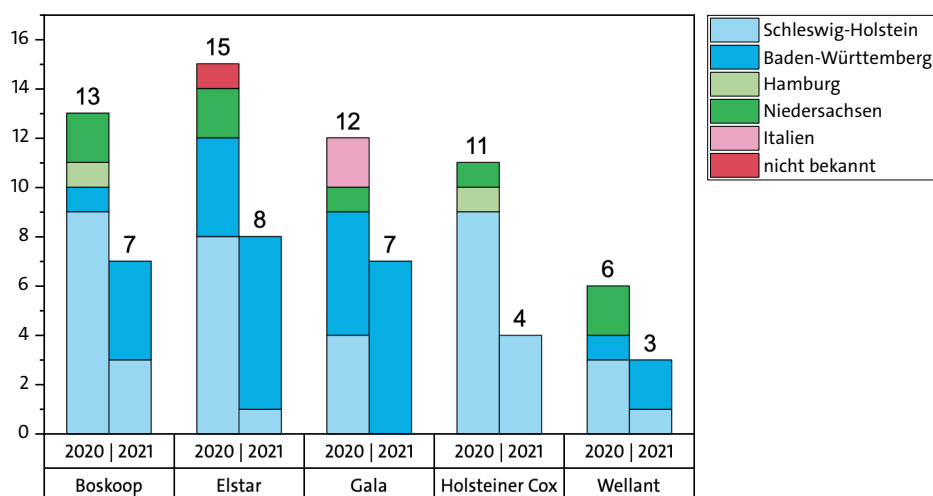


Abbildung 132: Anzahl der verwendeten 86 Apfelproben pro Sorte, Erntejahr und Herkunft unter Angabe des Bundeslandes für die Proben aus Deutschland, sofern bekannt.

6.7.1 Phosphatpuffer (3.0 M, pH 8.0, 10 mM TSP- d_4 , 3 mM NaN_3)

Für die Vorbereitung der Apfel-Proben wurde ein 3.0 M Phosphatpuffer in Deuteriumoxid hergestellt. Dazu wurden 19.6 g (113 mmol) Dikaliumhydrogenphosphat, 5.20 g (38.2 mmol)

Kaliumdihydrogenphosphat, 88.0 mg (0.511 mmol) TSP- d_4 und 10.6 mg (0.163 mmol) Natriumazid in 50 mL Deuteriumoxid gelöst und mit Hilfe von deuterierter Natronlauge (4 M) und deuterierter Salzsäure (4 M) ein pH-Wert von 8.0 eingestellt.

6.7.2 Probenvorbereitung

Die für diese Arbeit verwendeten Apfel-Proben wurden durch das Landeslabor Schleswig-Holstein zur Verfügung gestellt. Zur Probenvorbereitung wurde der gesamte Apfel, inkl. Kerngehäuse und Schale, gepresst und der Apfelsaft bei $-20\text{ }^{\circ}\text{C}$ gelagert. Insgesamt standen 86 Proben aus den Erntejahren 2020 und 2021 der Sorten Boskoop, Elstar, Gala, Wellant und Holsteiner Cox zur Verfügung. Die genauen Informationen der Apfelproben können der Tabelle 37 im Anhang unter Abschnitt 8.2 entnommen werden. Zur Aufnahme der ^1H -NMR-Spektren wurden die Apfelsäfte am Tag der Messung bei Raumtemperatur aufgetaut. Es wurden 900 μL der Apfel-Probe mit 100 μL Phosphatpuffer (3.0 M, pH 8.0, 10 mM TSP- d_4 , 3 mM NaN_3) versetzt, für 10 Sekunden mit dem Vortexmischer durchmischt und anschließend bei $4\text{ }^{\circ}\text{C}$ und 14 000 rpm für 10 Minuten zentrifugiert. Es wurden 600 μL des Überstandes in ein 5 mm NMR-Röhrchen überführt. Die ^1H -NMR-Messung erfolgte innerhalb von 18 Stunden nach der Probenvorbereitung. Jede Apfel-Probe wurde in einer Dreifachbestimmung aufgearbeitet und gemessen, wobei eine randomisierte Reihenfolge gewählt wurde.

6.7.3 NMR-Spektroskopie

Die NMR-Spektren der 86 Apfel-Proben, die zur multivariaten Datenanalyse genutzt wurden, wurden an einem Avance III 400 MHz Ultra Shield Spektrometer der Firma BRUKER unter Verwendung der Software TopSpin 3.6.2 generiert. Es wurde ein 5 mm BBO-Probenkopf verwendet. Die Spektren wurden bei einer Temperatur von 300 K aufgenommen. Die NMR-Spektren wurden mit Hilfe der *noesygppr1d* Pulssequenz unter Verwendung des *baseopt* Filters und einem Relaxationsdelays von 4 Sekunden akquiriert. Die spektrale Weite betrug 21.04 ppm und es wurden 65 536 Datenpunkte aufgenommen. Der *Receiver Gain* betrug 63.79 und der *Transmitter Frequency Offset* wurde auf 4.704 ppm eingestellt. Es wurden 4 *Dummy Scans* sowie 128 *Scans* aufgenommen, wodurch sich eine Gesamtmesszeit von 17 Minuten und 28 Sekunden pro ^1H -NMR-Spektrum ergibt.

Zur Prozessierung der erhaltenen Spektren wurde der FID (*free induction decay*) in TopSpin 4.1.4 mit einer exponentiellen Fensterfunktion mit einem *line broadening* Faktor von 0.3 Hz multipliziert. Nach der Fourier-Transformation erfolgte sowohl eine automatische Phasenkorrektur 0. Ordnung (Befehl *apk0*) als auch eine automatische Basislinienkorrektur (Befehl *absn*). Die Spektren wurden anschließend gegen TSP-*d*₄ ($\delta = 0.00$ ppm) kalibriert.

6.8 Multivariate Datenanalyse

Zur multivariaten Datenanalyse der ¹H-NMR-Spektren der 86 Apfelproben in Hinblick auf die Sortenunterscheidung wurden Hauptkomponentenanalysen (siehe Abschnitt 6.8.1) sowie Klassifizierungsmodelle auf Basis des *Support Vector Machine* Algorithmus (siehe Abschnitt 6.8.2) erstellt.

6.8.1 Hauptkomponentenanalyse

Das *Bucketing* der NMR-Spektren der Apfelproben wurde mit Hilfe der Software *Amix* der Firma BRUKER durchgeführt. Es wurde ein *Buckettable* mit 159 *Buckets* variabler Größe erstellt. Die Bereiche der gewählten *Buckets* können dem Anhang unter Abschnitt 8.3.1 entnommen werden. Es erfolgte eine Normalisierung der *Buckets* auf das Gesamtintegral der Signale. Die Hauptkomponentenanalysen (PCAs) wurden in OriginPro 2022 der Firma ORIGINLAB CORPORATION mit Hilfe der App *Principal Component Analysis* durchgeführt. Die Variablen wurden auf die Einheitsvarianz skaliert. Das Standard-Konfidenzintervall wurde auf 95 % eingestellt und die Anzahl der Hauptkomponenten wurde auf eine minimal erklärte Varianz von 95 % festgelegt. Nachdem die Dreifachbestimmungen der einzelnen Sorten auf Ausreißer untersucht wurden und diese aus der weiteren Betrachtung ausgeschlossen wurden, wurden die jeweiligen *Buckets* einer Mehrfachbestimmung gemittelt. Zusätzlich fand eine Variablenselektion durch Verwendung des *Boruta*-Pakets in *R* statt. Dadurch wurde die Anzahl der Variablen auf 82 reduziert. Für die Variablenselektion mittels *Boruta* wurde eine maximale Iteration von 1000 festgelegt und Variablen, die als *tentatives* bezeichnet wurden, wurden nicht inkludiert. Die selektierten *Buckets* zur Differenzierung aller fünf Apfelsorten sind im Anhang unter Abschnitt 8.3.2 aufgeführt.

6.8.2 Klassifizierungsmodelle

Die Klassifizierungsmodelle wurden mit Hilfe der Software MATLAB R2022a der Firma THE MATHWORKS unter Verwendung der *Classification Learner* App erstellt. Es wurde ein *Support Vector Machine* Algorithmus mit quadratischer Kernel-Funktion genutzt. Das *Box-Constraint-Level* wurde auf eins festgelegt und es wurde die Multiklassenmethode *one-vs-one* ausgewählt. Die erhaltenen Modelle wurden mittels fünffacher *repeated nested cross-validation* (RNCV) validiert. Dazu wurde der Datensatz zufällig in fünf Teile unterteilt, wobei vier Teile jeweils das Trainingsset bilden. Mit Hilfe des Trainingssets wurde das Modell in einer fünffachen inneren Kreuzvalidierung trainiert. Der fünfte Teil bildete das zugehörige Testset, mit dem das Modell anschließend getestet wurde. Innerhalb einer fünffachen äußeren Kreuzvalidierung wurde dieser Prozess nun noch weitere vier Mal wiederholt, wobei sich das Trainingsset aus jeweils anderen Kombinationen der fünf Teile ergab, d.h. jedes der fünf Sets fungierte einmal als Testset. Dieser gesamte Prozess der inneren und äußeren Kreuzvalidierung wurde daraufhin mit unterschiedlich neu gemischten Einteilungen des Datensatzes weitere vier Mal wiederholt. Die erhaltenen Konfusionsmatrizen wurden anschließend zu einer Konfusionsmatrix fusioniert und gemittelt, wobei die zugeordneten Proben als Zahl und die Genauigkeiten in Prozent angegeben wurden. Die randomisierte Einteilung der Proben in die unterschiedlichen Splits mit den jeweils unterschiedlichen Trainings- und Testdatensätzen erfolgte über ein im Arbeitskreis entwickeltes Python-Skript.^[45]

6.9 **Apfel-Proben für die Identifizierung von Biomarkern**

Anschließend an die Untersuchung der Sortenunterscheidung verschiedener Apfelproben mittels multivariater Datenanalyse, sollten spezifische Metaboliten identifiziert werden. Dazu wurde eine der 86 Apfelproben ausgewählt und die Metaboliten mit Hilfe von *SCORE-metabolite-ID* analysiert.

6.9.1 Flash-Chromatographie

Die Probe 20-E-SH-9 wurde exemplarisch ausgewählt. Diese Apfelprobe der Sorte Elstar aus dem Jahr 2020 wurde, wie unter Abschnitt 6.7.2 erläutert, gepresst und bis zur Analyse bei $-20\text{ }^{\circ}\text{C}$ gelagert. Nach dem Auftauen des gepressten Apfelsaftes, der noch Reste des Fruchtfleisches

enthielt, wurde die Probe für 10 Minuten bei 4 °C und 8000 rpm zentrifugiert. Es wurden 47.0 g des Überstandes abgenommen und der Apfelsaft wurde gefriergetrocknet. Es verblieben 7.64 g des getrockneten Apfelsaftes als Rückstand. Für die Fraktionierung der Probe wurden 5.00 g der getrockneten Probe in 15 mL Wasser (bidest.) gelöst und an 13.4 g des 3-Amino-funktionalisierten Kieselgels adsorbiert. Die Fraktionierung erfolgte analog zur Fraktionierung des Pinienkernextraktes, wie unter Abschnitt 6.6.3.2 beschrieben.

6.9.2 Probenvorbereitung

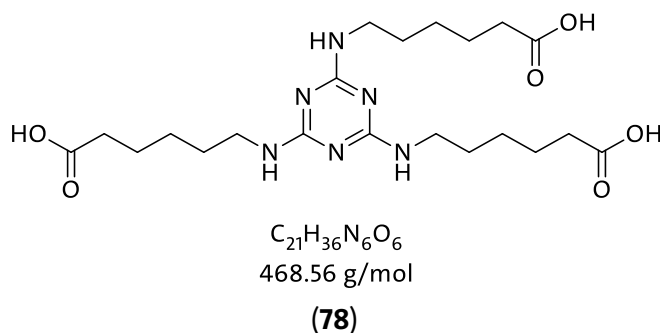
Nach der Fraktionierung der Apfelprobe wurden die Fraktionen jeweils gefriergetrocknet. Der Rückstand wurde in 1000 µL Wasser (LC-MS *grade*) aufgenommen. Die Lösung wurde filtriert und es wurden jeweils 50 µL für die folgenden massenspektrometrischen Messungen abgenommen. Die Proben jeder Fraktion wurden 1000-fach mit Wasser (LC-MS *grade*) verdünnt. Das Injektionsvolumen für die Aufnahme der Massenspektren betrug 5 µL dieser Probe. Das Lösungsmittel der restlichen 950 µL jeder Fraktion wurde mittels *Speed Vacuum Concentrator* sowie mittels Gefriertrocknung entfernt und daraufhin für die Messungen der NMR-Spektren in 700 µL Phosphatpuffer (1 M, pH 7.0, 2 mM TSP-*d*₄, 3 mM NaN₃; siehe Abschnitt 6.5.2 zur Herstellung des Phosphatpuffers) aufgenommen. Es wurden jeweils 600 µL der Lösung in ein 5 mm NMR-Röhrchen überführt.

6.9.3 NMR-Spektroskopie

Die NMR-Spektren der Fraktionen der Apfelprobe wurden an einem Avance NEO 600 MHz NMR-Spektrometer der Firma BRUKER unter Verwendung der Software TopSpin 4.1.3 aufgenommen. Es wurde ein 5 mm TCI-Cryoprobekopf bei einer Basisfrequenz von 600.25 MHz verwendet. Für jede Fraktion wurde ein wasserunterdrücktes ¹H-NMR-Spektrum mit Hilfe der Pulssequenz *zgesgp* akquiriert. Die Temperatur betrug 298 K. Es wurden jeweils 65 536 Datenpunkte bei einer spektralen Weite von 16.0 ppm und einem *transmitter frequency offset* von 4.728 ppm aufgenommen. Es wurden 128 *Scans* sowie 4 *Dummy Scans* mit einem Relaxationsdelay von 1 Sekunde akquiriert. Der *Receiver Gain* betrug 12.7. Es ergab sich eine Messdauer pro Experiment von 9 Minuten und 50 Sekunden.

Die Prozessierung der NMR-Spektren erfolgte in TopSpin 4.1.4. Die FIDs wurden mit einer Exponentialfunktion mit einem *line broadening* Faktor von 0.3 Hz multipliziert. Nach der Fourier-Transformation erfolgte die Referenzierung gegen TSP-*d*₄ ($\delta = 0.00$ ppm) sowie eine automatische Phasenkorrektur 0. Ordnung (Befehl *apk0*) und eine automatische Basislinienkorrektur (Befehl *absn*).

6.10 Synthese von 2,4,6-Tri-(6-aminocaprinsäure)-1,3,5-triazin (TACT)



Die Synthese von 2,4,6-Tri-(6-aminocaprinsäure)-1,3,5-triazin (TACT, **78**) erfolgte im Reaktionskalorimeter RC 1 MX der Firma METTLER TOLEDO. Dabei wurde die Reaktionstemperatur aktiv über den Reaktormantel gesteuert und kontinuierlich mit einer Rührgeschwindigkeit von 500 rpm gerührt. Es wurden 300 mL einer Lösung aus Natrium-6-Aminocapronat (**77**) (0.76 mol/L, 0.228 mol) in Natronlauge (1.52 mol/L) vorgelegt. Nach Temperierung auf 20 °C wurden 13.14 g (71.3 mmol) Cyanurchlorid (**75**) zugegeben, wodurch die Temperatur im Reaktor über einen Zeitraum von ca. 2 Minuten um 16 °C stieg. Entsprechend des zuvor programmierten Temperaturprofils wurde die Reaktionslösung auf eine Temperatur von 25 °C temperiert und für 60 Minuten unter diesen Bedingungen gerührt. Daraufhin wurde die Reaktionslösung innerhalb von 10 Minuten auf eine Temperatur von 50 °C erwärmt und weitere 120 Minuten gerührt. Im nächsten Intervall wurde die Lösung innerhalb von 10 Minuten auf 85 °C erwärmt und weitere 90 Minuten gerührt. Zum Zeitpunkt der Zugabe des Cyanurchlorids (**75**) wurde eine farblose Suspension erhalten, welche sich innerhalb von ca. 3 Minuten klärte. Die Probenentnahme erfolgte innerhalb der ersten 60 Minuten im Abstand von 5 Minuten und im weiteren Verlauf der Reaktion im Abstand von 10 Minuten. Über den gesamten Reaktionszeitraum von 5 Stunden wurden insgesamt 36 Proben entnommen.

¹H-NMR (H₂O/D₂O (9:1), 700 MHz, 298 K): δ [ppm] = 3.30 (s (br.), 6 H, H-6'), 2.21 (t, J = 7.6 Hz, 6 H, H-2'), 1.63-1.53 (m, 12 H, H-5', H-3'), 1.35 (tt, J = 7.2 Hz, 8.1 Hz, 6 H, H-4').

¹³C-NMR (H₂O/D₂O (9:1), 176 MHz, 298 K): δ [ppm] = 186.5 (C-1'), 168.0 (C-1), 43.0 (C-6'), 40.2 (C-2'), 31.3 (C-5'), 28.8 (C-4'), 28.3 (C-3').

HR-ESI-MS (ESI-(+)): gefunden: m/z = 469.2759 ([M+H]⁺), berechnet für [C₂₁H₃₆N₆O₆+H]⁺: m/z = 469.2769.

6.10.1 Probenvorbereitung

Während der Synthese von 2,4,6-Tri-(6-aminocapronsäure)-1,3,5-triazin (**78**) wurden Proben entnommen, die im Anschluss per NMR-Spektroskopie und Massenspektrometrie untersucht werden sollten. Dazu wurden jeweils ca. 2 mL Reaktionslösung entnommen, wovon 1500 μ L sofort mit 500 μ L Wasser verdünnt wurden. Für die Vorbereitung der NMR-Proben wurden 630 μ L der mit Wasser verdünnten Proben der Reaktionslösung mit 70 μ L Deuteriumoxid (+ 0.05 M TSP-*d*₄) versetzt und mittels Vortexmischer durchmischt. Daraufhin wurden 600 μ L der Lösung in ein 5 mm NMR-Röhrchen überführt und bis zur Messung bei –80 °C gelagert. Die Proben für die massenspektrometrischen Messungen wurden in zwei Schritten 100 000-fach in Wasser (LC-MS *grade*) verdünnt. Das Injektionsvolumen für die Aufnahme der Massenspektren betrug 5 μ L dieser Probe.

6.10.2 NMR-Spektroskopie

Die eingefrorenen Proben der Synthese von TACT (**78**) für die NMR-Spektroskopie wurden kurz vor der Messung aufgetaut. Die Spektren wurden an einem AVANCE I HD 700 MHz Ultra Shield Spektrometer der Firma BRUKER mit einem 5 mm TXI Probenkopf bei einer Basisfrequenz von 700.13 MHz und einer Temperatur von 298 K aufgenommen. Die Aufnahme der Spektren erfolgte mittels der Software TopSpin 2.1. Zunächst wurde mit Hilfe des *Inversion-Recovery* Experiments mit dem Pulsprogramm *t1irpg* die T1-Relaxationszeit von den ¹³C-Kernen des TACTs bestimmt. Dazu wurden jeweils 128 *Scans* sowie 2 *Dummy Scans* aufgenommen und der Relaxationsdelay betrug 45 Sekunden. Es wurden folgende Werte für τ [μ s] gewählt: 0.01, 0.06, 0.12, 0.30, 0.70, 1.3, 3.0, 7.0, 15, 25. Für die Reaktionsverfolgung wurde anschließend von jeder

Probe ein ^{13}C -NMR-Spektrum mit der Pulssequenz *zgpg* aufgenommen. Dabei wurden jeweils 256 *Scans* sowie 4 *Dummy Scans* aufgenommen. Der Relaxationsdelay wurde auf 15 Sekunden festgelegt. Es wurde ein fester Wert für den *Receiver gain* von 13004 eingestellt. Insgesamt ergab sich somit für die ^{13}C -NMR-Spektren eine Messdauer von 1 Stunde 8 Minuten und 39 Sekunden pro Probe. Zur Prozessierung der Spektren in TopSpin 4.1.4 wurde eine Multiplikation der FIDs mit einer Exponentialfunktion mit *line broadening* Faktor von 1.0 Hz durchgeführt. Nach Fourier-Transformation erfolgte die Kalibrierung gegen TSP- d_4 ($\delta = 0.00$ ppm) sowie die manuelle Phasenkorrektur und automatische Basislinienkorrektur (Befehl *absn*). Zusätzlich erfolgte anschließend ein *Alignment* für einzelne Signale im Verlauf der Proben mit Hilfe von MestReNova, um minimale Unterschiede in den spezifischen Signallagen auszugleichen.

6.11 Massenspektrometrie

Mit Ausnahme der Aufnahme der Massenspektren zur Reaktionsverfolgung der Synthese von TACT (**78**) (siehe Abschnitt 6.10) wurden alle Massenspektren am maXis ESI-Q-ToF Massenspektrometer der Firma BRUKER DALTONICS, gekoppelt an die DIONEX Ultimate 3000 UPLC der Firma THERMO FISHER SCIENTIFIC, aufgenommen. Aufgrund eines Defektes am oben genannten Gerät wurden die Massenspektren der Reaktionsverfolgung von TACT (**78**) am maXis II (ETD) ESI-Q-ToF Massenspektrometer der Firma BRUKER DALTONICS, gekoppelt an die BRUKER DALTONICS Elute HPLC, unter den gleichen unten aufgeführten Messparametern mit Ausnahme einer geänderten *Transfer time* aufgenommen.

6.11.1 Messungen per Direkteinlass

Die Akquisition der per Direkteinlass (DI) gemessenen Massenspektren des artifiziellen Gemisches A und B, der Pinienkernextrakt-Proben, der Apfel-Proben und der Reaktionsverfolgungs-Proben erfolgte mit Wasser + 0.1 % Ameisensäure und Acetonitril + 0.1 % Ameisensäure als Eluenten im Verhältnis 50:50 bei einer Flussrate von 0.3 mL/min. Für die Messung der Proben per Direkteinlass wurde im Regelfall ein reguläres Blindstück verwendet. Für die Aufnahme der Massenspektren der Apfel-Proben wurde statt des Blindstücks eine Edelstahl-Kapillare mit einem Innendurchmesser von 0.05 mm und einer Länge von 500 mm verwendet. Die Gesamtdauer der Messung einer Probe lag bei 3 Minuten. Das jeweils

verwendete Injektionsvolumen ist in den vorherigen Abschnitten der jeweiligen Probenvorbereitung angegeben. Die Massenspektren wurden sowohl im positiven als auch im negativen Ionenmodus in einem Massenbereich von $m/z = 50$ -2300 aufgezeichnet. Die Massenspektren im positiven (negativen) Ionenmodus wurden unter folgenden ESI-Quellenparametern aufgenommen: Kapillarspannung: 4500 V (3000 V), *end plate offset*: -500 V (-500 V), Trocknungsgasfluss: 8.0 L/min (8.0 L/min), Temperatur Ionenquelle: 200 °C (200 °C), Zerstäubergas: 5.0 bar (4.0 bar), Spektrenrate: 1.0 Spek./min (1.0 Spek./min), *Transfer time*: 50 µs (am maXis ESI-Q-ToF Massenspektrometer) bzw. 70 µs (am maXis II (ETD) ESI-Q-ToF Massenspektrometer). Die Kalibrierung des Massenspektrometers wurde vor Beginn der Messungen mit einer Natriumformiat-Cluster-Lösung sowie am Ende jeder einzelnen DI-MS-Messung jeder Probe durch Umschalten eines Ventils und der Spritzenpumpe durchgeführt. Die Flussrate der Spritzenpumpe betrug 0.1 mL/h. Zur Herstellung der Natriumformiat-Cluster-Lösung wurden zu einer Mischung von 50 mL Wasser (LC-MS *grade*) und 50 mL 2-Propanol (LC-MS *grade*) 100 µL Ameisensäure sowie 1 mL Natriumhydroxid-Lösung (1 M) zugegeben. Anschließend wurde die Lösung 1:1000 in Wasser/2-Propanol (1:1) verdünnt.

Die DI-Massenspektren wurden mit Compass Data Analysis 4.2 (BRUKER DALTONICS GmbH) prozessiert und analysiert. Für alle Massenspektren, sowohl im positiven als auch im negativen Ionenmodus, wurde ein gemittelttes Massenspektrum im Retentionszeitbereich von Minute 2.3 bis 2.4 berechnet, das nur Signale des Kalibrierstandards (Natriumformiat-Cluster) enthält. Dieses Massenspektrum wurde für die automatische interne Kalibrierung der einzelnen Probe verwendet. Zusätzlich wurde für alle Massenspektren ein weiteres gemittelttes Massenspektrum aus dem *total ion current* (TIC) im Retentionszeitbereich von Minute 0.2 bis 0.6 für Messungen mit dem Blindstück bzw. im Retentionszeitbereich von Minute 0.1 bis 0.4 für Messungen mit der Edelstahl-Kapillare für jede Probe jeder Fraktion berechnet. Das Skript für die automatische Prozessierung der Spektren ist im Anhang in Abschnitt 8.6.3 angegeben. Für die weitere Analyse wurde eine Textdatei bestehend aus den m/z -Werten (zwei Dezimalstellen) und den zugehörigen Intensitäten für jede Probe exportiert.

6.11.2 LC-MS-Messungen

Für die weiterführende Analyse einzelner Fraktionen wurden diese in einzelnen Fällen durch chromatographische Trennung und Analyse mittels LC-MS untersucht. Dazu wurde eine

NUCLEOSHELL HILIC-Säule (150 mm x 2 mm ID, Partikelgröße 2.7 µm) von MACHEREY-NAGEL verwendet. Als Lösungsmittel wurden Wasser (A) und Acetonitril (B) (je LC-MS *grade*) eingesetzt. Die Flussrate betrug 0.3 mL/min. Der eingesetzte Gradient kann der Tabelle 34 entnommen werden.

Tabelle 34: Gradient für LC-MS-Messungen einzelner Fraktionen unter HILIC-Bedingungen auf der NUCLEOSHELL HILIC-Säule. B: Acetonitril.

Zeit [min]	B [%]
0	95
20	40
21	40
23	95
30	95

Die Einstellung der ESI-Quellenparameter erfolgte analog zu den Messungen per Direkteinlass. Die qualitative Auswertung der Daten wurde mittels Compass Data Analysis 4.2 (BRUKER DALTONICS GmbH) durchgeführt.

6.12 Zweidimensionale und selektive NMR-Experimente

Zusätzlich zu den in den Abschnitten 6.5 bis 6.10 beschriebenen NMR-Experimenten wurden zur weiteren Strukturaufklärung teilweise selektive NMR-Experimente sowie 2D homo- oder heteronukleare NMR-Experimente durchgeführt. Zweidimensionale und selektive NMR-Experimente wurden an einem BRUKER Avance III HD 600 MHz NMR-Spektrometer unter Verwendung der Software TopSpin 3.6.4, ausgestattet mit einem 5 mm BBFO-Probenkopf, bei einer Basisfrequenz von 600.13 MHz und einer Temperatur von 298 K akquiriert. Die im Rahmen dieser Arbeit verwendeten Pulssequenzen sind in Tabelle 35 angegeben.

Tabelle 35: Angabe der im Rahmen der Arbeit verwendeten Pulssequenzen zur Aufnahme von zweidimensionalen sowie selektiven NMR-Experimenten unter Angabe der Aufnahmeparameter.

NMR-Experiment	Verwendete Pulssequenz	TD2	TD1	SW2 [ppm]	SW1 [ppm]	DS	D1 [s]
2D HSQC	<i>hsqcedetgpsp.3</i>	2048	128	15.1	180	32	1
2D HMBC	<i>hmbcetgpl3nd</i>	4096	256	15.1	230	16	1

NMR-Experiment	Verwendete Pulssequenz	TD2	TD1	SW2 [ppm]	SW1 [ppm]	DS	D1 [s]
2D COSY	<i>cosygpmfppqf</i>	4096	256	12.0	12.0	16	1
2D TOCSY	<i>dipsi2esgpph</i>	2048	256	10.0	10.0	16	1.5
1D selective NOESY	<i>seldigpzs</i>	-	32768	-	14.0	4	1.5
1D selective TOCSY	<i>selnogpzs</i>	-	32768	-	14.0	4	1.5

Zusätzlich sind in Tabelle 35 die Aufnahmeparameter, wie z.B. die Anzahl der Datenpunkte, die spektrale Weite, die Anzahl der *Dummy Scans* sowie der Relaxationsdelay angegeben. Der *Receiver gain* sowie die Anzahl der *Scans* wurde individuell für die zu untersuchende Probe angepasst. Für die selektiven TOCSY-NMR-Experimente wurde eine *mixing time* von 0.08 Sekunden eingestellt und der GPZ0-Wert des z-Filters auf 9% eingestellt. Für das selektive NOESY-Experiment wurde eine *mixing time* von 0.8 Sekunden gewählt. Für alle erhaltenen Spektren erfolgte nach der (zweifachen) Fourier-Transformation die Kalibrierung gegen TSP- d_4 ($\delta_{1H} = 0.00$ ppm, $\delta_{13C} = 0.00$ ppm), eine automatische Phasenkorrektur sowie eine automatische Basislinienkorrektur.

7 Literaturverzeichnis

- [1] E. Cubero-Leon, R. Peñalver, A. Maquet, Review on metabolomics for food authentication, *Food Res. Int.* **2014**, 60, 95-107.
- [2] J. Hoffbauer, Das nationale System zur Bekämpfung von Lebensmittelbetrug, *J. Consum. Prot. Food Saf.* **2017**, 12, 23-26.
- [3] X. Pei, A. Tandon, A. Alldrick, L. Giorgi, W. Huang, R. Yang, The China melamine milk scandal and its implications for food safety regulation, *Food Policy* **2011**, 36, 412-420.
- [4] S. W. Drew, A. L. Demain, Effect of primary metabolites on secondary metabolism, *Annu. Rev. Microbiol.* **1977**, 31, 343-356.
- [5] S. G. Villas-Bôas, U. Roessner, M. A. E. Hansen, J. Smedsgaard, J. Nielsen, *Metabolome analysis: An introduction*, John Wiley & Sons, Inc., New Jersey, **2007**.
- [6] J. J. J. van der Hooft, N. Rankin, *Metabolite Identification in Complex Mixtures Using Nuclear Magnetic Resonance Spectroscopy* in: *Modern Magnetic Resonance* (Hrsg.: G. A. Webb), Springer International Publishing, Cham, **2018**, S. 1309-1341.
- [7] K. Bingol, R. Brüscheiler, Two elephants in the room: new hybrid nuclear magnetic resonance and mass spectrometry approaches for metabolomics, *Current Opinion* **2015**, 18, 471-477.
- [8] J. R. Idle, F. J. Gonzalez, Metabolomics, *Cell Metab.* **2007**, 6, 348-351.
- [9] A. Klassen, A. T. Faccio, G. A. Baptista Canuto, P. L. Rocha da Cruz, H. Caracho Ribeiro, M. F. M. Tavares, A. Sussulini in *Metabolomics: From Fundamentals to Clinical Applications* (Hrsg.: A. Sussulini), Springer International Publishing, Cham, **2017**, S. 3-18.
- [10] C. K. Larive, G. A. Barding, M. M. Dinges, NMR spectroscopy for metabolomics and metabolic profiling, *Anal. Chem.* **2015**, 87, 133-146.
- [11] J. L. Markley, R. Brüscheiler, A. S. Edison, H. R. Eghbalnia, R. Powers, D. Raftery, D. S. Wishart, The future of NMR-based metabolomics, *Curr. Opin. Biotechnol.* **2017**, 43, 34-40.
- [12] J. Sotelo, C. M. Slupsky, *Metabolomics using nuclear magnetic resonance (NMR) in Metabolomics in food and nutrition* (Hrsg.: B.C. Weimer, C. Slupsky), Woodhead Publishing Limited, Davis, **2013**, S. 29-43.

- [13] N. Vinayavekhin, A. Saghatelian, Untargeted Metabolomics, *Curr. Protoc. Mol. Biol.* **2010**, 1-24.
- [14] A. Böttger, U. Vothknecht, C. Bolle, A. Wolf, *Plant Secondary Metabolites and their general function in plants* in *Lessons on Caffeine, Cannabis & Co.*, Springer International Publishing, Cham, **2018**, S. 3-17.
- [15] A.-H. M. Emwas, R. M. Salek, J. L. Griffin, J. Merzaban, NMR-based metabolomics in human disease diagnosis: applications, limitations, and recommendations, *Metabolomics* **2013**, 9, 1048-1072.
- [16] E. Hatzakis, Nuclear Magnetic Resonance (NMR) Spectroscopy in Food Science: A Comprehensive Review, *Compr. Rev. Food Sci. Food Saf.* **2019**, 18, 189-220.
- [17] S. Medina, R. Perestrelo, P. Silva, J. A. Pereira, J. S. Câmara, Current trends and recent advances on food authenticity technologies and chemometric approaches, *Trends Food Sci.* **2019**, 85, 163-176.
- [18] R. Matissek, M. Fischer, G. Steiner, *Lebensmittelanalytik*, 6. Auflage, Springer, Berlin, Heidelberg, **2018**, S. 687-699.
- [19] P. S. Belton, I. J. Colquhoun, E. Kemsley, I. Delgadillo, P. Roma, M. Dennis, M. Sharman, E. Holmes, J. K. Nicholson, M. Spraul, Application of chemometrics to the ^1H NMR spectra of apple juices: discrimination between apple varieties, *Food Chem.* **1998**, 61, 207-213.
- [20] G. Di Matteo, M. Spano, C. Esposito, C. Santarcangelo, A. Baldi, M. Daglia, L. Mannina, C. Ingallina, A. P. Sobolev, NMR Characterization of Ten Apple Cultivars from the Piedmont Region, *Foods* **2021**, 10.
- [21] P. Eisenmann, M. Ehlers, C. H. Weinert, P. Tzvetkova, M. Silber, M. J. Rist, B. Luy, C. Muhle-Goll, Untargeted NMR Spectroscopic Analysis of the Metabolic Variety of New Apple Cultivars, *Metabolites* **2016**, 6.
- [22] C. Li, L. Li, Y. Wu, M. Lu, Y. Yang, L. Li, Apple Variety Identification Using Near-Infrared Spectroscopy, *J. Spectrosc.* **2018**, 1, 1-7.
- [23] C. H. Johnson, J. Ivanisevic, G. Siuzdak, Metabolomics: beyond biomarkers and towards mechanisms, *Nat. Rev. Mol. Cell Biol.* **2016**, 17, 451-459.

- [24] S. Tomita, T. Nemoto, Y. Matsuo, T. Shoji, F. Tanaka, H. Nakagawa, H. Ono, J. Kikuchi, M. Ohnishi-Kameyama, Y. Sekiyama, A NMR-based, non-targeted multistep metabolic profiling revealed L-rhamnitol as a metabolite that characterised apples from different geographic origins, *Food Chem.* **2015**, 174, 163-172.
- [25] X. Qi, X. Chen, Y. Wang, *Plant Metabolomics: Methods and Applications*, Springer Netherlands, Dordrecht, **2015**.
- [26] W. Kessler, *Multivariate Datenanalyse für die Pharma-, Bio- und Prozessanalytik*, 1. Auflage, Wiley-VCH, Weinheim, **2006**.
- [27] S. Sousa, A. Magalhães, M. M. C. Ferreira, Optimized bucketing for NMR spectra: Three case studies, *Chemometr. Intell. Lab. Syst.* **2013**, 122, 93-102.
- [28] B. Worley, R. Powers, Multivariate Analysis in Metabolomics, *Curr. Metabolomics* **2013**, 1, 92-107.
- [29] R. G. Brereton, Consequences of sample size, variable selection, and model validation and optimisation, for predicting classification ability from analytical data, *Trends Anal. Chem.* **2006**, 25, 1103-1111.
- [30] C. V. Di Anibal, M. P. Callao, I. Ruisánchez, ^1H NMR variable selection approaches for classification. A case study: the determination of adulterated foodstuffs, *Talanta* **2011**, 86, 316-323.
- [31] M. B. Kursu, R. W. Rudnicki, Feature Selection with the Boruta Package, *J. Stat. Softw.* **2010**, 36(11), 1-13.
- [32] R. Schlittgen, *Multivariate Statistik*, Oldenbourg Wissenschaftsverlag, München, **2009**.
- [33] R. Rinne, *Statistische Analyse multivariater Daten*, Oldenbourg Wissenschaftsverlag, München, **2000**.
- [34] K.-L. Du, M. N. S. Swamy, *Neural Networks and Statistical Learning*, 2. Auflage, Springer, London, **2019**.
- [35] G. James, D. Witten, T. Hastie, R. Tibshirani, *An Introduction to Statistical Learning*, 2. Auflage, Springer, New York, **2013**.

- [36] S. Esslinger, J. Riedl, C. Fauhl-Hassek, Potential and limitations of non-targeted fingerprinting for authentication of food in official control, *Food Res. Int.* **2014**, 60, 189-204.
- [37] I. Steinwart, A. Christmann, *Support Vector Machines*, Springer, New York, **2008**.
- [38] O. A. Montesinos López, A. Montesinos López, J. Crossa, *Multivariate Statistical Machine Learning Methods for Genomic Prediction*, Springer International Publishing, Cham, **2022**.
- [39] S. Abe, *Support Vector Machines for Pattern Classification*, 2. Auflage, Springer, London, **2010**.
- [40] C. J. C. Burges, A Tutorial on Support Vector Machines for Pattern Recognition, *Data Min. Knowl. Discov.* **1998**, 2, 121-167.
- [41] K. Backhaus, B. Erichson, W. Plinke, R. Weiber, R., *Multivariate Analysemethoden*, 15. Auflage, Springer, Berlin, Heidelberg, **2018**.
- [42] B. Richter, M. Rurik, S. Gurk, O. Kohlbacher, M. Fischer, Food monitoring: Screening of the geographical origin of white asparagus using FT-NIR and machine learning, *Food Control* **2019**, 104, 318-325.
- [43] S. Varma, R. Simon, Bias in error estimation when using cross-validation for model selection, *BMC bioinformatics* **2006**, 7(91) 1-8.
- [44] D. Krstajic, L. J. Buturovic, D. E. Leahy, S. Thomas, Cross-validation pitfalls when selecting and assessing regression and classification models, *J. Cheminf.* **2014**, 6(10), 1-15.
- [45] C. Schmitt, Geographische Herkunftsbestimmung von Walnüssen (*Juglans regia* L.) und Detektion von Erdnüssen (*Arachis hypogaea* L.) in Lebensmitteln mittels NMR-Spektroskopie, *Dissertation* **2020**, Universität Hamburg.
- [46] M. M. Rinschen, J. Ivanisevic, M. Giera, G. Siuzdak, G, Identification of bioactive metabolites using activity metabolomics, *Nat. Rev. Mol. Cell Biol.* **2019**, 20, 353-367.
- [47] M. Yang, T. Soga, P. J. Pollard, Oncometabolites: linking altered metabolism with cancer, *J. Clin. Invest.* **2013**, 123, 3652-3658.
- [48] D. S. Wishart, Computational strategies for metabolite identification in metabolomics. *Bioanalysis* **2009**, 1(9), 1579-1596.

- [49] L. W. Sumner, A. Amberg, D. Barrett, M. H. Beale, R. Beger, C. A. Daykin, T. W.-M. Fan, O. Fiehn, R. Goodacre, J. L. Griffin *et al.*, Proposed minimum reporting standards for chemical analysis, *Metabolomics* **2007**, 3, 211–221.
- [50] I. Garcia-Perez, J. M. Posma, J. I. Serrano-Contreras, C. L. Boulangé, Q. Chan, G. Frost, J. Stamler, P. Elliott, J. C. Lindon, E. Holmes, *et al.*, Identifying unknown metabolites using NMR-based metabolic profiling techniques, *Nat. Protoc.* **2020**, 15, 2538-2567.
- [51] B. Buszewski, S. Noga, S, Hydrophilic interaction liquid chromatography (HILIC) – a powerful separation technique, *Anal. Bioanal. Chem.* **2012**, 402, 231-247.
- [52] I. Kohler, M. Verhoeven, R. Haselberg, A. F. Gargano, Hydrophilic interaction chromatography – mass spectrometry for metabolomics and proteomics: state-of-the-art and current trends, *Microchem. J.* **2022**, 175, 106986.
- [53] M. Da Pereira Costa, C. A. Conte-Junior, Chromatographic Methods for the Determination of Carbohydrates and Organic Acids in Foods of Animal Origin, *Compr. Rev. Food Sci. Food Saf.* **2015**, 14, 586-600.
- [54] K. Spagou, H. Tsoukali, N. Raikos, H. Gika, I. D. Wilson, G. Theodoridis, Hydrophilic interaction chromatography coupled to MS for metabonomic/metabolomic studies, *J. Sep. Sci.* **2010**, 33, 716-727.
- [55] A. J. Alpert, Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds, *J. Chromatogr.* **1990**, 499, 177-196.
- [56] P. Hemström, K. Irgum, K, Hydrophilic interaction chromatography, *J. Sep. Sci.* **2006**, 29, 1784-1821.
- [57] D. Rojo, C. Barbas, F. J. Rupérez, LC-MS metabolomics of polar compounds, *Bioanalysis* **2012**, 4, 1235-1243.
- [58] B. A. Olsen, Hydrophilic interaction chromatography using amino and silica columns for the determination of polar pharmaceuticals and impurities, *J. Chromatogr. A.* **2001**, 913, 113-122.
- [59] T. D. Claridge, *High-Resolution NMR Techniques in Organic Chemistry*, 3. Auflage, Elsevier, Amsterdam, Oxford, Cambridge, **2016**.

- [60] A.-H. Emwas, R. Roy, R. T. McKay, L. Tenori, E. Saccenti, G. A. N. Gowda, D. Raftery, F. Alahmari, L. Jaremko, M. Jaremko, *et al.*, NMR Spectroscopy for Metabolomics Research, *Metabolites* **2019**, 9(7), 123.
- [61] R. M. Gathungu, R. Kautz, B. S. Kristal, S. S. Bird, P. Vouros, The integration of LC-MS and NMR for the analysis of low molecular weight trace analytes in complex matrices, *Mass Spectrom. Rev.* **2020**, 39, 35-54.
- [62] G. A. Nagana Gowda, D. Raftery, Can NMR solve some significant challenges in metabolomics?, *J. Magn. Reson.* **2015**, 260, 144-160.
- [63] G. Dal Poggetto, L. Castañar, R. W. Adams, G. A. Morris, M. Nilsson, Relaxation-encoded NMR experiments for mixture analysis: REST and beer, *Chem. Comm.* **2017**, 53, 7461-7464.
- [64] C. S. Johnson Jr., Diffusion ordered nuclear magnetic resonance spectroscopy: principles and applications, *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, 34, 203-256.
- [65] K. F. Morris, S. C. Johnson Jr., Resolution of discrete and continuous molecular size distributions by means of diffusion-ordered 2D NMR spectroscopy, *J. Am. Chem. Soc.* **1993**, 115, 4291-4299.
- [66] M. J. Smith, L. Castañar, R. W. Adams, G. A. Morris, M. Nilsson, Giving Pure Shift NMR Spectroscopy a REST – Ultrahigh-Resolution Mixture Analysis, *Anal. Chem.* **2022**, 94, 12757-12761.
- [67] G. Dal Poggetto, L. Castañar, R. W. Adams, G. A. Morris, M. Nilsson, Dissect and Divide: Putting NMR Spectra of Mixtures under the Knife, *J. Am. Chem. Soc.* **2019**, 141, 5766-5771.
- [68] K. Bingol, R. Brüschweiler, Multidimensional approaches to NMR-based metabolomics, *Anal. Chem.* **2014**, 86, 47-57.
- [69] H. Friebolin, *Ein- und zweidimensionale NMR-Spektroskopie*, 5. Auflage, Wiley-VCH, Weinheim, **2014**.
- [70] H. Günther, *NMR Spectroscopy: Basic Principles, Concepts, and Applications in Chemistry*, 3. Auflage, Wiley-VCH, Weinheim, **2013**.
- [71] J. Keeler, *Understanding NMR Spectroscopy*, 2. Auflage, John Wiley & Sons, Chichester, **2010**.

- [72] M. Hesse, H. Meier, B. Zeeh, B., *Spektroskopische Methoden in der organischen Chemie*, 8. überarb. und erw. Auflage, Thieme, Stuttgart, New York, **2012**.
- [73] A. Le Guennec, P. Giraudeau, S. Caldarelli, Evaluation of fast 2D NMR for metabolomics. *Anal. Chem.* **2014**, *86*, 5946-5954.
- [74] D. Schulze-Sünninghausen, J. Becker, B. Luy, Rapid heteronuclear single quantum correlation NMR spectra at natural abundance, *J. Am. Chem. Soc.* **2014**, *136*, 1242-1245.
- [75] P. Schanda, E. Kupce, B. Brutscher, SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds, *J. Biomol. NMR* **2005**, *33*, 199-211.
- [76] B. Vitorge, G. Bodenhausen, P. Pelupessy, Speeding up nuclear magnetic resonance spectroscopy by the use of SMALL Recovery Times – SMART NMR, *J. Magn. Reson.* **2010**, *207*, 149-152.
- [77] P. Schanda, H. van Melckebeke, B. Brutscher, B., Speeding up three-dimensional protein NMR experiments to a few minutes, *J. Am. Chem. Soc.* **2006**, *128*, 9042-9043.
- [78] P. Giraudeau, L. Frydman, Ultrafast 2D NMR: an emerging tool in analytical spectroscopy, *Annu. Rev. Anal. Chem.* **2014**, *7*, 129-161.
- [79] M. Mobli, J. C. Hoch, Nonuniform sampling and non-Fourier signal processing methods in multidimensional NMR, *Prog. Nucl. Magn. Reson. Spectrosc.* **2014**, *83*, 21-41.
- [80] F. Delaglio, G. S. Walker, K. A. Farley, R. Sharma, J. C. Hoch, L. W. Arbogast, R. G. Brinson, K. P. Marino, Non-uniform sampling for all: More NMR spectral quality, less measurement time, *Am. Pharm. Rev.* **2017**, *21*, 339681.
- [81] R. Dass, Ž, K. Grudzia, T. Ishikawa, M. Nowakowski, R. Dębowska, K. Kazimierczuk, Fast 2D NMR spectroscopy for In vivo monitoring of bacterial metabolism in complex mixtures, *Front. Microbiol.* **2017**, *8*, 1306.
- [82] S. G. Hyberts, H. Arthanari, G. Wagner, Applications of non-uniform sampling and processing, *Top. Curr. Chem.* **2012**, *316*, 125-148.
- [83] D. Schulze-Sünninghausen, J. Becker, M. R. M. Koos, B. Luy, B., Improvements, extensions, and practical aspects of rapid ASAP-HSQC and ALSOFAST-HSQC pulse sequences for studying small molecules at natural abundance, *J. Magn. Reson.* **2017**, *281*, 151-161.

- [84] M. R. M. Koos, B. Luy, Polarization recovery during ASAP and SOFAST/ALSOFAST-type experiments, *J. Magn. Reson.* **2019**, *300*, 61-75.
- [85] L. Mueller, Alternate HMQC experiments for recording HN and HC-correlation spectra in proteins at high throughput, *J. Biomol. NMR* **2008**, *42*, 129-137.
- [86] E. Kupce, R. Freeman, Fast multidimensional NMR by polarization sharing, *Magnetic resonance in chemistry* **2007**, *45*, 2-4.
- [87] O. Cloarec, M.-E. Dumas, A. Craig, R. H. Barton, J. Trygg, J. Hudson, C. Blancher, D. Gauguier, J. C. Lindon, E. Holmes, *et al.*, Statistical Total Correlation Spectroscopy: An Exploratory Approach for Latent Biomarker Identification from Metabolic ^1H NMR Data Sets, *Anal. Chem.* **2005**, 1282-1289.
- [88] J. M. Poma, I. Garcia-Perez, M. De Iorio, J. C. Lindon, P. Elliott, E. Holmes, T. M. D. Ebbels, J. K. Nicholson, Subset optimization by reference matching (STORM): an optimized statistical approach for recovery of metabolic biomarker structural information from ^1H NMR spectra of biofluids, *Anal. Chem.* **2012**, *84*, 10694-10701.
- [89] T. R. Rudd, E. Macchi, L. Muzi, M. Ferro, D. Gaudesi, G. Torri, B. Casu, M. Guerrini, E. A. Yates, Unravelling structural information from complex mixtures utilizing correlation spectroscopy applied to HSQC spectra, *Anal. Chem.* **2013**, *85*, 7487-7493.
- [90] C. Zhang, Y. Idelbayev, N. Roberts, Y. Tao, Y. Nannapaneni, B. M. Duggan, J. Min, E. C. Lin, E. C. Gerwick, G. W. Cottrell, *et al.*, Small Molecule Accurate Recognition Technology (SMART) to Enhance Natural Products Research, *Sci. Rep.* **2017**, *7*, 14243.
- [91] R. Reher, H. W. Kim, C. Zhang, H. H. Mao, M. Wang, L.-F. Nothias A. M. Caraballo-Rodriguez E. Glukhov, B. Teke, T. Leao, *et al.*, A Convolutional Neural Network-Based Approach for the Rapid Annotation of Molecularly Diverse Natural Products, *J. Am. Chem. Soc.* **2020**, *142*, 4114-4120.
- [92] H. Budzikiewicz, M. Schäfer, *Massenspektrometrie: Eine Einführung*, 6. Auflage, Wiley-VCH, Weinheim, **2012**.
- [93] C. S. Ho, C. W. Lam, M. H. M. Chan R. C. Cheung, L. K. Law, L. C. Lit, K. F. Ng, M. W. M. Suen, H. L. Tai, *Electrospray Ionisation Mass Spectrometry: Principles and Clinical Applications*, *Clin. Biochem. Rev.* **2003**, *24*, 3-12.

- [94] T. Kind, O. Fiehn, Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry, *BMC Bioinformatics* **2007**, *8*, 1-20.
- [95] L. Lin, Q. Yu, X. Yan, W. Hang, J. Zheng, J. Xing, B. Huang, Direct infusion mass spectrometry or liquid chromatography mass spectrometry for human metabonomics? A serum metabonomic study of kidney cancer, *Analyst* **2010**, *135*, 2970-2978.
- [96] D. D. Marshall, C. Lei, B. Worley, Y. Huang, A. Garcia-Garcia, R. Franco, E. D. Dodds, R. Powers, Combining DI-ESI-MS and NMR datasets for metabolic profiling, *Metabolomics* **2015**, *11*, 391-402.
- [97] J. Draper, A. J. Lloyd, R. Goodacre, M. Beckmann, Flow infusion electrospray ionisation mass spectrometry for high throughput, non-targeted metabolite fingerprinting: a review, *Metabolomics* **2013**, *9*, 4–29.
- [98] G. A. Nagana Gowda, D. Djukovic, Overview of mass spectrometry-based metabolomics: opportunities and challenges, *Methods Mol. Biol.* **2014**, *1198*, 3-12.
- [99] I. Erngren, J. Haglöf, M. K. R. Engskog, M. Nestor, M. Hedeland, T. Arvidsson, C. Pettersson, Adduct formation in electrospray ionisation-mass spectrometry with hydrophilic interaction liquid chromatography is strongly affected by the inorganic ion concentration of the samples, *J. Chromatogr. A.* **2019**, *1600*, 174-182.
- [100] N. B. Cech, C. G. Enke, Practical implications of some recent studies in electrospray ionization fundamentals, *Mass Spec. Rev.* **2001**, *20*, 362-387.
- [101] A. Krueve, K. Kaupmees, Adduct Formation in ESI/MS by Mobile Phase Additives, *J. Am. Soc. Mass Spectrom.* **2017**, *28*, 887-894.
- [102] I. Ferrer, E. M. Thurman, *Liquid Chromatography Time-of-Flight Mass Spectrometry: Principles, Tools, and Applications for Accurate Mass Analysis in Chemical Analysis: A Series of Monographs on Analytical Chemistry and its Applications* (Hrsg: J. D. Winefordner), John Wiley & Sons, Hoboken, New Jersey, **2009**.
- [103] H. Gu, G. A. Nagana Gowda, F. C. Neto, M. R. Opp, D. Raftery, RAMSY: ratio analysis of mass spectrometry to improve compound identification, *Anal. Chem.* **2013**, *85*, 10771-10779.

- [104] M. Brown, W. B. Dunn, P. Dobson, Y. Patel, C. L. Winder, S. Francis-McIntyre, P. Begley, K. Carroll, D. Broadhurst, A. Tseng, *et al.*, Mass spectrometry tools and metabolite-specific databases for molecular identification in metabolomics, *The Analyst* **2009**, *134*, 1322-1332.
- [105] J. H. Gross, *Massenspektrometrie*, 1. Auflage, Springer, Berlin, Heidelberg, **2013**.
- [106] T. Kind, O. Fiehn, *Advances in structure elucidation of small molecules using mass spectrometry* in *Frontiers of Bioanalytical Chemistry* (Hrsg.: F.-M. Matysik, J. Wegener), Springer, Berlin, Heidelberg, **2013**, S. 129-166.
- [107] A. T. Aron, E. C. Gentry, K. L. McPhail, L.-F. Nothias, M. Nothias-Esposito, A. Bouslimani, D. Petras, J. M. Gauglitz, N. Sikora, F. Vargas *et al.*, Reproducible molecular networking of untargeted mass spectrometry data using GNPS, *Nat. Protoc.* **2020**, *15*, 1954-1991.
- [108] M. Wang, J. J. Carver, V. V. Phelan, L. M. Sanchez, N. Garg, Y. Peng, D. D. Nguyen, J. Watrous, C. A. Kapon, T. Luzzatto-Knaan *et al.*, Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking, *Nat. Biotechnol.* **2016**, *34*, 828-837.
- [109] F. Tayyari, G. A. Nagana Gowda, H. Gu, D. Raftery, ¹⁵N-cholamine – a smart isotope tag for combining NMR- and MS-based metabolite profiling, *Anal. Chem.* **2013**, *85*, 8715-8721.
- [110] T. W.-M. Fan, P. K. Lorkiewicz, K. Sellers, H. N. B. Moseley, R. M. Higashi, A. N. Lane, Stable isotope-resolved metabolomics and applications for drug development. *Pharmacology & therapeutics* **2012**, *133*, 366-391.
- [111] K. Bingol, L. Bruschweiler-Li, C. Yu, A. Somogyi, F. Zhang, R. Bruschweiler, Metabolomics beyond spectroscopic databases: a combined MS/NMR strategy for the rapid identification of new metabolites in complex mixtures, *Anal. Chem.* **2015**, *87*, 3864-3870.
- [112] K. Bingol, R. Bruschweiler, NMR/MS Translator for the enhanced simultaneous analysis of metabolomics mixtures by NMR spectroscopy and mass spectrometry: application to human urine, *J. Proteome Res.* **2015**, *14*, 2642-2648.
- [113] D. J. Crockford, E. Holmes, J. C. Lindon, R. S. Plumb, S. Zirah, S. J. Bruce, P. Rainville, C. L. Stumpf, J. K. Nicholson, Statistical heterospectroscopy, an approach to the integrated analysis of NMR and UPLC-MS data sets: application in metabonomic toxicology studies, *Anal. Chem.* **2006**, *78*, 363-371.

- [114] D. Dai, J. He, R. Sun, R. Zhang, J. A. Aisa, Z. Abliz, Nuclear magnetic resonance and liquid chromatography-mass spectrometry combined with an incompleated separation strategy for identifying the natural products in crude extract, *Anal. Chim. Acta* **2009**, 632, 221-228.
- [115] X.-X. Wang, J.-M. He, C.-L. Wang, R.-P. Zhang, W.-Y. He, S.-X. Guo, R.-X. Sun, Z. Abliz, Simultaneous structural identification of natural products in fractions of crude extract of the rare endangered plant *Anoectochilus roxburghii* using ^1H NMR/RRLC-MS parallel dynamic spectroscopy, *Int. J. Mol. Sci.* **2011**, 12, 2556-2571.
- [116] H. N. Behnken, M. Fellenberg, M. P. Koetzler, R. Jirmann, T. Nagel, B. Meyer, Resolving the problem of chromatographic overlap by 3D cross correlation (3DCC) processing of LC, MS and NMR data for characterization of complex glycan mixtures, *Anal. Bioanal. Chem.* **2012**, 404, 1427-1437.
- [117] M. Fellenberg, H. N. Behnken, T. Nagel, A. Wiegandt, M. Bärenfänger, B. Meyer, Glycan analysis: scope and limitations of different techniques – a case for integrated use of LC-MS(/MS) and NMR techniques, *Anal. Bioanal. Chem.* **2013**, 405, 7291-7305.
- [118] A. Wiegandt, B. Meyer, Unambiguous characterization of *N*-glycans of monoclonal antibody cetuximab by integration of LC-MS/MS and ^1H NMR spectroscopy, *Anal. Chem.* **2014**, 86, 4807-4814.
- [119] E. Reh, *Chemometrie: Grundlagen der Statistik, numerischen Mathematik und Software-Anwendung in der Chemie*, De Gruyter, Berlin, Boston, **2017**.
- [120] H.-F. Eckey, R. Kosfeld, M. Türrck, *Deskriptive Statistik: Grundlagen - Methoden - Beispiele*, 4. Auflage, Gabler Verlag, Wiesbaden, **2005**.
- [121] J. Hartung, B. Elpelt, K.-H. Klösener, *Statistik – Lehr- und Handbuch der angewandten Statistik*, 15. Auflage, Oldenbourg Wissenschaftsverlag, München, **2009**.
- [122] J. N. Miller, J. C. Miller, *Statistics and Chemometrics for Analytical Chemistry*, 6. Auflage, Pearson Education Limited, Harlow, Essex, **2010**.
- [123] J. F. G. Vliegenthart, L. Dorland, H. van Halbeek, High-resolution, ^1H -Nuclear Magnetic Resonance Spectroscopy as a tool in the structural analysis of carbohydrates related to glycoproteins, *Adv. Carbohydr. Chem. Biochem.* **1983**, 41, 209-374.

- [124] R. T. McKay, How the 1D-NOESY suppresses solvent signal in metabonomics NMR spectroscopy: An examination of the pulse sequence components and evolution, *Concepts Magn. Reson. Part A* **2011**, 38A, 197-220.
- [125] R. Kostianen, A. P. Bruins, Effect of Solvent on Dynamic Range and Sensitivity in Pneumatically-assisted Electrospray (Ion Spray) Mass Spectrometry, *Rapid Commun. Mass Spectrom.* **1996**, 10, 1393-1399.
- [126] F. Savorani, G. Tomasi, S. B. Engelsen, Icoshift: A versatile tool for the rapid alignment of 1D NMR spectra, *J. Magn. Reson.* **2010**, 202, 190-202.
- [127] N. Nyberg, *MATLAB Central File Exchange: RBNMR*, zu finden unter <https://de.mathworks.com/matlabcentral/fileexchange/40332-rbnmr>, letzter Zugriff am **03.01.2024**.
- [128] T. N. Vu, K. Laukens, Getting your peaks in line: a review of alignment methods for NMR spectral data, *Metabolites* **2013**, 3, 259-276.
- [129] A. Edison, *Edison Lab Metabolomics Toolbox* zu finden unter https://github.com/edisonomics/metabolomics_toolbox, letzter Zugriff am **06.07.2023**.
- [130] *Biological Magnetic Resonance Data Bank (BMRB) – Phenylalanin* zu finden unter https://bmrb.io/metabolomics/mol_summary/show_data.php?id=bmse000045, letzter Zugriff am **15.09.2024**.
- [131] M. Mamani-Huanca, A. Gradillas, A. G. de La Fuente, Á. López-González, C. Barbas, Unveiling the fragmentation mechanisms of modified amino acids as the key for their targeted identification, *Anal. Chem.* **2020**, 92, 4848-4857.
- [132] C. Alasalvar, F. Shahidi, *Tree nuts: Composition, phytochemicals, and health effects*, 1. Auflage, CRC Press, Boca Raton, **2009**.
- [133] M. Vega-Castellote, M.-T. Sánchez, I. Torres-Rodríguez, J.-A. Entrenas, D. Pérez-Marín, NIR sensing technologies for the detection of fraud in nuts and nut products: A review, *Foods* **2024**, 13.

- [134] Europol/INTERPOL, *Operation OPSON VI - Targeting Counterfeit and Substandard Foodstuff and Beverage*, zu finden unter https://www.bvl.bund.de/SharedDocs/Downloads/01_Lebensmittel/FoodFraud/Public_Report_OPSON_VI.pdf, letzter Zugriff am **29.09.2024**.
- [135] A. Srbinovska, L. Gasparotto, C. Conchione, L. Menegoz Ursol, F. Lambertini, M. Suman, S. Moret, Mineral oil contamination in basil pesto from the Italian market: Ingredient contribution and market survey, *J. Food Compos. Anal.* **2023**, *115*, 104914.
- [136] N. Holz, Evaluierung neuer *sel*-TOCSY-NOESY und *sel*-NOESY-TOCSY Pulssequenzen und ihre Anwendung für die Strukturaufklärung von Metaboliten aus Cashewkernen, *Masterarbeit* **2021**, Universität Hamburg.
- [137] C. Schmitt, T. Bastek, A. Stelzer, T. Schneider, M. Fischer, T. Hackl, Detection of peanut adulteration in food samples by nuclear magnetic resonance spectroscopy, *J. Agric. Food Chem.* **2020**, *68*, 14364-14373.
- [138] A. B. Anderson, D. L. MacDonald, H. O. L. Fischer, The Structure of Pinitol, *J. Am. Chem. Soc.* **1952**, *74*, 1479-1480.
- [139] T. Posternak, Recherches dans la série des cyclites III. Sur la configuration des inosites actives, *Helv. Chim. Acta* **1936**, *19*, 1007–1010.
- [140] L. Ruiz-Aceituno, L. Ramos, I. Martinez-Castro, M. L. Sanz, Low molecular weight carbohydrates in pine nuts from *Pinus pinea* L., *J. Agric. Food Chem.* **2012**, *60*, 4957-4959.
- [141] F. A. Loewus, P. P. N. Murthy, *myo*-Inositol metabolism in plants, *Plant Sci.* **2000**, *150*, 1-19.
- [142] M. Ligor, I.-A. Rațiu, H. Al-Suod, A. Owczarczyk-Saczonek, L. Lahuta, R. Górecki, B. Buszewski, *Cyclitols – Determination in Food and Bioactivity in the Human Organism in Analytical Methods in the Determination of Bioactive Compounds and Elements in Food* (Hrsg.: M. Jeszka-Skowron, A. Zgoła-Grześkowiak, T. Grześkowiak, A. Ramakrishna), Springer International Publishing, Cham, **2021**, S. 163-191.
- [143] Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway Database, *Fructose and mannose metabolism – Reference pathway*, zu finden unter <https://www.kegg.jp/pathway/map00051>, letzter Zugriff am **16.10.2023**.

- [144] *Biological Magnetic Resonance Data Bank (BMRB) – L-(–)-Glutathione, Oxidized*, zu finden unter https://bmr.io/metabolomics/mol_summary/show_data.php?id=bmse000170, letzter Zugriff am **17.10.2023**.
- [145] P. Pallerla, S. R. Bhumireddy, S. S. Lingampally, N. C. Ragi, P. Sripadi, P. ESI-MS/MS analysis of protonated *N*-methyl amino acids and their immonium ions, *J. Mass Spectrom.* **2019**, *54*, 761-771.
- [146] V. Naresh Chary, C. Dinesh Kumar, M. Vairamani, S. Prabhakar, Characterization of amino acid-derived betaines by electrospray ionization tandem mass spectrometry, *J. Mass Spectrom.* **2012**, *47*, 79-88.
- [147] M. Lever, S. Slow, The clinical significance of betaine, an osmolyte with a key role in methyl group metabolism, *Clin. Biochem.* **2010**, *43*, 732-744.
- [148] K. Kai, K. Wakasa, H. Miyagawa, Metabolism of indole-3-acetic acid in rice: identification and characterization of *N*-beta-D-glucopyranosyl indole-3-acetic acid and its conjugates, *Phytochem.* **2007**, *68*, 2512-2522.
- [149] K. Ljung, A. Östin, L. Lioussanne, G. Sandberg, Development regulation of indole-3-acetic acid turnover in scots pine seedlings, *Plant Physiol.* **2001**, *125*, 464-475.
- [150] J.-T. Cheng, C. Guo, W.-J. Cui, Q. Zhang, S.-H. Wang, Q.-H. Zhao, D.-W. Liu, J. Zhang, S. Chen, C. Chen, C. *et al.*, Isolation of two rare *N*-glycosides from *Ginkgo biloba* and their anti-inflammatory activities, *Sci. Rep.* **2020**, *10*, 5994.
- [151] P. Shu, M. Sun, J. Li, L. Zhang, H. Xu, Y. Lou, Z. Ju, X. Wei, W. Wu, N. Sun, Chemical constituents from *Ginkgo biloba* leaves and their cytotoxicity activity, *J. Nat. Med.* **2020**, *74*, 269-274.
- [152] J. Normanly, Auxin metabolism, *Physiol. Plant.* **1997**, *100*, 431-442.
- [153] J. P. Glusker, Citrate conformation and chelation: Enzymatic implications, *Acc. Chem. Res.* **1980**, *13*, 345-352.
- [154] K. Takanori, S. Yoshinari, U. Katsuhito, S. Sadao, Acidic *N*^α-acylarginine derivatives in arginine-accumulating plant tissues, *Phytochem.* **1983**, *22*, 147-149.
- [155] T. Kasai, S. Sakamura, Acidic *N*^α-acylarginine derivatives in apple and pear trees, *Phytochem.* **1984**, *23*, 19-22.

- [156] D. L. Callahan, U. Roessner, V. Dumontet, N. Perrier, A. G. Wedd, R. A. J. O'Hair, A. J. M. Baker, S. D. Kolev, LC-MS and GC-MS metabolite profiling of nickel(II) complexes in the latex of the nickel-hyperaccumulating tree *Sebertia acuminata* and identification of methylated aldaric acid as a new nickel(II) ligand, *Phytochem.* **2008**, 69, 240-251.
- [157] T. Hunscha, Sortendifferenzierung von Apfelproben mittels NMR-Spektroskopie und multivariater Datenanalyse, *Bachelorarbeit* **2022**, Universität Hamburg.
- [158] Bundesministerium für Ernährung und Landwirtschaft (Statista), *Pro-Kopf-Konsum von Obst in Deutschland nach Art*, zu finden unter <https://de.statista.com/statistik/daten/studie/247425/umfrage/die-beliebtesten-obstsorten-der-deutschen/>, letzter Zugriff am **15.08.2023**.
- [159] Food and Agriculture Organization of the United Nations (FAOSTAT), *Crops and livestock products*, zu finden unter <https://www.fao.org/faostat>, letzter Zugriff am **15.08.2023**.
- [160] M. E. Dasenaki, N. S. Thomaidis, Quality and Authenticity Control of Fruit Juices – A Review, *Molecules* **2019**, 24.
- [161] J. C. Moore, J. Spink, M. Lipp, Development and application of a database of food ingredient fraud and economically motivated adulteration from 1980 to 2010, *J. Food Sci.* **2012**, 77, 118-126.
- [162] N. P. Bondonno, C. P. Bondonno, N. C. Ward, J. M. Hodgson, K. D. Croft, The cardiovascular health benefits of apples: Whole fruit vs. isolated compounds, *Trends Food Sci. Technol.* **2017**, 69, 243-256.
- [163] P. Kumar, S. Sethi, R. R. Sharma, S. Singh, S. Saha, V. K. Sharma, M. K. Verma, S. K. Sharma, Nutritional characterization of apple as a function of genotype, *J. Food Sci. Technol.* **2018**, 55, 2729-2738.
- [164] A. Alberti, A. A. F. Zielinski, M. Couto, P. Judacewski, L. I. Mafra, A. Nogueira, Distribution of phenolic compounds and antioxidant capacity in apples tissues during ripening, *J. Food Sci. Technol.* **2017**, 54, 1511-1518.
- [165] H. Zhang, R. Tsao, Dietary polyphenols, oxidative stress and antioxidant and anti-inflammatory effects, *Curr. Opin. Food Sci.* **2016**, 8, 33-42.
- [166] G. Antranikian, *Angewandte Mikrobiologie*, Springer, Berlin, Heidelberg, **2006**.

- [167] Kyoto Encyclopedia of Genes and Genomes (KEGG) Pathway Database, *Butanoate metabolism – Reference pathway*, zu finden unter https://www.kegg.jp/kegg-bin/show_pathway?map00650, letzter Zugriff am **21.09.2023**.
- [168] D. Radoš, D. L. Turner, T. Catarino, E. Hoffart, A. R. Neves, B. J. Eikmanns, B. Blombach, H. Santos, Stereospecificity of *Corynebacterium glutamicum* 2,3-butanediol dehydrogenase and implications for the stereochemical purity of bioproduced 2,3-butanediol, *Appl. Microbiol. Biotechnol.* **2016**, *100*, 10573-10583.
- [169] C. Chen, D. Wei, J. Shi, M. Wang, J. Hao, Mechanism of 2,3-butanediol stereoisomer formation in *Klebsiella pneumoniae*, *Appl. Microbiol. Biotechnol.* **2014**, *98*, 4603-4613.
- [170] L. Zhang, Q. Xu, S. Zhan, Y. Li, H. Lin, S. Sun, L. Sha, K. Hu, X. Guan, Y. Shen, A new NAD(H)-dependent *meso*-2,3-butanediol dehydrogenase from an industrially potential strain *Serratia marcescens* H30. *Appl. Microbiol. Biotechnol.* **2014**, *98*, 1175-1184.
- [171] E. González, M. R. Fernández, C. Larroy, L. Solà, M. A. Pericàs, X. Parés, J. A. Biosca, Characterization of a (2*R*,3*R*)-2,3-butanediol dehydrogenase as the *Saccharomyces cerevisiae* YAL060W gene product. Disruption and induction of the gene, *J. Biol. Chem.* **2000**, *275*, 35876-35885.
- [172] J. Gao, H.-H. Yang, X.-H. Feng, S. Li, H. Xu, A 2,3-butanediol dehydrogenase from *Paenibacillus polymyxa* ZJ-9 for mainly producing *R,R*-2,3-butanediol: Purification, characterization and cloning, *J. Basic Microbiol.* **2013**, *53*, 733-741.
- [173] K. J. R. Rosman, P. D. P. Taylor, Isotopic Compositions of the Elements 1997, *Pure Appl. Chem.* **1998**, *70*, 217–235.
- [174] S. H. Jeon, W. Chun, Y. J. Choi, Y. S. Kwon, Cytotoxic constituents from the bark of *salix hulteni*, *Arch. Pharm. Res.* **2008**, *31*, 978-982.
- [175] G. Cefarelli, B. D'Abrosca, A. Fiorentino, A. Izzo, C. Mastellone, S. Pacifico, V. Piscopo, Free-radical-scavenging and antioxidant activities of secondary metabolites from reddened cv. annurca apple fruits, *J. Agric. Food Chem.* **2006**, *54*, 803-809.
- [176] J. Dou, P. Ilina, J. Hemming, K. Malinen, H. Mäkkylä, N. Oliveira De Farias, P. Tammela, G. de Aragao Umbuzeiro, R. Räisänen, T. Vuorinen, Effect of hybrid type and harvesting

- season on phytochemistry and antibacterial activity of extracted metabolites from *salix* bark, *J. Agric. Food Chem.* **2022**, *70*, 2948-2956.
- [177] C.-C. Zhao, J. Shen, J. Chen, J.-H. Shao, K.-H. Li, W.-Y. Gu, B.-J. Miao, Phenolic glycoside constituents from *Brassica rapa* flowers and their α -glucosidase inhibitory activity, *Nat. Prod. Res.* **2019**, *33*, 3398-3403.
- [178] A. Tolonen, Analysis of secondary metabolites in plant and cell culture tissue of *Hypericum perforatum* L. and *Rhodiola rosea* L., *Dissertation* **2003**, University of Oulu.
- [179] H. Ito, E. Kobayashi, S.-H. Li, T. Hatano, D. Sugita, N. Kubo, S. Shimura, Y. Itoh, T. Yoshida, Megastigmane glycosides and an acylated triterpenoid from *Eriobotrya japonica*, *J. Nat. Prod.* **2001**, *64*, 737-740.
- [180] W. Schwab, P. Schreier, Vomifoliol 1-O- β -D-Xylopyranosyl-6-O- β -D-glucopyranoside: A disaccharide glycoside from apple fruit, *Phytochem.* **1990**, *29*, 161-164.
- [181] M. Herderich, W. Feser, P. Schreier, Vomifoliol 9-O- β -D-Glucopyranosyl-4-O- β -D-xylopyranosyl-6-O- β -D-glucopyranosid: A trisaccharide glycoside from apple fruit, *Phytochem.* **1992**, *31*, 895-897.
- [182] M. Matsuda, K. Isawa, M. Kikuchi, Megastigmane glycosides from *Lonicera gracilipes* var. *Glandulosa*, *Phytochem.* **1997**, *45*, 777-779.
- [183] V. A. Marinos, M. E. Tate, P. J. Williams, Protocol for FAB MS/MS characterization of terpene disaccharides of wine, *J. Agric. Food Chem.* **1994**, *42*, 2486-2492.
- [184] N. De Tommasi, R. Aquino, F. de Simone, C. Pizza, Plant metabolites. New sesquiterpene and ionone glycosides from *Eriobotrya japonica*, *J. Nat. Prod.* **1992**, *55*, 1025-1032.
- [185] H. Higuchi, K. Fukui, J. Kinjo, T. Nohara, Four New Glycosides from *Albizia Cortex*. III, *Chem. Pharm. Bull.* **1992**, *40*, 534-535.
- [186] L. Yu, J. Liu, L. Yu, L. Chen, F. Qiu, Chemical constituents of seed oil leavings of *Xanthoceras sorbifolia*, *Chem. Nat. Compd.* **2018**, *54*, 769-771.
- [187] K. Weinges, H. Schick, M. Lautenschläger, G. Schilling, Isolierung und NMR-spektroskopische Untersuchungen von (-)-1-O-Benzyl- β -D-primverosid aus grünen Früchten von *Prunus laurocerasus*, *Liebigs Ann. Chem.* **1991**, *7*, 703-705.

- [188] K. L. Stuart, L. B. Coke, The Effect of vomifoliol on stomatal aperture, *Planta* **1975**, 122, 307-310.
- [189] S. Hasegawa, S. M. Poling, V. P. Maier, R. D. Bennett, Metabolism of Absciscic acid: Bacterial conversion to Dehydrovomifoliol and Vomifoliol Dehydrogenase activity, *Phytochem.* **1984**, 23, 2769-2771.
- [190] B. Gouilleux, B. Charrier, E. Danieli, J.-N. Dumez, S. Akoka, F.-X. Felpin, M. Rodriguez-Zubiri, P. Giraudeau, Real-time reaction monitoring by ultrafast 2D NMR on a benchtop spectrometer, *Analyst* **2015**, 140, 7854-7585.
- [191] S.-H. Yoo, Y.-W. Kim, J. Shin, N.-K. Kim, J.-S. Kim, Effects of the chain length of Tris(carboxyalkylamino)triazine on corrosion inhibition properties, *Bull. Korean Chem. Soc.* **2015**, 36, 346-355.
- [192] K. Bretterbauer, C. Schwarzingner, Melamine derivatives – A review on synthesis and application, *Curr. Org. Synth.* **2012**, 9, 342–356.
- [193] E. M. Smolin, L. Rapoport, *s-Triazines and derivatives* in *The chemistry of heterocyclic compounds* (Hrsg.: A. Weissberger), Interscience Publishers, New York, London, **1959**.
- [194] H. Nestler, H. Fürst, Über die Darstellung von *N(s-Triazinyl)*-amino-säurederivaten, *J. prakt. Chem.* **1963**, 22, 173-185.
- [195] J. B. Lambert, S. Gronert, H. Shurvell, D. A. Lightner, *Spektroskopie: Strukturaufklärung in der Organischen Chemie*, 2. Auflage, Pearson Deutschland, München, **2012**.
- [196] F. Puig-Castellví, Y. Pérez, B. Piña, R. Tauler, I. Alfonso, Comparative analysis of ^1H NMR and ^1H - ^{13}C HSQC NMR metabolomics to understand the effects of medium composition in yeast growth, *Anal. Chem.* **2018**, 90, 12422-12430.
- [197] J. Meiler, R. Meusinger, M. Will, Fast determination of ^{13}C NMR chemical shifts using artificial neural networks, *J. Chem. Inf. Comput. Sci.* **2000**, 40, 1169-1176.
- [198] C. S. Clendinen, G. S. Stupp, R. Ajredini, B. Lee-McMullen, C. Beecher, A. S. Edison, An overview of methods using ^{13}C for improved compound identification in metabolomics and natural products, *Front Plant Sci.* **2015**, 6, 1–13.
- [199] A. Shaban, *2D NMR spectra fitter* zu finden unter <https://www.mathworks.com/matlabcentral/fileexchange/52300-2d-nmr-spectra-fitter>, letzter Zugriff am **13.09.2023**.

- [200] SMART NMR, zu finden unter <http://smart.ucsd.edu/classic>, letzter Zugriff am **21.02.2024**.
- [201] M. Szultka, M. Buszewska-Forajita, R. Kaliszan, B. Buszewski, Determination of ascorbic acid and its degradation products by high-performance liquid chromatography-triple quadrupole mass spectrometry, *Electrophoresis* **2013**, 35, 585-592.
- [202] MERCK KGaA, *Sicherheitsdatenblätter*, zu finden unter <https://www.sigmaaldrich.com/>, letzter Zugriff am **30.05.2024**.
- [203] METALL-CHEMIE GmbH & Co. KG, TC (50, 65, 85) *Sicherheitsdatenblatt: gemäß Verordnung (EG) Nr. 1907/2006 (REACH)* **2021**.
- [204] MACHEREY-NAGEL GmbH & Co. KG, *Sicherheitsdatenblatt: gemäß REACH-Verordnung 1907/2006/EG 2024*, zu finden unter https://www.mn-net.com/files/digitalAssets/sds/de-DE/711014.100_de-DE_GHS.pdf, letzter Zugriff am **30.05.2024**.

8 Anhang

8.1 Verwendete Chemikalien

In der folgenden Tabelle 36 sind die in dieser Arbeit verwendeten Chemikalien mit den zugehörigen H- und P-Sätzen sowie den Gefahrenpiktogrammen aufgeführt.^[202–204]

Tabelle 36: In dieser Arbeit verwendete Chemikalien unter Angabe der zugehörigen Gefahrenpiktogramme sowie der H- und P-Sätze.

Substanz	Gefahren- piktogramm und Signalwort	H-Sätze	P-Sätze
Acetonitril	 Gefahr	H225, H302+H312+H332, H319	P210, P280, P301+P312, P303+P361+P353, P304+P340+P312, P305+P351+P338
Acetonitril- d_3	 Gefahr	H225, H302+H312+H332, H319	P210, P280, P301+P312, P303+P361+P353, P304+P340+P312, P305+P351+P338
Adenosin	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
L-Alanin	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
β -Alanin	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Ameisensäure	 Gefahr	H226, H302, H314, H331	P210, P280, P303+P361+P353, P304+P340+P310, P305+P350+P338
6-Aminocapronsäure	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
3-Aminopropyl- funktionalisiertes Kieselgel	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		

Substanz	Gefahren- piktogramm und Signalwort	H-Sätze	P-Sätze
Äpfelsäure	 Achtung	H319	P264, P280, P305+P351+P338, P337+P313
Ascorbinsäure	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
L-Asparagin	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
L-Asparaginsäure	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
2,3-Butandiol	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Calcium-D- Pantothenat	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
ε-Caprolactam	 Achtung	H302+H332, H315, H319, H335	P261, P264, P301+P312, P302+P352, P403+P340+P312, P305+P351+P338
(±)-Carnitin- Hydrochlorid	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Chinasäure	 Achtung	H319	P264, P280, P305+P351+P338, P337+P313
Chloroform	 Gefahr	H302, H315, H319, H331, H336, H351, H361d, H372	P202, P301+P312, P302+P352, P304+P340+P311, P305+P351+P338, P308+P313
Chloroform-d	 Gefahr	H302, H315, H319, H331, H336, H351, H361d, H372	P202, P301+P312, P302+P352, P304+P340+P311, P305+P351+P338, P308+P313

Substanz	Gefahren- piktogramm und Signalwort	H-Sätze	P-Sätze
Chlorogensäure	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Cholinchlorid	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Citronensäure	 Achtung	H319, H335	P261, P264, P271, P280, P304+P340+P312, P305+P351+P338
Creatinin	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Cyanurchlorid	 Gefahr	H302, H314, H317, H330, H335	P260, P280, P301+P312, P303+P361+P353, P304+P340+P310, P305+P351+P338
Deuteriumchlorid	 Gefahr	H290, H314, H335	P280, P301+P330+P331, P303+P361+P353, P305+P351+P338+ P310
Deuteriumoxid	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Dikaliumhydrogen- phosphat	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Ethanol	 Gefahr	H225, H319	P210, P233, P240, P241, P242, P305+P351+P338
Fumarsäure	 Achtung	H319	P264, P280, P305+P351+P338, P337+P313
D-Glucose Monohydrat	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		

Substanz	Gefahren- piktogramm und Signalwort	H-Sätze	P-Sätze
L-Glutaminsäure	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Histamin	 Gefahr	H301, H315, H317, H319, H334, H335	P280, P301+P310+P330, P302+P352, P305+P351+P338
L-Histidin	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
L-Isoleucin	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Isovaleriansäure	 Gefahr	H314	P280, P303+P361+P353, P304+P340+P310, P305+P351+P338, P363, P405
Kaliumdihydrogen- phosphat	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
L-Leucin	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Methanol	 Gefahr	H225, H301+H311+H331, H370	P210, P233, P280, P301+P310, P303+P361+P353, P304+P340+P311
Methanol-<i>d</i>₄	 Gefahr	H225, H301+H311+H331, H370	P210, P233, P280, P301+P310, P303+P361+P353, P304+P340+P311
Milchsäure	 Gefahr	H314	P280, P303+P361+P353, P304+P340+P310, P305+P351+P338, P363, P405

Substanz	Gefahren- piktogramm und Signalwort	H-Sätze	P-Sätze
Natriumazid	 Gefahr	H300+H310+H330, H373, H410	P262, P264, P273, P280, P302+P352+P310, P304+P340+P310
Natriumhydroxid	 Gefahr	H290, H314	P234, P260, P280, P303+P361+P353, P304+P340+P310, P305+P351+P338
Nicotinamid	 Achtung	H319	P264, P280, P305+P351+P338, P337+P313
Nicotinsäure	 Achtung	H319	P264, P280, P305+P351+P338, P337+P313
L-Phenylalanin	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
D-Pinitol	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
2-Propanol	 Gefahr	H225, H319, H336	P210, P233, P240, P241, P242, P305+P351+P338
D-Raffinose	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Saccharose	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Sorbitol	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Stachyose	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Thiamin- Hydrochlorid	 Achtung	H319	P264, P280, P305+P351+P338, P337+P313

Substanz	Gefahren- piktogramm und Signalwort	H-Sätze	P-Sätze
L-Threonin	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Tiglinsäure	 Achtung	H315, H319, H335	P261, P264, P271, P280, P302+P352, P305+P351+P338
2,4,6-Tri-(6-aminocapronsäure)-1,3,5-triazin	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Tricarallylsäure	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
3-(Trimethylsilyl)-propionsäure-2,2,3,3-d₄-Natriumsalz	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
L-Tryptophan	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Uridin	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Valeriansäure	 Gefahr	H314, H412	P273, P280, P303+P361+P353, P304+P340+P310, P305+P351+P338, P363
L-Valin	Diese Substanz ist gemäß Verordnung (EG) Nr. 1272/2008 nicht als gefährlich eingestuft.		
Vanillin	 Achtung	H319	P264, P280, P305+P351+P338, P337+P313

8.2 Apfelproben

In Tabelle 37 sind die Angaben zu der in dieser Arbeit verwendeten Apfelprobe bzgl. Sorte, Erntejahr, Herkunftsland, Bundesland und Kreis gegeben. Alle 86 Proben wurden vom Landeslabor Schleswig-Holstein zur Verfügung gestellt.

Tabelle 37: Informationen über die in dieser Arbeit verwendeten Apfelproben mit Angaben zur Probenbezeichnung, der Sorte, dem Erntejahr, dem Herkunftsland inkl. Bundesland und Kreis. Fehlende Angaben sind mit NB (nicht bekannt) gekennzeichnet. DE: Deutschland, IT: Italien, BW: Baden-Württemberg, HH: Hamburg, NI: Niedersachsen, SH: Schleswig-Holstein.

Probenbezeichnung	Sorte	Erntejahr	Herkunftsland	Bundesland	Kreis
20-B-BW-1	Boskoop	2020	DE	BW	Rems-Murr
20-B-HH-2	Boskoop	2020	DE	HH	Hamburg
20-B-NI-3	Boskoop	2020	DE	NI	Stade
20-B-NI-4	Boskoop	2020	DE	NI	Stade
20-B-SH-5	Boskoop	2020	DE	SH	Ostholstein
20-B-SH-6	Boskoop	2020	DE	SH	Pinneberg
20-B-SH-7	Boskoop	2020	DE	SH	Pinneberg
20-B-SH-8	Boskoop	2020	DE	SH	Pinneberg
20-B-SH-9	Boskoop	2020	DE	SH	Plön
20-B-SH-10	Boskoop	2020	DE	SH	Plön
20-B-SH-11	Boskoop	2020	DE	SH	Rendsburg-Eckernförde
20-B-SH-12	Boskoop	2020	DE	SH	Segeberg
20-B-SH-13	Boskoop	2020	DE	SH	Stormarn
21-B-SH-14	Boskoop	2021	DE	SH	Rendsburg-Eckernförde
21-B-SH-15	Boskoop	2021	DE	SH	Rendsburg-Eckernförde
21-B-SH-16	Boskoop	2021	DE	SH	Rendsburg-Eckernförde
21-B-BW-17	Boskoop	2021	DE	BW	Karlsruhe
21-B-BW-18	Boskoop	2021	DE	BW	Rems-Murr
21-B-BW-19	Boskoop	2021	DE	BW	Rems-Murr
21-B-BW-20	Boskoop	2021	DE	BW	Weinstadt
20-E-BW-1	Elstar	2020	DE	BW	Heilbronn
20-E-BW-2	Elstar	2020	DE	BW	Konstanz
20-E-BW-3	Elstar	2020	DE	BW	Karlsruhe
20-E-BW-4	Elstar	2020	DE	BW	Rems-Murr

Proben- bezeichnung	Sorte	Ernte- jahr	Herkunfts- land	Bundes- land	Kreis
20-E-NB-5	Elstar	2020	DE	NB	NB
20-E-NI-6	Elstar	2020	DE	NI	Stade
20-E-NI-7	Elstar	2020	DE	NI	Stade
20-E-SH-8	Elstar	2020	DE	SH	Ostholstein
20-E-SH-9	Elstar	2020	DE	SH	Ostholstein
20-E-SH-10	Elstar	2020	DE	SH	Pinneberg
20-E-SH-11	Elstar	2020	DE	SH	Rendsburg-Eckernförde
20-E-SH-12	Elstar	2020	DE	SH	Segeberg
20-E-SH-13	Elstar	2021	DE	SH	Segeberg
20-E-SH-14	Elstar	2020	DE	SH	Steinburg
20-E-SH-15	Elstar	2020	DE	SH	Stormarn
21-E-BW-16	Elstar	2021	DE	BW	Hohenlohe-Kreis
21-E-BW-17	Elstar	2021	DE	BW	Karlsruhe
21-E-BW-18	Elstar	2021	DE	BW	NB
21-E-BW-19	Elstar	2021	DE	BW	NB
21-E-BW-20	Elstar	2021	DE	BW	Weinstadt
21-E-BW-21	Elstar	2021	DE	BW	Waldshut
21-E-BW-22	Elstar	2021	DE	BW	Rems-Murr
21-E-SH-23	Elstar	2021	DE	SH	Plön
20-G-BW-1	Gala	2020	DE	BW	Heilbronn
20-G-BW-2	Gala	2020	DE	BW	Konstanz
20-G-BW-3	Gala	2020	DE	BW	Karlsruhe
20-G-BW-4	Gala	2020	DE	BW	Rems-Murr
20-G-BW-5	Gala	2020	DE	BW	Rems-Murr
20-G-NI-6	Gala	2020	DE	NI	Stade
20-G-SH-7	Gala	2020	DE	SH	Pinneberg
20-G-SH-8	Gala	2020	DE	SH	Pinneberg
20-G-SH-9	Gala	2020	DE	SH	Segeberg
20-G-SH-10	Gala	2020	DE	SH	Steinburg
20-G-IT-11	Gala	2020	IT	NB	NB
20-G-IT-12	Gala	2020	IT	NB	NB
21-G-BW-13	Gala	2021	DE	BW	Karlsruhe

Proben- bezeichnung	Sorte	Ernte- jahr	Herkunfts- land	Bundes- land	Kreis
21-G-BW-14	Gala	2021	DE	BW	NB
21-G-BW-15	Gala	2021	DE	BW	NB
21-G-BW-16	Gala	2021	DE	BW	NB
21-G-BW-17	Gala	2021	DE	BW	Weinstadt
21-G-BW-18	Gala	2021	DE	BW	Ravensburg
21-G-BW-19	Gala	2021	DE	BW	Stuttgart
20-H-SH-1	Holsteiner Cox	2020	DE	SH	Kiel
20-H-HH-2	Holsteiner Cox	2020	DE	HH	Hamburg
20-H-NI-3	Holsteiner Cox	2020	DE	NI	Stade
20-H-SH-4	Holsteiner Cox	2020	DE	SH	Ostholstein
20-H-SH-5	Holsteiner Cox	2020	DE	SH	Pinneberg
20-H-SH-6	Holsteiner Cox	2020	DE	SH	Pinneberg
20-H-SH-7	Holsteiner Cox	2020	DE	SH	Pinneberg
20-H-SH-8	Holsteiner Cox	2020	DE	SH	Rendsburg-Eckernförde
20-H-SH-9	Holsteiner Cox	2020	DE	SH	Rendsburg-Eckernförde
20-H-SH-10	Holsteiner Cox	2020	DE	SH	Segeberg
20-H-SH-11	Holsteiner Cox	2020	DE	SH	Steinburg
21-H-SH-12	Holsteiner Cox	2021	DE	SH	Neumünster
21-H-SH-13	Holsteiner Cox	2021	DE	SH	Rendsburg-Eckernförde
21-H-SH-14	Holsteiner Cox	2021	DE	SH	Rendsburg-Eckernförde
21-H-SH-15	Holsteiner Cox	2021	DE	SH	Silberstedt
20-W-BW-1	Wellant	2020	DE	BW	Konstanz
20-W-NI-2	Wellant	2020	DE	NI	Stade
20-W-NI-3	Wellant	2020	DE	NI	Stade
20-W-SH-4	Wellant	2020	DE	SH	Herzogtum Lauenburg
20-W-SH-5	Wellant	2020	DE	SH	Ostholstein
20-W-SH-7	Wellant	2020	DE	SH	Steinburg
21-W-BW-8	Wellant	2021	DE	BW	Rems-Murr
21-W-BW-9	Wellant	2021	DE	BW	Weinstadt
21-W-SH-10	Wellant	2021	DE	SH	Rendsburg-Eckernförde

8.3 Multivariate Datenanalyse

8.3.1 Liste der variablen Buckets

Bereiche der 159 variablen *Buckets* zur Erstellung des Mehrklassenmodells zur Unterscheidung aller Apfelsorten.

[ppm]: 0.695 – 0.7207; 0.8382 – 0.8661; 0.8668 – 0.8825; 0.8828 – 0.8931; 0.8934 – 0.9001;
 0.9001 – 0.9088; 0.9091 – 0.9181; 0.9184 – 0.9281; 0.9284 – 0.9454; 0.9457 – 0.9727;
 0.9781 – 0.9919; 0.9935 – 1.0074; 1.0080 – 1.0240; 1.0247 – 1.0388; 1.0391 – 1.0706;
 1.1104 – 1.1601; 1.1605 – 1.1948; 1.1948 – 1.2006; 1.2009 – 1.2224; 1.2231 – 1.2609;
 1.2613 – 1.2914; 1.2918 – 1.3229; 1.3229 – 1.3489; 1.3492 – 1.3649; 1.3653 – 1.3871;
 1.3877 – 1.3970; 1.4192 – 1.4288; 1.4567 – 1.5004; 1.5007 – 1.5521; 1.5527 – 1.5688;
 1.5691 – 1.5842; 1.5848 – 1.6140; 1.6144 – 1.6311; 1.6317 – 1.6497; 1.6500 – 1.6673;
 1.6676 – 1.6850; 1.6853 – 1.7058; 1.7062 – 1.7248; 1.7251 – 1.7434; 1.7437 – 1.7617;
 1.7620 – 1.7797; 1.7800 – 1.7986; 1.8330 – 1.9151; 1.9158 – 1.9331; 1.9411 – 2.0012;
 2.0018 – 2.0272; 2.0275 – 2.0926; 2.1013 – 2.1138; 2.1142 – 2.1344; 2.1350 – 2.1453;
 2.1456 – 2.1527; 2.1530 – 2.1710; 2.2352 – 2.2397; 2.2679 – 2.2891; 2.2910 – 2.3103;
 2.3138 – 2.3244; 2.3254 – 2.4557; 2.4560 – 2.4762; 2.4888 – 2.6281; 2.6306 – 2.7552;
 2.7555 – 2.7754; 2.7757 – 2.8065; 2.8072 – 2.8258; 2.8264 – 2.8374; 2.8377 – 2.8768;
 2.8772 – 2.9202; 2.9215 – 2.9616; 2.9626 – 2.9940; 2.9943 – 3.0415; 3.1625 – 3.1757;
 3.1818 – 3.1917; 3.1927 – 3.2075; 3.2078 – 3.2200; 3.2203 – 3.2864; 3.2868 – 3.3063;
 3.3067 – 3.3166; 3.3173 – 3.3336; 3.3346 – 3.3561; 3.3757 – 3.3917; 3.3920 – 3.4026;
 3.4030 – 3.4158; 3.4161 – 3.4267; 3.4270 – 3.4392; 3.4396 – 3.4723; 3.4726 – 3.4903;
 3.4906 – 3.5176; 3.5188 – 3.5336; 3.5339 – 3.5673; 3.5676 – 3.5904; 3.5907 – 3.6049;
 3.6164 – 3.6334; 3.6052 – 3.6161; 3.6338 – 3.6546; 3.6549 – 3.6790; 3.6793 – 3.7037;
 3.7037 – 3.7166; 3.7169 – 3.7211; 3.7214 – 3.7346; 3.7346 – 3.7602; 3.7830 – 3.8045;
 3.7609 – 3.7827; 3.8049 – 3.8405; 3.8405 – 3.8774; 3.8777 – 3.9063; 3.9063 – 3.9329;
 3.9333 – 3.9779; 3.9804 – 4.0244; 4.0247 – 4.0633; 4.0642 – 4.0899; 4.0934 – 4.1538;
 4.1952 – 4.2482; 4.2796 – 4.3489; 4.1541 – 4.1917; 4.2488 – 4.2777; 4.3794 – 4.4237;
 4.4314 – 4.4921; 4.5695 – 4.5830; 4.5830 – 4.5897; 4.5900 – 4.6019; 4.6019 – 4.6106;
 4.6131 – 4.7088; 4.9874 – 5.0202; 5.1116 – 5.1354; 5.1797 – 5.1954; 5.1954 – 5.2195;
 5.2198 – 5.2689; 5.2692 – 5.2981; 5.2985 – 5.3241; 5.3245 – 5.3636; 5.3639 – 5.4911;
 5.4933 – 5.5276; 5.8111 – 5.8554; 5.8846 – 5.9414; 6.0271 – 6.0567; 6.0782 – 6.1019;

6.1401 – 6.1851; 6.1035 – 6.1266; 6.2059 – 6.3170; 6.3289 – 6.3767; 6.4515 – 6.4720;
 6.4836 – 6.5436; 6.8197 – 6.8742; 7.0505 – 7.0855; 7.1131 – 7.1914; 7.2052 – 7.2626;
 7.4093 – 7.4488; 7.3163 – 7.4023; 7.4922 – 7.5271; 7.5284 – 7.5676; 7.5894 – 7.6385;
 7.6389 – 7.7153; 7.7557 – 7.8193; 7.8555 – 7.8969; 7.9265 – 7.9605; 8.2202 – 8.2728;
 8.0565 – 8.2006; 8.4179 – 8.4474; 8.8111 – 8.8830; 9.1061 – 9.1540.

8.3.2 Liste der variablen, selektierten *Buckets*

Bereiche der 82 variablen *Buckets*, die für das Mehrklassenmodell zur Unterscheidung aller fünf Apfelsorten sowie für das Zweiklassenmodell zur Unterscheidung der Proben Boskoop und Gala verwendet wurden.

[ppm]: 0.8382 – 0.8661; 0.8934 – 0.9001; 0.9001 – 0.9088; 0.9091 – 0.9181; 0.9184 – 0.9281;
 1.0080 – 1.0240; 1.0247 – 1.0388; 1.0391 – 1.0706; 1.2918 – 1.3229; 1.3229 – 1.3489;
 1.4192 – 1.4288; 1.5007 – 1.5521; 1.5527 – 1.5688; 1.5691 – 1.5842; 1.6144 – 1.6311;
 1.6317 – 1.6497; 1.6500 – 1.6673; 1.6676 – 1.6850; 1.8330 – 1.9151; 1.9158 – 1.9331;
 1.9411 – 2.0012; 2.0018 – 2.0272; 2.0275 – 2.0926; 2.1530 – 2.1710; 2.3254 – 2.4557;
 2.4560 – 2.4762; 2.4888 – 2.6281; 2.6306 – 2.7552; 2.7555 – 2.7754; 2.7757 – 2.8065;
 2.8072 – 2.8258; 2.8264 – 2.8374; 3.2078 – 3.2200; 3.2203 – 3.2864; 3.2868 – 3.3063;
 3.3067 – 3.3166; 3.3173 – 3.3336; 3.3346 – 3.3561; 3.3757 – 3.3917; 3.3920 – 3.4026;
 3.4161 – 3.4267; 3.4270 – 3.4392; 3.4396 – 3.4723; 3.4726 – 3.4903; 3.4906 – 3.5176;
 3.5339 – 3.5673; 3.5676 – 3.5904; 3.6338 – 3.6546; 3.6549 – 3.6790; 3.7037 – 3.7166;
 3.7214 – 3.7346; 3.7830 – 3.8045; 3.7609 – 3.7827; 3.8405 – 3.8774; 3.8777 – 3.9063;
 3.9063 – 3.9329; 3.9333 – 3.9779; 3.9804 – 4.0244; 4.0934 – 4.1538; 4.2796 – 4.3489;
 4.1541 – 4.1917; 4.2488 – 4.2777; 4.4314 – 4.4921; 4.5830 – 4.5897; 4.5900 – 4.6019;
 4.6019 – 4.6106; 4.6131 – 4.7088; 5.2198 – 5.2689; 5.2692 – 5.2981; 5.2985 – 5.3241;
 5.3245 – 5.3636; 5.8846 – 5.9414; 7.2052 – 7.2626; 7.4093 – 7.4488; 7.3163 – 7.4023;
 7.5284 – 7.5676; 7.5894 – 7.6385; 7.6389 – 7.7153; 7.7557 – 7.8193; 7.8555 – 7.8969;
 8.0565 – 8.2006; 8.4179 – 8.4474.

8.3.3 PCA Score Plots

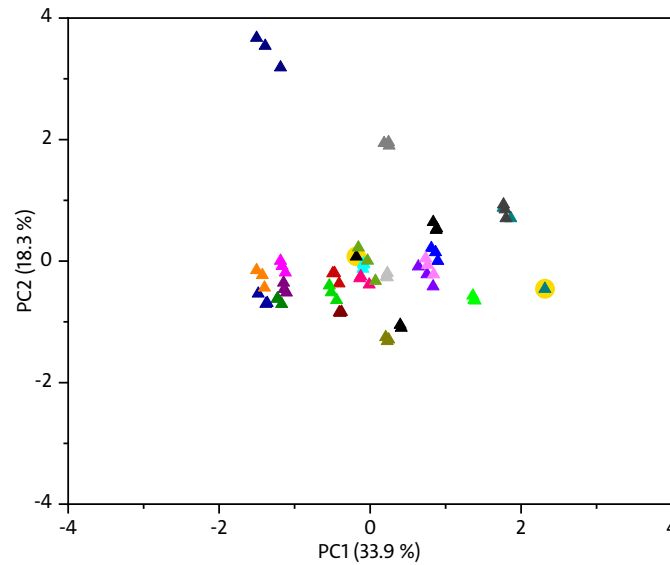


Abbildung 133: PCA Score Plot der im Triplikate aufgearbeiteten Proben der Sorte Elstar zur Detektion von Ausreißern. Messungen derselben Probe sind jeweils in der gleichen Farbe dargestellt. Ausreißer, die aus folgenden Analysen eliminiert wurden, sind gelb markiert. In diesem Fall tritt dies auf jeweils eine Messung der Probe 20-E-SH-15 (dunkelgrün) und 20-E-NI-7 (schwarz) zu.

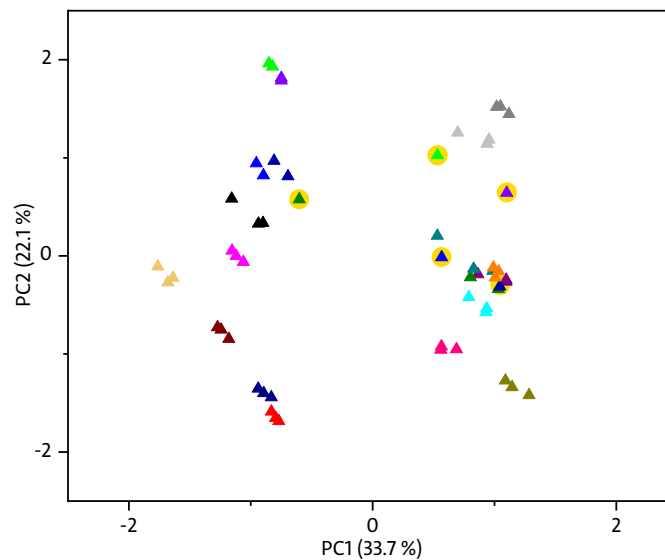


Abbildung 134: PCA Score Plot der im Triplikate aufgearbeiteten Proben der Sorte Gala zur Detektion von Ausreißern. Messungen derselben Probe sind jeweils in der gleichen Farbe dargestellt. Ausreißer, die aus folgenden Analysen eliminiert wurden, sind gelb markiert. In diesem Fall tritt dies auf jeweils eine Messung der Probe 20-G-BW-3 (hellgrün), 20-G-BW-4 (blau), 20-G-SH-9 (dunkelgrün), 21-G-BW-14 (dunkelblau) und 21-G-BW-16 (violett) zu.

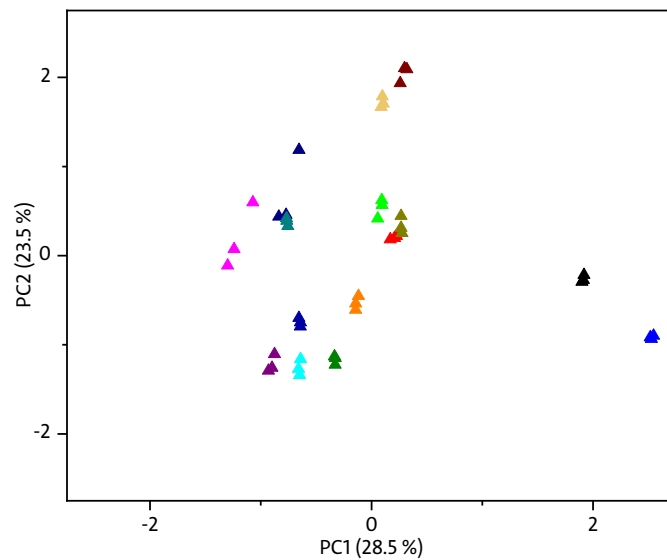


Abbildung 135: PCA Score Plot der im Triplikate aufgearbeiteten Proben der Sorte Holsteiner Cox zur Detektion von Ausreißern. Messungen derselben Probe sind jeweils in der gleichen Farbe dargestellt. Es wurden keine Ausreißer detektiert.

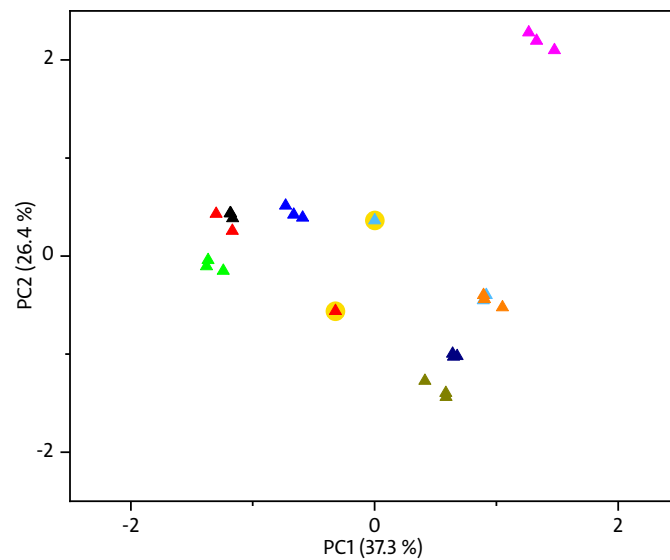


Abbildung 136: PCA Score Plot der im Triplikate aufgearbeiteten Proben der Sorte Wellant zur Detektion von Ausreißern. Messungen derselben Probe sind jeweils in der gleichen Farbe dargestellt. Ausreißer, die aus folgenden Analysen eliminiert wurden, sind gelb markiert. In diesem Fall tritt dies auf jeweils eine Messung der Probe 20-W-NI-2 (rot) und 20-W-SH-5 (hellblau) zu.

8.4 Korrelationskoeffizienten

8.4.1 Korrelationskoeffizienten zu Metaboliten in Pinienkernen

Tabelle 38: Korrelationskoeffizienten der im polaren Extrakt der Pinienkerne identifizierten Kohlenhydrate Saccharose (**2**) und Raffinose (**19**) sowie des Alditols Glycerol (**18**) und des Cyclitols Pinitol (**16**). In der linken Spalte sind die Metaboliten unter Angabe der monoisotopischen Masse sowie des Fraktionsbereiches, über den die Korrelationskoeffizienten berechnet wurden, aufgelistet. In der rechten Spalte sind die Korrelationskoeffizienten zwischen EDCs und EMCs oder zwischen EDCs untereinander angegeben.

Metabolit	Korrelationskoeffizienten				
D-Pinitol (16) 194.0790 Da F1-37					
	ESI-(+)	195.09 ([M+H] ⁺)	217.07 ([M+Na] ⁺)	109.03 ([M+H+Na] ²⁺)	411.15 ([2M+Na] ⁺)
	4.013 ppm	0.9859	0.9861	0.9865	0.9874
	3.809 ppm	0.966	0.9733	0.9831	0.9648
	3.772 ppm	0.9426	0.9557	0.9739	0.9388
	3.656 ppm	0.9804	0.9861	0.9894	0.9811
	3.599 ppm	0.9757	0.9845	0.9887	0.9758
	3.349 ppm	0.9810	0.9828	0.9851	0.9824
	ESI(-)	193.07 ([M-H] ⁻)	239.08 ([M+FA-H] ⁻)		
	4.013 ppm	0.9933	0.9854		
	3.809 ppm	0.9757	0.9696		
	3.772 ppm	0.9552	0.9509		
	3.656 ppm	0.9916	0.9847		
	3.599 ppm	0.9895	0.9839		
	3.349 ppm	0.9898	0.9815		
Glycerol (18) 92.0473 Da F1-30					
	EDC [ppm]	3.779	3.665	3.641	3.546
	3.779	1.0000	0.9125	0.9179	0.9129
	3.665	0.9125	1.0000	0.9985	0.9995
	3.641	0.9179	0.9985	1.0000	0.9987
	3.546	0.9129	0.9995	0.9984	1.0000

Metabolit	Korrelationskoeffizienten				
Saccharose* (2) 342.1162 Da F1-80	ESI-(+)		365.11 ([M+Na] ⁺)	707.23 ([2M+Na] ²⁺)	
	5.412 ppm		0.9531	0.9929	
	4.226 ppm		0.9524	0.9869	
	4.051 ppm		0.8678	0.9294	
	3.896 ppm		0.9406	0.9509	
	3.857 ppm		0.9688	0.9673	
	3.821 ppm		0.9630	0.9928	
	3.767 ppm		0.9629	0.9510	
	3.676 ppm		0.8638	0.9251	
	3.567 ppm		0.9479	0.9774	
	3.473 ppm		0.8919	0.9525	
	ESI-(-)	341.10 ([M-H] ⁻)	387.11 ([M+FA-H] ⁻)	683.21 ([2M-H] ²⁻)	729.21 ([2M+FA-H] ²⁻)
	5.412 ppm	0.9812	0.9833	0.9967	0.9974
	4.226 ppm	0.9781	0.9780	0.9862	0.9874
	4.051 ppm	0.8992	0.9013	0.9681	0.9644
	3.896 ppm	0.9488	0.9448	0.9486	0.9495
	3.857 ppm	0.9840	0.9808	0.9430	0.9477
	3.821 ppm	0.9846	0.9858	0.9935	0.9945
	3.767 ppm	0.9738	0.9691	0.9178	0.9236
	3.676 ppm	0.8883	0.8911	0.9614	0.9571
	3.567 ppm	0.9691	0.9679	0.9800	0.9807
	3.473 ppm	0.922	0.9257	0.9832	0.9802
*Die Massenspektren wurden in 3000-facher Verdünnung gemessen.					
Raffinose (19) 504.1690 Da F1-80	ESI-(+)		527.16 ([M+Na] ⁺)	543.13 ([M+K] ⁺)	
	5.435 ppm		0.9807	0.9693	
	4.996 ppm		0.9749	0.9653	
	4.236 ppm		0.9797	0.9771	
	4.086 ppm		0.9503	0.9367	
	ESI-(-)	503.17 ([M-H] ⁻)	549.17 ([M+FA-H] ⁻)	617.16 ([M+TFA-H] ⁻)	
	5.435 ppm	0.9802	0.9837	0.9942	
	4.996 ppm	0.9703	0.9746	0.9839	
	4.236 ppm	0.9755	0.9761	0.9593	
	4.086 ppm	0.9495	0.9548	0.9807	

Tabelle 39: Korrelationskoeffizienten der im polaren Extrakt der Pinienkerne identifizierten Aminosäuren. In der linken Spalte sind die Metaboliten unter Angabe der monoisotopischen Masse sowie des Fraktionsbereiches, über den die Korrelationskoeffizienten berechnet wurden, aufgelistet. In der rechten Spalte sind die Korrelationskoeffizienten zwischen EDCs und EMCs oder zwischen EDCs untereinander angegeben.

Metabolit	Korrelationskoeffizienten			
L-Alanin (10) 89.0477 Da F1-80	ESI-(+)	90.05 ([M+H] ⁺)		
	1.489 ppm	0.9618		
L-Arginin (34) 174.1117 Da F35-80	ESI-(+)	175.12 ([M+H] ⁺)	157.11 ([M-NH ₃ +H] ⁺)	60.05 (protoniertes Guanidin)
	3.246 ppm	0.9789	0.9789	0.9903
	1.931 ppm	0.9817	0.9750	0.9849
	1.654 ppm	0.9724	0.9688	0.9823
L-Asparagin (31) 132.0535 Da F40-50	ESI-(+)	133.06 ([M+H] ⁺)	87.05 ([M-H ₂ O-CO+H] ⁺)	
	2.940 ppm	0.9888	0.9864	
	2.875 ppm	0.9888	0.9862	
L-Asparagin-säure (33) 133.0375 Da F46-66	ESI-(+)	134.04 ([M+H] ⁺)	116.03 ([M-H ₂ O+H] ⁺)	88.04 ([M-H ₂ O-CO+H] ⁺)
	2.799 ppm	0.9948	0.9719	0.9720
	2.704 ppm	0.9948	0.9721	0.9738
L-Glutamin-säure (32) 147.0532 Da F46-80	ESI-(+)	148.06 ([M+H] ⁺)	102.05 ([M-H ₂ O-CO+H] ⁺)	84.04 ([M-2H ₂ O-CO+H] ⁺)
	2.354 ppm	0.9844	0.9812	0.9808
	2.119 ppm	0.9865	0.9829	0.9830
	2.065 ppm	0.9920	0.9882	0.9883
L-Glutamin (30) 146.0691 Da F41-47	ESI-(+)	147.08 ([M+H] ⁺)	101.07 ([M-H ₂ O-CO+H] ⁺)	
	2.464 ppm	0.9411	0.9583	
	2.150 ppm	0.8691	0.8939	

Metabolit	Korrelationskoeffizienten				
L-Histidin (27) 155.0695 Da F40-50	ESI-(+)	156.08 ([M+H] ⁺)	178.06 ([M+Na] ⁺)	110.07 ([M-H ₂ O-CO+H] ⁺)	
	7.769 ppm	0.9325	0.9704	0.9432	
	7.122 ppm	0.9920	0.9626	0.9872	
L-Isoleucin (6) 131.0946 Da F30-37	EDC [ppm]	1.4768	1.9867	1.0097	0.9437
	1.4768	-	0.9977	0.9942	0.9836
	1.9867	0.9977	-	0.9858	0.9721
	1.0097	0.9942	0.9858	-	0.9956
	0.9437	0.9836	0.9721	0.9956	-
L-Leucin (5) 131.0946 Da F30-37	EDC [ppm]	1.7068	0.9662		
	1.7068	-	0.9967		
	0.9662	0.9967	-		
L-Lysin (29) 146.1055 Da F48-62	ESI-(+)	147.11 ([M+H] ⁺)	130.09 ([M-NH ₃ +H] ⁺)	84.08 ([M-NH ₃ -H ₂ O-CO+H] ⁺)	
	3.041 ppm	0.9554	0.9605	0.9711	
	1.734 ppm	0.9138	0.9160	0.9336	
	1.469 ppm	0.8452	0.8574	0.8720	
L-Methionin (23) 149.0510 Da F34-41	ESI-(+)	150.06 ([M+H] ⁺)			
	2.646 ppm	0.9938			
	2.139 ppm	0.9472			
L-Phenylalanin (4) 165.0790 Da F1-41	ESI-(+)	166.09 ([M+H] ⁺)	120.08 ([M-H ₂ O-CO+H] ⁺)		
	7.435 ppm	0.9787	0.9777		
	7.380 ppm	0.9945	0.9930		
	7.340 ppm	0.9962	0.9974		

Metabolit	Korrelationskoeffizienten				
L-Prolin (24) 115.0633 Da F37-43	ESI-(+)	116.07 ([M+H] ⁺)	70.06 ([M-H ₂ O-CO+H] ⁺)		
	4.136 ppm	0.9946	0.9867		
	3.351 ppm	0.9394	0.9537		
	3.413 ppm	0.9490	0.9258		
	2.336 ppm	0.9927	0.9832		
	2.071 ppm	0.9893	0.9909		
	1.991 ppm	0.9834	0.9689		
L-Threonin (28) 119.0582 Da F40-48	ESI-(+)	74.06 ([M-H ₂ O-CO+H] ⁺)			
	1.340 ppm	0.9666			
L-Tryptophan (25) 204.0899 Da F1-39	ESI-(+)	205.10 ([M+H] ⁺)	227.08 ([M+Na] ⁺)	188.07 ([M-NH ₃ +H] ⁺)	146.07
	7.735 ppm	0.9907	0.9822	0.9975	0.9966
	7.540 ppm	0.9915	0.9901	0.9960	0.9957
	7.290 ppm	0.9920	0.9840	0.9980	0.9972
	7.205 ppm	0.9923	0.9821	0.9981	0.9969
L-Tyrosin (26) 181.0739 Da F36-80	ESI-(+)	182.08 ([M+H] ⁺)	165.05 ([M-NH ₃ +H] ⁺)	136.08 ([M-H ₂ O-CO+H] ⁺)	119.05 ([M-H ₂ O-CO-NH ₃ +H] ⁺)
	7.190 ppm	0.9814	0.9951	0.9743	0.9956
	6.910 ppm	0.9825	0.9989	0.9786	0.9956
L-Valin (22) 117.0790 Da F36-41	ESI-(+)	118.09 ([M+H] ⁺)	72.08 ([M-H ₂ O-CO+H] ⁺)		
	2.280 ppm	0.9826	0.8073		
	1.039 ppm	0.9681	0.9596		
	1.000 ppm	0.9838	0.9414		

Metabolit	Korrelationskoeffizienten				
L-Glutathion Disulfid (Glutathion oxidiert) (38) 612.1520 Da F65-72	ESI-(+)		613.15 ([M+H] ⁺)	307.08 ([M+2H] ²⁺)	307.58 ([M+2H] ²⁺)- Isotopenpeak
	3.770 ppm		0.8139	0.8440	0.8452
	2.976 ppm		0.9321	0.9569	0.9584
	2.175 ppm		0.9007	0.9283	0.9295
	ESI-(-)		611.15 ([M-H] ⁻)		
	3.770 ppm		0.8854		
	2.976 ppm		0.9807		
	2.175 ppm		0.9549		

Tabelle 40: Korrelationskoeffizienten der im polaren Extrakt der Pinienkerne identifizierten Betaine. In der linken Spalte sind die Metaboliten unter Angabe der monoisotopischen Masse sowie des Fraktionsbereiches, über den die Korrelationskoeffizienten berechnet wurden, aufgelistet. In der rechten Spalte sind die Korrelationskoeffizienten zwischen EDCs und EMCs angegeben.

Metabolit	Korrelationskoeffizienten					
Betain (39) 117.0790 Da F21-34	ESI-(+)		118.09 ([M+H] ⁺)			
	3.269 ppm		0.9503			
Cholin (8) 104.1070 Da F1-36	ESI-(+)		104.11 ([M] ⁺)			
	4.069 ppm		0.9983			
	3.531 ppm		0.9695			
	3.206 ppm		0.9894			
Ergothionein (42) 229.0885 Da F1-80	ESI-(+)		230.10 ([M+H] ⁺)	252.08 ([M+Na] ⁺)	268.05 ([M+K] ⁺)	274.06 ([M+2Na-H] ⁺)
	6.804 ppm		0.9989	0.9866	0.9541	0.9600
	3.276 ppm		0.9958	0.9834	0.9606	0.9527
	ESI-(-)		228.08 ([M-H] ⁻)			
	6.804 ppm		0.9942			
	3.276 ppm		0.9856			

Metabolit	Korrelationskoeffizienten			
Phosphorylcholin (40) 183.0660 Da F49-80	ESI-(+)	184.07 ([M+H] ⁺)	206.05 ([M+Na] ⁺)	228.04 ([M+2Na-H] ⁺)
	4.176 ppm	0.9792	0.9850	0.9729
	3.601 ppm	0.9754	0.9818	0.9710
	3.226 ppm	0.9874	0.9944	0.9914
Trigonellin (41) 137.0477 Da F27-38	ESI-(+)	138.06 ([M+H] ⁺)		
	9.13 ppm	0.9810		
	8.84 ppm	0.9851		
	8.08 ppm	0.9817		
	4.44 ppm	0.9680		

Tabelle 41: Korrelationskoeffizienten der im polaren Extrakt der Pinienkerne identifizierten Nucleobasen, Nucleoside, Nucleotide und Derivate. In der linken Spalte sind die Metaboliten unter Angabe der monoisotopischen Masse sowie des Fraktionsbereiches, über den die Korrelationskoeffizienten berechnet wurden, aufgelistet. In der rechten Spalte sind die Korrelationskoeffizienten zwischen EDCs und EMCs angegeben.

Metabolit	Korrelationskoeffizienten		
Adenin* (43) 135.0545 Da F1-37	ESI-(+)	136.06 ([M+H] ⁺)	158.04 ([M+Na] ⁺)
	8.254 ppm	0.9506	0.9751
	8.201 ppm	0.9509	0.9805
	ESI-(-)	134.05 ([M-H] ⁻)	
	8.254 ppm	0.9905	
	8.201 ppm	0.9928	
	* Aufgrund von Signalverschiebungen bzgl. der chemischen Verschiebung wurde ein <i>Bucketing</i> mit einer Größe von 0.01 ppm angewendet.		

Metabolit	Korrelationskoeffizienten					
Adenosin (44) 267.0968 Da F1-28	ESI-(+)		268.10 ([M+H] ⁺)	290.09 ([M+Na] ⁺)		
	8.347 ppm		0.9943	0.9992		
	8.267 ppm		0.9750	0.9915		
	6.077 ppm		0.9791	0.9949		
	4.300 ppm		0.9005	0.9367		
	ESI-(-)		266.09 ([M-H] ⁻)	302.07 ([M+Cl] ⁻)		
			312.10 ([M+FA-H] ⁻)			
	8.347 ppm		0.9930	0.9978		
	8.267 ppm		0.9980	0.9990		
6.077 ppm		0.9980	0.9842			
4.300 ppm		0.9583	0.9081			
0.9585						
Adenosin-5'- monophosphat (45) 347.0631 Da F65-75	ESI-(+)		348.07 ([M+H] ⁺)	370.05 ([M+Na] ⁺)	392.03 ([M+2Na-H] ⁺)	136.06 (Adenin Fragment)
	8.602 ppm		0.9738	0.9499	0.9139	0.9742
	8.272 ppm		0.9856	0.9667	0.9298	0.9909
	6.152 ppm		0.9822	0.9628	0.9254	0.9886
	4.511 ppm		0.9841	0.9674	0.9340	0.9888
	4.371 ppm		0.9734	0.9498	0.9122	0.9779
	4.031 ppm		0.9740	0.9577	0.9267	0.9794
	ESI-(-)		346.06 ([M-H] ⁻)			
	8.602 ppm		0.9764			
	8.272 ppm		0.9906			
	6.152 ppm		0.9903			
	4.511 ppm		0.9919			
	4.371 ppm		0.9812			
	4.031 ppm		0.9867			
Cytidin-5'- diphospho- cholin (46) 488.1073 Da F46-54	ESI-(+)		489.11 ([M+H] ⁺)	511.10 ([M+Na] ⁺)	533.08 ([M+2Na-H] ⁺)	
	7.942 ppm		0.8694	0.8828	0.8946	
	6.137 ppm		0.9102	0.9088	0.9120	
	5.997 ppm		0.9544	0.9754	0.9716	

Metabolit	Korrelationskoeffizienten			
Guanosin (47) 283.0917 Da F32-42	ESI-(+)		284.10 ([M+H] ⁺)	306.08 ([M+Na] ⁺)
				152.06 (Guanin Fragment)
	8.005 ppm		0.9518	0.9783
				0.9920
	ESI-(-)		282.09 ([M-H] ⁻)	
	8.005 ppm		0.9822	
Nicotinamid- adenin- dinucleotid (NAD) (49) 663.1091 Da F46-65	ESI-(+)		664.12 ([M+H] ⁺)	332.56 ([M+2H] ²⁺)
	9.343 ppm		0.9380	0.9848
	9.142 ppm		0.9451	0.9878
	8.847 ppm		0.9256	0.9396
	8.432 ppm		0.9501	0.9908
	8.196 ppm		0.9246	0.9680
	8.179 ppm		0.9381	0.9849
	6.122 ppm		0.9541	0.9093
	6.007 ppm		0.9587	0.9296
Uridin (7) 244.0695 Da F28-34	ESI-(+)		267.06 ([M+Na] ⁺)	289.04 ([M+2Na-H] ⁺)
	7.887 ppm		0.9916	0.9982
	5.916 ppm		0.9918	0.9968
	4.360 ppm		0.9382	0.9647
	ESI-(-)		243.07 ([M-H] ⁻)	279.04 ([M+Cl] ⁻)
				289.07 ([M+FA-H] ⁻)
	7.887 ppm		0.9969	0.9909
	5.916 ppm		0.9948	0.9903
	4.360 ppm		0.9653	0.9717
				0.9613

Metabolit	Korrelationskoeffizienten		
Uridin-5'-monophosphat (48) 324.0359 Da F64-77	ESI-(+)	325.04 ([M+H] ⁺)	347.02 ([M+Na] ⁺)
	8.102 ppm	0.9063	0.8965
	5.997 ppm	0.9279	0.8909
	5.982 ppm	0.9246	0.8380
	4.423 ppm	0.9315	0.9003
	ESI-(-)	611.15 ([M-H] ⁻)	
	8.102 ppm	0.9311	
	5.997 ppm	0.9377	
	5.982 ppm	0.9027	
	4.423 ppm	0.9487	

Tabelle 42: Korrelationskoeffizienten der im polaren Extrakt der Pinienkerne identifizierten organischen Säuren. In der linken Spalte sind die Metaboliten unter Angabe der monoisotopischen Masse sowie des Fraktionsbereiches, über den die Korrelationskoeffizienten berechnet wurden, aufgelistet. In der rechten Spalte sind die Korrelationskoeffizienten zwischen EDCs und EMCs angegeben.

Metabolit	Korrelationskoeffizienten			
Äpfelsäure (51) 134.0215 Da F50-67	ESI-(+)	157.01 ([M+Na] ⁺)	172.98 ([M+K] ⁺)	178.99 ([M+2Na-H] ⁺)
	4.289 ppm	0.9837	0.9800	0.9634
	2.659 ppm	0.9834	0.9765	0.9579
	2.384 ppm	0.9780	0.9633	0.9447
	ESI-(-)	133.01 ([M-H] ⁻)	155.00 ([M+Na-2H] ⁻)	115.00 ([M-H ₂ O-H] ⁻)
	4.289 ppm	0.9817	0.9850	0.9794
	2.659 ppm	0.9787	0.9878	0.9829
	2.384 ppm	0.9714	0.9755	0.9868
Fumarsäure (53) 116.0110 Da F46-68	ESI-(+)	99.01 ([M-OH] ⁺)		
	6.522 ppm	0.9741		

Metabolit	Korrelationskoeffizienten			
Nicotinsäure (50) 123.0320 Da F39-47	ESI-(+)	124.04 ([M+H] ⁺)	ESI-(-)	122.03 ([M-H] ⁻)
	8.942 ppm	0.9874	8.942 ppm	0.9820
	8.612 ppm	0.9957	8.612 ppm	0.9639
	8.250 ppm	0.9932	8.250 ppm	0.9552
	7.542 ppm	0.9584	7.542 ppm	0.8968
Zitronensäure* (52) 192.0270 Da F60-80	ESI-(-)	191.02 ([M-H] ⁻)	213.00 ([M+Na-2H] ⁻)	
	2.70 ppm	0.7956	0.7244	
	2.57 ppm	0.9015	0.8798	
	2.54 ppm	0.8805	0.8563	
	*Es wurden <i>Buckets</i> mit einer Größe von 0.03 ppm gewählt, da sich die chemischen Verschiebungen der Signale über die Fraktionen vmtl. aufgrund von pH-Wert-Änderungen verändern.			

8.4.2 Korrelationskoeffizienten zu Metaboliten in Apfelproben

Tabelle 43: Zehn intensivste EMCs im positiven Ionisationsmodus mit einem Korrelationskoeffizienten > 0.95 zu dem EDC bei einer chemischen Verschiebung von 7.439 ppm in den Fraktionen 1-30 nach Fraktionierung der Apfelprobe 20-E-SH-9.

EMC [m/z]	Korrelationskoeffizient	max. Intensität [abs.]
74.10	0.9648	29218
133.06	0.9855	32522
329.16	0.9636	30899
335.11	0.9896	94858
336.11	0.9917	17430
337.12	0.9937	35244
343.06	0.9954	10251
365.12	0.9924	24510
447.22	0.9676	10706
501.37	0.9942	9597

8.5 Zusätzliche NMR-Spektren

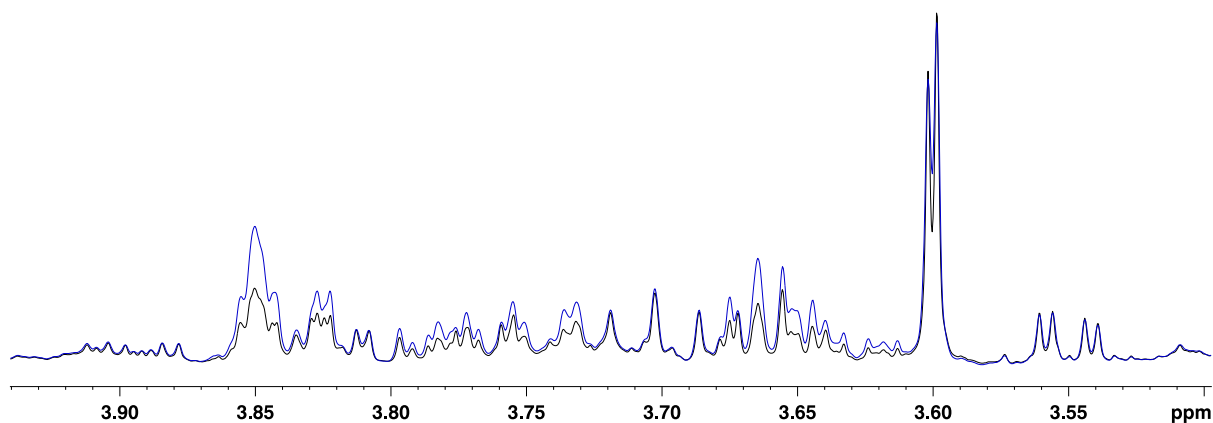


Abbildung 137: 1D-NOESY-Spektrum der Fraktion 36 nach *Spike-In* Experiment mit einer 0.1 M Lösung von Sorbitol (**17**) in D₂O. Es wurden 5 µL der Lösung zugegeben (blaues Spektrum).

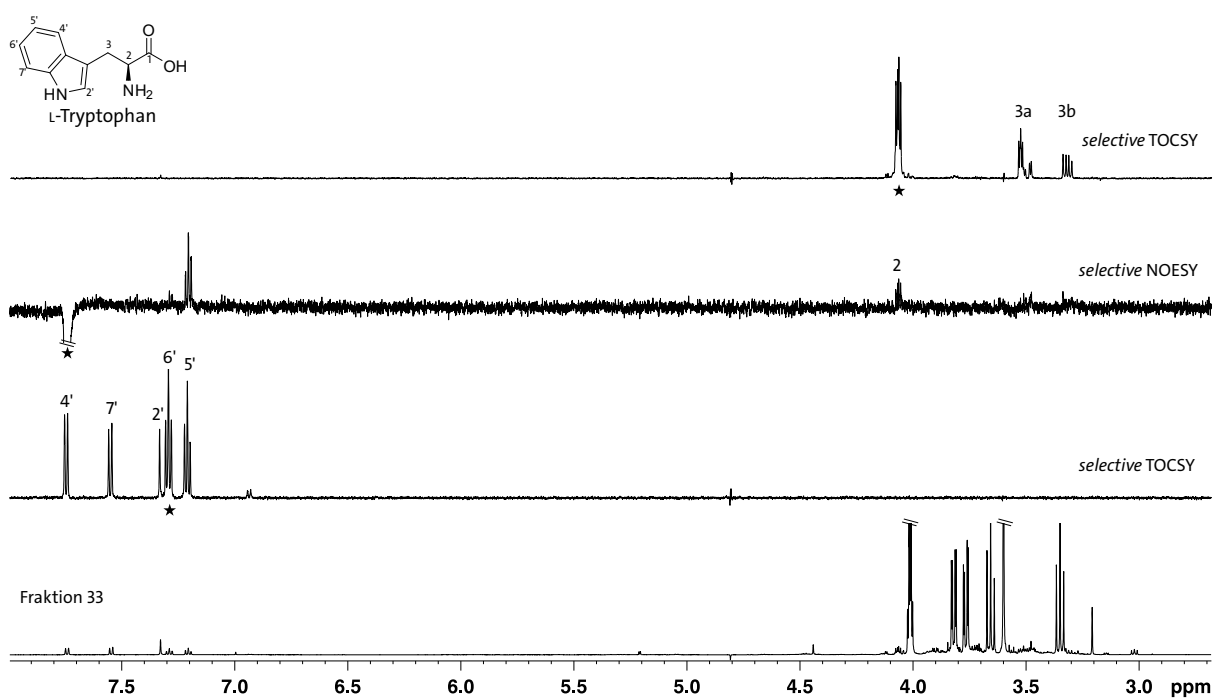


Abbildung 138: ¹H-NMR-Spektrum der Fraktion 33 des polaren Pinienkernextraktes (unten) sowie selektive TOCSY- und NOESY-NMR-Spektren. Es wurde jeweils auf das mit einem Sternchen markierte Signal eingestrahlt. Dies entspricht dem H-6' bei 7.29 ppm im *selective* TOCSY (zweites Spektrum von unten) sowie dem H-4' bei 7.74 ppm im *selective* NOESY (drittes Spektrum von unten) und dem H-2 bei 4.07 ppm im obersten *selective* TOCSY-Spektrum. Die Spektren zeigen die Signale der Aminosäure Tryptophan (**25**).

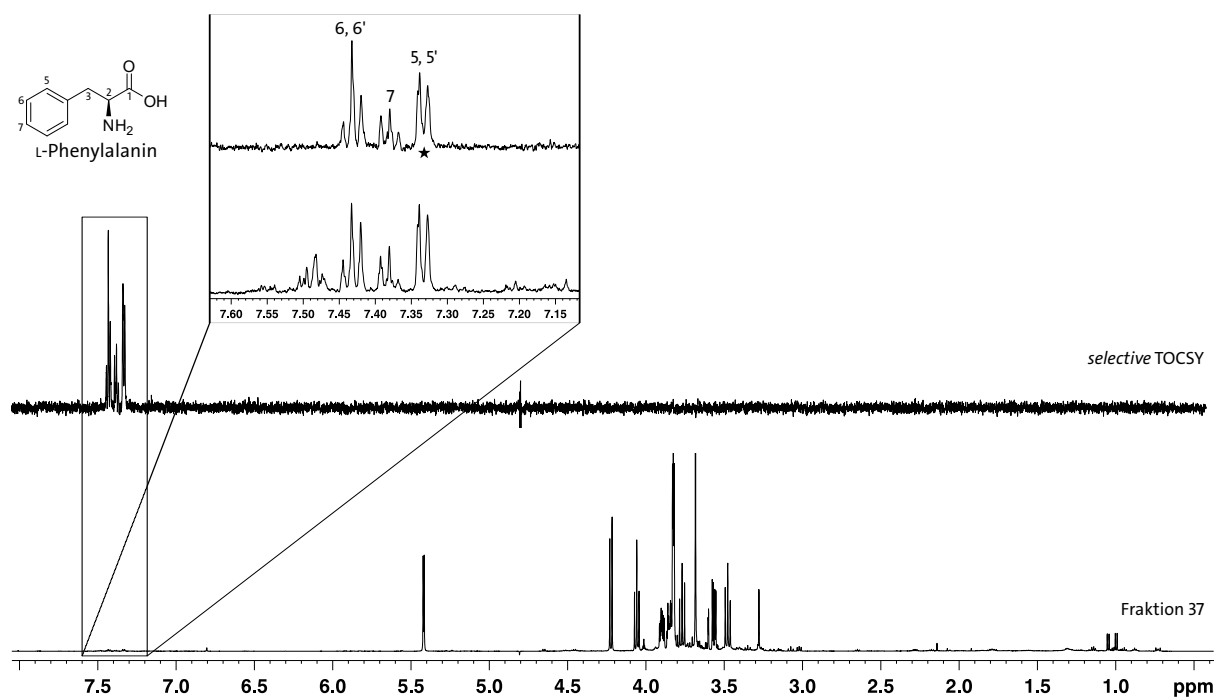


Abbildung 139: ¹H-NMR-Spektrum der Fraktion 37 des polaren Pinienkernextraktes (unten) sowie selektives TOCSY-Experiment (oben). Es wurde auf das mit einem Sternchen markierte Signal bei 7.34 ppm eingestrahlt. Das Spektrum zeigt die aromatischen Protonen der Aminosäure Phenylalanin (**4**).

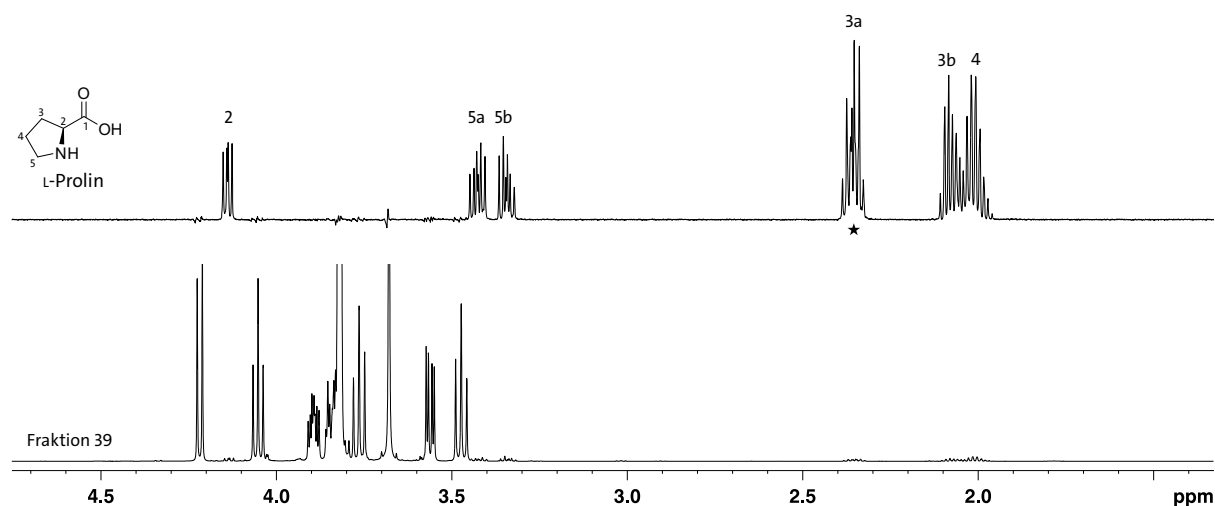


Abbildung 140: ¹H-NMR-Spektrum der Fraktion 39 des polaren Pinienkernextraktes (unten) sowie selektives TOCSY-Experiment (oben). Es wurde auf das mit einem Sternchen markierte Signal bei 2.35 ppm eingestrahlt. Das Spektrum zeigt die Aminosäure Prolin (**24**).

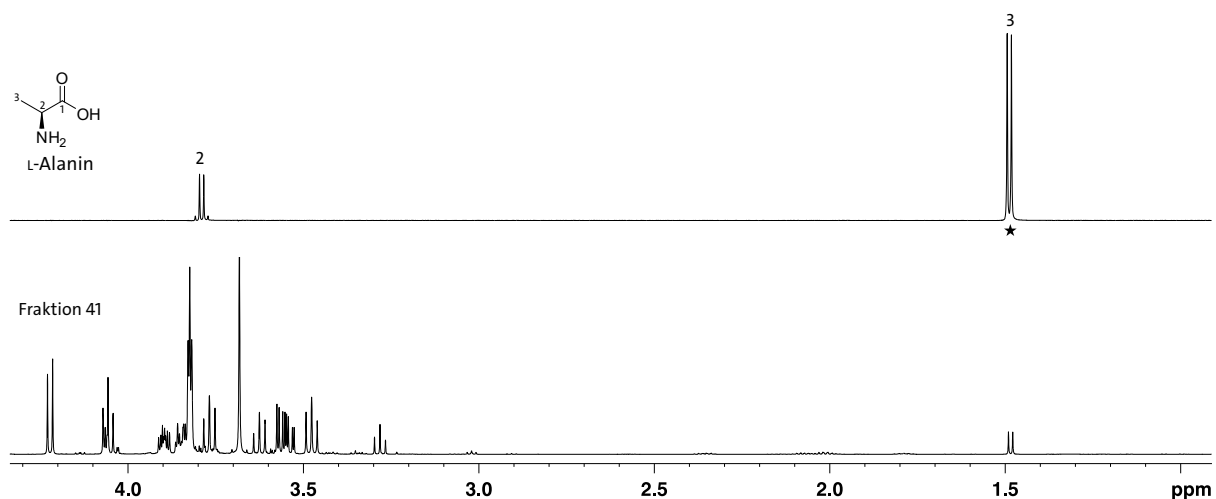


Abbildung 141: ¹H-NMR-Spektrum der Fraktion 41 des polaren Pinienkernextraktes (unten) sowie selektives TOCSY-Experiment (oben). Es wurde auf das mit einem Sternchen markierte Signale bei 1.49 ppm eingestrahlt. Das Spektrum zeigt die Aminosäure Alanin (**10**).

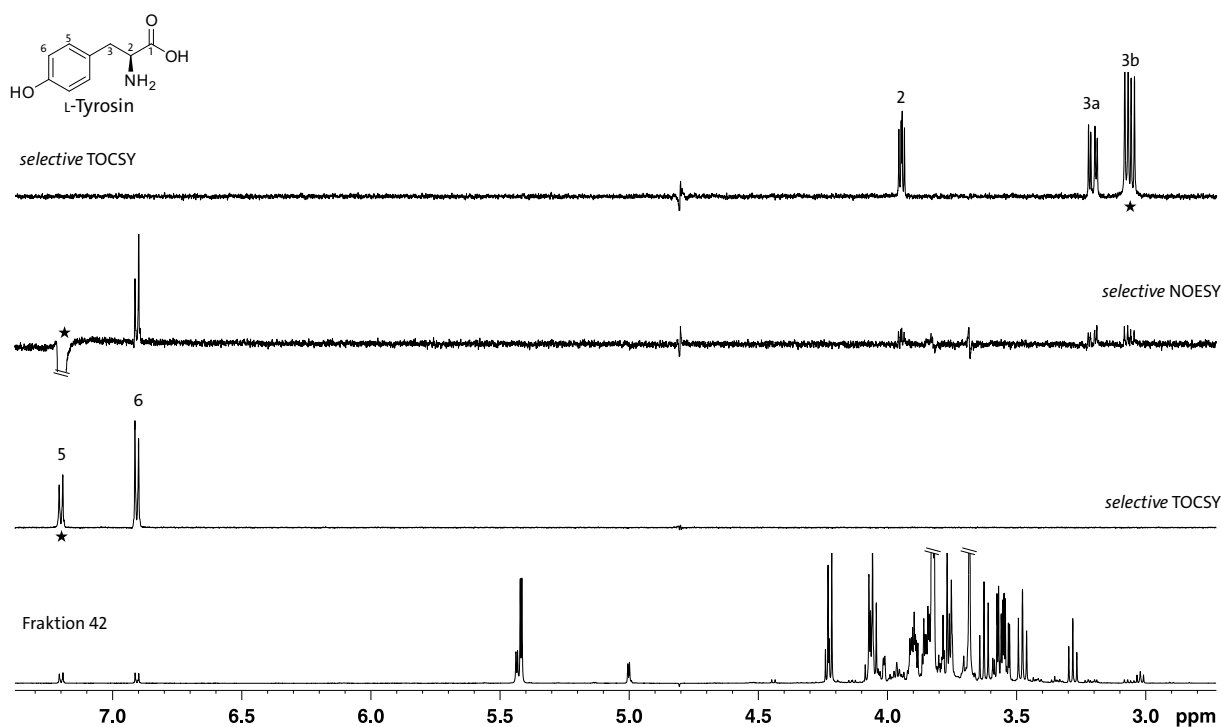


Abbildung 142: ¹H-NMR-Spektrum der Fraktion 42 des polaren Pinienkernextraktes (unten) sowie selektive TOCSY- und NOESY-Spektren (oben). Es wurde jeweils auf das mit einem Sternchen markierte Signale eingestrahlt. Dies entspricht dem H-5 bei 7.20 ppm im *selective* TOCSY und *selective* NOESY (zweites und drittes Spektrum von unten) sowie dem H-3b bei 3.06 ppm im obersten *selective* TOCSY. Die Spektren zeigen die Signale der Aminosäure Tyrosin (**26**).

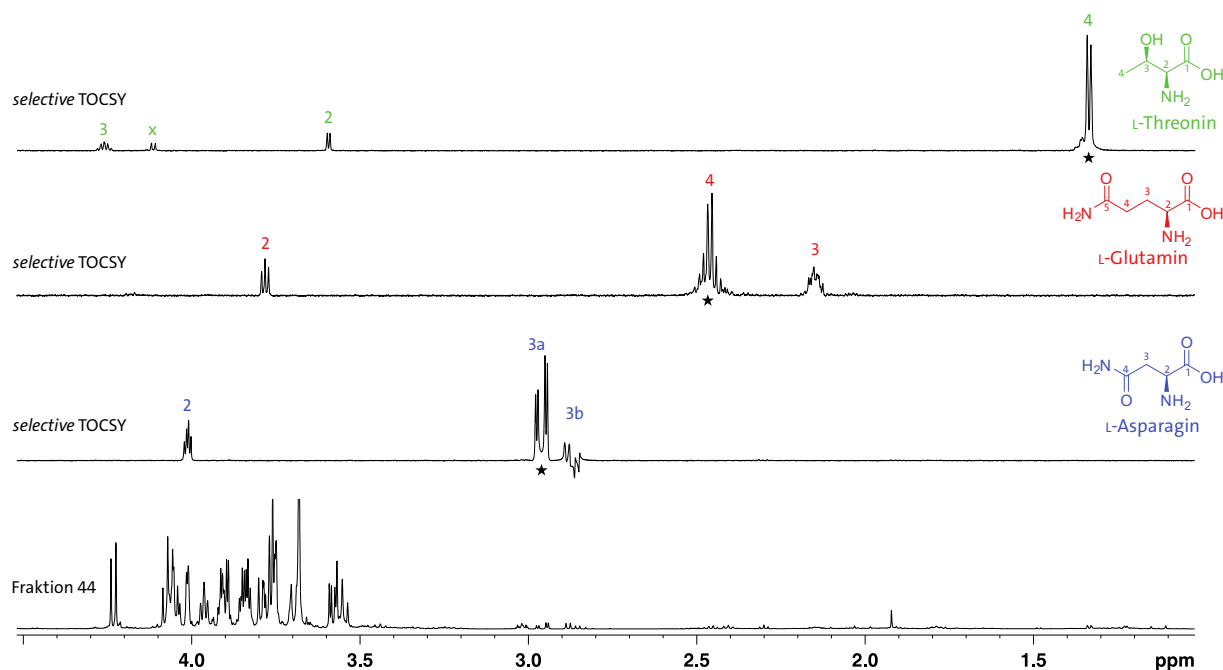


Abbildung 143: ^1H -NMR-Spektrum der Fraktion 44 des polaren Pinienkernextraktes (unten) sowie selektive TOCSY-Spektren (oben). Es wurde jeweils auf das mit einem Sternchen markierte Signale eingestrahlt. Die Spektren zeigen die Signale der Aminosäuren Asparagin (31), Glutamin (30) und Threonin (28). Das mit einem „x“ gekennzeichnete Signal bei einer chemischen Verschiebung von 4.11 ppm im selektiven TOCSY-Spektrum von Threonin (28) stammt vom α -Proton von Lactat (54). Dieses tritt im selektiven TOCSY-Spektrum auf, da das Dublett der Methylgruppe von Lactat (54) mit dem Dublett von Threonin (28) bei 1.34 ppm teilweise überlagert und deshalb ebenfalls angeregt wurde.

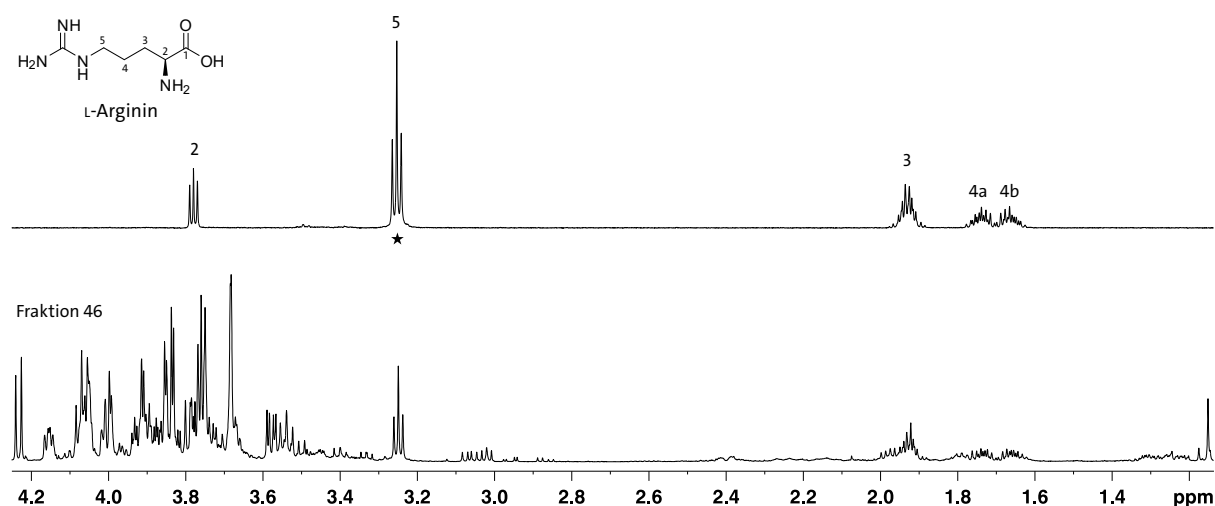


Abbildung 144: ^1H -NMR-Spektrum der Fraktion 46 des polaren Pinienkernextraktes (unten) sowie selektives TOCSY-Experiment (oben). Es wurde auf das mit einem Sternchen markierte Signale bei 3.25 ppm eingestrahlt. Das Spektrum zeigt die Aminosäure Arginin (34).

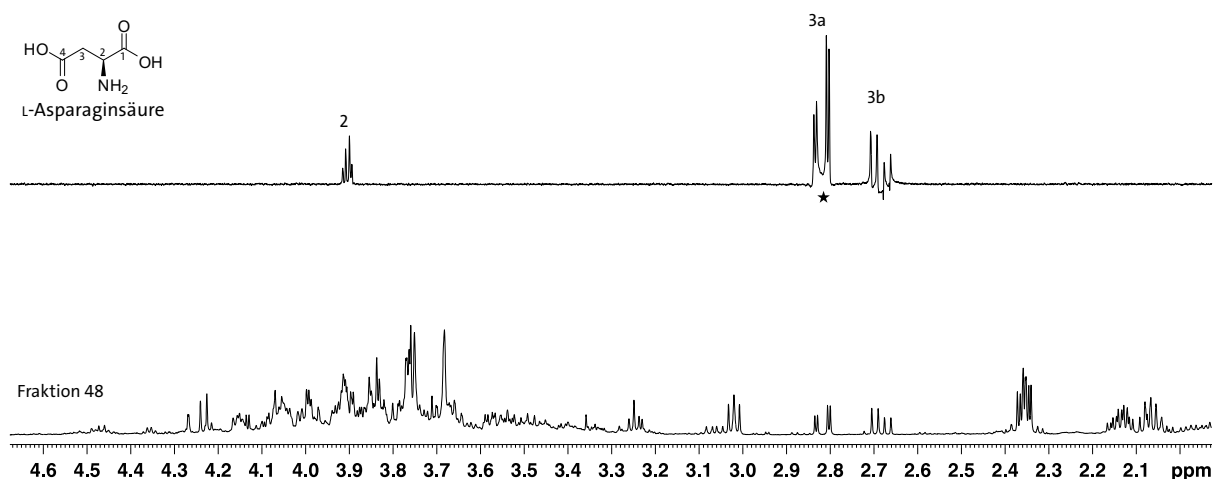


Abbildung 145: ¹H-NMR-Spektrum der Fraktion 48 des polaren Pinienkernextraktes (unten) sowie selektives TOCSY-Experiment (oben). Es wurde auf das mit einem Sternchen markierte Signale bei 2.82 ppm eingestrahlt. Das Spektrum zeigt die Aminosäure Asparaginsäure (33).

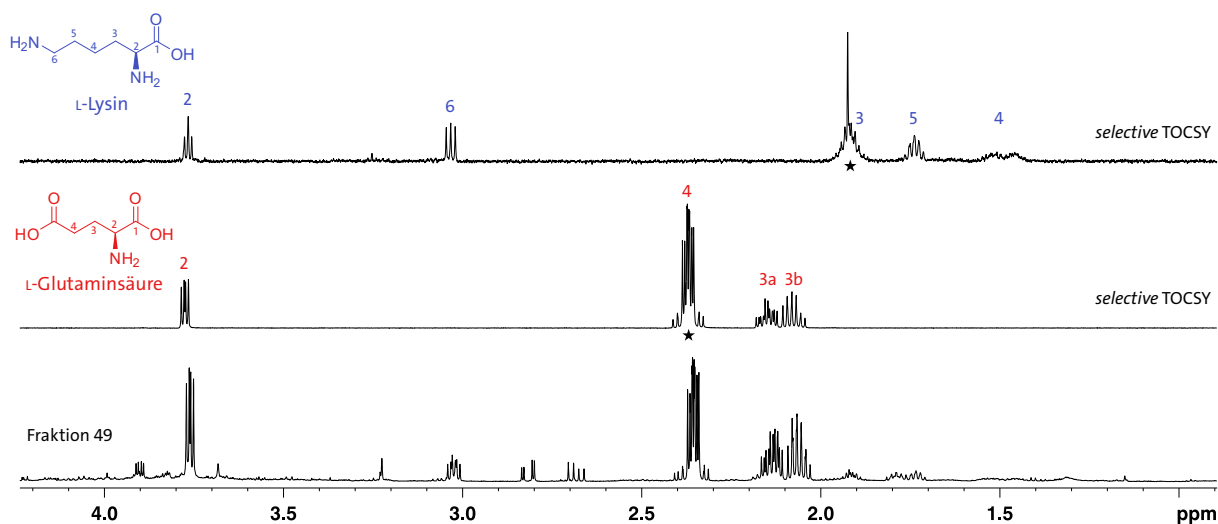


Abbildung 146: ¹H-NMR-Spektrum der Fraktion 49 des polaren Pinienkernextraktes (unten) sowie selektive TOCSY- und NOESY-Spektren (oben). Es wurde jeweils auf das mit einem Sternchen markierte Signale eingestrahlt. Dies entspricht dem H-4 bei 2.36 ppm (zweites Spektrum von unten) sowie dem H-3 bei 1.92 ppm im oberen *selective* TOCSY. Die Spektren zeigen die Signale der Aminosäure Glutaminsäure (32) und Lysin (29).

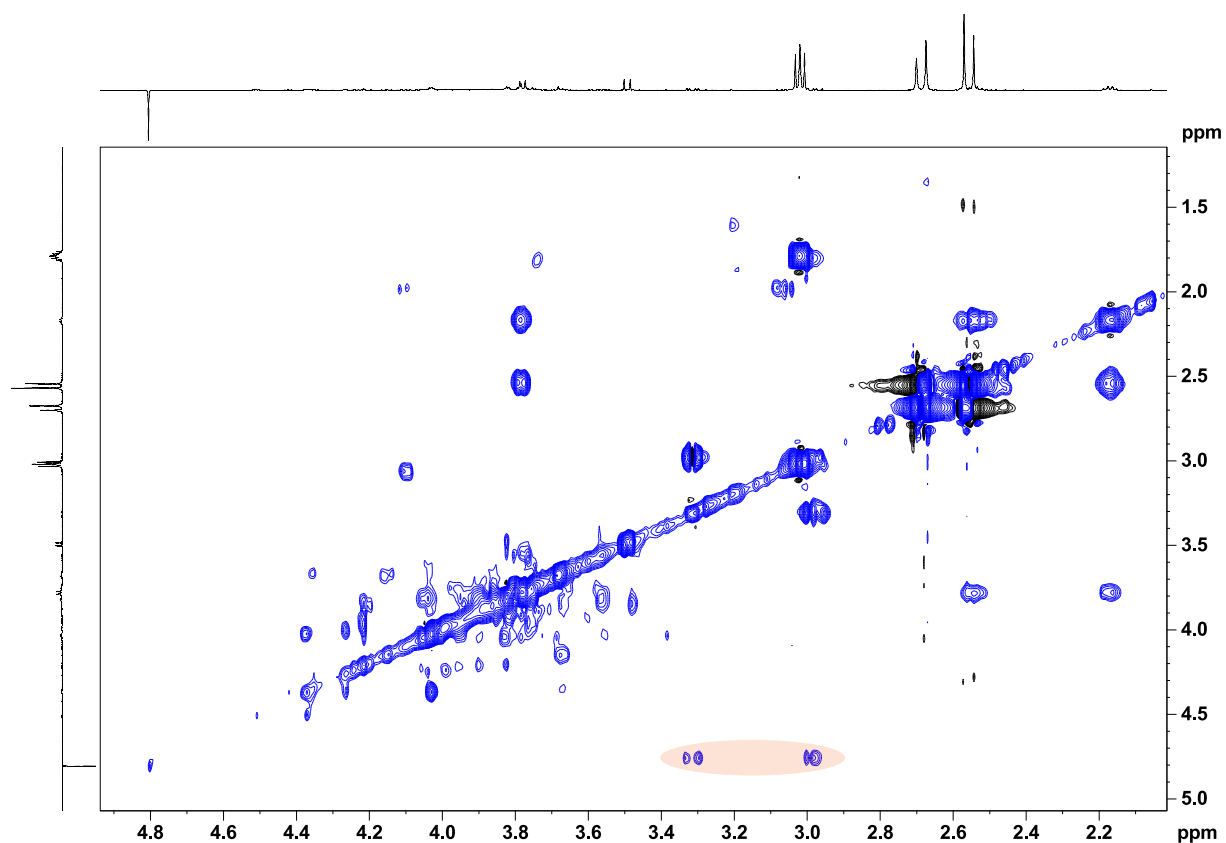


Abbildung 147: 2D-TOCSY-Spektrum der Fraktion 71 des polaren Pinienkernextraktes. Die Kopplung der diastereotopen Protonen 3'a und 3'b des Cystein-Spinsystems im Glutathion Disulfid (**38**) zu dem α -Proton des Cystein-Spinsystems ist farblich hervorgehoben.

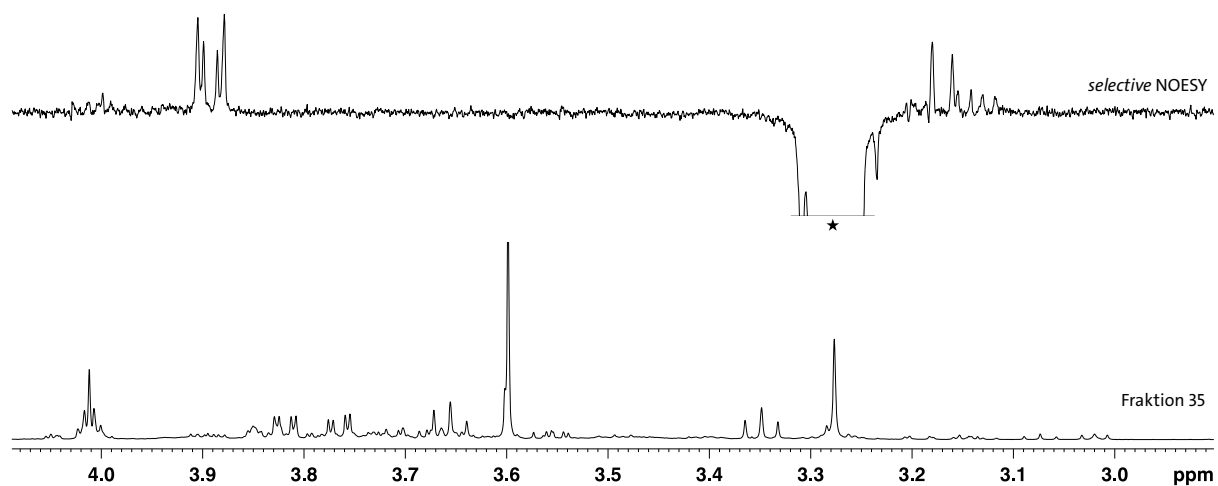


Abbildung 148: ^1H -NMR-Spektrum der Fraktion 35 des polaren Pinienkernextraktes (unten) und selektives NOESY-Spektrum (oben). Es wurde selektive auf das mit einem Sternchen markierte Signal bei 3.276 ppm eingestrahlt. Die Signale gehören zu dem Metaboliten Ergothionein (**42**).

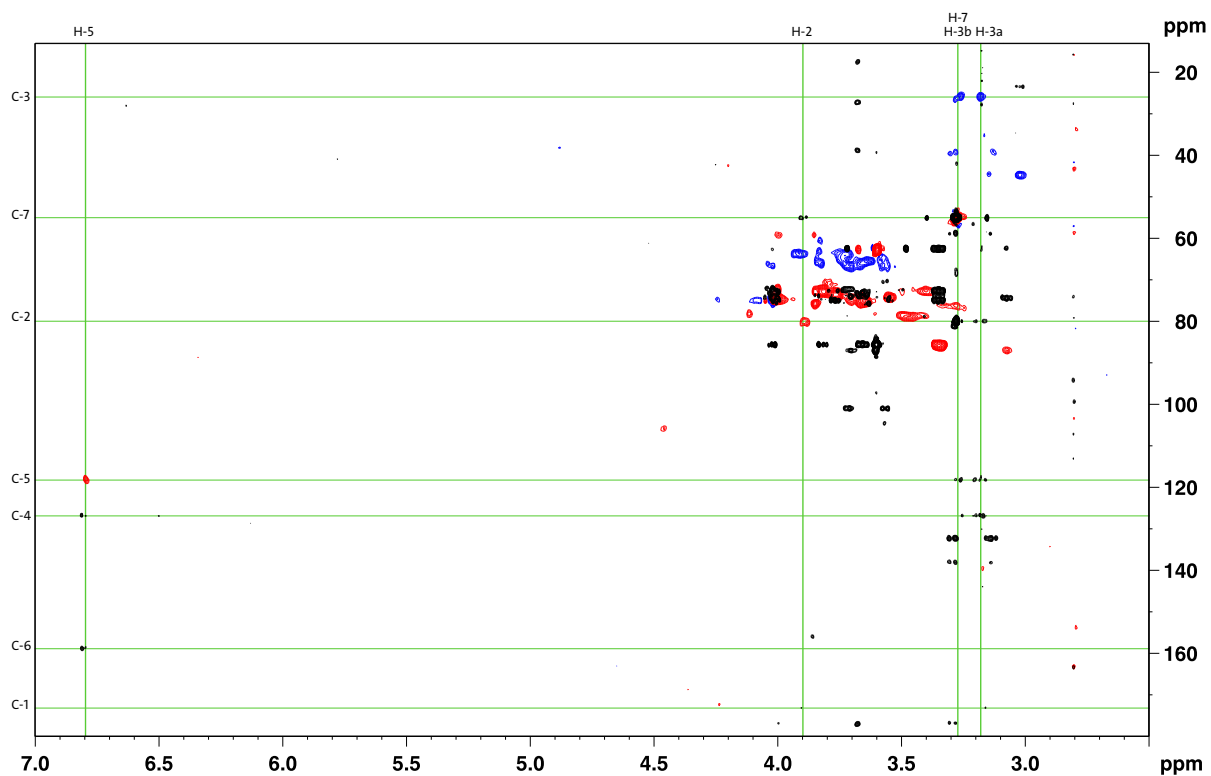


Abbildung 149: 2D HSQC- (rote und blaue Signale) und HMBC-Spektrum (schwarze Signale) der Fraktion 35 des polaren Pinienkernextraktes, das den Metaboliten Ergothionein (**42**) enthält. Die grünen Linien implizieren die chemischen Verschiebungen der Signale dieses Metaboliten sowie deren Kopplungen.

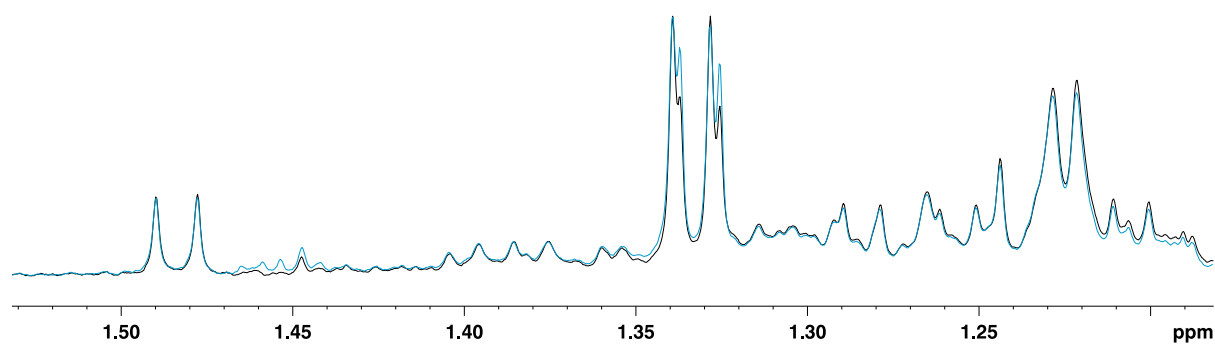


Abbildung 150: 1D-NOESY-Spektrum der Fraktion 44 nach *Spike-In* Experiment mit einer 0.1M Lösung von Milchsäure (**54**) in D₂O. Es wurden 1 μ L der Lösung zugegeben (blaues Spektrum).

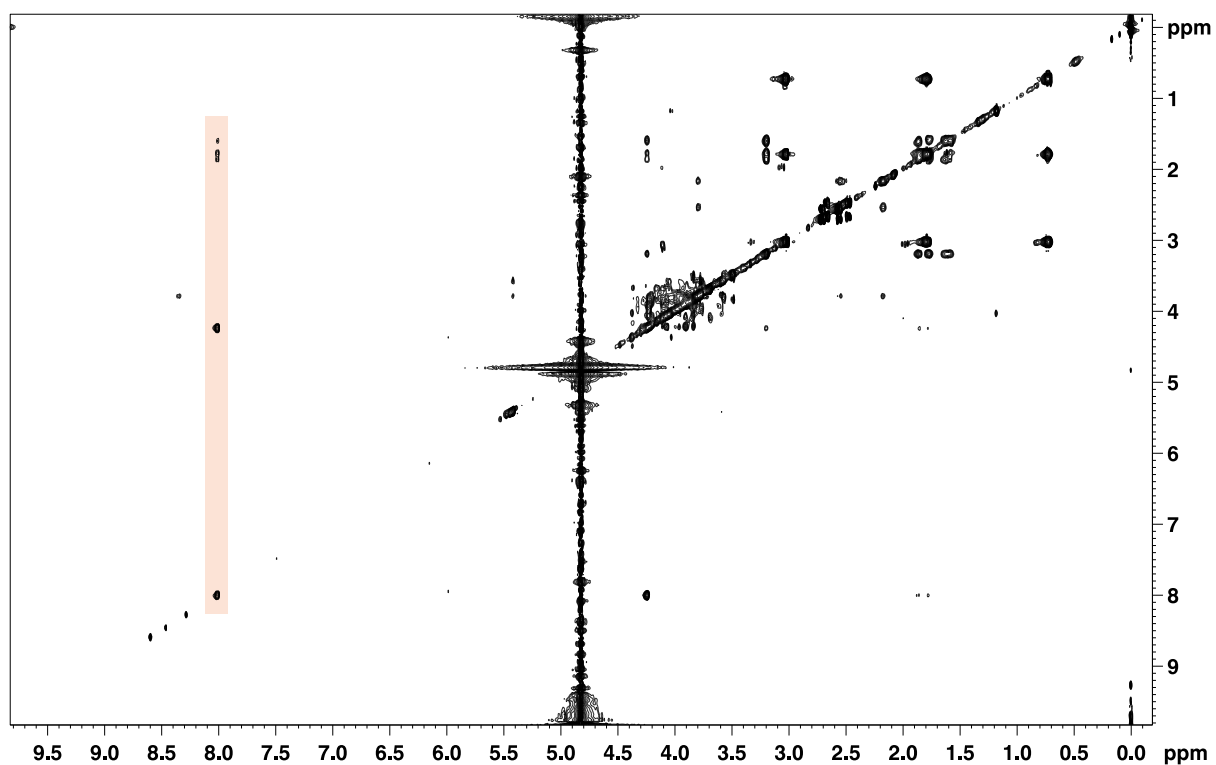


Abbildung 151: 2D-TOCSY-Spektrum der Fraktion 67 des polaren Pinienkernextraktes, das den Metaboliten N^{α} -(2-Hydroxy-2-carboxymethylsuccinyl)-L-arginin (**57**) enthält, in $\text{H}_2\text{O}/\text{D}_2\text{O}$ (Verhältnis 9:1). Das Amid-Proton bei einer chemischen Verschiebung von 8.01 ppm weist eine Kopplung zu den Protonen H-2' bei 4.24 ppm sowie zu den Protonen H-3'a/b (1.76-1.90 ppm) und H-4'a/b (1.50-1.65 ppm) auf. Diese Kopplungen sind farblich hervorgehoben.

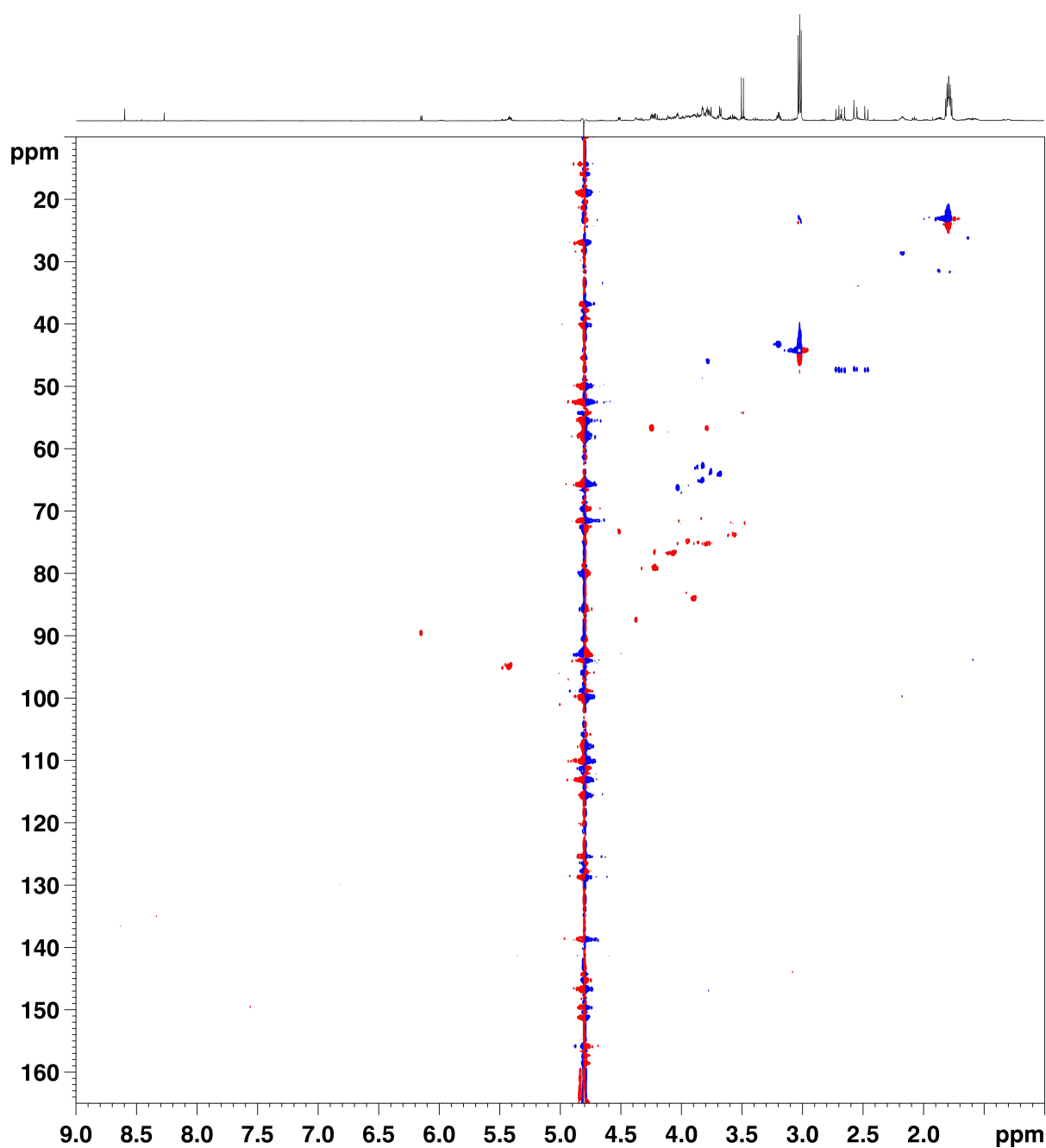


Abbildung 152: HSQC-Spektrum der Fraktion 68 des polaren Pinienkernextraktes, die den Metaboliten N^{α} -(2-Hydroxy-2-carboxymethylsuccinyl)-L-arginin (**57**) enthält.

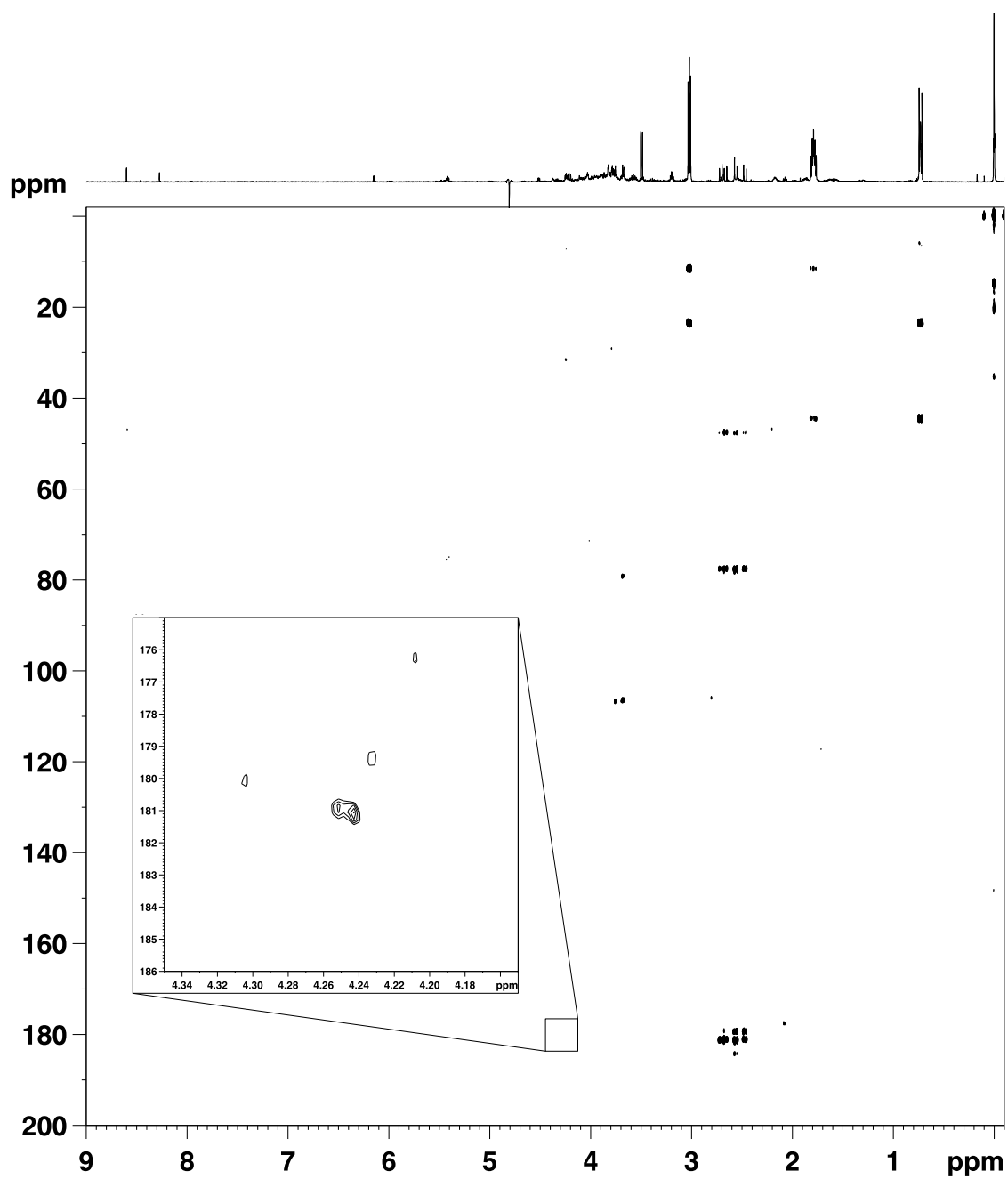


Abbildung 153: HMBC-Spektrum der Fraktion 68 des polaren Pinienkernextraktes, die den Metaboliten N^{α} -(2-Hydroxy-2-carboxymethylsuccinyl)-L-arginin (**57**) enthält. Das vergrößerte Signal zeigt die Kopplung von H-2' zu C-1' mit einem Signal-zu-Rausch-Verhältnis von 3.3.

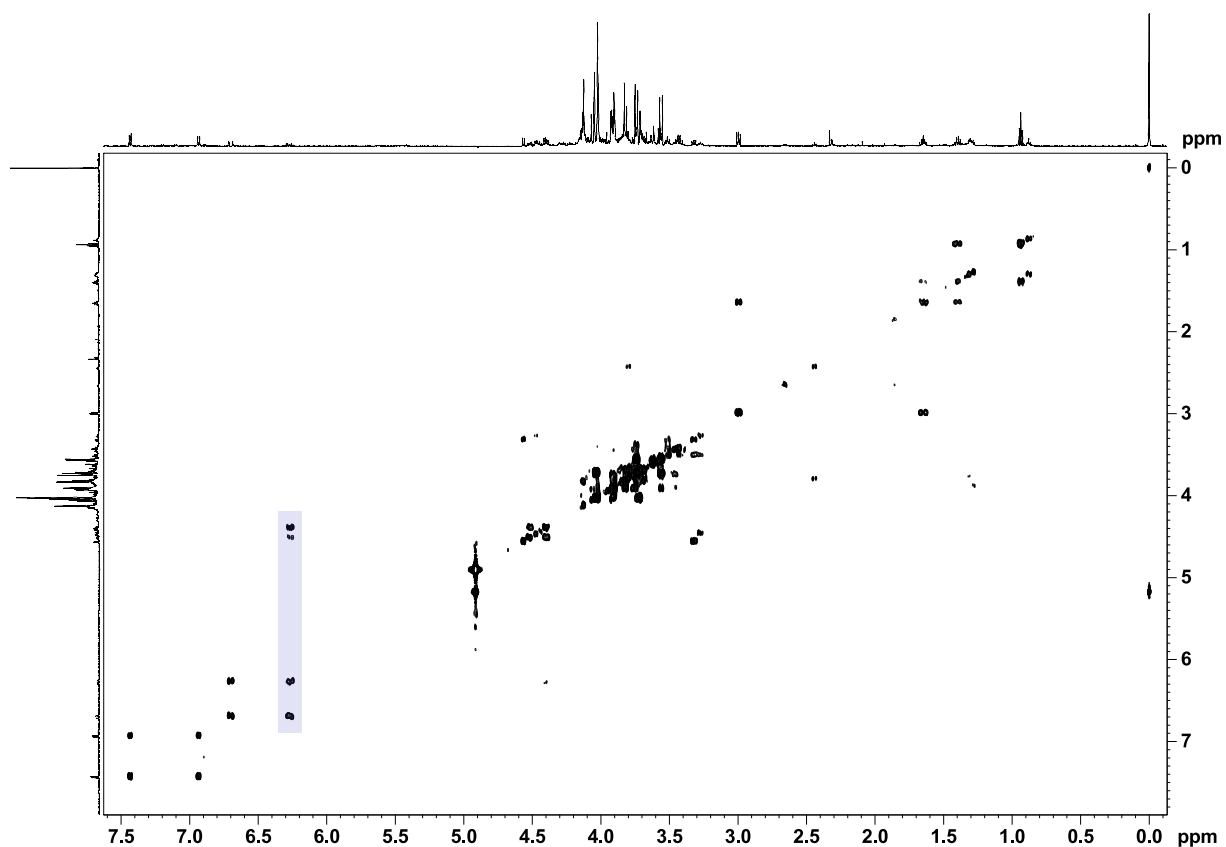


Abbildung 154: 2D-COSY-Spektrum der Fraktion 28 der Apfelprobe 20-E-SH-9, in der der Metabolit Triandrin (**66**) enthalten ist. Die Kopplung des Protons H-8 (6.27 ppm) zu den Protonen H-9a (4.52 ppm), H-9b (4.41 ppm) und H-7 (6.70 ppm) ist farblich hervorgehoben.

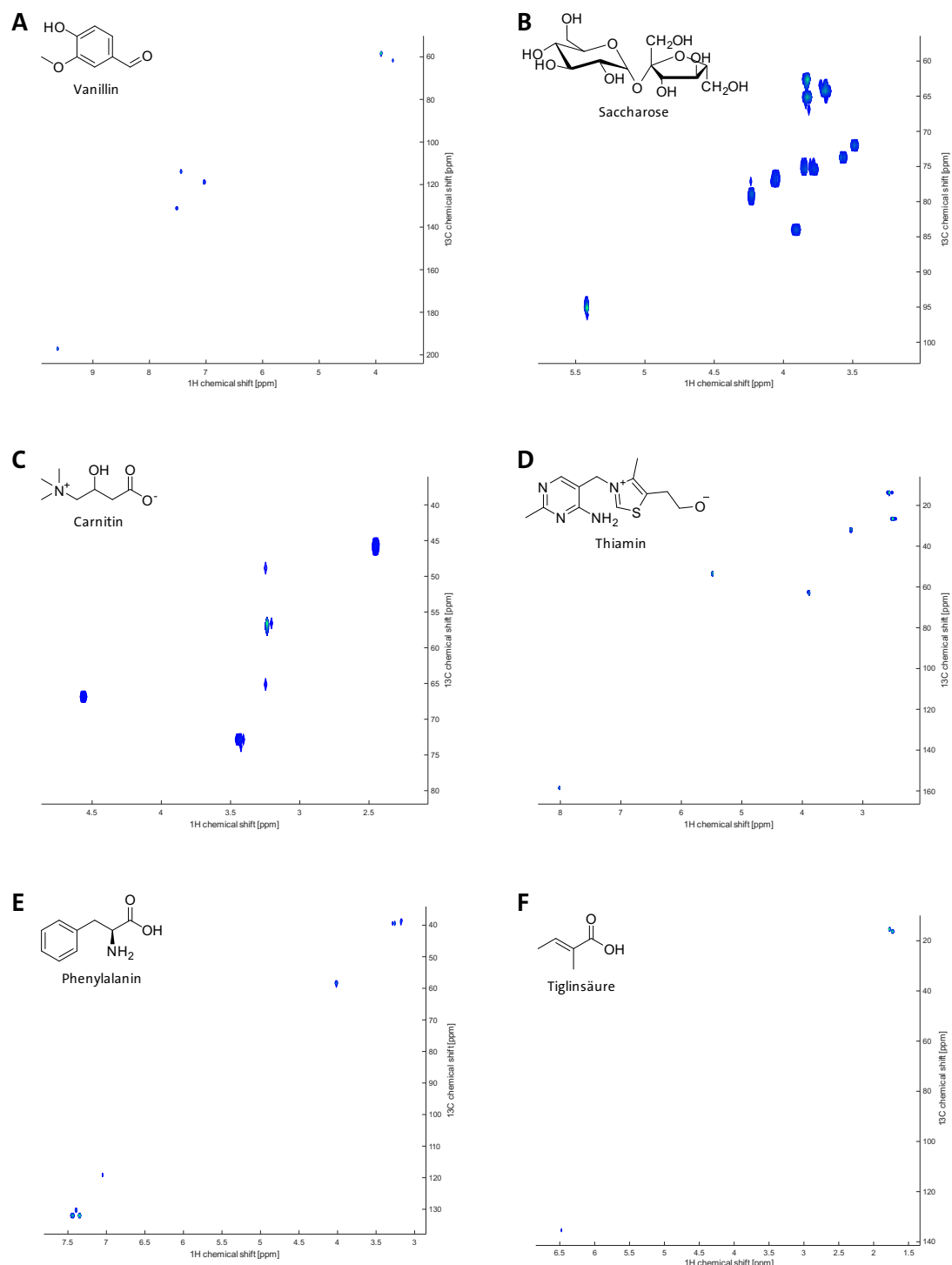


Abbildung 155: Über die semi-automatische Korrelationsanalyse berechnete HSQC-Spektren verschiedener Analyten des artifiziiellen Gemisches B. Für alle Analyten wurde die Korrelation über alle 30 Fraktionen durchgeführt. Die Spektren **A** bis **E** entsprechen den EDDCs, die einen Korrelationskoeffizienten > 0.95 (bzw. > 0.90 für Vanillin (**87**)) zu einem EMC des Analyten im positiven Ionisationsmodus aufweisen ($m/z = 153.05$ für Vanillin (**87**) ($[M+H]^+$), 366.11 für Saccharose (**2**) ($[M+Na]^+$ -Isotopenpeak), 162.11 für Carnitin (**94**) ($[M+H]^+$), 133.06 für Thiamin (**95**) ($[M+2H]^{2+}$) und 166.09 für Phenylalanin (**4**) ($[M+H]^+$). Da Tiglinsäure (**90**) in den Massenspektren nicht detektierbar ist, zeigt das Spektrum **F** diejenigen EDDCs, die einen Korrelationskoeffizienten > 0.95 zu dem EDDC bei $1.732/16.28$ ppm aufweisen.

8.6 Skripte

8.6.1 Skript zu *SCORE-metabolite-ID*

Im Folgenden ist das Skript der MATLAB Applikation *SCORE-metabolite-ID* angegeben, das unter Verwendung des AppDesigners in MATLAB R2020b entwickelt wurde.

```
classdef SCOREmetaboliteID < matlab.apps.AppBase

% Properties that correspond to app components
properties (Access = public)
    SCOREmetaboliteIDApp UIFigure matlab.ui.Figure
    GridLayout matlab.ui.container.GridLayout
    LeftPanel matlab.ui.container.Panel
    ChangeTheBucketSizeButton matlab.ui.control.Button
    LengthMSEditField matlab.ui.control.EditField
    SmoothingLengthLabel_2 matlab.ui.control.Label
    SwitchSmoothMSData matlab.ui.control.Switch
    LengthNMREditField matlab.ui.control.EditField
    SmoothingLengthLabel matlab.ui.control.Label
    SwitchSmoothNMRdata matlab.ui.control.Switch
    SmoothTheMSdataButton matlab.ui.control.Button
    SmoothTheNMRdataButton matlab.ui.control.Button
    NormalizeMSSSpectrumButton matlab.ui.control.Button
    BucketSizeEditField matlab.ui.control.EditField
    BucketSizeEditFieldLabel matlab.ui.control.Label
    ESionizationModeButtonGroup matlab.ui.container.ButtonGroup
    negativeButton matlab.ui.control.RadioButton
    positiveButton matlab.ui.control.RadioButton
    MSSSpectrumLabel matlab.ui.control.Label
    NMRSpectrumLabel matlab.ui.control.Label
    mzvalueofEMCEditFieldLabel matlab.ui.control.Label
    mzvalueofEMCEditField matlab.ui.control.NumericEditField
    chemicalshiftofEDCEditFieldLabel matlab.ui.control.Label
    chemicalshiftofEDCEditField matlab.ui.control.NumericEditField
    ImportNMRSpectraButton matlab.ui.control.Button
    CalibrateNMRSpectraButton matlab.ui.control.Button
    ImportMSSSpectraButton matlab.ui.control.Button
    NormalizeNMRSpectraButton matlab.ui.control.Button
    SubtractBlankMSSSpectrumButton matlab.ui.control.Button
    ShowEDCButton matlab.ui.control.Button
    ShowEMCButton matlab.ui.control.Button
    BucketingofNMRSpectraButton matlab.ui.control.Button
    CenterPanel matlab.ui.container.Panel
    uppermzEditField matlab.ui.control.NumericEditField
    lowermzEditField matlab.ui.control.NumericEditField
    mzregionLabel matlab.ui.control.Label
    mzLabel matlab.ui.control.Label
    ppmLabel matlab.ui.control.Label
    nmrupperlimitEditField matlab.ui.control.NumericEditField
    nmrlowerlimitEditField matlab.ui.control.NumericEditField
    chemicalshiftregionLabel matlab.ui.control.Label
    FractionstoplotListBoxLabel matlab.ui.control.Label
    FractionstoplotListBox matlab.ui.control.ListBox
    EditField_2Label matlab.ui.control.Label
    EditFieldLabel matlab.ui.control.Label
    NMRSpectra matlab.ui.control.UIAxes
    MSSSpectra matlab.ui.control.UIAxes
    RightPanel matlab.ui.container.Panel
    TypeofcorrelationLabel matlab.ui.control.Label
    ButtonGroup matlab.ui.container.ButtonGroup
    SpearmanButton matlab.ui.control.RadioButton
    KendallButton matlab.ui.control.RadioButton
    PearsonButton matlab.ui.control.RadioButton
    MSMSButton matlab.ui.control.Button
    CorrelationCoefficientsLabel matlab.ui.control.Label
    Label matlab.ui.control.Label
    NMRNMRButton matlab.ui.control.Button
    TabGroup matlab.ui.container.TabGroup
    NMRSSTab matlab.ui.container.Tab
    TabGroup_MS_NMR matlab.ui.container.TabGroup
    CoefficientsTab matlab.ui.container.Tab
    PearsonTableMS_NMR matlab.ui.container.Tab
    p_values matlab.ui.container.Tab
    p_valuesTable_MS_NMR matlab.ui.container.Tab
    NMRNMRTab matlab.ui.container.Tab
    TabGroup_NMR_NMR matlab.ui.container.TabGroup
    CoefficientsTab matlab.ui.container.Tab
    PearsonTableNMR_NMR matlab.ui.container.Tab
    pValuesTab matlab.ui.container.Tab
    p_valuesTable_NMR_NMR matlab.ui.container.Tab
    MSMSSTab matlab.ui.container.Tab
    TabGroup_MS_MS matlab.ui.container.TabGroup
    CoefficientsTab_2 matlab.ui.container.Tab
    PearsonTableMS_MS matlab.ui.container.Tab

    pValuesTab_2 matlab.ui.container.Tab
    p_valuesTable_MS_MS matlab.ui.control.Table
    ShowallcorrelatedEMCsstoEMCButton matlab.ui.control.Button
    ShowallcorrelatedEDCsstoEDCButton matlab.ui.control.Button
    EMCSLabel matlab.ui.control.Label
    EDCsLabel matlab.ui.control.Label
    ShowallcorrelatedEMCsButton matlab.ui.control.Button
    ShowallcorrelatedEDCsButton matlab.ui.control.Button
    ClearPlotButton matlab.ui.control.Button
    PlotEMCButton matlab.ui.control.Button
    PlotEDCButton matlab.ui.control.Button
    NMRMSButton matlab.ui.control.Button
    AddEMCButton matlab.ui.control.Button
    AddEDCButton matlab.ui.control.Button
    DeleteEMCButton matlab.ui.control.Button
    ListBoxEMCPlot matlab.ui.control.ListBox
    DeleteEDCButton matlab.ui.control.Button
    ListBoxEDCPlot matlab.ui.control.ListBox
    EDCandEMCPlot matlab.ui.control.UIAxes
    ContextMenu_NMR_MS matlab.ui.container.ContextMenu
    Copydatatoworkspace_NMRMS matlab.ui.container.Menu
    ContextMenu_NMR_NMR matlab.ui.container.ContextMenu
    Copydatatoworkspace_NMRNMR matlab.ui.container.Menu
    ContextMenu_CopyEDCandEMCPlot matlab.ui.container.ContextMenu
    CopyEDCandEMCPlot matlab.ui.container.Menu
    CreatePlotofEDCandEMCMenu matlab.ui.container.Menu
    NormalizeEDCIntensitiesinPlotMenu matlab.ui.container.Menu
    NormalizeEMCIntensitiesinPlotMenu matlab.ui.container.Menu
    ContextMenu_CopyMSSSpectra matlab.ui.container.ContextMenu
    CopyPlotofMSSSpectraMenu matlab.ui.container.Menu
    ContextMenu_CopyNMRSpectra matlab.ui.container.ContextMenu
    CopyPlotofNMRSpectraMenu matlab.ui.container.Menu
    ContextMenu_MS_MS matlab.ui.container.ContextMenu
    CopydatatoworkspaceMenu_MS_MS matlab.ui.container.Menu
end

% Properties that correspond to apps with auto-reflow
properties (Access = private)
    onePanelWidth = 576;
    twoPanelWidth = 768;
end

properties (Access = private)
    nmr % intensities of NMR spectra, one nmr spectra per
    column
    ppm % column vector of ppm values of NMR spectra
    Datatypeselection % data type of NMR data according to
    selection
    CAL %chemical shift value [ppm] of calibration standard
    nmldata % data structure of NMR spectra after import
    via rbnmr
    fractions % row vector from 1 to total number of
    fractions
    indexplotnmr % required for NMR plot
    XAxis % ppm values for all NMR spectra each (matrix)
    mz % m/z values after import as txt file (not sorted)
    intensity % intensity values belonging to m/z values
    indexplotms % required for plot of MS spectra
    mzall % column vector with all different m/z values
    intensityall % intensities of MS spectra sorted (one
    m/z value per row)
    intensitysub % intensities of MS spectra after blank
    subtraction (one m/z value per row)
    fractionslistbox %fractions to plot
    EDCvalue % Value of EDC in Edit Field Box for Plot
    ZeileEDCvalue % Zeilennummer, in der der nächste ppm-
    Wert bezogen auf eingegebenen EDC-Wert steht
    EMCvalue % eingegebener m/z-Wert in EMC-EditField Box
    für Plot
    ZeileEMCvalue % Zeilennummer, in der der nächste m/z-
    Wert bezogen auf eingegebenen EMC-Wert steht
    nmrbuckets % NMR-Daten (Intensitätswerte) nach
    Bucketing
    ppmbuckets % ppm-Skala nach Bucketing
end
```

```

        nmrlimit % lower ppm value (1,1) and upper limit (1,2)
    for NMR plot in specific ppm range
        mzlimit % lower m/z value (1,1) and upper limit (1,2)
    for MS plot in specific m/z range
        wholeEMC % EMC in "larger" m/z value field, from
        ZeileEMCvalue-1 to ZeileEMCvalue+1 (summed intensities)
        nmrbucketsOriginal
        intensitysubOriginal
        ppmOriginal
        nmrOriginal
        window_length
    end

% Callbacks that handle component events
methods (Access = private)
% Button pushed function: ImportNMRsSpectraButton
function ImportNMRsSpectraButtonPushed(app, event)
    title = 'NMR data format';
    msg = 'Please select format of NMR data';
    app.Datatypeselection = questdlg(msg, title,
    'Bruker format', 'Text files', 'Bruker format');

    dirnmr = uigetdir('', 'Please select 1D NMR data');

    d = uiprogressdlg(app.SCOREmetaboliteIDAppUIFigure,
    'Title', 'Please Wait', 'Message', 'Opening the NMR Spectra...
    (Confirmation in Matlab Command Window may be required)',
    'Indeterminate', 'on');
    drawnow

    switch app.Datatypeselection
        case 'Bruker format'
            app.nmrdata = rbnmr(dirnmr);
            for k = 1:size(app.nmrdata,1)
                app.ppm(:,k) = app.nmrdata{k,1}.XAxis;
            end

            for k = 1:size(app.nmrdata,1)
                app.nmr(:,k) = app.nmrdata{k,1}.Data;
            end
        case 'Text files'
            cd(dirnmr);
            numfiles = inputdlg({'Number of files used
            for correlation'}, 'Number of files', [1 50]);
            numfiles = str2num(numfiles{1});

            nmldata = cell(numfiles,1);

            txtfiles=dir('*.txt');

            for k = 1:numfiles
                nmldata{k} =
            importdata(txtfiles(k).name);
            end

            % Maximale Zeilen-Anzahl der Daten finden
            for k = 1:size(nmldata)
                x(k,:) = size(nmldata{k,1}.data);
            end

            y = max(x(:,1));

            % ppm-Werte in Matrix laden
            app.ppm = zeros(y,numfiles); % leere Matrix
            für ppm-Werte
            for k = 1:numfiles
                app.ppm(1:x(k,1), k) =
            nmldata{k,1}.data(:,1);
            end

            % Intensitätswerte in Matrix laden
            für Intensitätswerte der NMR-Spektren
            app.nmr = zeros(y,numfiles); % leere Matrix
            for k = 1:numfiles
                app.nmr(1:x(k,1), k) =
            nmldata{k,1}.data(:,2);
            end

            app.nmrdata = nmldata;

        end

        app.fractions = 1:size(app.nmr,2);
        app.indexplotnmr(1,1) = size(app.ppm, 1);
        app.indexplotnmr(1,2) = 1;

        [X, ~] = meshgrid(app.fractions,
        min(app.indexplotnmr):max(app.indexplotnmr));
        plot3(app.NMRSpectra, X,
        app.ppm(min(app.indexplotnmr):max(app.indexplotnmr)),
        app.nmr(min(app.indexplotnmr):max(app.indexplotnmr)),
        app.fractions);
        xticks(app.NMRSpectra, 1:size(app.nmr,2));

        set(app.CalibrateNMRsSpectraButton, 'Enable', 'on');
        set(app.ImportNMRsSpectraButton, 'Enable', 'off');
        set(app.ImportMSSSpectraButton, 'Enable', 'on');
        close(d);

    end

% Button pushed function: CalibrateNMRsSpectraButton
function CalibrateNMRsSpectraButtonPushed(app, event)
    calibration.nmr = app.nmr;
    calibration.ppm = app.ppm;
    calibration.cal = input

    dlg({'Enter chemical shift for calibration'}, 'Calibration', [1
    50], {'0.000'});

    switch app.Datatypeselection
        case 'Bruker format'
            SW = app.nmrdata{1,1}.Acqus.SW;
        case 'Text files'
            SW = (app.nmrdata{1,1}.data(1,1) -
            app.nmrdata{1,1}.data(end,1));
        end

        TD = size(app.nmr,1);
        calibration.RES = SW/TD; %RES = Abstand zwischen
        Datenpunkten (auf ppm-Skala)

        %Input-Dialoge für oberes und unteres Limit, in dem sich TSP-
        Signal befindet
        edit_up = inputdlg({'Upper limit for
        calibration'}, 'Upper Limit', [1 50], {'-0.20'});
        calibration.lim(1,1) = str2double(edit_up);
        calibration.a = find(calibration.ppm(:,1) <
        calibration.lim(1,1),1,'first');

        edit_low = inputdlg({'Lower limit for
        calibration'}, 'Lower Limit', [1 50], {'-0.20'});
        calibration.lim(1,2) = str2double(edit_low);
        calibration.b = find(calibration.ppm(:,1) >
        calibration.lim(1,2),1,'last');

        app.nmr = calibration.nmr;
        app.ppm = calibration.ppm;
        RES = calibration.RES;
        app.CAL = str2double(calibration.cal);

        lowlim = min(calibration.lim);
        uplim = max(calibration.lim);

        index = zeros(1, size(app.nmr,2));
        for k = 1:size(app.nmr,2)
            a1 = find(app.ppm(:,k) > lowlim,1,'last');
            b1 = find(app.ppm(:,k) < uplim,1,'first');
            [~, index(1,k)] = max(app.nmr(b1:a1,k)); %finde
            Maximalwert zwischen a1 und b1 pro Spektrum (k), suche aber nur
            ZWISCHEN a1 und b1, Zeilennummer in Matrix Index speichern
            index(1,k) = index(1,k) + b1 -1; %Anpassung der
            Index-Werte auf Gesamtskala und nicht nur zwischen a1 und b1
        end

        ND = max(index)+TD-min(index); %Gesamtanzahl
        benötigter Datenpunkte für gesamte x-Achse
        app.XAxis = zeros(ND,size(app.nmr,2));
        %Preallocation

        for k = 1:size(app.nmr,2)
            app.XAxis(1:(max(index)-1),k) =
            [(app.CAL+RES*(max(index)-1))-RES:(app.CAL+RES)]; %x-Achse vor
            Indexwert
            app.XAxis(max(index),k) = app.CAL; %Wert der X-
            Achse bei Kalibrierungswert (z.B. 0.00 ppm) dem Indexwert
            zuschreiben
            app.XAxis((max(index)+1):ND,k) = [(app.CAL-
            RES):-RES:(app.CAL-RES*(TD-min(index)))]; %x-Achse nach
            Indexwert
        end % -> alle x-Achsen jeweils gleich, aber nmr
        noch nicht zugeordnet

        nmr2 = zeros(size(app.XAxis, 1), size(app.nmr,2));

        for k = 1:size(app.nmr,2)
            nmr2((max(index)-index(1,k))+1:(TD-
            index(1,k)+max(index)),k) = app.nmr(:,k); %nmr-Werte ab
            bestimmter Zeilennummer zuordnen
        end

        app.nmr = nmr2;
        app.ppm = app.XAxis(:,1);
        calibration.nmr = nmr2;
        calibration.ppm = app.XAxis;

        app.indexplotnmr(1,1) = size(app.ppm, 1);

        [X, ~] = meshgrid(app.fractions,
        min(app.indexplotnmr):max(app.indexplotnmr));
        plot3(app.NMRSpectra, X,
        app.ppm(min(app.indexplotnmr):max(app.indexplotnmr)),
        app.nmr(min(app.indexplotnmr):max(app.indexplotnmr)),
        app.fractions);
        xticks(app.NMRSpectra, 1:size(app.nmr,2));

        set(app.NormalizeNMRsSpectraButton, 'Enable', 'on');

```

```

set(app.CalibrateNMRspectrumButton, 'Enable', 'off');

assignin('base', 'nmr', app.nmr);
assignin('base', 'ppm', app.ppm);
assignin('base', 'calibration', calibration);
end

% Button pushed function: NormalizeNMRspectrumButton
function NormalizeNMRspectrumButtonPushed(app, event)
    z = find(app.ppm(:,1) == app.CAL);
    %z = find(app.ppm(:,1) == 0);
    tmsp = app.nmr(z,:);
    nmr3 = zeros(size(app.XAxis,1), size(app.nmr,2));

    for k = 1:size(app.nmr,2)
        nmr3(:,k) = app.nmr(:,k)./tmsp(1,k);
    end

    app.nmr = nmr3;

    app.indexplotnmr(1,1) = size(app.ppm, 1);

    [X, ~] = meshgrid(app.fractions,
min(app.indexplotnmr):max(app.indexplotnmr));
    plot3(app.NMRspectrum, X,
app.ppm(min(app.indexplotnmr):max(app.indexplotnmr)),
app.nmr(min(app.indexplotnmr):max(app.indexplotnmr)),
app.fractions));
    xticks(app.NMRspectrum, 1:size(app.nmr,2));

    set(app.BucketingofNMRspectrumButton, 'Enable',
'on');
    set(app.NormalizeNMRspectrumButton, 'Enable',
'off');
end

% Button pushed function: BucketingofNMRspectrumButton
function BucketingofNMRspectrumButtonPushed(app, event)
    %Abfrage, ob Bucketing durchgeführt werden soll
    title = 'Bucketing';
    msg = 'Shall Bucketing of the NMR Spectra be
performed?';
    selection = questdlg(msg, title, 'Yes', 'No',
'Yes');
    app.ppmOriginal = app.ppm;
    app.nmrOriginal = app.nmr;
    switch selection
        case 'Yes'
            %Abfrage für Bucket-Size
            bucketsize = inputdlg({'Enter Bucket size
in ppm (e.g. 0.005 ppm)'}, 'Bucket Size', [1 50], {'0.005'});
            bucketsize = str2double(bucketsize);

            numberofbuckets = (max(app.ppm) -
min(app.ppm)) ./ bucketsize;
            numberofbuckets = round(numberofbuckets,
0); %auf ganze Zahl runden
            ppmbins = discretize(app.ppm,
numberofbuckets); %Spaltenvektor mit Index-Werten je nach
Bucketanzahl
            for i = 1:size(app.nmr,2)
                app.nmrbuckets(:,i) =
accumarray(ppmbins, app.nmr(:, i)); %Summiert Werte in NMR-
Daten entsprechend Index in ppmbins
            end
            app.ppmbuckets = accumarray(ppmbins,
app.ppm, [], @mean); %ppm-Skala-Erstellung durch
Mittelwertbildung der ppm-Werte entsprechend der Bin-Indizes
            set(app.BucketsizeEditField, 'Enable',
'on');
            app.BucketsizeEditField.Value =
(string(bucketsize) + ' ppm');
            assignin('base', 'BucketSize', bucketsize);
            case 'No'
                app.ppmOriginal = app.ppm;
                app.ppmbuckets = app.ppm;
                app.nmrbuckets = app.nmr;

                set(app.BucketsizeEditField, 'Enable',
'on');
                app.BucketsizeEditField.Value = ('--');
            end
            app.indexplotnmr(1,1) = size(app.ppmbuckets, 1);
            [X, ~] = meshgrid(app.fractions,
min(app.indexplotnmr):max(app.indexplotnmr));
            plot3(app.NMRspectrum, X,
app.ppmbuckets(min(app.indexplotnmr):max(app.indexplotnmr)),
app.nmrbuckets(min(app.indexplotnmr):max(app.indexplotnmr)),
app.fractions));
            xticks(app.NMRspectrum, 1:size(app.nmr,2));
            set(app.BucketingofNMRspectrumButton, 'Enable',
'off');
        end
    set(app.SmooththeNMRdataButton, 'Enable', 'on');
    set(app.ChangetheBucketsizeButton, 'Visible',
'on');
    set(app.BucketingofNMRspectrumButton, 'Visible',
'off');
    set(app.ChangetheBucketsizeButton, 'Enable', 'on');

    assignin('base', 'NMRspectrum_Intensity',
app.nmrbuckets);
    assignin('base', 'NMRspectrum_ppmValues',
app.ppmbuckets);
end

% Button pushed function: ChangetheBucketsizeButton
function ChangetheBucketsizeButtonPushed(app, event)
    %Abfrage für Bucket-Size
    bucketsize = inputdlg({'Enter Bucket size
in ppm (e.g. 0.005 ppm)'}, 'Bucket Size', [1 50], {'0.005'});
    bucketsize = str2double(bucketsize);

    numberofbuckets = (max(app.ppm) -
min(app.ppm)) ./ bucketsize;
    numberofbuckets = round(numberofbuckets,
0); %auf ganze Zahl runden
    ppmbins = discretize(app.ppm,
numberofbuckets); %Spaltenvektor mit Index-Werten je nach
Bucketanzahl
    for i = 1:size(app.nmr,2)
        app.nmrbuckets(:,i) =
accumarray(ppmbins, app.nmr(:, i)); %Summiert Werte in NMR-
Daten entsprechend Index in ppmbins
    end
    app.ppmbuckets = accumarray(ppmbins,
app.ppm, [], @mean); %ppm-Skala-Erstellung durch
Mittelwertbildung der ppm-Werte entsprechend der Bin-Indizes
    set(app.BucketsizeEditField, 'Enable',
'on');
    app.BucketsizeEditField.Value =
(string(bucketsize) + ' ppm');
    assignin('base', 'BucketSize', bucketsize);
    assignin('base', 'NMRspectrum_Intensity',
app.nmrbuckets);
    assignin('base', 'NMRspectrum_ppmValues',
app.ppmbuckets);

    %Plot
    app.indexplotnmr = [];
    app.indexplotnmr(1,1) =
size(app.ppmbuckets, 1);
    app.indexplotnmr(1,2) = 1;
    [X, ~] = meshgrid(app.fractions,
min(app.indexplotnmr):max(app.indexplotnmr));
    plot3(app.NMRspectrum, X,
app.ppmbuckets(min(app.indexplotnmr):max(app.indexplotnmr)),
app.nmrbuckets(min(app.indexplotnmr):max(app.indexplotnmr)),
app.fractions));
    xticks(app.NMRspectrum, 1:size(app.nmr,2));
    app.nmrbucketsOriginal = app.nmrbuckets;

    switch app.SwitchSmoothNMRdata.Value
        case 'On'
            app.nmrbuckets = smoothdata(app.nmrbuckets,
2, 'movmean', app.window_length);
            app.indexplotnmr(1,1) =
size(app.ppmbuckets, 1);
            [X, ~] = meshgrid(app.fractions,
min(app.indexplotnmr):max(app.indexplotnmr));
            plot3(app.NMRspectrum, X,
app.ppmbuckets(min(app.indexplotnmr):max(app.indexplotnmr)),
app.nmrbuckets(min(app.indexplotnmr):max(app.indexplotnmr)),
app.fractions));
            xticks(app.NMRspectrum, 1:size(app.nmr,2));
        end
    end

% Button pushed function: SmooththeNMRdataButton
function SmooththeNMRdataButtonPushed(app, event)
    %Abfrage, ob Smoothing durchgeführt werden soll
    title = 'Smoothing';
    msg = 'Shall data smoothing of the NMR Spectra be
performed?';
    selection = questdlg(msg, title, 'Yes', 'No',
'Yes');
    switch selection
        case 'Yes'
            app.window_length = inputdlg({'Please state
the length of the sliding window for data smoothing (e.g. 4).'},
'Sliding window length', [1 50], {'4'});
            app.window_length =
str2double(app.window_length);
            app.nmrbucketsOriginal = app.nmrbuckets;
            app.nmrbuckets = smoothdata(app.nmrbuckets,
2, 'movmean', app.window_length);
            app.indexplotnmr(1,1) =
size(app.ppmbuckets, 1);
            [X, ~] = meshgrid(app.fractions,
min(app.indexplotnmr):max(app.indexplotnmr));
            plot3(app.NMRspectrum, X,
app.ppmbuckets(min(app.indexplotnmr):max(app.indexplotnmr)),
app.nmrbuckets(min(app.indexplotnmr):max(app.indexplotnmr)),
app.fractions));
            xticks(app.NMRspectrum, 1:size(app.nmr,2));
        end
    end

```

```

        app.SwitchSmoothNMRdata.Value = {'On'};
        set(app.lengthNMREditField, 'Enable', 'on');
        app.lengthNMREditField.Value =
(string(app.window_length));
        case 'No'
            app.SwitchSmoothNMRdata.Value = {'Off'};
            set(app.lengthNMREditField, 'Enable', 'on');
            app.lengthNMREditField.Value = ('--');
        end
        set(app.SmooththeNMRdataButton, 'Enable', 'off');
        set(app.SwitchSmoothNMRdata, 'Enable', 'on');
        set(app.chemicalshiftofEDCEditField, 'Enable',
'on');
        set(app.ShowEDCButton, 'Enable', 'on');
        set(app.AddEDCButton, 'Enable', 'on');
        set(app.ListBoxEDCPlot, 'Enable', 'on');
        assignin('base', 'NMRspectra_intensity',
app.nmrbuckets);
        assignin('base', 'NMRspectra_ppmValues',
app.ppmbuckets);
        end

% Value changed function: SwitchSmoothNMRdata
function SwitchValueSmoothNMRdataChanged(app, event)
    switch app.SwitchSmoothNMRdata.Value
        case 'Off'
            app.lengthNMREditField.Value = ('--');
            app.nmrbuckets = app.nmrbucketsOriginal;
            app.indexplotnmr(1,1) =
size(app.ppmbuckets, 1);
            [X, ~] = meshgrid(app.fractions,
min(app.indexplotnmr):max(app.indexplotnmr));
            plot3(app.NMRspectra, X,
app.ppmbuckets(min(app.indexplotnmr):max(app.indexplotnmr)),
app.nmrbuckets(min(app.indexplotnmr):max(app.indexplotnmr)),
app.fractions));
            xticks(app.NMRspectra, 1:size(app.nmr,2));
        case 'On'
            app.window_length = inputdlg('Please state
the length of the sliding window for data smooting (e.g. 4).',
'Sliding window length', [1 50], {'4'});
            app.window_length =
str2double(app.window_length);
            app.nmrbucketsOriginal = app.nmrbuckets;
            app.nmrbuckets = smoothdata(app.nmrbuckets,
2, 'movmean', app.window_length);
            app.lengthNMREditField.Value =
(string(app.window_length));
            app.indexplotnmr(1,1) =
size(app.ppmbuckets, 1);
            [X, ~] = meshgrid(app.fractions,
min(app.indexplotnmr):max(app.indexplotnmr));
            plot3(app.NMRspectra, X,
app.ppmbuckets(min(app.indexplotnmr):max(app.indexplotnmr)),
app.nmrbuckets(min(app.indexplotnmr):max(app.indexplotnmr)),
app.fractions));
            xticks(app.NMRspectra, 1:size(app.nmr,2));
        end
    end

% Button pushed function: ImportMSSpectraButton
function ImportMSSpectraButtonPushed(app, event)
    dirms = uigetdir('', 'Please select MS data');
    cd(dirms);
    numfiles = inputdlg('Number of files used for
correlation', 'Number of files', [1 50]); % if-Schleife
einbauen, wenn hier anderer Wert, als bei NMR
    numfiles = str2num(numfiles{1});
    if numfiles ~= size(app.fractions)
        msgbox('An Error occured: Number of files for
NMR and MS must be the same', 'Error', 'error')
    end
    d = uiprogessdlg(app.SCOREmetaboliteIDAppUIFigure,
'Title', 'Please Wait', 'Message', 'Opening the MS Spectra',
'Indeterminate', 'on');
    drawnow
    txtfiles=dir('*.txt');
    msdata=cell(numfiles,1);
    for k = 1:numfiles
        msdata{k} = importdata(txtfiles(k).name);
    end
    % Maximale Zeilen-Anzahl der MS-Daten finden
    for k = 1:size(msdata)
        x(k,:) = size(msdata{k,1}.data);
    end
    y = max(x(:,1));
    % m/z-Werte in Matrix laden
    app.mz = zeros(y,numfiles); % leere Matrix für mz-
Werte
    for k = 1:numfiles
        app.mz(1:x(k,1), k) = msdata{k,1}.data(:,1);
    end
    % Intensitätswerte in Matrix laden
    app.intensity = zeros(y,numfiles); % leere Matrix
für Intensitätswerte der MS-Spektren
    for k = 1:numfiles
        app.intensity(1:x(k,1), k) =
msdata{k,1}.data(:,2);
    end

    end

%Anpassung der MS-Spektren, sodass pro Zeile die
%Intensitätswerte der verschiedenen Fraktionen bei jeweils
%gleichen m/z-Werten stehen
    app.mzall = unique(app.mz); %Erstellung eines
Spaltenvektors mit allen versch. m/z-Werten
    app.mzall(app.mzall == 0, :) = []; %Löschen der
Zeile mit m/z = 0
    %Zuordnung der Intensitätswerte zu den m/z-Werten
    %m/z-Wert in Spalte 1, Fraktionen in Spalten 2 bis x
    [~, ia] = ismember(app.mz, app.mzall);
    highestvalue = max(ia, [], 'all');
    ia(ia == 0) = highestvalue + 1;
    app.intensityall = zeros((size(app.mzall,1)+1),
size(app.intensity,2));
    for i = 1:size(ia)
        for j = 1:size(ia,2)
            if app.intensityall(ia(i,j), j) == 0
                app.intensityall(ia(i,j), j) =
app.intensity(i,j);
            else
                app.intensityall(ia(i,j), j) =
(app.intensityall(ia(i,j), j) + app.intensity(i,j));
            end
        end
    end
    app.intensityall(highestvalue+1,:) = [];
    %mzall nicht als Vektor, sondern in so vielen Spalten, wie auch
%Fraktionen vorhanden sind
    for i = 1:size(app.intensityall, 2)
        app.mzall(:,i) = app.mzall(:,1);
    end
    %Plot
    app.indexplotms(1,1) = size(app.intensityall,1);
    app.indexplotms(1,2) = 1;
    [X, ~] = meshgrid(app.fractions,
min(app.indexplotms):max(app.indexplotms));
    plot3(app.MSSpectra, X,
app.mzall(min(app.indexplotms):max(app.indexplotms)),
app.intensityall(min(app.indexplotms):max(app.indexplotms),
app.fractions));
    xticks(app.MSSpectra, 1:size(app.intensityall, 2));
    set(app.SubtractBlankMSSpectrumButton, 'Enable',
'on');
    set(app.ImportMSSpectraButton, 'Enable', 'off');
    %List Box Fractions to plot
    set(app.FractionstoplotListBox, 'Enable', 'on');
    app.fractionstoplot = string(app.fractions);
    app.FractionstoplotListBox.Items =
app.fractionstoplot;
    app.FractionstoplotListBox.ItemsData =
app.fractions;
    close(d);
    end

% Button pushed function: SubtractBlankMSSpectrumButton
function SubtractBlankMSSpectrumButtonPushed(app,
event)
    d = uiprogessdlg(app.SCOREmetaboliteIDAppUIFigure,
'Title', 'Please Wait', 'Message', 'Preparing the MS Spectra',
'Indeterminate', 'on');
    drawnow
    title = 'MS Blank Subtraction';
    msg = 'Shall a subtraction of a MS Blank Spectrum
be performed?';
    selection = questdlg(msg, title, 'Yes', 'No',
'Yes');
    switch selection
        case 'Yes'
            blank = uiimport('-file');
            [C, ia] = ismember(app.mzall,
blank.data(:,1));
            app.intensitysub =
zeros(size(app.intensityall, 1), size(app.intensityall, 2));
            %Intensitysub = Intensitätswerte nach Subtraktion des Blanks
            %Subtraktion der Intensitätswerte des Blanks der
Intensitätswerte bei den
            %zugehörigen m/z-Verhältnissen
            for i = 1:size(app.intensityall,1)
                for k = 1:size(app.intensityall,2)
                    if C(i,k) == 1
                        app.intensitysub(i,k) =
app.intensityall(i,k) - blank.data(ia(i,k), 2);
                    else
                        app.intensitysub(i,k) =
app.intensityall(i,k);
                    end
                end
            end
            %wenn Werte in "intensitysub" kleiner 0, dann gleich 0 setzen
            app.intensitysub(app.intensitysub < 0) = 0;
        case 'No'
            app.intensitysub = app.intensityall;
        end
    end
    %Plot
    app.indexplotms(1,1) = size(app.intensitysub,1);
    [X, ~] = meshgrid(app.fractions,
min(app.indexplotms):max(app.indexplotms));

```

```

        plot3(app.MSSpectra, X,
app.mzall(min(app.indexplotms):max(app.indexplotms)),
app.intensitysub(min(app.indexplotms):max(app.indexplotms),
app.fractions));
        xticks(app.MSSpectra, 1:size(app.intensitysub, 2));
        assignin('base', 'MSSpectra_Intensity',
app.intensitysub);
        assignin('base', 'MSSpectra_mzValues', app.mzall);
        title = 'ESI ionization mode';
        msg = 'Please specify the ESI ionization mode';
        selection = questdlg(msg, title, 'positive',
'negative', 'positive');
        switch selection
            case 'positive'
                app.negativeButton.Value = false;
                app.positiveButton.Value = true;
            case 'negative'
                app.negativeButton.Value = true;
                app.positiveButton.Value = false;
        end

        set(app.SubtractBlankMSSpectrumButton, 'Enable',
'off');
        set(app.NormalizeMSSpectraButton, 'Enable', 'on');

    end

    close(d);

% Button pushed function: NormalizeMSSpectraButton
function NormalizeMSSpectraButtonPushed(app, event)
    title = 'MS Spectra Normalization';
    msg = 'Do you want to normalize the MS spectra?';
    selection = questdlg(msg, title, 'Yes', 'No',
'Yes');

    switch selection
        case 'Yes'
            edit_Std = inputdlg({'Specify the m/z of
the internal standard'}, 'Internal Standard', [1 50]);
            Standard(1,1) = str2double(edit_Std);

            [I ZeileStandard] = min(abs(app.mzall(:,1)
- Standard));

            for i = 1:size(app.intensitysub, 2)
                InternalStandard =
app.intensitysub(ZeileStandard, i);
                for j = 1:size(app.intensitysub,1)
                    intensity_norm(j,i) =
app.intensitysub(j,i)./InternalStandard;
                end
            end
            app.intensitysub = intensity_norm;
            %Plot
            app.indexplotms(1,1) =
size(app.intensitysub,1);
            [X, ~] = meshgrid(app.fractions,
min(app.indexplotms):max(app.indexplotms));
            plot3(app.MSSpectra, X,
app.mzall(min(app.indexplotms):max(app.indexplotms)),
app.intensitysub(min(app.indexplotms):max(app.indexplotms),
app.fractions));
            xticks(app.MSSpectra,
1:size(app.intensitysub, 2));
            case 'No'
                end
                assignin('base', 'MSSpectra_Intensity',
app.intensitysub);
                assignin('base', 'MSSpectra_mzValues', app.mzall);
                set(app.NormalizeMSSpectraButton, 'Enable', 'off');
                set(app.SmooththeMSdataButton, 'Enable', 'on');
            end

% Button pushed function: SmooththeMSdataButton
function SmooththeMSdataButtonPushed(app, event)
    title = 'Smoothing';
    msg = 'Do you want to smooth the MS data?';
    selection = questdlg(msg, title, 'Yes', 'No',
'Yes');

    switch selection
        case 'Yes'
            window_length = inputdlg({'Please state the
length of the sliding window for data smoothing (e.g. 4).'},
'Sliding window length', [1 50], {'4'});
            window_length = str2double(window_length);
            app.intensitysubOriginal =
app.intensitysub;
            app.intensitysub =
smoothdata(app.intensitysub, 2, 'movmean', window_length);
            %Plot
            app.indexplotms(1,1) =
size(app.intensitysub,1);
            [X, ~] = meshgrid(app.fractions,
min(app.indexplotms):max(app.indexplotms));
            plot3(app.MSSpectra, X,
app.mzall(min(app.indexplotms):max(app.indexplotms)),
app.intensitysub(min(app.indexplotms):max(app.indexplotms),
app.fractions));
            xticks(app.MSSpectra, 1:size(app.intensitysub, 2));
            end

% Value changed function: SwitchSmoothMSdata
function SwitchSmoothMSdataValueChanged(app, event)
    switch app.SwitchSmoothMSdata.Value
        case 'Off'
            app.lengthMSEditField.Value = ('--');
            app.intensitysubOriginal =
app.intensitysub;
            %Plot
            app.indexplotms(1,1) =
size(app.intensitysub,1);
            [X, ~] = meshgrid(app.fractions,
min(app.indexplotms):max(app.indexplotms));
            plot3(app.MSSpectra, X,
app.mzall(min(app.indexplotms):max(app.indexplotms)),
app.intensitysub(min(app.indexplotms):max(app.indexplotms),
app.fractions));
            xticks(app.MSSpectra,
1:size(app.intensitysub, 2));
            case 'On'
                window_length = inputdlg({'Please state the
length of the sliding window for data smoothing (e.g. 4).'},
'Sliding window length', [1 50], {'4'});
                window_length = str2double(window_length);
                app.intensitysubOriginal =
app.intensitysub;
                app.intensitysub =
smoothdata(app.intensitysub, 2, 'movmean', window_length);
                app.lengthMSEditField.Value =
(string(window_length));
                %Plot
                app.indexplotms(1,1) =
size(app.intensitysub,1);
                [X, ~] = meshgrid(app.fractions,
min(app.indexplotms):max(app.indexplotms));
                plot3(app.MSSpectra, X,
app.mzall(min(app.indexplotms):max(app.indexplotms)),
app.intensitysub(min(app.indexplotms):max(app.indexplotms),
app.fractions));
                xticks(app.MSSpectra,
1:size(app.intensitysub, 2));
            end

% Value changed function: FractionstoplotListBox
function FractionstoplotListBoxValueChanged(app, event)
    %einstellen, welche Fraktionen für NMR und MS
geplottet werden
    app.fractions = app.FractionstoplotListBox.Value;
    assignin('base', 'fractions', app.fractions);
    [X, ~] = meshgrid(app.fractions,
min(app.indexplotnmr):max(app.indexplotnmr));
    plot3(app.NMRSpectra, X,
app.ppbuckets(min(app.indexplotnmr):max(app.indexplotnmr)),
app.nmrbuckets(min(app.indexplotnmr):max(app.indexplotnmr),
app.fractions));
    xticks(app.NMRSpectra, 1:size(app.nmr,2));
    [X, ~] = meshgrid(app.fractions,
min(app.indexplotms):max(app.indexplotms));
    plot3(app.MSSpectra, X,
app.mzall(min(app.indexplotms):max(app.indexplotms)),
app.intensitysub(min(app.indexplotms):max(app.indexplotms),
app.fractions));
    xticks(app.MSSpectra, 1:size(app.intensitysub, 2));
    end

% Value changed function: chemicalshiftofEDCeditField
function chemicalshiftofEDCeditFieldValueChanged(app,
event)
    app.EDCvalue =
app.chemicalshiftofEDCeditField.Value;
    %Finde nächsten ppm-Wert, für den ein Datenpunkt in der
%NMR-Matrix vorhanden ist
    [I app.ZeileEDCvalue] = min(abs(app.ppbuckets(:,1)
- app.EDCvalue));
    end

```



```
% Button pushed function: ShowEDCButton
function ShowEDCButtonPushed(app, event)
% ppm nicht nur als Spaltenvektor, sondern "ppmperfraction" als
Matrix in Größe der Anzahl der Fraktionen
ppmperfraction = app.ppmbuckets; %nur für
Preallocation
for i = 1:size(app.nmrbuckets, 2)
    ppmperfraction(:,i) = app.ppmbuckets(:,i);
end
%regulärer Plot
[X, ~] = meshgrid(app.fractions,
min(app.indexplotnmr):max(app.indexplotnmr));
plot3(app.NMRSpectra, X,
app.ppmbuckets(min(app.indexplotnmr):max(app.indexplotnmr)),
app.nmrbuckets(min(app.indexplotnmr):max(app.indexplotnmr)),
app.fractions);
xticks(app.NMRSpectra, 1:size(app.nmr,2));
%Hinzufügen des EDC-Plots
hold(app.NMRSpectra, 'on')
plot3(app.NMRSpectra, app.fractions,
ppmperfraction(app.ZeileEDCvalue, app.fractions),
app.nmrbuckets(app.ZeileEDCvalue, app.fractions), 'LineWidth',
2);
hold(app.NMRSpectra, 'off')
end

% Value changed function: mzvalueofEMCEditField
function mzvalueofEMCEditFieldValueChanged(app, event)
app.EMCvalue = app.mzvalueofEMCEditField.Value;
%Finde nächsten m/z-Wert, für den ein Datenpunkt in der
%'intensitysub'-Matrix vorhanden ist
[I app.ZeileEMCvalue] = min(abs(app.mzall(:,1) -
app.EMCvalue));
End

% Button pushed function: ShowEMCButton
function ShowEMCButtonPushed(app, event)
%regulärer Plot der MS-Spektren nach Blank-Subtraktion
[X, ~] = meshgrid(app.fractions,
min(app.indexplotms):max(app.indexplotms));
plot3(app.MSSpectra, X,
app.mzall(min(app.indexplotms):max(app.indexplotms)),
app.intensitysub(min(app.indexplotms):max(app.indexplotms),
app.fractions));
xticks(app.MSSpectra, 1:size(app.intensitysub, 2));

%wholeEMC-Berechnung rausgenommen
%Erstellung des wholeEMC Zeilenvektors, der die summierten
%Intensitätswerte des EMC-Werts +/- 0.01 m/z enthält;
%if-Schleife, wenn app.ZeileEMCvalue-1 (oder +1) nicht um 0.01
%abweicht, sondern mehr (gespeichert in 'distances')
%distances(1,1) = (app.mzall(app.ZeileEMCvalue, 1)-
app.mzall(app.ZeileEMCvalue-1, 1));
%distances(1,2) = (app.mzall(app.ZeileEMCvalue +1,
1) - app.mzall(app.ZeileEMCvalue, 1));
%distances = round(distances, 2); %Ergebnis ist
sonst "zu genau"
%if distances(1,1) == 0.01 && distances (1,2) ==
0.01
% wholeEMC(1,app.fractions) =
sum(app.intensitysub(app.ZeileEMCvalue-1:app.ZeileEMCvalue+1,
app.fractions));
%elseif distances(1,1) == 0.01 && distances(1,2) >
0.01
% wholeEMC(1,app.fractions) =
sum(app.intensitysub(app.ZeileEMCvalue-1:app.ZeileEMCvalue,
app.fractions));
%elseif distances(1,1) > 0.01 && distances(1,2) ==
0.01
% wholeEMC(1,app.fractions) =
sum(app.intensitysub(app.ZeileEMCvalue:app.ZeileEMCvalue+1,
app.fractions));
%elseif distances(1,1) > 0.01 && distances(1,2) >
0.01
% wholeEMC(1,app.fractions) =
sum(app.intensitysub(app.ZeileEMCvalue, app.fractions));
%end

%Hinzufügen der EMC-Plots
hold(app.MSSpectra, 'on')
%plot3(app.MSSpectra, app.fractions,
app.mzall(app.ZeileEMCvalue, app.fractions), wholeEMC(1,
app.fractions), 'LineWidth', 2);
plot3(app.MSSpectra, app.fractions,
app.mzall(app.ZeileEMCvalue, app.fractions),
app.intensitysub(app.ZeileEMCvalue, app.fractions),
'LineWidth', 2);
hold(app.MSSpectra, 'off')
end

% Button pushed function: AddEDCButton
function AddEDCButtonPushed(app, event)
% Wert des EDC (im Edit Field) in ListBox des EDC/EMC-Plots
% übernehmen
temp = app.ListBoxEDCPlot.Items;
temp = string(temp);
b = string(app.chemicalshiftofEDCEditField.Value);
c = [temp, b];

app.ListBoxEDCPlot.Items = string(c);
temp2 = app.ListBoxEDCPlot.ItemsData;
b2 = app.ZeileEDCvalue;
c2 = [temp2, b2];
app.ListBoxEDCPlot.ItemsData = c2;
set(app.DeleteEDCButton, 'Enable', 'on');
set(app.PlotEDCButton, 'Enable', 'on');
set(app.ClearPlotButton, 'Enable', 'on');
if size(app.ListBoxEDCPlot.Items) >= 1 &
size(app.ListBoxEMCPlot.Items) >= 1
set(app.NMRMSButton, 'Enable', 'on');
set(app.PearsonTableMS_NMR, 'Enable', 'on');
end
if size(app.ListBoxEDCPlot.Items) >= 1
set(app.NMRNMRButton, 'Enable', 'on');
set(app.PearsonTableNMR_NMR, 'Enable', 'on');
set(app.SpearmanButton, 'Enable', 'on');
set(app.KendallButton, 'Enable', 'on');
set(app.PearsonButton, 'Enable', 'on');
end
end

% Button pushed function: AddEMCButton
function AddEMCButtonPushed(app, event)
% Wert des EMC (im Edit Field) in ListBox des EDC/EMC-Plots
% übernehmen
temp = app.ListBoxEMCPlot.Items;
temp = string(temp);
b = string(app.mzvalueofEMCEditField.Value);
c = [temp, b];
app.ListBoxEMCPlot.Items = string(c);
temp2 = app.ListBoxEMCPlot.ItemsData;
b2 = app.ZeileEMCvalue;
c2 = [temp2, b2];
app.ListBoxEMCPlot.ItemsData = c2;
set(app.DeleteEMCButton, 'Enable', 'on');
set(app.PlotEMCButton, 'Enable', 'on');
set(app.ClearPlotButton, 'Enable', 'on');
if size(app.ListBoxEDCPlot.Items) >= 1 &
size(app.ListBoxEMCPlot.Items) >= 1
set(app.NMRMSButton, 'Enable', 'on');
set(app.PearsonTableMS_NMR, 'Enable', 'on');
end
if size(app.ListBoxEMCPlot.Items) >= 1
set(app.MSMSButton, 'Enable', 'on');
set(app.PearsonTableMS_MS, 'Enable', 'on');
set(app.SpearmanButton, 'Enable', 'on');
set(app.KendallButton, 'Enable', 'on');
set(app.PearsonButton, 'Enable', 'on');
end
end

% Button pushed function: DeleteEDCButton
function DeleteEDCButtonPushed(app, event)
% Löschen des hervorgehobenen EDC-Wertes in ListBox
[~, idx] = ismember(app.ListBoxEDCPlot.Value,
app.ListBoxEDCPlot.ItemsData);
app.ListBoxEDCPlot.Items(idx) = [];
app.ListBoxEDCPlot.ItemsData(idx) = [];
end

% Button pushed function: DeleteEMCButton
function DeleteEMCButtonPushed(app, event)
% Löschen des hervorgehobenen EMC-Wertes in ListBox
[~, idx] = ismember(app.ListBoxEMCPlot.Value,
app.ListBoxEMCPlot.ItemsData);
app.ListBoxEMCPlot.Items(idx) = [];
app.ListBoxEMCPlot.ItemsData(idx) = [];
end

% Button pushed function: PlotEDCButton
function PlotEDCButtonPushed(app, event)
plot(app.EDCandEMCPlot, app.fractions,
(app.nmrbuckets(app.ListBoxEDCPlot.ItemsData,
app.fractions))');
legend(app.EDCandEMCPlot, app.ListBoxEDCPlot.Items,
'location', 'northeast');
set(app.NormalizeEDCIntensitiesinPlotMenu,
'Enable', 'on');
set(app.NormalizeEMCIntensitiesinPlotMenu,
'Enable', 'off');
end

% Button pushed function: PlotEMCButton
function PlotEMCButtonPushed(app, event)
%wholeEMC-Berechnung rausgenommen
%app.wholeEMC = zeros(size(app.intensitysub, 1),
size(app.intensitysub, 2)); %preallocation
%Erstellung des wholeEMC Zeilenvektors, der die summierten
%Intensitätswerte des EMC-Werts +/- 0.01 m/z enthält;
%mit eingebauter if-Schleife, wenn Abweichung zwischen Zeilen
%größer als 0.01 (Werte in distances)
%distances(1,1) = (app.mzall(app.ZeileEMCvalue, 1)-
app.mzall(app.ZeileEMCvalue-1, 1));
%distances(1,2) = (app.mzall(app.ZeileEMCvalue +1,
1) - app.mzall(app.ZeileEMCvalue, 1));
%distances = round(distances, 2); %Ergebnis ist
sonst "zu genau"
```

```

        %for i = 1:size(app.ListBoxEMCPlot.ItemsData, 2)
        % if distances(1,1) == 0.01 && distances (1,2)
== 0.01
        %
        app.wholeEMC(app.ListBoxEMCPlot.ItemsData(i),app.fractions) =
sum(app.intensitysub(app.ListBoxEMCPlot.ItemsData(i)-
1:app.ListBoxEMCPlot.ItemsData(i)+1, app.fractions));
        % elseif distances(1,1) == 0.01 &&
distances(1,2) > 0.01
        %
        app.wholeEMC(app.ListBoxEMCPlot.ItemsData(i),app.fractions) =
sum(app.intensitysub(app.ListBoxEMCPlot.ItemsData(i)-
1:app.ListBoxEMCPlot.ItemsData(i), app.fractions));
        % elseif distances(1,1) > 0.01 && distances(1,2)
== 0.01
        %
        app.wholeEMC(app.ListBoxEMCPlot.ItemsData(i),app.fractions) =
sum(app.intensitysub(app.ListBoxEMCPlot.ItemsData(i):app.ListBox
xEMCPlot.ItemsData(i)+1, app.fractions));
        % elseif distances(1,1) > 0.01 && distances(1,2)
> 0.01
        %
        app.wholeEMC(app.ListBoxEMCPlot.ItemsData(i),app.fractions) =
app.intensitysub(app.ListBoxEMCPlot.ItemsData(i),
app.fractions);
        % end
        %end
        plot(app.EDCandEMCPlot, app.fractions,
(app.intensitysub(app.ListBoxEMCPlot.ItemsData,
app.fractions)));
        legend(app.EDCandEMCPlot, app.ListBoxEMCPlot.Items,
'location', 'northeast');
        set(app.NormalizeEMCIntensitiesinPlotMenu,
'Enable', 'on');
        set(app.NormalizeEDCIntensitiesinPlotMenu,
'Enable', 'off');
    end

% Button pushed function: ClearPlotButton
function ClearPlotButtonPushed(app, event)
cla(app.EDCandEMCPlot);
set(app.NormalizeEDCIntensitiesinPlotMenu,
'Enable', 'off');
set(app.NormalizeEMCIntensitiesinPlotMenu,
'Enable', 'off');
end

% Button pushed function: NMRMSButton
function NMRMSButtonPushed(app, event)
%Person-Table etc. löschen (sonst bleiben Werte bestehen, wenn
%sie aus Listbox gelöscht werden)
app.PearsonTableMS_NMR.Data = [];
app.PearsonTableMS_NMR.ColumnName = [];
app.PearsonTableMS_NMR.RowName = [];
clear [row, col]
removeStyle(app.PearsonTableMS_NMR);
app.p_ValuesTable_MS_NMR.Data = [];
removeStyle(app.p_ValuesTable_MS_NMR);
app.p_ValuesTable_MS_NMR.ColumnName = [];
app.p_ValuesTable_MS_NMR.RowName = [];
app.wholeEMC = [];
%distances = [];
%Erstellung des wholeEMC Zeilenvektors, der die summierten
%Intensitätswerte des EMC-Werts +/- 0.01 m/z enthält
%mit eingebauter if-Schleife, wenn Abweichung zwischen Zeilen
%größer als 0.01 (Werte in distances)
%for i = 1:size(app.ListBoxEMCPlot.ItemsData, 2)
% distances(i,1) =
(app.mzall(app.ListBoxEMCPlot.ItemsData(i), 1)-
app.mzall(app.ListBoxEMCPlot.ItemsData(i)-1, 1));
% distances(i,2) =
(app.mzall(app.ListBoxEMCPlot.ItemsData(i) +1, 1) -
app.mzall(app.ListBoxEMCPlot.ItemsData(i), 1));
% distances = round(distances, 2); %Ergebnis ist
sonst "zu genau"
% if distances(i,1) == 0.01 && distances (i,2)
== 0.01
        %
        app.wholeEMC(app.ListBoxEMCPlot.ItemsData(i),app.fractions) =
sum(app.intensitysub(app.ListBoxEMCPlot.ItemsData(i)-
1:app.ListBoxEMCPlot.ItemsData(i)+1, app.fractions));
        % elseif distances(i,1) == 0.01 &&
distances(i,2) > 0.01
        %
        app.wholeEMC(app.ListBoxEMCPlot.ItemsData(i),app.fractions) =
sum(app.intensitysub(app.ListBoxEMCPlot.ItemsData(i)-
1:app.ListBoxEMCPlot.ItemsData(i), app.fractions));
        % elseif distances(i,1) > 0.01 && distances(i,2)
== 0.01
        %
        app.wholeEMC(app.ListBoxEMCPlot.ItemsData(i),app.fractions) =
sum(app.intensitysub(app.ListBoxEMCPlot.ItemsData(i):app.ListBox
xEMCPlot.ItemsData(i)+1, app.fractions));
        % elseif distances(i,1) > 0.01 && distances(i,2)
> 0.01
        %
        app.wholeEMC(app.ListBoxEMCPlot.ItemsData(i),app.fractions) =
app.intensitysub(app.ListBoxEMCPlot.ItemsData(i),
app.fractions);
        % end
        %end
        %if-Schleife, wenn Abstand größer als 0.01 m/z
        %for i = 1:size(app.ListBoxEMCPlot.ItemsData, 2)
        %
        app.wholeEMC(app.ListBoxEMCPlot.ItemsData(i),app.fractions) =
sum(app.intensitysub(app.ListBoxEMCPlot.ItemsData(i) -
1:app.ListBoxEMCPlot.ItemsData(i) +1, app.fractions));
        %end
        if app.PearsonButton.Value == true
            type = 'Pearson';
        elseif app.KendallButton.Value == true
            type = 'Kendall';
        elseif app.SpearmanButton.Value == true
            type = 'Spearman';
        end
        for i = 1:size(app.ListBoxEDCPlot.Items, 2)
            for j = 1:size(app.ListBoxEMCPlot.Items, 2)
                [tempR, tempP] =
corr(app.nmrbuckets(app.ListBoxEDCPlot.ItemsData(i),
app.fractions),
app.intensitysub(app.ListBoxEMCPlot.ItemsData(j),
app.fractions), 'Type', type);
                app.PearsonTableMS_NMR.Data(i,j) =
tempR(1,1);
                app.p_ValuesTable_MS_NMR.Data(i,j) =
tempP(1,1);
            end
        end
        for i = 1:size(app.p_ValuesTable_MS_NMR.Data,2)
            app.p_ValuesTable_MS_NMR.ColumnFormat{1,i} =
{'shortG'};
            end
            app.PearsonTableMS_NMR.ColumnName =
app.ListBoxEMCPlot.Items;
            app.PearsonTableMS_NMR.RowName =
app.ListBoxEDCPlot.Items;
            app.p_ValuesTable_MS_NMR.ColumnName =
app.ListBoxEMCPlot.Items;
            app.p_ValuesTable_MS_NMR.RowName =
app.ListBoxEDCPlot.Items;
            s = uistyle('BackgroundColor', 'green');
            [row, col] = find(app.PearsonTableMS_NMR.Data >
0.95);
            addStyle(app.PearsonTableMS_NMR, s, 'cell', [row,
col]);
            addStyle(app.p_ValuesTable_MS_NMR, s, 'cell', [row,
col]);
            %assignin('base', 'PearsonTableMS_NMR',
app.PearsonTableMS_NMR);
            %assignin('base', 'p_ValuesTableMS_NMR',
app.p_ValuesTable_MS_NMR);
            app.TabGroup.SelectedTab = app.NMRMSTab;
        end

% Button pushed function: ShowallcorrelatedEDCsButton
function ShowallcorrelatedEDCsButtonPushed(app, event)
%Ziel: Suche nach passenden EDCs (mit Korrelationskoeffizient >
%0.95) für einen ausgewählten MS-Peak
%Button ist nur aktiv, wenn genau ein Wert in EMC-Listbox steht
ALLEMCCorrelations = [];
HighEMCCorrelations = [];
ind = [];
if app.PearsonButton.Value == true
    type = 'Pearson';
elseif app.KendallButton.Value == true
    type = 'Kendall';
elseif app.SpearmanButton.Value == true
    type = 'Spearman';
end
%Berechnung aller möglichen Korrelationskoeffizienten zwischen
%EMC und allen NMR-Signalen in nmrbuckets:
[~, idx] = ismember(app.ListBoxEMCPlot.Value,
app.ListBoxEMCPlot.ItemsData); %idx = Index des in der Listbox
ausgewählten Eintrags
for j = 1:size(app.nmrbuckets,1)
    [tempR, tempP] = corr(app.nmrbuckets(j,
app.fractions),
app.intensitysub(app.ListBoxEMCPlot.ItemsData(idx),
app.fractions), 'type', type);
    ALLEMCCorrelations(j,1) = tempR(1,1); %Spalte 1
= Korrelationskoeffizienten
    ALLEMCCorrelations(j,2) = tempP(1,1); %Spalte 2
= p-Werte
end
%Abfrage, ab welchem Wert die Korrelationskoeffizienten
%angezeigt werden sollen
edit_corrcoef = inputdlg('Please enter the minimum
value of correlation coefficients to be shown','Limit for
correlation coefficient',[1 50], {'0.95'});
corrcoef_limit = str2double(edit_corrcoef);
%Zeige nur Korrelationen an, bei denen der
%Korrelationskoeffizient größer ist als der Wert, der im Input-
Dlg angegeben wurde
for i = 1:size(ALLEMCCorrelations, 1)
    if ALLEMCCorrelations(i,1) >= corrcoef_limit

```

```

        HighEMCCorrelations(i,2) =
AllEMCCorrelations(i,1);
        HighEMCCorrelations(i,1) =
app.ppmbuckets(i,1);
        HighEMCCorrelations(i,3) =
max(app.nmrbuckets(i, app.fractions));
        HighEMCCorrelations(i,4) =
AllEMCCorrelations(i,2);%p-Werte
    end
end
%Erstellung eines neuen Fensters mit neuer Tabelle
fig = uifigure;
fig.Name = ('Correlating EDCs to m/z ' +
string(app.ListBoxEMCPlot.Items(idx)) + ' in fractions ' +
string(min(app.fractions)) + ' - ' +
string(max(app.fractions)));
s1 = uistyle('BackgroundColor', 'green');
s2 = uistyle('BackgroundColor', [0.65 1 0.65]);
%mittleres grün
s3 = uistyle('BackgroundColor', [0.85 1 0.85]);
%sehr abgeschwächtes grün
%lösche die übrigen Zeilen (bestehend aus Nullen)
if size(HighEMCCorrelations) ~= 0 %um Fehlermeldung
zu vermeiden, wenn keine Korrelationen gefunden werden können
ind = HighEMCCorrelations(:,1) ~= 0;
HighEMCCorrelations =
HighEMCCorrelations(ind,:);
uit = uitable(fig, 'Data', HighEMCCorrelations,
'ColumnName', {'chemical shift [ppm]', 'correlation
coefficient', 'maximal intensity [rel]', 'p-value'},
'Position', [50 50 450 350], 'ColumnSortable', true,
'ColumnFormat', {'shortG'});
if size(HighEMCCorrelations,1) >= 15
[~, I] = maxk(HighEMCCorrelations(:,3),
15); %finde die 15 NMR-Signale, die maximale Intensität haben
addStyle(uit, s1, 'row', I(1:5));
addStyle(uit, s2, 'row', I(6:10));
addStyle(uit, s3, 'row', I(11:15));
elseif size(HighEMCCorrelations,1) >= 5
[~, I] = maxk(HighEMCCorrelations(:,3), 5);
%finde die 5 NMR-Signale, die maximale Intensität haben
addStyle(uit, s1, 'row', I(1:5));
end
else
uialert(fig, 'No correlating EDCs found!', 'EDC
Correlations');
end
%Context-Menü im öffnenden Fenstern mit Tabelle
cm = uicontextmenu(fig);
m1 = uimenu(cm, 'Text', 'Copy data to workspace');
m2 = uimenu(cm, 'Text', 'Create plot of
correlations');
fig.ContextMenu = cm;
m1.MenuSelectedFcn = @MenuSelectedFcn;
m2.MenuSelectedFcn = @MenuSelectedFcn2;

%Funktion, um Daten aus Tabelle in MATLAB Workspace zu
%übertragen
function MenuSelectedFcn(src, event)
assignin('base', 'CorrelatingEDCstoEMC',
HighEMCCorrelations);
assignin('base', 'AllCorrelatingEDCstoEMC',
AllEMCCorrelations);
end

%Funktion, um STOCSY-like Plot zu erstellen
function MenuSelectedFcn2(src, event)
VisLoadings1D_NMR(app.nmrbuckets(:,
app.fractions), AllEMCCorrelations(:,1),
app.ppmbuckets(:,1));
end
end

% Button pushed function: ShowallcorrelatedEDCstoEDCButton
function ShowallcorrelatedEDCstoEDCButtonPushed(app,
event)
%Ziel: Suche nach passenden EDCs (mit Korrelationskoeffizient >
%0.95) für einen ausgewählten NMR-Peak
%Button ist nur aktiv, wenn genau ein Wert in EMC-ListBox steht
AllEDCCorrelations = [];
HighEDCCorrelations = [];
ind = [];
if app.PearsonButton.Value == true
type = 'Pearson';
elseif app.KendallButton.Value == true
type = 'Kendall';
elseif app.SpearmanButton.Value == true
type = 'Spearman';
end
%Berechnung aller möglichen Korrelationskoeffizienten zwischen
%EMC und allen NMR-Signalen in nmrbuckets:
[~, idx] = ismember(app.ListBoxEDCPlot.Value,
app.ListBoxEDCPlot.ItemsData); %idx = Index des in der Listbox
ausgewählten Eintrags
for j = 1:size(app.nmrbuckets,1)
[tempR, tempP] = corr(app.nmrbuckets(j,
app.fractions),
app.nmrbuckets(app.ListBoxEDCPlot.ItemsData(idx),
app.fractions), 'Type', type);
AllEDCCorrelations(j,1) = tempR(1,1);
AllEDCCorrelations(j,2) = tempP(1,1);
end
%Abfrage, ab welchem Wert die Korrelationskoeffizienten
%angezeigt werden sollen
edit_corrcoef = inputdlg('Please enter the minimum
value of correlation coefficients to be shown','Limit for
correlation coefficient',[1 50], {'0.95'});
corrcoef_limit = str2double(edit_corrcoef);
%Zeige nur Korrelationen an, bei denen der
%Korrelationskoeffizient größer ist als der Wert, der im Input-
Dlg angegeben wurde
for i = 1:size(AllEDCCorrelations, 1)
if AllEDCCorrelations(i,1) >= corrcoef_limit
HighEDCCorrelations(i,2) =
AllEDCCorrelations(i,1);
HighEDCCorrelations(i,1) =
app.ppmbuckets(i,1);
HighEDCCorrelations(i,3) =
max(app.nmrbuckets(i, app.fractions));
HighEDCCorrelations(i,4) =
AllEDCCorrelations(i,2);
end
end
%Erstellung eines neuen Fensters mit neuer Tabelle
fig = uifigure;
fig.Name = ('Correlating EDCs to EDC at ' +
string(app.ListBoxEDCPlot.Items(idx)) + ' ppm in fractions ' +
string(min(app.fractions)) + ' - ' +
string(max(app.fractions)));
s1 = uistyle('BackgroundColor', 'green');
s2 = uistyle('BackgroundColor', [0.65 1 0.65]);
%mittleres grün
s3 = uistyle('BackgroundColor', [0.85 1 0.85]);
%sehr abgeschwächtes grün
%lösche die übrigen Zeilen (bestehend aus Nullen)
if size(HighEDCCorrelations) ~= 0 %um Fehlermeldung
zu vermeiden, wenn keine Korrelationen gefunden werden können
ind = HighEDCCorrelations(:,1) ~= 0;
HighEDCCorrelations =
HighEDCCorrelations(ind,:);
uit = uitable(fig, 'Data', HighEDCCorrelations,
'ColumnName', {'chemical shift [ppm]', 'correlation
coefficient', 'maximal intensity [rel]', 'p-value'},
'Position', [50 50 450 350], 'ColumnSortable', true,
'ColumnFormat', {'shortG'});
if size(HighEDCCorrelations,1) >= 15
[~, I] = maxk(HighEDCCorrelations(:,3),
15); %finde die 15 NMR-Signale, die maximale Intensität haben
addStyle(uit, s1, 'row', I(1:5));
addStyle(uit, s2, 'row', I(6:10));
addStyle(uit, s3, 'row', I(11:15));
elseif size(HighEDCCorrelations,1) >= 5
[~, I] = maxk(HighEDCCorrelations(:,3), 5);
%finde die 5 NMR-Signale, die maximale Intensität haben
addStyle(uit, s1, 'row', I(1:5));
end
else
uialert(fig, 'No correlating EDCs found!', 'EDC
Correlations');
end
%Context-Menü im öffnenden Fenstern mit Tabelle
cm = uicontextmenu(fig);
m1 = uimenu(cm, 'Text', 'Copy data to workspace');
m2 = uimenu(cm, 'Text', 'Create plot of
correlations');
fig.ContextMenu = cm;
m1.MenuSelectedFcn = @MenuSelectedFcn;
m2.MenuSelectedFcn = @MenuSelectedFcn2;

%Funktion, um Daten aus Tabelle in MATLAB Workspace zu
%übertragen
function MenuSelectedFcn(src, event)
assignin('base', 'CorrelatingEDCstoEDC',
HighEDCCorrelations);
assignin('base', 'AllCorrelatingEDCstoEDC',
AllEDCCorrelations);
end

%Funktion, um STOCSY-like Plot zu erstellen
function MenuSelectedFcn2(src,event)
VisLoadings1D_NMR(app.nmrbuckets(:,
app.fractions), AllEDCCorrelations(:,1),
app.ppmbuckets(:,1));
end
end

% Value changed function: nmrlowerlimitEditField
function nmrlowerlimitEditFieldValueChanged(app, event)
app.nmrlimit(1,1) =
app.nmrlowerlimitEditField.Value;
if app.ppmbuckets ~= 0
app.indexplotnmr(1,1) = find(app.ppmbuckets >
min(app.nmrlimit, 1, 'last'));

```



```

        app.indexplotnmr(1,2) = find(app.ppbuckets <
max(app.nmrlimit), 1, 'first');
[X, ~] = meshgrid(app.fractions,
min(app.indexplotnmr):max(app.indexplotnmr));
plot3(app.NMRSpectra, X,
app.ppbuckets(min(app.indexplotnmr):max(app.indexplotnmr)),
app.nmrbuckets(min(app.indexplotnmr):max(app.indexplotnmr)),
app.fractions));
xticks(app.NMRSpectra, 1:size(app.nmr,2));
    else
        app.indexplotnmr(1,1) = find(app.ppm(:,1) >
min(app.nmrlimit), 1, 'last');
        app.indexplotnmr(1,2) = find(app.ppm(:,1) <
max(app.nmrlimit), 1, 'first');
[X, ~] = meshgrid(app.fractions,
min(app.indexplotnmr):max(app.indexplotnmr));
plot3(app.NMRSpectra, X,
app.ppm(min(app.indexplotnmr):max(app.indexplotnmr)),
app.nmr(min(app.indexplotnmr):max(app.indexplotnmr)),
app.fractions));
xticks(app.NMRSpectra, 1:size(app.nmr,2));
    end
end

% Value changed function: nmrupperlimitEditField
function nmrupperlimitEditFieldValueChanged(app, event)
    app.nmrlimit(1,2) =
app.nmrupperlimitEditField.Value;
    if app.ppbuckets ~= 0
        app.indexplotnmr(1,2) = find(app.ppbuckets >
max(app.nmrlimit), 1, 'first');
        app.indexplotnmr(1,1) = find(app.ppbuckets <
min(app.nmrlimit), 1, 'last');
[X, ~] = meshgrid(app.fractions,
min(app.indexplotnmr):max(app.indexplotnmr));
plot3(app.NMRSpectra, X,
app.ppbuckets(min(app.indexplotnmr):max(app.indexplotnmr)),
app.nmrbuckets(min(app.indexplotnmr):max(app.indexplotnmr)),
app.fractions));
xticks(app.NMRSpectra, 1:size(app.nmr,2));
    else
        app.indexplotnmr(1,2) = find(app.ppm(:,1) >
max(app.nmrlimit), 1, 'last');
        app.indexplotnmr(1,1) = find(app.ppm(:,1) <
min(app.nmrlimit), 1, 'first');
[X, ~] = meshgrid(app.fractions,
min(app.indexplotnmr):max(app.indexplotnmr));
plot3(app.NMRSpectra, X,
app.ppm(min(app.indexplotnmr):max(app.indexplotnmr)),
app.nmr(min(app.indexplotnmr):max(app.indexplotnmr)),
app.fractions));
xticks(app.NMRSpectra, 1:size(app.nmr,2));
    end
end

% Value changed function: lowermzEditField
function lowermzEditFieldValueChanged(app, event)
    app.mzlimit(1,1) = app.lowermzEditField.Value;
    app.indexplots(1,1) = find(app.mzall(:,1) >
min(app.mzlimit), 1, 'first');
    app.indexplots(1,2) = find(app.mzall(:,1) <
max(app.mzlimit), 1, 'last');
[X, ~] = meshgrid(app.fractions,
min(app.indexplots):max(app.indexplots));
plot3(app.MSSpectra, X,
app.mzall(min(app.indexplots):max(app.indexplots)),
app.intensitysub(min(app.indexplots):max(app.indexplots)),
app.fractions));
xticks(app.MSSpectra, 1:size(app.intensitysub, 2));
end

% Value changed function: uppermzEditField
function uppermzEditFieldValueChanged(app, event)
    app.mzlimit(1,2) = app.uppermzEditField.Value;
    app.indexplots(1,2) = find(app.mzall(:,1) >
max(app.mzlimit), 1, 'first');
    app.indexplots(1,1) = find(app.mzall(:,1) <
min(app.mzlimit), 1, 'last');
[X, ~] = meshgrid(app.fractions,
min(app.indexplots):max(app.indexplots));
plot3(app.MSSpectra, X,
app.mzall(min(app.indexplots):max(app.indexplots)),
app.intensitysub(min(app.indexplots):max(app.indexplots)),
app.fractions));
xticks(app.MSSpectra, 1:size(app.intensitysub, 2));
end

% Button pushed function: NMRNMRButton
function NMRNMRButtonPushed(app, event)
%Person-Table etc. löschen (sonst bleiben Werte bestehen, wenn
%sie aus Listbox gelöscht werden)
    app.PearsonTableNMR_NMR.Data = [];
    app.PearsonTableNMR_NMR.ColumnName = [];
    app.PearsonTableNMR_NMR.RowName = [];
    clear [row, col]
    removeStyle(app.PearsonTableNMR_NMR);
    app.p_valuesTable_NMR_NMR.Data = [];
    app.p_valuesTable_NMR_NMR.ColumnName = [];

        app.p_valuesTable_NMR_NMR.RowName = [];
        removeStyle(app.p_valuesTable_NMR_NMR);
        if app.PearsonButton.Value == true
            type = 'Pearson';
        elseif app.KendallButton.Value == true
            type = 'Kendall';
        elseif app.SpearmanButton.Value == true
            type = 'Spearman';
        end
        for i = 1:size(app.ListBoxEDCPlot.Items, 2)
            for j = 1:size(app.ListBoxEDCPlot.Items, 2)
                [tempR, tempP] =
corr(app.nmrbuckets(app.ListBoxEDCPlot.ItemsData(i),
app.fractions)',
app.nmrbuckets(app.ListBoxEDCPlot.ItemsData(j),
app.fractions)'), 'Type', type);
                app.PearsonTableNMR_NMR.Data(i,j) =
tempR(1,1);
                app.p_valuesTable_NMR_NMR.Data(i,j) =
tempP(1,1);
            end
        end
        app.PearsonTableNMR_NMR.ColumnName =
app.ListBoxEDCPlot.Items;
        app.PearsonTableNMR_NMR.RowName =
app.ListBoxEDCPlot.Items;
        app.p_valuesTable_NMR_NMR.ColumnName =
app.ListBoxEDCPlot.Items;
        app.p_valuesTable_NMR_NMR.RowName =
app.ListBoxEDCPlot.Items;
        s = uistyle('BackgroundColor', 'green');
        s1 = uistyle('FontColor', 'white',
'BackgroundColor', 'white');
        %Style für Korrelationskoeffizienten > 0.95
        [row, col] = find(app.PearsonTableNMR_NMR.Data >
0.95);
        addStyle(app.PearsonTableNMR_NMR, s, 'cell',
[row, col]);
        addStyle(app.p_valuesTable_NMR_NMR, s, 'cell',
[row, col]);
        %Tabelleneinträge in Diagonale weiß einfärben (Style s1)
        for i =
1:size(app.PearsonTableNMR_NMR.ColumnName)
            for j =
1:size(app.PearsonTableNMR_NMR.RowName)
                if i == j
                    addStyle(app.PearsonTableNMR_NMR, s1,
'cell', [i, j]);
                    addStyle(app.p_valuesTable_NMR_NMR,
s1, 'cell', [i, j]);
                end
            end
        end
        app.p_valuesTable_NMR_NMR.ColumnFormat =
{'shortE'};
        app.TabGroup.SelectedTab = app.NMRNMRTab;
    end

% Button pushed function: MSMSButton
function MSMSButtonPushed(app, event)
    app.PearsonTableMS_MS.Data = [];
    app.PearsonTableMS_MS.ColumnName = [];
    app.PearsonTableMS_MS.RowName = [];
    clear [row, col]
    removeStyle(app.PearsonTableMS_MS);
    app.p_valuesTable_MS_MS.Data = [];
    removeStyle(app.p_valuesTable_MS_MS);
    app.p_valuesTable_MS_MS.ColumnName = [];
    app.p_valuesTable_MS_MS.RowName = [];
    if app.PearsonButton.Value == true
        type = 'Pearson';
    elseif app.KendallButton.Value == true
        type = 'Kendall';
    elseif app.SpearmanButton.Value == true
        type = 'Spearman';
    end
    for i = 1:size(app.ListBoxEMCPlot.Items, 2)
        for j = 1:size(app.ListBoxEMCPlot.Items, 2)
            [tempR, tempP] =
corr(app.intensitysub(app.ListBoxEMCPlot.ItemsData(i),
app.fractions)',
app.intensitysub(app.ListBoxEMCPlot.ItemsData(j),
app.fractions)'), 'Type', type);
            app.PearsonTableMS_MS.Data(i,j) =
tempR(1,1);
            app.p_valuesTable_MS_MS.Data(i,j) =
tempP(1,1);
        end
    end
    app.p_valuesTable_MS_MS.Data(2) =
app.p_valuesTable_MS_MS.ColumnFormat{1,i} =
{'shortG'};
    end
    app.PearsonTableMS_MS.ColumnName =
app.ListBoxEMCPlot.Items;
    app.PearsonTableMS_MS.RowName =
app.ListBoxEMCPlot.Items;

```

```

        app.p_valuesTable_MS_MS.ColumnName =
app.ListBoxEMCPlot.Items;
        app.p_valuesTable_MS_MS.RowName =
app.ListBoxEMCPlot.Items;
        s = uistyle('BackgroundColor', 'green');
        s1 = uistyle('FontColor', 'white',
'BackgroundColor', 'white');
        [row, col] = find(app.PearsonTableMS_MS.Data >
0.95);
        addStyle(app.PearsonTableMS_MS, s, 'cell', [row,
col]);
        addStyle(app.p_valuesTable_MS_MS, s, 'cell', [row,
col]);
%Tabelleneinträge in Diagonale weiß einfärben (Style s1)
        for i = 1:size(app.PearsonTableMS_MS.ColumnName)
            for j = 1:size(app.PearsonTableMS_MS.RowName)
                if i == j
                    addStyle(app.PearsonTableMS_MS, s1,
'cell', [i, j]);
                    addStyle(app.p_valuesTable_MS_MS, s1,
'cell', [i, j]);
                end
            end
        end
        app.p_valuesTable_MS_MS.ColumnFormat =
{'shortE'};
        app.TabGroup.SelectedTab = app.MSMTTab;
    end

% Button pushed function: ShowallcorrelatedEMCsButton
function ShowallcorrelatedEMCsButtonPushed(app, event)
%Ziel: Suche nach passenden EMCs (mit Korrelationskoeffizient >
%0.95) für einen ausgewählten NMR-Peak (= EDC)
%Button ist nur aktiv, wenn genau ein Wert in EDC-ListBox steht
    AllEDCCorrelations = [];
    HighEDCCorrelations = [];
    ind = [];
    if app.PearsonButton.Value == true
        type = 'Pearson';
    elseif app.KendallButton.Value == true
        type = 'Kendall';
    elseif app.SpearmanButton.Value == true
        type = 'Spearman';
    end
%Berechnung aller möglichen Korrelationskoeffizienten zwischen
%EDC und allen MS-Signalen in intensitysub
    [~, idx] = ismember(app.ListBoxEDCPlot.Value,
app.ListBoxEDCPlot.ItemsData); %idx = Index des in der ListBox
ausgewählten Eintrags
    for j = 1:size(app.intensitysub,1)
        [tempR, tempP] = corr(app.intensitysub(j,
app.fractions)',
app.nmrbuckets(app.ListBoxEDCPlot.ItemsData(idx),
app.fractions)', 'Type', type);
        AllEDCCorrelations(j,1) = tempR(1,1);
        AllEDCCorrelations(j,2) = tempP(1,1);
    end
%Abfrage, ab welchem Wert die Korrelationskoeffizienten
%angezeigt werden sollen
    edit_corrcoef = inputdlg({'Please enter the minimum
value of correlation coefficients to be shown'}, 'Limit for
correlation coefficient', [1 50], {'0.95'});
    corrcoef_limit = str2double(edit_corrcoef);
%Zeige nur Korrelationen an, bei denen der
%Korrelationskoeffizient größer ist als der Wert, der im
%Input-Dlg angegeben wurde
    for i = 1:size(AllEDCCorrelations, 1)
        if AllEDCCorrelations(i,1) >= corrcoef_limit
            HighEDCCorrelations(i,1) = app.mzall(i,1);
            HighEDCCorrelations(i,2) = app.mzall(i,1);
            HighEDCCorrelations(i,3) =
max(app.intensitysub(i, app.fractions));
            HighEDCCorrelations(i,4) =
AllEDCCorrelations(i,2);
        end
    end
%Erstellung eines neuen Fensters mit neuer Tabelle
    fig = uifigure;
    fig.Name = ('Correlating EMCs to EDC at ' +
string(app.ListBoxEDCPlot.Items(idx)) + ' ppm in fractions ' +
string(min(app.fractions)) + ' to ' +
string(max(app.fractions))));
    s1 = uistyle('BackgroundColor', 'green');
    s2 = uistyle('BackgroundColor', [0.65 1 0.65]);
%mittleres grün
    s3 = uistyle('BackgroundColor', [0.85 1 0.85]);
%sehr abgeschwächtes grün
%Lösche die übrigen Zeilen (bestehend aus Nullen)
    if size(HighEDCCorrelations) ~= 0 %um Fehlermeldung
zu vermeiden, wenn keine Korrelationen gefunden werden können
        ind = HighEDCCorrelations(:,1) ~= 0;
        HighEDCCorrelations =
HighEDCCorrelations(ind,:);
        uit = uitable(fig, 'Data', HighEDCCorrelations,
'ColumnName', {'m/z', 'correlation coefficient', 'maximal
intensity', 'p-value'}, 'Position', [50 50 450 350],
'ColumnSortable', true, 'ColumnFormat', {'shortG'});

        if size(HighEDCCorrelations,1) >= 15
            [~, I] = maxk(HighEDCCorrelations(:,3),
15); %finde die 15 m/z-Werte, die maximale Intensität haben
            addStyle(uit, s1, 'row', I(1:5));
            addStyle(uit, s2, 'row', I(6:10));
            addStyle(uit, s3, 'row', I(11:15));
        elseif size(HighEDCCorrelations,1) >= 5
            [~, I] = maxk(HighEDCCorrelations(:,3), 5);
            %finde die 15 m/z-Werte, die maximale Intensität haben
            addStyle(uit, s1, 'row', I(1:5));
        end
    else
        uialert(fig, 'No correlating EMCs found!', 'EMC
Correlations');
    end

%Context-Menü im öffnenden Fenster mit Tabelle
    cm = uicontextmenu(fig);
    m1 = uimenu(cm, 'Text', 'Copy data to workspace');
    m2 = uimenu(cm, 'Text', 'Create plot of
correlations');
    fig.ContextMenu = cm;
    m1.MenuSelectedFcn = @MenuSelectedFcn;
    m2.MenuSelectedFcn = @MenuSelectedFcn2;

%Funktion, um Daten aus Tabelle in MATLAB Workspace zu
%übertragen
    function MenuSelectedFcn(src, event)
        assignin('base', 'CorrelatingEMCsToEDC',
HighEDCCorrelations);
        assignin('base', 'AllCorrelatingEMCsToEDC',
AllEDCCorrelations);
    end

%Funktion, um STOCYSY/SHY-like Plot zu erstellen
    function MenuSelectedFcn2(src, event)
        Visloadings1D_MS(app.intensitysub(:,
app.fractions)', AllEDCCorrelations(:,1)', app.mzall(:,1)');
    end

% Button pushed function: ShowallcorrelatedEMCsToEMCButton
function ShowallcorrelatedEMCsToEMCButtonPushed(app,
event)
%Zeige alle hoch-korrelierenden MS-Signale zu einem
%ausgewählten MS-Signal
%Button ist nur aktiv, wenn genau ein Wert in EMC-ListBox steht
    AllEMCCorrelations = [];
    HighEMCCorrelations = [];
    ind = [];
    if app.PearsonButton.Value == true
        type = 'Pearson';
    elseif app.KendallButton.Value == true
        type = 'Kendall';
    elseif app.SpearmanButton.Value == true
        type = 'Spearman';
    end
%Berechnung aller möglichen Korrelationskoeffizienten zwischen
%EMC und allen NMR-Signalen in nmrbuckets:
    [~, idx] = ismember(app.ListBoxEMCPlot.Value,
app.ListBoxEMCPlot.ItemsData); %idx = Index des in der ListBox
ausgewählten Eintrags
    for j = 1:size(app.intensitysub,1)
        [tempR, tempP] = corr(app.intensitysub(j,
app.fractions)',
app.intensitysub(app.ListBoxEMCPlot.ItemsData(idx),
app.fractions)', 'Type', type);
        AllEMCCorrelations(j,1) = tempR(1,1);
        AllEMCCorrelations(j,2) = tempP(1,1);
    end
%Abfrage, ab welchem Wert die Korrelationskoeffizienten
%angezeigt werden sollen
    edit_corrcoef = inputdlg({'Please enter the minimum
value of correlation coefficients to be shown'}, 'Limit for
correlation coefficient', [1 50], {'0.95'});
    corrcoef_limit = str2double(edit_corrcoef);
%Zeige nur Korrelationen an, bei denen der
%Korrelationskoeffizient größer ist als der Wert, der im Input-
Dlg angegeben wurde
    for i = 1:size(AllEMCCorrelations, 1)
        if AllEMCCorrelations(i,1) >= corrcoef_limit
            HighEMCCorrelations(i,1) = app.mzall(i,1);
            HighEMCCorrelations(i,2) = app.mzall(i,1);
            HighEMCCorrelations(i,3) =
max(app.intensitysub(i, app.fractions));
            HighEMCCorrelations(i,4) =
AllEMCCorrelations(i,2);
        end
    end
%Erstellung eines neuen Fensters mit neuer Tabelle
    fig = uifigure;
    fig.Name = ('Correlating EMCs to EMC at m/z ' +
string(app.ListBoxEMCPlot.Items(idx)) + ' in fractions ' +
string(min(app.fractions)) + ' - ' +
string(max(app.fractions))));

```

```

s1 = uistyle('BackgroundColor', 'green');
s2 = uistyle('BackgroundColor', [0.65 1 0.65]);
%mittleres grün
s3 = uistyle('BackgroundColor', [0.85 1 0.85]);
%sehr abgeschwächtes grün
%Lösche die übrigen Zeilen (bestehend aus Nullen)
if size(HighEMCCorrelations) ~= 0 %um Fehlermeldung
zu vermeiden, wenn keine Korrelationen gefunden werden können
ind = HighEMCCorrelations(:,1) ~= 0;
HighEMCCorrelations =
HighEMCCorrelations(ind,:);
uit = uitable(fig, 'Data', HighEMCCorrelations,
'ColumnName', {'m/z', 'correlation coefficient', 'maximal
intensity', 'p-value'}, 'Position', [50 50 450 350],
'ColumnSortTable', true, 'ColumnFormat', {'shortG'});
if size(HighEMCCorrelations,1) >= 15
[~, I] = maxk(HighEMCCorrelations(:,3),
15); %finde die 15 NMR-Signale, die maximale Intensität haben
addStyle(uit, s1, 'row', I(1:5));
addStyle(uit, s2, 'row', I(6:10));
addStyle(uit, s3, 'row', I(11:15));
elseif size(HighEMCCorrelations,1) >= 5
[~, I] = maxk(HighEMCCorrelations(:,3), 5);
%finde die 5 NMR-Signale, die maximale Intensität haben
addStyle(uit, s1, 'row', I(1:5));
end
else
uialert(fig, 'No correlating EMCs found!', 'EMC
Correlations');
end
%Context-Menü im öffnenden Fenster mit Tabelle
cm = uicontextmenu(fig);
m1 = uimenu(cm, 'Text', 'Copy data to workspace');
m2 = uimenu(cm, 'Text', 'Create plot of
correlations');
fig.ContextMenu = cm;
m1.MenuSelectedFcn = @MenuSelectedFcn;
m2.MenuSelectedFcn = @MenuSelectedFcn2;

%Funktion, um Daten aus Tabelle in MATLAB Workspace zu
%übertragen
function MenuSelectedFcn(src, event)
assignin('base', 'CorrelatingEMCstoEMC',
HighEMCCorrelations);
assignin('base', 'AllCorrelatingEMCstoEMC',
AllEMCCorrelations);
end

%Funktion, um STOCYSY/SHY-like Plot zu erstellen
function MenuSelectedFcn2(src, event)
VisLoadings1D_MS(app.intensitysub(:,
app.fractions), AllEMCCorrelations(:,1), app.mzall(:,1));
end
end

% Menu selected function: Copydatatoworkspace_NMRMS
function Copydatatoworkspace_NMRMSSelected(app, event)
assignin('base', 'fractions_min',
min(app.fractions));
assignin('base', 'fractions_max',
max(app.fractions));
for i = 1:size(app.PearsonTableMS_NMR.Data,1)
PearsonTableMS_NMR((i+1),1) =
app.PearsonTableMS_NMR.RowName(i);
PValuesMS_NMR((i+1),1) =
app.PearsonTableMS_NMR.RowName(i);
end
for i = 1:size(app.PearsonTableMS_NMR.Data,2)
PearsonTableMS_NMR(1,(i+1)) =
app.PearsonTableMS_NMR.ColumnName(i);
PValuesMS_NMR(1,(i+1)) =
app.PearsonTableMS_NMR.ColumnName(i);
end
for i = 1:size(app.PearsonTableMS_NMR.Data,1)
for j = 1:size(app.PearsonTableMS_NMR.Data,2)
PearsonTableMS_NMR((i+1), (j+1)) =
app.PearsonTableMS_NMR.Data(i,j);
PValuesMS_NMR{(i+1), (j+1)} =
app.p_ValuesTable_MS_NMR.Data(i,j);
end
end
assignin('base', 'PearsonTableMS_NMR',
PearsonTableMS_NMR);
assignin('base', 'PValuesMS_NMR', PValuesMS_NMR);
end

% Menu selected function: Copydatatoworkspace_NMRNMR
function Copydatatoworkspace_NMRNMRMenuSelected(app,
event)
assignin('base', 'fractions_min',
min(app.fractions));
assignin('base', 'fractions_max',
max(app.fractions));
for i = 1:size(app.PearsonTableNMR_NMR.Data,1)
PearsonTableNMR_NMR((i+1),1) =
app.PearsonTableNMR_NMR.RowName(i); %ok<*ADPROPLC>
PValuesNMR_NMR((i+1), 1) =
app.PearsonTableNMR_NMR.RowName(i);
end
for i = 1:size(app.PearsonTableNMR_NMR.Data,2)
PearsonTableNMR_NMR(1,(i+1)) =
app.PearsonTableNMR_NMR.ColumnName(i);
PValuesNMR_NMR(1,(i+1)) =
app.PearsonTableNMR_NMR.ColumnName(i);
end
for i = 1:size(app.PearsonTableNMR_NMR.Data,1)
for j = 1:size(app.PearsonTableNMR_NMR.Data,2)
PearsonTableNMR_NMR((i+1), (j+1)) =
app.PearsonTableNMR_NMR.Data(i,j);
PValuesNMR_NMR{(i+1), (j+1)} =
app.p_valuesTable_NMR_NMR.Data(i,j);
if (i+1) == (j+1)
PearsonTableNMR_NMR{(i+1), (j+1)} = [];
PValuesNMR_NMR{(i+1), (j+1)} = [];
end
end
end
assignin('base', 'PearsonTableNMR_NMR',
PearsonTableNMR_NMR);
assignin('base', 'PValuesNMR_NMR', PValuesNMR_NMR);
end

% Menu selected function: CopydatatoworkspaceMenu_MS_MS
function CopydatatoworkspaceMenu_MS_MSSelected(app,
event)
assignin('base', 'fractions_min',
min(app.fractions));
assignin('base', 'fractions_max',
max(app.fractions));
for i = 1:size(app.PearsonTableMS_MS.Data,1)
PearsonTableMS_MS((i+1),1) =
app.PearsonTableMS_MS.RowName(i);
PValuesMS_MS((i+1),1) =
app.PearsonTableMS_MS.RowName(i);
end
for i = 1:size(app.PearsonTableMS_MS.Data,2)
PearsonTableMS_MS(1,(i+1)) =
app.PearsonTableMS_MS.ColumnName(i);
PValuesMS_MS(1,(i+1)) =
app.PearsonTableMS_MS.ColumnName(i);
end
for i = 1:size(app.PearsonTableMS_MS.Data,1)
for j = 1:size(app.PearsonTableMS_MS.Data,2)
PearsonTableMS_MS((i+1), (j+1)) =
app.PearsonTableMS_MS.Data(i,j);
PValuesMS_MS{(i+1), (j+1)} =
app.p_ValuesTable_MS_MS.Data(i,j);
end
end
assignin('base', 'PearsonTableMS_MS',
PearsonTableMS_MS);
assignin('base', 'PValuesMS_MS', PValuesMS_MS);
end

% Value changed function: ListBoxEMCPlot
function ListBoxEMCPlotValueChanged(app, event)
if size(app.ListBoxEMCPlot.Value) == 1 %wenn ein
Wert in der EMC-Listbox ausgewählt wurde,
set(app.ShowallcorrelatedEDCsButton, 'Enable',
'on') %dann aktiviere den Button, damit alle Korrelationen mit
EDCs berechnet werden können
set(app.ShowallcorrelatedEMCstoEMCButton,
'Enable', 'on');
else
set(app.ShowallcorrelatedEDCsButton, 'Enable',
'off') %wenn kein oder mehr Werte ausgewählt wurden, können
keine automatische Korrelationen berechnet werden
set(app.ShowallcorrelatedEMCstoEMCButton,
'Enable', 'off');
end
end

% Value changed function: ListBoxEDCPlot
function ListBoxEDCPlotValueChanged(app, event)
if size(app.ListBoxEDCPlot.Value) == 1 %wenn ein
Wert in der EDC-Listbox ausgewählt wurde,
set(app.ShowallcorrelatedEMCsButton, 'Enable',
'on') %dann aktiviere den Button, damit alle Korrelationen mit
EMCs berechnet werden können
set(app.ShowallcorrelatedEDCsstoEDCButton,
'Enable', 'on');
else
set(app.ShowallcorrelatedEMCsButton, 'Enable',
'off') %wenn kein oder mehr Werte ausgewählt wurden, können
keine automatische Korrelationen berechnet werden
set(app.ShowallcorrelatedEDCsstoEDCButton,
'Enable', 'on');
end
end

% Menu selected function: CopyEDCorEMCPlot
function CopyEDCorEMCPlotMenuSelected(app, event)
fig1 = figure;
ax = app.EDCandEMCPlot;
copyobj(ax, fig1);
end

```

```

% Menu selected function: CopyPlotofMSSpectraMenu
function CopyPlotofMSSpectraMenuSelected(app, event)
    fig1 = figure;
    ax = app.MSSpectra;
    copyobj(ax, fig1);
end

% Menu selected function: CopyPlotofNMRSpectraMenu
function CopyPlotofNMRSpectraMenuSelected(app, event)
    fig1 = figure;
    ax = app.NMRSpectra;
    copyobj(ax, fig1);
end

% Menu selected function: CreatePlotofEDCandEMCMenu
function CreatePlotofEDCandEMCMenuSelected(app, event)
    fig2 = figure('Position', [950 300 800 700]);
    ax = uiaxes(fig2, 'Position', [50 50 700 600]);
    yyaxis(ax, 'left');
    plot(ax, app.fractions,
    (app.nmrbuckets(app.ListBoxEDCPlot.ItemsData,
    app.fractions)));
    yyaxis(ax, 'right')
    plot(ax, app.fractions,
    (app.intensitysub(app.ListBoxEMCPlot.ItemsData,
    app.fractions)));
end

% Menu selected function: NormalizeEDCIntensitiesinPlotMenu
function NormalizeEDCIntensitiesinPlotMenuSelected(app,
event)
    for i = 1:size(app.ListBoxEDCPlot.Items, 2)
        maxValue =
max(app.nmrbuckets(app.ListBoxEDCPlot.ItemsData,
app.fractions), [], 2);
        EDC_norm =
(app.nmrbuckets(app.ListBoxEDCPlot.ItemsData,
app.fractions))./maxValue;
        end
        cla(app.EDCandEMCPlot);
        plot(app.EDCandEMCPlot, app.fractions,
(EDC_norm));
        legend(app.EDCandEMCPlot, app.ListBoxEDCPlot.Items,
'location', 'northeast');
    end

% Menu selected function: NormalizeEMCIntensitiesinPlotMenu
function NormalizeEMCIntensitiesinPlotMenuSelected(app,
event)
    for i = 1:size(app.ListBoxEMCPlot.Items, 2)
        maxValue =
max(app.intensitysub(app.ListBoxEMCPlot.ItemsData,
app.fractions), [], 2);
        EMC_norm =
(app.intensitysub(app.ListBoxEMCPlot.ItemsData,
app.fractions))./maxValue;
        end
        cla(app.EDCandEMCPlot);
        plot(app.EDCandEMCPlot, app.fractions,
(EMC_norm));
        legend(app.EDCandEMCPlot, app.ListBoxEMCPlot.Items,
'location', 'northeast');
    end

% Changes arrangement of the app based on UIFigure width
function updateAppLayout(app, event)
    currentFigureWidth =
app.SCOREmetaboliteIDAppUIFigure.Position(3);
    if(currentFigureWidth <= app.onePanelWidth)
        % Change to a 3x1 grid
        app.GridLayout.RowHeight = {1029, 1029, 1029};
        app.GridLayout.ColumnWidth = {'1x', '1x'};
        app.CenterPanel.Layout.Row = 1;
        app.CenterPanel.Layout.Column = 1;
        app.LeftPanel.Layout.Row = 2;
        app.LeftPanel.Layout.Column = 1;
        app.RightPanel.Layout.Row = 3;
        app.RightPanel.Layout.Column = 1;
    elseif (currentFigureWidth > app.onePanelWidth &&
currentFigureWidth <= app.twoPanelWidth)
        % Change to a 2x2 grid
        app.GridLayout.RowHeight = {1029, 1029};
        app.GridLayout.ColumnWidth = {'1x', '1x'};
        app.CenterPanel.Layout.Row = 1;
        app.CenterPanel.Layout.Column = [1,2];
        app.LeftPanel.Layout.Row = 2;
        app.LeftPanel.Layout.Column = 1;
        app.RightPanel.Layout.Row = 2;
        app.RightPanel.Layout.Column = 2;
    else
        % Change to a 1x3 grid
        app.GridLayout.RowHeight = {'1x'};
        app.GridLayout.ColumnWidth = {245, '1x', 516};
        app.LeftPanel.Layout.Row = 1;
        app.LeftPanel.Layout.Column = 1;
        app.CenterPanel.Layout.Row = 1;
        app.CenterPanel.Layout.Column = 2;

        app.RightPanel.Layout.Row = 1;
        app.RightPanel.Layout.Column = 3;
    end
end

% Component initialization
methods (Access = private)

% Create UIFigure and components
function createComponents(app)
% Create SCOREmetaboliteIDAppUIFigure and hide until all
components are created
    app.SCOREmetaboliteIDAppUIFigure =
uifigure('Visible', 'off');
    app.SCOREmetaboliteIDAppUIFigure.AutoResizeChildren
= 'off';
    app.SCOREmetaboliteIDAppUIFigure.Position = [100
100 1666 1029];
    app.SCOREmetaboliteIDAppUIFigure.Name = 'SCORE-
metabolite-ID App';
    app.SCOREmetaboliteIDAppUIFigure.Icon =
'Lupe_Icon.png';
    app.SCOREmetaboliteIDAppUIFigure.SizeChangedFcn =
createCallbackFcn(app, @updateAppLayout, true);
% Create GridLayout
    app.GridLayout =
uigridlayout(app.SCOREmetaboliteIDAppUIFigure);
    app.GridLayout.ColumnWidth = {245, '1x', 516};
    app.GridLayout.RowHeight = {'1x'};
    app.GridLayout.ColumnSpacing = 0;
    app.GridLayout.RowSpacing = 0;
    app.GridLayout.Padding = [0 0 0 0];
    app.GridLayout.Scrollable = 'on';
% Create LeftPanel
    app.LeftPanel = uipanel(app.GridLayout);
    app.LeftPanel.Layout.Row = 1;
    app.LeftPanel.Layout.Column = 1;
% Create BucketingofNMRSpectraButton
    app.BucketingofNMRSpectraButton =
uibutton(app.LeftPanel, 'push');
    app.BucketingofNMRSpectraButton.ButtonPushedFcn =
createCallbackFcn(app, @BucketingofNMRSpectraButtonPushed,
true);
    app.BucketingofNMRSpectraButton.Enable = 'off';
    app.BucketingofNMRSpectraButton.Position = [26 728
104 49];
    app.BucketingofNMRSpectraButton.Text = {'Bucketing
of ' ; 'NMR Spectra'};
% Create ShowEMCButton
    app.ShowEMCButton = uibutton(app.LeftPanel,
'push');
    app.ShowEMCButton.ButtonPushedFcn =
createCallbackFcn(app, @ShowEMCButtonPushed, true);
    app.ShowEMCButton.Enable = 'off';
    app.ShowEMCButton.Position = [25 104 187 48];
    app.ShowEMCButton.Text = 'Show EMC';
% Create ShowEDCButton
    app.ShowEDCButton = uibutton(app.LeftPanel,
'push');
    app.ShowEDCButton.ButtonPushedFcn =
createCallbackFcn(app, @ShowEDCButtonPushed, true);
    app.ShowEDCButton.Enable = 'off';
    app.ShowEDCButton.Position = [23 529 187 49];
    app.ShowEDCButton.Text = 'Show EDC';
% Create SubtractBlankMSSpectrumButton
    app.SubtractBlankMSSpectrumButton =
uibutton(app.LeftPanel, 'push');
    app.SubtractBlankMSSpectrumButton.ButtonPushedFcn =
createCallbackFcn(app, @SubtractBlankMSSpectrumButtonPushed,
true);
    app.SubtractBlankMSSpectrumButton.Enable = 'off';
    app.SubtractBlankMSSpectrumButton.Position = [24
351 187 49];
    app.SubtractBlankMSSpectrumButton.Text = 'Subtract
Blank MS Spectrum';
% Create NormalizeNMRSpectraButton
    app.NormalizeNMRSpectraButton =
uibutton(app.LeftPanel, 'push');
    app.NormalizeNMRSpectraButton.ButtonPushedFcn =
createCallbackFcn(app, @NormalizeNMRSpectraButtonPushed, true);
    app.NormalizeNMRSpectraButton.Enable = 'off';
    app.NormalizeNMRSpectraButton.Position = [24 795
187 49];
    app.NormalizeNMRSpectraButton.Text = 'Normalize NMR
Spectra';
% Create ImportMSSpectraButton
    app.ImportMSSpectraButton = uibutton(app.LeftPanel,
'push');
    app.ImportMSSpectraButton.ButtonPushedFcn =
createCallbackFcn(app, @ImportMSSpectraButtonPushed, true);
    app.ImportMSSpectraButton.Enable = 'off';
    app.ImportMSSpectraButton.Position = [24 412 187
49];
    app.ImportMSSpectraButton.Text = 'Import MS
Spectra';
% Create CalibrateNMRSpectraButton
    app.CalibrateNMRSpectraButton =
uibutton(app.LeftPanel, 'push');

```

```

        app.CalibrateNMRsSpectraButton.ButtonPushedFcn =
createCallbackFcn(app, @CalibrateNMRsSpectraButtonPushed, true);
        app.CalibrateNMRsSpectraButton.Enable = 'off';
        app.CalibrateNMRsSpectraButton.Position = [24 860
187 49];
        app.CalibrateNMRsSpectraButton.Text = 'Calibrate NMR
Spectra';
        % Create ImportNMRsSpectraButton
        app.ImportNMRsSpectraButton =
uibutton(app.LeftPanel, 'push');
        app.ImportNMRsSpectraButton.ButtonPushedFcn =
createCallbackFcn(app, @ImportNMRsSpectraButtonPushed, true);
        app.ImportNMRsSpectraButton.Position = [24 927 187
49];
        app.ImportNMRsSpectraButton.Text = 'Import NMR
Spectra';
        % Create chemicalshiftofEDCEditField
        app.chemicalshiftofEDCEditField =
uieditfield(app.LeftPanel, 'numeric');
        app.chemicalshiftofEDCEditField.ValueChangedFcn =
createCallbackFcn(app,
@chemicalshiftofEDCEditFieldValueChanged, true);
        app.chemicalshiftofEDCEditField.Enable = 'off';
        app.chemicalshiftofEDCEditField.Position = [130 601
80 46];
        % Create chemicalshiftofEDCEditFieldLabel
        app.chemicalshiftofEDCEditFieldLabel =
uicontrol(app.LeftPanel);
        app.chemicalshiftofEDCEditFieldLabel.HorizontalAlignment =
'right';
        app.chemicalshiftofEDCEditFieldLabel.Position = [23
601 88 46];
        app.chemicalshiftofEDCEditFieldLabel.Text =
{'chemical'; 'shift of EDC'};
        % Create mzvalueofEMCEditField
        app.mzvalueofEMCEditField =
uieditfield(app.LeftPanel, 'numeric');
        app.mzvalueofEMCEditField.ValueChangedFcn =
createCallbackFcn(app, @mzvalueofEMCEditFieldValueChanged,
true);
        app.mzvalueofEMCEditField.Enable = 'off';
        app.mzvalueofEMCEditField.Position = [132 168 80
46];
        % Create mzvalueofEMCEditFieldLabel
        app.mzvalueofEMCEditFieldLabel =
uicontrol(app.LeftPanel);
        app.mzvalueofEMCEditFieldLabel.HorizontalAlignment =
'right';
        app.mzvalueofEMCEditFieldLabel.Position = [22 167
100 47];
        app.mzvalueofEMCEditFieldLabel.Text = 'm/z value of
EMC';
        % Create NMRsSpectraLabel
        app.NMRsSpectraLabel = uicontrol(app.LeftPanel);
        app.NMRsSpectraLabel.FontSize = 16;
        app.NMRsSpectraLabel.FontWeight = 'bold';
        app.NMRsSpectraLabel.Position = [25 990 105 22];
        app.NMRsSpectraLabel.Text = 'NMR Spectra';
        % Create MSSSpectraLabel
        app.MSSSpectraLabel = uicontrol(app.LeftPanel);
        app.MSSSpectraLabel.FontSize = 16;
        app.MSSSpectraLabel.FontWeight = 'bold';
        app.MSSSpectraLabel.Position = [25 479 93 22];
        app.MSSSpectraLabel.Text = 'MS Spectra';
        % Create ESIionizationmodeButtonGroup
        app.ESIionizationmodeButtonGroup =
uibuttongroup(app.LeftPanel);
        app.ESIionizationmodeButtonGroup.Enable = 'off';
        app.ESIionizationmodeButtonGroup.Title = 'ESI
ionization mode';
        app.ESIionizationmodeButtonGroup.BackgroundColor =
[0.9608 0.9608 0.9412];
        app.ESIionizationmodeButtonGroup.Position = [25 35
187 48];
        % Create positiveButton
        app.positiveButton =
uicontrol(app.ESIionizationmodeButtonGroup);
        app.positiveButton.Text = 'positive';
        app.positiveButton.Position = [15 2 62 22];
        app.positiveButton.Value = true;
        % Create negativeButton
        app.negativeButton =
uicontrol(app.ESIionizationmodeButtonGroup);
        app.negativeButton.Text = 'negative';
        app.negativeButton.Position = [96 2 67 22];
        % Create BucketsizeEditFieldLabel
        app.BucketsizeEditFieldLabel =
uicontrol(app.LeftPanel);
        app.BucketsizeEditFieldLabel.HorizontalAlignment =
'center';
        app.BucketsizeEditFieldLabel.Enable = 'off';
        app.BucketsizeEditFieldLabel.Position = [143 728 67
21];
        app.BucketsizeEditFieldLabel.Text = 'Bucket size';
        % Create BucketsizeEditField
        app.BucketsizeEditField =
uieditfield(app.LeftPanel, 'text');
        app.BucketsizeEditField.Editable = 'off';
        app.BucketsizeEditField.Enable = 'off';
        app.BucketsizeEditField.Position = [143 752 65 24];
        % Create NormalizeMSSSpectraButton
        app.NormalizeMSSSpectraButton =
uibutton(app.LeftPanel, 'push');
        app.NormalizeMSSSpectraButton.ButtonPushedFcn =
createCallbackFcn(app, @NormalizeMSSSpectraButtonPushed, true);
        app.NormalizeMSSSpectraButton.Enable = 'off';
        app.NormalizeMSSSpectraButton.Position = [25 289 187
49];
        app.NormalizeMSSSpectraButton.Text = 'Normalize MS
Spectra';
        % Create SmooththeNMRdataButton
        app.SmooththeNMRdataButton =
uibutton(app.LeftPanel, 'push');
        app.SmooththeNMRdataButton.ButtonPushedFcn =
createCallbackFcn(app, @SmooththeNMRdataButtonPushed, true);
        app.SmooththeNMRdataButton.Enable = 'off';
        app.SmooththeNMRdataButton.Position = [26 664 104
49];
        app.SmooththeNMRdataButton.Text = 'Data Smoothing';
        % Create SmooththeMSdataButton
        app.SmooththeMSdataButton = uicontrol(app.LeftPanel,
'push');
        app.SmooththeMSdataButton.ButtonPushedFcn =
createCallbackFcn(app, @SmooththeMSdataButtonPushed, true);
        app.SmooththeMSdataButton.Enable = 'off';
        app.SmooththeMSdataButton.Position = [26 228 105
49];
        app.SmooththeMSdataButton.Text = 'Data Smoothing';
        % Create SwitchSmoothNMRdata
        app.SwitchSmoothNMRdata = uicontrol(app.LeftPanel,
'slider');
        app.SwitchSmoothNMRdata.Orientation = 'vertical';
        app.SwitchSmoothNMRdata.ValueChangedFcn =
createCallbackFcn(app, @SwitchValueSmoothNMRdataChanged, true);
        app.SwitchSmoothNMRdata.Enable = 'off';
        app.SwitchSmoothNMRdata.Position = [148 677 10 23];
        % Create SmoothinglengthLabel
        app.SmoothinglengthLabel = uicontrol(app.LeftPanel);
        app.SmoothinglengthLabel.HorizontalAlignment =
'center';
        app.SmoothinglengthLabel.Enable = 'off';
        app.SmoothinglengthLabel.Position = [170 663 38
22];
        app.SmoothinglengthLabel.Text = 'length';
        % Create lengthNMReditField
        app.lengthNMReditField = uieditfield(app.LeftPanel,
'text');
        app.lengthNMReditField.Editable = 'off';
        app.lengthNMReditField.Enable = 'off';
        app.lengthNMReditField.Position = [178 684 21 24];
        % Create SwitchSmoothMSdata
        app.SwitchSmoothMSdata = uicontrol(app.LeftPanel,
'slider');
        app.SwitchSmoothMSdata.Orientation = 'vertical';
        app.SwitchSmoothMSdata.ValueChangedFcn =
createCallbackFcn(app, @SwitchValueSmoothMSdataValue, true);
        app.SwitchSmoothMSdata.Enable = 'off';
        app.SwitchSmoothMSdata.Position = [148 239 10 23];
        % Create SmoothinglengthLabel_2
        app.SmoothinglengthLabel_2 =
uicontrol(app.LeftPanel);
        app.SmoothinglengthLabel_2.HorizontalAlignment =
'center';
        app.SmoothinglengthLabel_2.Enable = 'off';
        app.SmoothinglengthLabel_2.Position = [170 227 38
22];
        app.SmoothinglengthLabel_2.Text = 'length';
        % Create lengthMSEditField
        app.lengthMSEditField = uieditfield(app.LeftPanel,
'text');
        app.lengthMSEditField.Editable = 'off';
        app.lengthMSEditField.Enable = 'off';
        app.lengthMSEditField.Position = [178 248 21 24];
        % Create ChangetheBucketsizeButton
        app.ChangetheBucketsizeButton =
uibutton(app.LeftPanel, 'push');
        app.ChangetheBucketsizeButton.ButtonPushedFcn =
createCallbackFcn(app, @ChangetheBucketsizeButtonPushed, true);
        app.ChangetheBucketsizeButton.Enable = 'off';
        app.ChangetheBucketsizeButton.Visible = 'off';
        app.ChangetheBucketsizeButton.Position = [26 728
104 49];
        app.ChangetheBucketsizeButton.Text = {'Change the
'; 'Bucket size'};
        % Create CenterPanel
        app.CenterPanel = uipanel(app.GridLayout);
        app.CenterPanel.Layout.Row = 1;
        app.CenterPanel.Layout.Column = 2;
        % Create MSSSpectra
        app.MSSSpectra = uiaxes(app.CenterPanel);
        title(app.MSSSpectra, 'MS Spectra');
        xlabel(app.MSSSpectra, 'fraction');
        ylabel(app.MSSSpectra, 'm/z');
        zlabel(app.MSSSpectra, 'intensity');
        app.MSSSpectra.YDir = 'reverse';

```



```

app.MSSpectra.SortMethod = 'depth';
app.MSSpectra.Position = [31 38 692 460];
% Create NMR Spectra
app.NMRSpectra = uiaxes(app.CenterPanel);
title(app.NMRSpectra, 'NMR Spectra')
xlabel(app.NMRSpectra, 'fraction')
ylabel(app.NMRSpectra, 'chemical shift [ppm]')
zlabel(app.NMRSpectra, 'intensity')
app.NMRSpectra.SortMethod = 'depth';
app.NMRSpectra.Position = [31 545 692 460];
% Create EditFieldLabel
app.EditFieldLabel = uilabel(app.CenterPanel);
app.EditFieldLabel.HorizontalAlignment = 'right';
app.EditFieldLabel.Position = [775 929 25 22];
app.EditFieldLabel.Text = '-';
% Create EditField_2Label
app.EditField_2Label = uilabel(app.CenterPanel);
app.EditField_2Label.HorizontalAlignment = 'right';
app.EditField_2Label.Position = [778 94 25 22];
app.EditField_2Label.Text = '-';
% Create FractionstoplotListBox
app.FractionstoplotListBox =
uilistbox(app.CenterPanel);
app.FractionstoplotListBox.Items = {};
app.FractionstoplotListBox.Multiselect = 'on';
app.FractionstoplotListBox.ValueChangedFcn =
createCallbackFcn(app, @FractionstoplotListBoxValueChanged,
true);
app.FractionstoplotListBox.Enable = 'off';
app.FractionstoplotListBox.Position = [747 186 129
682];
app.FractionstoplotListBox.Value = {};
% Create FractionstoplotListBoxLabel
app.FractionstoplotListBoxLabel =
uilabel(app.CenterPanel);
app.FractionstoplotListBoxLabel.HorizontalAlignment
= 'right';
app.FractionstoplotListBoxLabel.FontWeight =
'bold';
app.FractionstoplotListBoxLabel.Position = [762 867
99 22];
app.FractionstoplotListBoxLabel.Text = 'Fractions
to plot';
% Create chemicalshiftregionLabel
app.chemicalshiftregionLabel =
uilabel(app.CenterPanel);
app.chemicalshiftregionLabel.Position = [740 963
133 28];
app.chemicalshiftregionLabel.Text = 'chemical shift
region: ';
% Create nmrlowerlimitEditField
app.nmrlowerlimitEditField =
uieditfield(app.CenterPanel, 'numeric');
app.nmrlowerlimitEditField.ValueChangedFcn =
createCallbackFcn(app, @nmrlowerlimitEditFieldValueChanged,
true);
app.nmrlowerlimitEditField.Position = [740 922 52
34];
% Create nmrupperlimitEditField
app.nmrupperlimitEditField =
uieditfield(app.CenterPanel, 'numeric');
app.nmrupperlimitEditField.ValueChangedFcn =
createCallbackFcn(app, @nmrupperlimitEditFieldValueChanged,
true);
app.nmrupperlimitEditField.Position = [806 922 52
34];
% Create ppmLabel
app.ppmLabel = uilabel(app.CenterPanel);
app.ppmLabel.Position = [872 929 29 22];
app.ppmLabel.Text = 'ppm';
% Create mzLabel
app.mzLabel = uilabel(app.CenterPanel);
app.mzLabel.Position = [865 94 25 22];
app.mzLabel.Text = 'm/z';
% Create mzregionLabel
app.mzregionLabel = uilabel(app.CenterPanel);
app.mzregionLabel.Position = [740 130 65 22];
app.mzregionLabel.Text = 'm/z region: ';
% Create lowermzEditField
app.lowermzEditField = uieditfield(app.CenterPanel,
'numeric');
app.lowermzEditField.ValueChangedFcn =
createCallbackFcn(app, @lowermzEditFieldValueChanged, true);
app.lowermzEditField.Position = [741 88 52 34];
% Create uppermzEditField
app.uppermzEditField = uieditfield(app.CenterPanel,
'numeric');
app.uppermzEditField.ValueChangedFcn =
createCallbackFcn(app, @uppermzEditFieldValueChanged, true);
app.uppermzEditField.Position = [807 88 52 34];
% Create RightPanel
app.RightPanel = uipanel(app.GridLayout);
app.RightPanel.Layout.Row = 1;
app.RightPanel.Layout.Column = 3;
% Create EDCandEMCPlot
app.EDCandEMCPlot = uiaxes(app.RightPanel);
title(app.EDCandEMCPlot, 'EDC & EMC')
xlabel(app.EDCandEMCPlot, 'fraction')
ylabel(app.EDCandEMCPlot, 'intensity')
zlabel(app.EDCandEMCPlot, 'Z')
app.EDCandEMCPlot.VColor = [0.149 0.149 0.149];
app.EDCandEMCPlot.NextPlot = 'add';
app.EDCandEMCPlot.Position = [11 726 283 261];
% Create ListBoxEDCPlot
app.ListBoxEDCPlot = uilistbox(app.RightPanel);
app.ListBoxEDCPlot.Items = {};
app.ListBoxEDCPlot.Multiselect = 'on';
app.ListBoxEDCPlot.ValueChangedFcn =
createCallbackFcn(app, @ListBoxEDCPlotValueChanged, true);
app.ListBoxEDCPlot.Enable = 'off';
app.ListBoxEDCPlot.Position = [312 786 86 141];
app.ListBoxEDCPlot.Value = {};
% Create DeleteEDCButton
app.DeleteEDCButton = uibutton(app.RightPanel,
'push');
app.DeleteEDCButton.ButtonPushedFcn =
createCallbackFcn(app, @DeleteEDCButtonPushed, true);
app.DeleteEDCButton.Enable = 'off';
app.DeleteEDCButton.Position = [312 726 86 44];
app.DeleteEDCButton.Text = 'Delete EDC';
% Create ListBoxEMCPlot
app.ListBoxEMCPlot = uilistbox(app.RightPanel);
app.ListBoxEMCPlot.Items = {};
app.ListBoxEMCPlot.Multiselect = 'on';
app.ListBoxEMCPlot.ValueChangedFcn =
createCallbackFcn(app, @ListBoxEMCPlotValueChanged, true);
app.ListBoxEMCPlot.Enable = 'off';
app.ListBoxEMCPlot.Position = [405 786 86 141];
app.ListBoxEMCPlot.Value = {};
% Create DeleteEMCButton
app.DeleteEMCButton = uibutton(app.RightPanel,
'push');
app.DeleteEMCButton.ButtonPushedFcn =
createCallbackFcn(app, @DeleteEMCButtonPushed, true);
app.DeleteEMCButton.Enable = 'off';
app.DeleteEMCButton.Position = [405 726 86 44];
app.DeleteEMCButton.Text = 'Delete EMC';
% Create AddEDCButton
app.AddEDCButton = uibutton(app.RightPanel,
'push');
app.AddEDCButton.ButtonPushedFcn =
createCallbackFcn(app, @AddEDCButtonPushed, true);
app.AddEDCButton.Enable = 'off';
app.AddEDCButton.Position = [312 943 86 44];
app.AddEDCButton.Text = 'Add EDC';
% Create AddEMCButton
app.AddEMCButton = uibutton(app.RightPanel,
'push');
app.AddEMCButton.ButtonPushedFcn =
createCallbackFcn(app, @AddEMCButtonPushed, true);
app.AddEMCButton.Enable = 'off';
app.AddEMCButton.Position = [405 943 86 44];
app.AddEMCButton.Text = 'Add EMC';
% Create NMRMSButton
app.NMRMSButton = uibutton(app.RightPanel, 'push');
app.NMRMSButton.ButtonPushedFcn =
createCallbackFcn(app, @NMRMSButtonPushed, true);
app.NMRMSButton.Wrap = 'on';
app.NMRMSButton.FontWeight = 'bold';
app.NMRMSButton.Enable = 'off';
app.NMRMSButton.Position = [134 523 106 46];
app.NMRMSButton.Text = 'NMR + MS';
% Create PlotEDCButton
app.PlotEDCButton = uibutton(app.RightPanel,
'push');
app.PlotEDCButton.ButtonPushedFcn =
createCallbackFcn(app, @PlotEDCButtonPushed, true);
app.PlotEDCButton.Enable = 'off';
app.PlotEDCButton.Position = [30 660 106 46];
app.PlotEDCButton.Text = 'Plot EDC';
% Create PlotEMCButton
app.PlotEMCButton = uibutton(app.RightPanel,
'push');
app.PlotEMCButton.ButtonPushedFcn =
createCallbackFcn(app, @PlotEMCButtonPushed, true);
app.PlotEMCButton.Enable = 'off';
app.PlotEMCButton.Position = [178 660 106 46];
app.PlotEMCButton.Text = 'Plot EMC';
% Create ClearPlotButton
app.ClearPlotButton = uibutton(app.RightPanel,
'push');
app.ClearPlotButton.ButtonPushedFcn =
createCallbackFcn(app, @ClearPlotButtonPushed, true);
app.ClearPlotButton.Enable = 'off';
app.ClearPlotButton.Position = [114 601 106 46];
app.ClearPlotButton.Text = 'Clear Plot';
% Create ShowallcorrelatedEDCsButton
app.ShowallcorrelatedEDCsButton =
uibutton(app.RightPanel, 'push');
app.ShowallcorrelatedEDCsButton.ButtonPushedFcn =
createCallbackFcn(app, @ShowallcorrelatedEDCsButtonPushed,
true);
app.ShowallcorrelatedEDCsButton.Wrap = 'on';
app.ShowallcorrelatedEDCsButton.Enable = 'off';
app.ShowallcorrelatedEDCsButton.Position = [405 651
58 58];

```

```

        app.ShowallcorrelatedEDCsButton.Text = 'Show all
correlated EDCs';
% Create ShowallcorrelatedEMCsButton
app.ShowallcorrelatedEMCsButton =
uibutton(app.RightPanel, 'push');
app.ShowallcorrelatedEMCsButton.ButtonPushedFcn =
createCallbackFcn(app, @ShowallcorrelatedEMCsButtonPushed,
true);

    app.ShowallcorrelatedEMCsButton.WordWrap = 'on';
    app.ShowallcorrelatedEMCsButton.Enable = 'off';
    app.ShowallcorrelatedEMCsButton.Position = [312 651
86 58];
    app.ShowallcorrelatedEMCsButton.Text = 'Show all
correlated EMCs';
% Create EDCsLabel
app.EDCsLabel = uilabel(app.RightPanel);
app.EDCsLabel.FontSize = 16;
app.EDCsLabel.FontWeight = 'bold';
app.EDCsLabel.Position = [331 990 48 22];
app.EDCsLabel.Text = 'EDCs';
% Create EMCsLabel
app.EMCsLabel = uilabel(app.RightPanel);
app.EMCsLabel.FontSize = 16;
app.EMCsLabel.FontWeight = 'bold';
app.EMCsLabel.Position = [423 990 50 22];
app.EMCsLabel.Text = 'EMCs';
% Create ShowallcorrelatedEDCstoEDCButton
app.ShowallcorrelatedEDCstoEDCButton =
uibutton(app.RightPanel, 'push');

app.ShowallcorrelatedEDCstoEDCButton.ButtonPushedFcn =
createCallbackFcn(app, @ShowallcorrelatedEDCstoEDCButtonPushed,
true);
    app.ShowallcorrelatedEDCstoEDCButton.WordWrap =
'off';
    app.ShowallcorrelatedEDCstoEDCButton.Enable =
'off';
    app.ShowallcorrelatedEDCstoEDCButton.Position =
[312 582 86 58];
    app.ShowallcorrelatedEDCstoEDCButton.Text = 'Show
all correlated EDCs';
% Create ShowallcorrelatedEMCstoEMCButton
app.ShowallcorrelatedEMCstoEMCButton =
uibutton(app.RightPanel, 'push');

app.ShowallcorrelatedEMCstoEMCButton.ButtonPushedFcn =
createCallbackFcn(app, @ShowallcorrelatedEMCstoEMCButtonPushed,
true);
    app.ShowallcorrelatedEMCstoEMCButton.WordWrap =
'off';
    app.ShowallcorrelatedEMCstoEMCButton.Enable =
'off';
    app.ShowallcorrelatedEMCstoEMCButton.Position =
[405 582 86 58];
    app.ShowallcorrelatedEMCstoEMCButton.Text = 'Show
all correlated EMCs';
% Create TabGroup
app.TabGroup = uitabgroup(app.RightPanel);
app.TabGroup.Position = [19 25 472 446];
% Create NMRMSTab
app.NMRMSTab = uitab(app.TabGroup);
app.NMRMSTab.Title = 'NMR + MS';
% Create TabGroup_MS_NMR
app.TabGroup_MS_NMR = uitabgroup(app.NMRMSTab);
app.TabGroup_MS_NMR.Position = [0 -2 470 424];
% Create Coefficients
app.Coefficients = uitab(app.TabGroup_MS_NMR);
app.Coefficients.Title = 'Coefficients';
% Create PearsonTableMS_NMR
app.PearsonTableMS_NMR = uitable(app.Coefficients);
app.PearsonTableMS_NMR.ColumnName = '';
app.PearsonTableMS_NMR.RowName = {};
app.PearsonTableMS_NMR.Position = [1 -1 469 400];
% Create p_values
app.p_values = uitab(app.TabGroup_MS_NMR);
app.p_values.Title = 'p-Values';
% Create p_valuesTable_MS_NMR
app.p_valuesTable_MS_NMR = uitable(app.p_values);
app.p_valuesTable_MS_NMR.ColumnName = '';
app.p_valuesTable_MS_NMR.RowName = {};
app.p_valuesTable_MS_NMR.Position = [1 0 467 399];
% Create NMRNMRTab
app.NMRNMRTab = uitab(app.TabGroup);
app.NMRNMRTab.Title = 'NMR + NMR';
% Create TabGroup_NMR_NMR
app.TabGroup_NMR_NMR = uitabgroup(app.NMRNMRTab);
app.TabGroup_NMR_NMR.Position = [0 -2 471 424];
% Create CoefficientsTab
app.CoefficientsTab = uitab(app.TabGroup_NMR_NMR);
app.CoefficientsTab.Title = 'Coefficients';
% Create PearsonTableNMR_NMR
app.PearsonTableNMR_NMR =
uitable(app.CoefficientsTab);
app.PearsonTableNMR_NMR.ColumnName = '';
app.PearsonTableNMR_NMR.RowName = {};
app.PearsonTableNMR_NMR.Position = [1 0 469 398];
% Create pValuesTab
app.pValuesTab = uitab(app.TabGroup_NMR_NMR);

app.pValuesTab.Title = 'p-Values';
% Create p_valuesTable_NMR_NMR
app.p_valuesTable_NMR_NMR =
uitable(app.pValuesTab);
app.p_valuesTable_NMR_NMR.ColumnName = '';
app.p_valuesTable_NMR_NMR.RowName = {};
app.p_valuesTable_NMR_NMR.Position = [1 0 470 398];
% Create MSMSTab
app.MSMSTab = uitab(app.TabGroup);
app.MSMSTab.Title = 'MS + MS';
% Create TabGroup_MS_MS
app.TabGroup_MS_MS = uitabgroup(app.MSMSTab);
app.TabGroup_MS_MS.Position = [1 -2 471 425];
% Create CoefficientsTab_2
app.CoefficientsTab_2 = uitab(app.TabGroup_MS_MS);
app.CoefficientsTab_2.Title = 'Coefficients';
% Create PearsonTableMS_MS
app.PearsonTableMS_MS =
uitable(app.CoefficientsTab_2);
app.PearsonTableMS_MS.ColumnName = '';
app.PearsonTableMS_MS.RowName = {};
app.PearsonTableMS_MS.Position = [1 1 469 398];
% Create pValuesTab_2
app.pValuesTab_2 = uitab(app.TabGroup_MS_MS);
app.pValuesTab_2.Title = 'p-Values';
% Create p_valuesTable_MS_MS
app.p_valuesTable_MS_MS =
uitable(app.pValuesTab_2);
app.p_valuesTable_MS_MS.ColumnName = '';
app.p_valuesTable_MS_MS.RowName = {};
app.p_valuesTable_MS_MS.Position = [1 0 470 399];
% Create NMRNMRButton
app.NMRNMRButton = uibutton(app.RightPanel,
'push');
    app.NMRNMRButton.ButtonPushedFcn =
createCallbackFcn(app, @NMRNMRButtonPushed, true);
    app.NMRNMRButton.WordWrap = 'on';
    app.NMRNMRButton.FontWeight = 'bold';
    app.NMRNMRButton.Enable = 'off';
    app.NMRNMRButton.Position = [259 523 106 46];
    app.NMRNMRButton.Text = 'NMR + NMR';
% Create Label
app.Label = uilabel(app.RightPanel);
app.Label.HorizontalAlignment = 'center';
app.Label.Position = [7 571 502 22];
app.Label.Text =
'
';
% Create CorrelationCoefficientsLabel
app.CorrelationCoefficientsLabel =
uilabel(app.RightPanel);

app.CorrelationCoefficientsLabel.HorizontalAlignment =
'center';
    app.CorrelationCoefficientsLabel.FontSize = 16;
    app.CorrelationCoefficientsLabel.FontWeight =
'bold';
    app.CorrelationCoefficientsLabel.Position = [19 525
97 39];
    app.CorrelationCoefficientsLabel.Text =
{'Correlation'; 'Coefficients'};
% Create MSMSButton
app.MSMSButton = uibutton(app.RightPanel, 'push');
app.MSMSButton.ButtonPushedFcn =
createCallbackFcn(app, @MSMSButtonPushed, true);
app.MSMSButton.WordWrap = 'on';
app.MSMSButton.FontWeight = 'bold';
app.MSMSButton.Enable = 'off';
app.MSMSButton.Position = [386 523 106 46];
app.MSMSButton.Text = 'MS + MS';
% Create ButtonGroup
app.ButtonGroup = uibuttongroup(app.RightPanel);
app.ButtonGroup.Position = [135 479 357 36];
% Create PearsonButton
app.PearsonButton = uiradiobutton(app.ButtonGroup);
app.PearsonButton.Enable = 'off';
app.PearsonButton.Text = 'Pearson';
app.PearsonButton.Position = [9 7 67 22];
app.PearsonButton.Value = true;
% Create KendallButton
app.KendallButton = uiradiobutton(app.ButtonGroup);
app.KendallButton.Enable = 'off';
app.KendallButton.Text = 'Kendall';
app.KendallButton.Position = [137 7 65 22];
% Create SpearmanButton
app.SpearmanButton =
uiradiobutton(app.ButtonGroup);
app.SpearmanButton.Enable = 'off';
app.SpearmanButton.Text = 'Spearman';
app.SpearmanButton.Position = [262 7 77 22];
% Create TypeofcorrelationLabel
app.TypeofcorrelationLabel =
uilabel(app.RightPanel);
app.TypeofcorrelationLabel.HorizontalAlignment =
'center';
    app.TypeofcorrelationLabel.Position = [19 486 104
22];

```

```

        app.TypeOfcorrelationLabel.Text = 'Type of
correlation';
% Create ContextMenu_NMR_MS
app.ContextMenu_NMR_MS =
uicontextmenu(app.SCOREmetaboliteIDAppUIFigure);
% Create Copydatatoworkspace_NMRMS
app.Copydatatoworkspace_NMRMS =
uimenu(app.ContextMenu_NMR_MS);
app.Copydatatoworkspace_NMRMS.MenuSelectedFcn =
createCallbackFcn(app, @Copydatatoworkspace_NMRMSSelected,
true);
app.Copydatatoworkspace_NMRMS.Text = 'Copy data to
workspace';
% Assign app.ContextMenu_NMR_MS
app.PearsonTableMS_NMR.ContextMenu =
app.ContextMenu_NMR_MS;
app.p_ValuesTable_MS_NMR.ContextMenu =
app.ContextMenu_NMR_MS;
% Create ContextMenu_NMR_NMR
app.ContextMenu_NMR_NMR =
uicontextmenu(app.SCOREmetaboliteIDAppUIFigure);
% Create Copydatatoworkspace_NMRNMR
app.Copydatatoworkspace_NMRNMR =
uimenu(app.ContextMenu_NMR_NMR);
app.Copydatatoworkspace_NMRNMR.MenuSelectedFcn =
createCallbackFcn(app, @Copydatatoworkspace_NMRNMRMenuSelected,
true);
app.Copydatatoworkspace_NMRNMR.Text = 'Copy data to
workspace';
% Assign app.ContextMenu_NMR_NMR
app.PearsonTableNMR_NMR.ContextMenu =
app.ContextMenu_NMR_NMR;
app.p_valuesTable_NMR_NMR.ContextMenu =
app.ContextMenu_NMR_NMR;
app.PearsonTableMS_MS.ContextMenu =
app.ContextMenu_NMR_NMR;
% Create ContextMenu_CopyEDCandEMCPlot
app.ContextMenu_CopyEDCandEMCPlot =
uicontextmenu(app.SCOREmetaboliteIDAppUIFigure);
% Create CopyEDCorEMCPlot
app.CopyEDCorEMCPlot =
uimenu(app.ContextMenu_CopyEDCandEMCPlot);
app.CopyEDCorEMCPlot.MenuSelectedFcn =
createCallbackFcn(app, @CopyEDCorEMCPlotMenuSelected, true);
app.CopyEDCorEMCPlot.Text = 'Copy EDC or EMC Plot';
% Create CreatePlotofEDCandEMCMenu
app.CreatePlotofEDCandEMCMenu =
uimenu(app.ContextMenu_CopyEDCandEMCPlot);
app.CreatePlotofEDCandEMCMenu.MenuSelectedFcn =
createCallbackFcn(app, @CreatePlotofEDCandEMCMenuSelected,
true);
app.CreatePlotofEDCandEMCMenu.Text = 'Create Plot
of EDC and EMC';
% Create NormalizeEDCIntensitiesinPlotMenu
app.NormalizeEDCIntensitiesinPlotMenu =
uimenu(app.ContextMenu_CopyEDCandEMCPlot);
app.NormalizeEDCIntensitiesinPlotMenu.MenuSelectedFcn =
createCallbackFcn(app,
@NormalizeEDCIntensitiesinPlotMenuSelected, true);
app.NormalizeEDCIntensitiesinPlotMenu.Enable =
'off';
app.NormalizeEDCIntensitiesinPlotMenu.Text =
'Normalize EDC Intensities in Plot';
% Create NormalizeEMCIntensitiesinPlotMenu
app.NormalizeEMCIntensitiesinPlotMenu =
uimenu(app.ContextMenu_CopyEDCandEMCPlot);
app.NormalizeEMCIntensitiesinPlotMenu.MenuSelectedFcn =
createCallbackFcn(app,
@NormalizeEMCIntensitiesinPlotMenuSelected, true);
app.NormalizeEMCIntensitiesinPlotMenu.Enable =
'off';

        app.NormalizeEMCIntensitiesinPlotMenu.Text =
'Normalize EMC Intensities in Plot';
% Assign app.ContextMenu_CopyEDCandEMCPlot
app.EDCandEMCPlot.ContextMenu =
app.ContextMenu_CopyEDCandEMCPlot;
% Create ContextMenu_CopyMSSpectra
app.ContextMenu_CopyMSSpectra =
uicontextmenu(app.SCOREmetaboliteIDAppUIFigure);
% Create CopyPlotofMSSpectraMenu
app.CopyPlotofMSSpectraMenu =
uimenu(app.ContextMenu_CopyMSSpectra);
app.CopyPlotofMSSpectraMenu.MenuSelectedFcn =
createCallbackFcn(app, @CopyPlotofMSSpectraMenuSelected, true);
app.CopyPlotofMSSpectraMenu.Text = 'Copy Plot of MS
Spectra';
% Assign app.ContextMenu_CopyMSSpectra
app.MSSpectra.ContextMenu =
app.ContextMenu_CopyMSSpectra;
% Create ContextMenu_CopyNMRSpectra
app.ContextMenu_CopyNMRSpectra =
uicontextmenu(app.SCOREmetaboliteIDAppUIFigure);
% Create CopyPlotofNMRSpectraMenu
app.CopyPlotofNMRSpectraMenu =
uimenu(app.ContextMenu_CopyNMRSpectra);
app.CopyPlotofNMRSpectraMenu.MenuSelectedFcn =
createCallbackFcn(app, @CopyPlotofNMRSpectraMenuSelected,
true);
app.CopyPlotofNMRSpectraMenu.Text = 'Copy Plot of
NMR Spectra';
% Assign app.ContextMenu_CopyNMRSpectra
app.NMRSpectra.ContextMenu =
app.ContextMenu_CopyNMRSpectra;
% Create ContextMenu_MS_MS
app.ContextMenu_MS_MS =
uicontextmenu(app.SCOREmetaboliteIDAppUIFigure);
% Create CopydatatoworkspaceMenu_MS_MS
app.CopydatatoworkspaceMenu_MS_MS =
uimenu(app.ContextMenu_MS_MS);
app.CopydatatoworkspaceMenu_MS_MS.MenuSelectedFcn =
createCallbackFcn(app, @CopydatatoworkspaceMenu_MS_MSSelected,
true);
app.CopydatatoworkspaceMenu_MS_MS.Text = 'Copy data
to workspace';
% Assign app.ContextMenu_MS_MS
app.p_valuesTable_MS_MS.ContextMenu =
app.ContextMenu_MS_MS;
% Show the figure after all components are created
app.SCOREmetaboliteIDAppUIFigure.Visible = 'on';
end
% App creation and deletion
methods (Access = public)

% Construct app
function app = SCOREmetaboliteID
% Create UIFigure and components
createComponents(app)
% Register the app with App Designer
registerApp(app, app.SCOREmetaboliteIDAppUIFigure)
if nargin == 0
    clear app
end
end
% Code that executes before app deletion
function delete(app)
% Delete UIFigure when app is deleted
delete(app.SCOREmetaboliteIDAppUIFigure)
end
end
end

```

8.6.2 Skript zu *SCORE-metabolite-ID*²

Im Folgenden ist das Skript der MATLAB Applikation *SCORE-metabolite-ID*² angegeben, das unter Verwendung des AppDesigners in MATLAB R2022b entwickelt wurde.

```

classdef SCORE_Metabolite_ID_Squared < matlab.apps.AppBase

% Properties that correspond to app components
properties (Access = public)
    SCOREmetaboliteIDSquaredUIFigure matlab.ui.Figure
    GridLayout matlab.ui.container.GridLayout
    LeftPanel matlab.ui.container.Panel
    ChangetheBucketSizeButton matlab.ui.control.Button
    ESIionizationmodeButtonGroup
    matlab.ui.container.ButtonGroup
    negativeButton matlab.ui.control.RadioButton
    positiveButton matlab.ui.control.RadioButton
    BucketSizeEditField matlab.ui.control.EditField
    BucketSizeLabel matlab.ui.control.Label
    NormalizationButton matlab.ui.control.Button
    CalibrationButton matlab.ui.control.Button
    BucketingButton matlab.ui.control.Button
    mzvalueofEMCEditFieldLabel matlab.ui.control.Label
    mzvalueofEMCEditField
    matlab.ui.control.NumericEditField
end

```



```

CchemicalshiftEditFieldLabel matlab.ui.control.Label
CchemicalshiftEditField
matlab.ui.control.NumericEditField
HchemicalshiftEditFieldLabel matlab.ui.control.Label
HchemicalshiftEditField
matlab.ui.control.NumericEditField
MSSpectraLabel matlab.ui.control.Label
NMRspectraLabel matlab.ui.control.Label
ImportHSQCSpectraButton matlab.ui.control.Button
ImportMSSpectraButton matlab.ui.control.Button
SubtractBlankMSSpectrumButton matlab.ui.control.Button
ShowEDDCButton matlab.ui.control.Button
ShowEMCButton matlab.ui.control.Button
EDDCPlot matlab.ui.control.UIAxes
CenterPanel matlab.ui.container.Panel
PearsonTableNMR_NMR_alt matlab.ui.control.Table
PearsonTableMS_NMR_alt matlab.ui.control.Table
ContourlevelsSlider matlab.ui.control.Slider
ContourlevelsLabel matlab.ui.control.Label
uppermzEditField matlab.ui.control.NumericEditField
lowermzEditField matlab.ui.control.NumericEditField
mzregionLabel matlab.ui.control.Label
mzLabel matlab.ui.control.Label
FractionstoplotListBoxLabel matlab.ui.control.Label
FractionstoplotListBox matlab.ui.control.ListBox
EditField_2Label matlab.ui.control.Label
NMRspectra matlab.ui.control.UIAxes
MSSpectra matlab.ui.control.UIAxes
RightPanel matlab.ui.container.Panel
TypeofcorrelationLabel matlab.ui.control.Label
ButtonGroup matlab.ui.container.ButtonGroup
SpearmanButton matlab.ui.control.RadioButton
KendallButton matlab.ui.control.RadioButton
PearsonButton matlab.ui.control.RadioButton
MSMSButton matlab.ui.control.Button
CorrelationCoefficientsLabel matlab.ui.control.Label
NMRNMRButton matlab.ui.control.Button
TabGroup matlab.ui.container.TabGroup
NMRMSTab matlab.ui.container.Tab
TabGroup_MS_NMR matlab.ui.container.TabGroup
Coefficients matlab.ui.container.Tab
PearsonTableMS_NMR matlab.ui.control.Table
p_values matlab.ui.container.Tab
p_valuesTable_MS_NMR matlab.ui.control.Table
NMRNMRTab matlab.ui.container.Tab
TabGroup_NMR_NMR matlab.ui.container.TabGroup
CoefficientsTab matlab.ui.container.Tab
PearsonTableNMR_NMR matlab.ui.control.Table
pValuesTab matlab.ui.container.Tab
p_valuesTable_NMR_NMR matlab.ui.control.Table
MSMSTab matlab.ui.container.Tab
TabGroup_MS_MS matlab.ui.container.TabGroup
CoefficientsTab_MS_MS matlab.ui.container.Tab
PearsonTableMS_MS matlab.ui.control.Table
pValuesTab_MSMS matlab.ui.container.Tab
p_valuesTable_MS_MS matlab.ui.control.Table
NMRMSButton matlab.ui.control.Button
Label matlab.ui.control.Label
EMCsLabel matlab.ui.control.Label
EDDCsLabel matlab.ui.control.Label
ShowallcorrelatedEMCstoEMCButton
matlab.ui.control.Button
ShowallcorrelatedEDDCstoEDDCButton
matlab.ui.control.Button
ShowallcorrelatedEMCsButton matlab.ui.control.Button
ShowallcorrelatedEDDCsButton matlab.ui.control.Button
ClearPlotButton matlab.ui.control.Button
PlotEMCButton matlab.ui.control.Button
PlotEDDCButton matlab.ui.control.Button
AddEMCButton matlab.ui.control.Button
AddEDDCButton matlab.ui.control.Button
DeleteEMCButton matlab.ui.control.Button
ListBoxEMCPlot matlab.ui.control.ListBox
DeleteEDDCButton matlab.ui.control.Button
ListBoxEDDCPlot matlab.ui.control.ListBox
EDDCandEMCPlot matlab.ui.control.UIAxes
ContextMenu_NMR_MS matlab.ui.container.ContextMenu
Copydatatoworkspace_NMRMS matlab.ui.container.Menu
ContextMenu_NMR_NMR matlab.ui.container.ContextMenu
Copydatatoworkspace_NMRNMR matlab.ui.container.Menu
ContextMenu_CopyNMRMSpectra matlab.ui.container.ContextMenu
CopyPlotofNMRMSpectraMenu matlab.ui.container.Menu
ContextMenu_CopyMSSpectra matlab.ui.container.ContextMenu
CopyPlotofMSSpectraMenu matlab.ui.container.Menu
ContextMenu_CopyEDDCandEMCPlot matlab.ui.container.ContextMenu
CopyEDDCorEMCPlotMenu matlab.ui.container.Menu
CreatePlotofEDDCandEMCMenu matlab.ui.container.Menu
NormalizeEDDCIntensitiesinPlotMenu
matlab.ui.container.Menu
NormalizeEMCIntensitiesinPlotMenu
matlab.ui.container.Menu
ContextMenu_CopyEDDCPlot matlab.ui.container.ContextMenu
CopyEDDCPlotMenu matlab.ui.container.Menu
ContextMenu_MS_MS matlab.ui.container.ContextMenu
CopydatatoworkspaceMenu_MS_MS matlab.ui.container.Menu
end
% Properties that correspond to apps with auto-reflow
properties (Access = private)
    onePanelWidth = 576;
    twoPanelWidth = 768;
end
properties (Access = private)
    nmr % intensities of NMR spectra, one nmr spectra per
column
    ppmProton % matrix of chemical shift values of 1H of
NMR spectra
    ppmCarbon % matrix of chemical shift values of 13C of
NMR spectra
    nmldata % data structure of NMR spectra after import
via rbnmr
    fractions % row vector from 1 to total number of
fractions
    Spectra % 3D array of size (13C, 1H, fraction)
containing the data points of HSQC spectra
    SpectraSum %Summe der HSQC-Datenpunkte über die
ausgewählten Fraktionen (size (256*2048))
    CAL1H %Chemische Verschiebung für Kalibrierungsstandard
in 1H-Domäne
    CAL13C %Chemische Verschiebung für
Kalibrierungsstandard in 13C-Domäne
    SpectraSumBuckets %Summe der HSQC-Datenpunkte über die
ausgewählten Fraktionen nach Bucketing
    indexplotnmr % required for NMR plot
    XAxis % ppm values for all NMR spectra each (matrix)
    mz % m/z values after import as txt file (not sorted)
    intensity % intensity values belonging to m/z values
    indexplotms % required for plot of MS spectra
    mzall % column vector with all different m/z values
    intensityall % intensities of MS spectra sorted (one
m/z value per row)
    intensitysub % intensities of MS spectra after blank
subtraction (one m/z value per row)
    fractionslistbox %fractions to plot
    EDCvalueProton % Value of EDC of 1H chemical shift in
Edit Field Box
    ZeileEDCvalueProton % Zeilennummer, in der der nächste
ppm-Wert bezogen auf eingegebenen EDC-Wert der 1H-Verschiebung
steht
    EDCvalueCarbon % Value of EDC of 13C chemical shift in
Edit Field Box
    EDDCPlotProton %Zeilennummern der Proton-Verschiebung
der Werte, die in EDDCListBox gelistet sind
    ZeileEDCvalueCarbon %Zeilennummer, in der der nächste
ppm-Wert bezogen auf eingegebenen EDC-Wert der 13C-Verschiebung
steht
    ListBoxEDDCPlotColumn %Spaltennummer, in der die 13C-
Zeilennummer in der EDDCListBox.ItemsData steht
    EMCvalue % eingegebener m/z-Wert in EMC-EditField Box
für Plot
    ZeileEMCvalue % Zeilennummer, in der der nächste m/z-
Wert bezogen auf eingegebenen EMC-Wert steht
    nmrbuckets % NMR-Daten (Intensitätswerte) nach
Bucketing
    ppmbuckets % ppm-Skala nach Bucketing
    nmrlimit % lower ppm value (1,1) and upper limit (1,2)
for NMR plot in specific ppm range
    mzlimit % lower m/z value (1,1) and upper limit (1,2)
for MS plot in specific m/z range
    wholeEMC % EMC in "larger" m/z value field, from
ZeileEMCvalue-1 to ZeileEMCvalue+1 (summed intensities)
    HSQC_Colormap %Colormap für HSQC Spektren: negativ =
orange, positiv = blau
end
% Callbacks that handle component events
methods (Access = private)

    Button pushed function: ImportHSQCSpectraButton
    function ImportHSQCSpectraButtonPushed(app, event)
        dirnmr = uigetdir('','Please select 2D NMR data');
        d1 =
uiprogressdlg(app.SCOREmetaboliteIDSquaredUIFigure, 'Title',
'Please Wait', 'Message', 'Opening the NMR Spectra...',
'Confirmation in Matlab Command Window may be required'),
'Indeterminate', 'on');
        drawnow
        app.nmrdata = rbnmr(dirnmr);
        for k = 1:size(app.nmrdata,1)
            app.ppmProton(:,k) = app.nmrdata{k,1}.XAxis;
        end
        for k = 1:size(app.nmrdata,1)
            app.ppmCarbon(:,k) = app.nmrdata{k,1}.YAxis;
        end
        app.fractions = 1:size(app.nmrdata,1);
        close(d1);
        d2 =
uiprogressdlg(app.SCOREmetaboliteIDSquaredUIFigure, 'Title',
'Please Wait', 'Message', 'Preparing the NMR data...',
'Indeterminate', 'on');
        %Bildung des multidimensional arrays (j,k,i), also (13C, 1H,
%Fraction)
        for i = 1:size(app.fractions,2) %i=Fraktionszähler

```

```

        for j = 1:size(app.nmrdata{1,1}.Data, 1) %j =
Zähler für y-Achse, also 13C Verschiebung
            for k = 1:size(app.nmrdata{1,1}.Data, 2) %k
= Zähler für x-Achse, also 1H Verschiebung
                app.Spectra(j,k,1) =
app.nmrdata{1,1}.Data(j,k); %Spectra = multidim. array of size
(256*2048*fraction)
            end
        end
        end
        close(d2);
        d3 =
uiprogressdlg(app.SCOREmetaboliteIDSquaredUIFigure, 'Title',
'Please Wait', 'Message', 'Plotting the HSQC spectra...',
'Indeterminate', 'on');
%Bildung der Summe der HSQC-Spektren über ausgewählten Bereich
%der Fraktionen:
        app.SpectraSum = sum(app.Spectra(:, :,
(min(app.fractions):max(app.fractions))),3); %matrix of size
(256*2048)
        assignin('base', 'SpectraSum', app.SpectraSum);
        assignin('base', 'Spectra', app.Spectra);
        assignin('base', 'ChemicalShiftCarbon',
app.ppmCarbon);
        assignin('base', 'ChemicalShiftProton',
app.ppmProton);
%Slider für Contour levels auf Start-Wert setzen:
        value = app.ContourlevelsSlider.Value;
%2D-Contour Plot der HSQC-Spektren:
        contour(app.NMRSpectra, app.ppmProton(:,1),
app.ppmCarbon(:,1), app.SpectraSum, value);
%HSQC-Colormap erstellen:
        app.HSQC_Colormap = [0 0.4470 0.7410; 0.8500 0.3250
0.0980]; %positiv = blau, negativ = orange
%set(app.NMRSpectra, 'colormap', app.HSQC_Colormap);
        set(app.NMRSpectra, 'colormap', winter);
        set(app.ImportHSQCSpectraButton, 'Enable', 'off');
        set(app.ImportMSSpectraButton, 'Enable', 'on');
        set(app.CalibrationButton, 'Enable', 'on');
        close(d3);
    end

% Button pushed function: CalibrationButton
function CalibrationButtonPushed(app, event)
    calibration.nmr = app.Spectra;
    calibration.ppmProton = app.ppmProton;
    calibration.ppmCarbon = app.ppmCarbon;
    calibration.cal1H = inputdlg({'Enter chemical shift
for calibration on 1H axis'}, 'Calibration', [1 50],
{'0.000'});
        calibration.cal13C = inputdlg({'Enter chemical
shift for calibration on 13C axis'}, 'Calibration', [1 50],
{'0.000'});
        SW1H = app.nmrdata{1,1}.Acqus.SW;
        SW13C = app.nmrdata{1,1}.Acqu2s.SW;
        TD1H = size(app.ppmProton,1);
        TD13C = size(app.ppmCarbon,1);
        calibration.RES1H = SW1H/TD1H;
        calibration.RES13C = SW13C/TD13C;

%Input-Dialoge für oberes und unteres Limit, in dem sich
%TSP-Signal befindet (1H-Domäne)
        edit_up1H = inputdlg({'Upper limit for calibration
on 1H axis'}, 'Upper Limit', [1 50], {'0.50'});
        calibration.lim1H(1,1) = str2double(edit_up1H);
        calibration.a = find(calibration.ppmProton(:,1) <
calibration.lim1H(1,1),1,'first');
        edit_low1H = inputdlg({'Lower limit for calibration
on 1H axis'}, 'Lower Limit', [1 50], {'-0.50'});
        calibration.lim1H(1,2) = str2double(edit_low1H);
        calibration.b = find(calibration.ppmProton(:,1) >
calibration.lim1H(1,2),1,'last');
%Input-Dialoge für oberes und unteres Limit, in dem sich
%TSP-Signal befindet (13C-Domäne)
        edit_up13C = inputdlg({'Upper limit for calibration
on 13C axis'}, 'Upper Limit', [1 50], {'1.50'});
        calibration.lim13C(1,1) = str2double(edit_up13C);
        calibration.c = find(calibration.ppmCarbon(:,1) <
calibration.lim13C(1,1),1,'first');
        edit_low13C = inputdlg({'Lower limit for
calibration on 13C axis'}, 'Lower Limit', [1 50], {'-1.50'});
        calibration.lim13C(1,2) = str2double(edit_low13C);
        calibration.d = find(calibration.ppmCarbon(:,1) >
calibration.lim13C(1,2),1,'last');
        app.CAL1H = str2double(calibration.cal1H);
        app.CAL13C = str2double(calibration.cal13C);
        lowlim1H = min(calibration.lim1H);
        uplim1H = max(calibration.lim1H);
        lowlim13C = min(calibration.lim13C);
        uplim13C = max(calibration.lim13C);

%Kalibrierung der Achsen (Maximalwert in 1H und 13C Dimension
%finden)
        %index = zeros(1, size(app.fractions,2));
        for k = 1:size(calibration.nmr,3) %k =
Fraktionszähler
            a1 = find(app.ppmProton(:,k) >
lowlim1H,1,'last');
            b1 = find(app.ppmProton(:,k) <
uplim1H,1,'first');
            c1 = find(app.ppmCarbon(:,k) >
lowlim13C,1,'last');
            d1 = find(app.ppmCarbon(:,k) <
uplim13C,1,'first');
            [~, index(1,k)] = max(calibration.nmr(d1:c1,
b1:a1, k), [], 'all', 'linear'); %finde Maximalwert zwischen a1
und b1 pro Spektrum (k), suche aber nur ZWISCHEN a1 und b1,
Zeilennummer in Matrix Index speichern
            [row(1,k), col(1,k)] =
ind2sub(size(calibration.nmr(d1:c1,b1:a1,k)), index(1,k));
            row(2,k) = (row(1,k)+d1-1);
            col(2,k) = (col(1,k)+b1-1); %Anpassung der
Index-Werte auf Gesamtskala und nicht nur zwischen a1 und b1
        end
        ND1H = max(col(2,:))+TD1H-min(col(2,:));
%Gesamtanzahl benötigter Datenpunkte für gesamte 1H-Achse
        XAxis = zeros(ND1H,size(app.fractions,2));
%Preallocation
        ND13C = max(row(2,:))+TD13C-min(row(2,:));
%Gesamtanzahl benötigter Datenpunkte für gesamte 13C-Achse
        YAxis = zeros(ND13C,size(app.fractions,2));
%Preallocation
        %für 1H-Domäne:
        for k = 1:size(calibration.nmr,3) %Fraktionszähler
            XAxis(1:(max(col(2,:))-1),k) =
[(app.CAL1H+calibration.RES1H*(max(col(2,:))-1))-
calibration.RES1H:(app.CAL1H+calibration.RES1H)]; %x-Achse vor
Indexwert
            XAxis((max(col(2,:))+1):ND1H,k) = [(app.CAL1H-
calibration.RES1H):-calibration.RES1H:(app.CAL1H-
calibration.RES1H*(TD1H-min(col(2,:))))]; %x-Achse nach
Indexwert
        end % -> alle x-Achsen jeweils gleich, aber
Intensitäten noch nicht zugeordnet
        %für 13C-Domäne:
        for k = 1:size(calibration.nmr,3) %Fraktionszähler
            YAxis(1:(max(row(2,:))-1),k) =
[(app.CAL13C+calibration.RES13C*(max(row(2,:))-1))-
calibration.RES13C:(app.CAL13C+calibration.RES13C)]; %x-Achse
vor Indexwert
            YAxis((max(row(2,:))+1):ND13C,k) = [(app.CAL13C-
calibration.RES13C):-calibration.RES13C:(app.CAL13C-
calibration.RES13C*(TD13C-
min(row(2,:))))]; %x-Achse nach Indexwert
        end
        nmr = zeros(size(calibration.nmr));
        for k = 1:size(calibration.nmr,3) %k =
Fraktionszähler
            nmr((max(row(2,:))-row(2,k))+1:(TD13C-
row(2,k)+max(row(2,:))),(max(col(2,:))-col(2,k))+1:(TD1H-
col(2,k)+max(col(2,:))),k) = calibration.nmr(:,k); %nmr-Werte
ab bestimmter Zeilennummer zuordnen
        end
        app.Spectra = nmr; %Überschreibung der app.Spectra-
Matrix mit der nmr-Matrix, die die kalibrierten Spektren
enthält
        app.ppmProton = XAxis;
        app.ppmCarbon = YAxis;
%Bildung der Summe der HSQC-Spektren über ausgewählten Bereich
%der Fraktionen:
        app.SpectraSum = sum(app.Spectra(:, :,
(min(app.fractions):max(app.fractions))),3);
        assignin('base', 'SpectraSum', app.SpectraSum);
        assignin('base', 'Spectra', app.Spectra);
        assignin('base', 'ChemicalShiftCarbon',
app.ppmCarbon);
        assignin('base', 'ChemicalShiftProton',
app.ppmProton);
%PLOT:
        %Slider für Contour levels auf Start-Wert setzen:
        value = app.ContourlevelsSlider.Value;
        %2D-Contour Plot der HSQC-Spektren:
        contour(app.NMRSpectra, app.ppmProton(:,1),
app.ppmCarbon(:,1), app.SpectraSum, value);
        %HSQC-Colormap erstellen:
        app.HSQC_Colormap = [0 0.4470 0.7410; 0.8500 0.3250
0.0980]; %positiv = blau, negativ = orange
        set(app.NMRSpectra, 'colormap', winter);
        set(app.CalibrationButton, 'Enable', 'off');
        set(app.NormalizationButton, 'Enable', 'on');
    end

% Button pushed function: NormalizationButton
function NormalizationButtonPushed(app, event)
    x = find(app.ppmProton(:,1) == app.CAL1H);
    y = find(app.ppmCarbon(:,1) == app.CAL13C);
    for i = 1:size(app.Spectra, 3)
        tsp(1,i) = app.Spectra(y,x,i);
    end
    SpectraNorm = zeros(size(app.Spectra));
%Preallocation
    for k = 1:size(app.Spectra,3)

```

```

        SpectraNorm(:, :, k) =
app.Spectra(:, :, k) ./ tsp(1, k);
    end
    app.Spectra = SpectraNorm;
    set(app.NormalizationButton, 'Enable', 'off');
    set(app.BucketingButton, 'Enable', 'on');
end

% Button pushed function: BucketingButton
function BucketingButtonPushed(app, event)
%Abfrage, ob Bucketing durchgeführt werden soll
    title = 'Bucketing';
    msg = 'Shall Bucketing of the NMR Spectra be
performed?';
    selection = questdlg(msg, title, 'Yes', 'No',
'Yes');
    switch selection
        case 'Yes'
            %Abfrage für Bucket-Size
            bucketsize1H = inputdlg({'Enter Bucket size
for 1H dimension in ppm (e.g. 0.01 ppm)'}, 'Bucket Size', [1
50], {'0.01'});
            bucketsize1H = str2double(bucketsize1H);
            numberOfBuckets = (max(app.ppmProton) -
min(app.ppmProton)) ./ bucketsize1H;
            numberOfBuckets = round(numberofBuckets,
0); %auf ganze Zahl runden
            for i = 1:size(app.fractions, 2)
                ppmbins(:, i) =
discretize(app.ppmProton(:, i), numberOfBuckets(1, i));
            %Spaltenvektor mit Index-Werten je nach Bucketanzahl
            end
            for i = 1:size(app.Spectra, 3) %Zähler über
Fraktionen
                for j = 1:size(app.Spectra, 1) %Zähler
über 13C
                    app.nmrBuckets(j, :, i) =
accumarray(ppmbins(:, i), app.Spectra(j, :, i)); %Summiert Werte
in 3-dimensionaler Matrix "spectra" entsprechend Index in
ppmbins
                end
            end
            for i = 1:size(app.ppmProton, 2)
                app.ppmBuckets(:, i) =
accumarray(ppmbins(:, i), app.ppmProton(:, i), [], @mean); %ppm-
Skala-Erstellung durch Mittelwertbildung der ppm-Werte
entsprechend der Bin-Indizes
            end
            app.SpectraSumBuckets =
sum(app.nmrBuckets(:, :,
(min(app.fractions):max(app.fractions))), 3); %Summe der
Spectra-Matrix über die Fraktionen zum Plotten
            set(app.BucketsizeEditField, 'Enable',
'on');
            app.BucketsizeEditField.Value =
(string(bucketsize1H) + ' ppm');
            assignin('base', 'BucketSize',
bucketsize1H);
            case 'No'
                app.nmrBuckets = app.Spectra;
                app.ppmBuckets = app.ppmProton;
                app.SpectraSumBuckets = app.SpectraSum;
                set(app.BucketsizeEditField, 'Enable',
'on');
                app.BucketsizeEditField.Value = ('--');
            end
            %Plot-Update
            value = app.ContourLevelsSlider.Value;
            contour(app.NMRSpectra, app.ppmBuckets(:, 1),
app.ppmCarbon(:, 1), app.SpectraSumBuckets, value);
            set(app.NMRSpectra, 'colormap', winter);
            assignin('base', 'SpectraAfterBucketing',
app.nmrBuckets);
            assignin('base', 'SpectraSumAfterBucketing',
app.SpectraSumBuckets);
            assignin('base',
'ChemicalShiftProtonAfterBucketing', app.ppmBuckets);
            set(app.HchemicalShiftEditField, 'Enable', 'on');
            set(app.CchemicalShiftEditField, 'Enable', 'on');
            set(app.ShowEDDCButton, 'Enable', 'on');
            set(app.AddEDDCButton, 'Enable', 'on');
            set(app.ListBoxEDDCPlot, 'Enable', 'on');
            set(app.ContourLevelsSlider, 'Enable', 'on');
            set(app.BucketingButton, 'Visible', 'off');
            set(app.ChangetheBucketsizeButton, 'Visible',
'on');
            set(app.ChangetheBucketsizeButton, 'Enable', 'on');
        end

% Button pushed function: ChangetheBucketsizeButton
function ChangetheBucketsizeButtonPushed(app, event)
    %Abfrage für Bucket-Size
    bucketsize1H = inputdlg({'Enter Bucket size
for 1H dimension in ppm (e.g. 0.01 ppm)'}, 'Bucket Size', [1
50], {'0.01'});
    bucketsize1H = str2double(bucketsize1H);
    numberOfBuckets = [];
    app.ppmBuckets = [];

        app.nmrBuckets = [];
        numberOfBuckets = (max(app.ppmProton) -
min(app.ppmProton)) ./ bucketsize1H;
        numberOfBuckets = round(numberofBuckets,
0); %auf ganze Zahl runden
        for i = 1:size(app.fractions, 2)
            ppmbins(:, i) =
discretize(app.ppmProton(:, i), numberOfBuckets(1, i));
        %Spaltenvektor mit Index-Werten je nach Bucketanzahl
        end
        for i = 1:size(app.Spectra, 3) %Zähler über
Fraktionen
            for j = 1:size(app.Spectra, 1) %Zähler
über 13C
                app.nmrBuckets(j, :, i) =
accumarray(ppmbins(:, i), app.Spectra(j, :, i)); %Summiert Werte
in 3-dimensionaler Matrix "spectra" entsprechend Index in
ppmbins
            end
        end
        for i = 1:size(app.ppmProton, 2)
            app.ppmBuckets(:, i) =
accumarray(ppmbins(:, i), app.ppmProton(:, i), [], @mean); %ppm-
Skala-Erstellung durch Mittelwertbildung der ppm-Werte
entsprechend der Bin-Indizes
        end
        app.SpectraSumBuckets =
sum(app.nmrBuckets(:, :,
(min(app.fractions):max(app.fractions))), 3); %Summe der
Spectra-Matrix über die Fraktionen zum Plotten
        set(app.BucketsizeEditField, 'Enable',
'on');
        app.BucketsizeEditField.Value =
(string(bucketsize1H) + ' ppm');
        assignin('base', 'BucketSize',
bucketsize1H);
        case 'No'
            app.nmrBuckets = app.Spectra;
            app.ppmBuckets = app.ppmProton;
            app.SpectraSumBuckets = app.SpectraSum;
            set(app.BucketsizeEditField, 'Enable',
'on');
            app.BucketsizeEditField.Value = ('--');
        end
        %Plot-Update
        value = app.ContourLevelsSlider.Value;
        contour(app.NMRSpectra, app.ppmBuckets(:, 1),
app.ppmCarbon(:, 1), app.SpectraSumBuckets, value);
        set(app.NMRSpectra, 'colormap', winter);
        assignin('base', 'SpectraAfterBucketing',
app.nmrBuckets);
        assignin('base', 'SpectraSumAfterBucketing',
app.SpectraSumBuckets);
        assignin('base',
'ChemicalShiftProtonAfterBucketing', app.ppmBuckets);
        set(app.HchemicalShiftEditField, 'Enable', 'on');
        set(app.CchemicalShiftEditField, 'Enable', 'on');
        set(app.ShowEDDCButton, 'Enable', 'on');
        set(app.AddEDDCButton, 'Enable', 'on');
        set(app.ListBoxEDDCPlot, 'Enable', 'on');
        set(app.ContourLevelsSlider, 'Enable', 'on');
        set(app.BucketingButton, 'Visible', 'off');
        set(app.ChangetheBucketsizeButton, 'Visible',
'on');
        set(app.ChangetheBucketsizeButton, 'Enable', 'on');
    end

% Button pushed function: ChangetheBucketsizeButton
function ChangetheBucketsizeButtonPushed(app, event)
    %Abfrage für Bucket-Size
    bucketsize1H = inputdlg({'Enter Bucket size
for 1H dimension in ppm (e.g. 0.01 ppm)'}, 'Bucket Size', [1
50], {'0.01'});
    bucketsize1H = str2double(bucketsize1H);
    numberOfBuckets = [];
    app.ppmBuckets = [];

        app.nmrBuckets = [];
        numberOfBuckets = (max(app.ppmProton) -
min(app.ppmProton)) ./ bucketsize1H;
        numberOfBuckets = round(numberofBuckets,
0); %auf ganze Zahl runden
        for i = 1:size(app.fractions, 2)
            ppmbins(:, i) =
discretize(app.ppmProton(:, i), numberOfBuckets(1, i));
        %Spaltenvektor mit Index-Werten je nach Bucketanzahl
        end
        for i = 1:size(app.Spectra, 3) %Zähler über
Fraktionen
            for j = 1:size(app.Spectra, 1) %Zähler
über 13C
                app.nmrBuckets(j, :, i) =
accumarray(ppmbins(:, i), app.Spectra(j, :, i)); %Summiert Werte
in 3-dimensionaler Matrix "spectra" entsprechend Index in
ppmbins
            end
        end
        for i = 1:size(app.ppmProton, 2)
            app.ppmBuckets(:, i) =
accumarray(ppmbins(:, i), app.ppmProton(:, i), [], @mean); %ppm-
Skala-Erstellung durch Mittelwertbildung der ppm-Werte
entsprechend der Bin-Indizes
        end
        app.SpectraSumBuckets =
sum(app.nmrBuckets(:, :,
(min(app.fractions):max(app.fractions))), 3); %Summe der
Spectra-Matrix über die Fraktionen zum Plotten
        set(app.BucketsizeEditField, 'Enable',
'on');
        app.BucketsizeEditField.Value =
(string(bucketsize1H) + ' ppm');
        assignin('base', 'BucketSize',
bucketsize1H);
        case 'No'
            app.nmrBuckets = app.Spectra;
            app.ppmBuckets = app.ppmProton;
            app.SpectraSumBuckets = app.SpectraSum;
            set(app.BucketsizeEditField, 'Enable',
'on');
            app.BucketsizeEditField.Value = ('--');
        end
        %Plot-Update
        value = app.ContourLevelsSlider.Value;
        contour(app.NMRSpectra, app.ppmBuckets(:, 1),
app.ppmCarbon(:, 1), app.SpectraSumBuckets, value);
        set(app.NMRSpectra, 'colormap', winter);
        assignin('base', 'SpectraAfterBucketing',
app.nmrBuckets);
        assignin('base', 'SpectraSumAfterBucketing',
app.SpectraSumBuckets);
        assignin('base',
'ChemicalShiftProtonAfterBucketing', app.ppmBuckets);
        set(app.HchemicalShiftEditField, 'Enable', 'on');
        set(app.CchemicalShiftEditField, 'Enable', 'on');
        set(app.ShowEDDCButton, 'Enable', 'on');
        set(app.AddEDDCButton, 'Enable', 'on');
        set(app.ListBoxEDDCPlot, 'Enable', 'on');
        set(app.ContourLevelsSlider, 'Enable', 'on');
        set(app.BucketingButton, 'Visible', 'off');
        set(app.ChangetheBucketsizeButton, 'Visible',
'on');
        set(app.ChangetheBucketsizeButton, 'Enable', 'on');
    end

% Button pushed function: ImportMSSpectraButton
function ImportMSSpectraButtonPushed(app, event)
    dirms = uigetdir('', 'Please select MS data');
    cd(dirms);

    numfiles = inputdlg({'Number of files used for
correlation'}, 'Number of files', [1 50]); % if-Schleife
    einbauen, wenn hier anderer Wert, als bei NMR
    numfiles = str2double(numfiles{1});
    if numfiles ~= size(app.fractions)
        msgbox('An Error occurred: Number of files for
NMR and MS must be the same', 'Error', 'error')
    end
    d =
uiprogressdlg(app.SCOREmetaboliteIDSquaredUIFigure, 'Title',
'Please Wait', 'Message', 'Opening the MS Spectra',
'Indeterminate', 'on');
    drawnow
    txtfiles=dir('*.txt');
    msdata=cell(numfiles, 1);
    for k = 1:numfiles
        msdata{k} = importdata(txtfiles(k).name);
    end
    % Maximale Zeilen-Anzahl der MS-Daten finden
    x(size(msdata, 1), 2) = zeros; %Preallocation
    for k = 1:size(msdata)
        x(k, :) = size(msdata{k, 1}.data);
    end
    y = max(x(:, 1));
    % m/z-Werte in Matrix laden
    app.mz = zeros(y, numfiles); % leere Matrix für mz-
Werte
    for k = 1:numfiles
        app.mz(1:x(k, 1), k) = msdata{k, 1}.data(:, 1);
    end
    % Intensitätswerte in Matrix laden
    app.intensity = zeros(y, numfiles); % leere Matrix
für Intensitätswerte der MS-Spektren
    for k = 1:numfiles

```

```

        app.intensity(1:x(k,1), k) =
msdata{k,1}.data(:,2);
    end
%Anpassung der MS-Spektren, sodass pro Zeile die
%Intensitätswerte der verschiedenen Fraktionen bei jeweils
%gleichen m/z-Werten stehen
app.mzall = unique(app.mz); %Erstellung eines
Spaltenvektors mit allen versch. m/z-Werten
app.mzall(app.mzall == 0, :) = []; %Löschen der
Zeile mit m/z = 0
%Zuordnung der Intensitätswerte zu den m/z-Werten
% m/z-Wert in Spalte 1, Fraktionen in Spalten 2 bis x
[~, ia] = ismember(app.mz, app.mzall);
highestvalue = max(ia, [], 'all');
ia(ia == 0) = highestvalue + 1;
app.intensityall = zeros((size(app.mzall,1)+1),
size(app.intensity,2));
for i = 1:size(ia)
    for j = 1:size(ia,2)
        if app.intensityall(ia(i,j), j) == 0
            app.intensityall(ia(i,j), j) =
app.intensity(i,j);
        else
            app.intensityall(ia(i,j), j) =
(app.intensityall(ia(i,j), j) + app.intensity(i,j));
        end
    end
end
app.intensityall(highestvalue+1,:) = [];
% mzall nicht als Vektor, sondern in so vielen Spalten, wie auch
% Fraktionen vorhanden sind
for i = 1:size(app.intensityall, 2)
    app.mzall(:,i) = app.mzall(:,1);
end
%Plot
app.indexplotms(1,1) = size(app.intensityall,1);
app.indexplotms(1,2) = 1;
[X, ~] = meshgrid(app.fractions,
min(app.indexplotms):max(app.indexplotms));
plot3(app.MSSpectra, X,
app.mzall(min(app.indexplotms):max(app.indexplotms)),
app.intensityall(min(app.indexplotms):max(app.indexplotms),
app.fractions));
xticks(app.MSSpectra, 1:size(app.intensityall, 2));
set(app.SubtractBlankMSSpectrumButton, 'Enable',
'on');
set(app.ImportMSSpectraButton, 'Enable', 'off');
%List Box Fractions to plot
set(app.FractionstoplotListBox, 'Enable', 'on');
app.fractionslistbox = string(app.fractions);
app.FractionstoplotListBox.Items =
app.fractionslistbox;
app.FractionstoplotListBox.ItemsData =
app.fractions;
close(d);
end

% Button pushed function: SubtractBlankMSSpectrumButton
function SubtractBlankMSSpectrumButtonPushed(app,
event)
    d =
uiprogressdlg(app.SCOREmetaboliteIDSquaredUIFigure, 'Title',
'Please Wait', 'Message', 'Preparing the MS Spectra',
'Indeterminate', 'on');
drawnow
title = 'MS Blank Subtraction';
msg = 'Should a subtraction of a MS Blank Spectrum
be performed?';
selection = questdlg(msg, title, 'Yes', 'No',
'Yes');
switch selection
    case 'Yes'
        blank = uimimport('-file');
        [C, ia] = ismember(app.mzall,
blank.data(:,1));
        app.intensitysub =
zeros(size(app.intensityall, 1), size(app.intensityall, 2));
%intensitysub = Intensitätswerte nach Subtraktion des Blanks
%Subtraktion der Intensitätswerte des Blanks der
Intensitätswerte bei den
%zugehörigen m/z-Verhältnissen
for i = 1:size(app.intensityall,1)
    for k = 1:size(app.intensityall,2)
        if C(i,k) == 1
            app.intensitysub(i,k) =
app.intensityall(i,k) - blank.data(ia(i,k), 2);
        else
            app.intensitysub(i,k) =
app.intensityall(i,k);
        end
    end
end

%wenn Werte in "intensitysub" kleiner 0, dann gleich 0 setzen
app.intensitysub(app.intensitysub < 0) = 0;
case 'No'
    app.intensitysub = app.intensityall;
end

%Plot
app.indexplotms(1,1) = size(app.intensitysub,1);
[X, ~] = meshgrid(app.fractions,
min(app.indexplotms):max(app.indexplotms));
plot3(app.MSSpectra, X,
app.mzall(min(app.indexplotms):max(app.indexplotms)),
app.intensitysub(min(app.indexplotms):max(app.indexplotms),
app.fractions));
xticks(app.MSSpectra, 1:size(app.intensitysub, 2));
title = 'ESI ionization mode';
msg = 'Please specify the ESI ionization mode';
selection = questdlg(msg, title, 'positive',
'negative', 'positive');
switch selection
    case 'positive'
        app.negativeButton.Value = false;
        app.positiveButton.Value = true;
    case 'negative'
        app.negativeButton.Value = true;
        app.positiveButton.Value = false;
end
set(app.SubtractBlankMSSpectrumButton, 'Enable',
'off');
set(app.mzvalueofEMCEditField, 'Enable', 'on');
set(app.ShowEMCButton, 'Enable', 'on');
set(app.AddEMCButton, 'Enable', 'on');
set(app.ListBoxEMCPlot, 'Enable', 'on');
assignin('base', 'MSSpectra_Intensity',
app.intensitysub);
assignin('base', 'MSSpectra_mzValues', app.mzall);
close(d);
end

% Value changed function: FractionstoplotListBox
function FractionstoplotListBoxValueChanged(app, event)
%einstellen, welche Fraktionen für NMR und MS geplottet werden
app.fractions = app.FractionstoplotListBox.Value;
%SpectraSum (summierte HSQC-Spektren) neu berechnen:
app.SpectraSumBuckets = sum(app.nmrbuckets(:, :,
(min(app.fractions):max(app.fractions))),3); %matrix of size
(256*2048)
%Slider für Contour levels ablesen:
value = app.ContourlevelsSlider.Value;
%2D-Contour Plot der HSQC-Spektren:
contour(app.NMRSpectra, app.ppmbuckets(:,1),
app.ppmCarbon(:,1), app.SpectraSumBuckets, value);
set(app.NMRSpectra, 'colormap', app.HSQC_Colormap);
%Plot der Blank-subtrahierten MS-Spektren
[X, ~] = meshgrid(app.fractions,
min(app.indexplotms):max(app.indexplotms));
plot3(app.MSSpectra, X,
app.mzall(min(app.indexplotms):max(app.indexplotms)),
app.intensitysub(min(app.indexplotms):max(app.indexplotms),
app.fractions));
xticks(app.MSSpectra, 1:size(app.intensitysub, 2));
end

% Value changed function: HchemicalshiftEditField
function HchemicalshiftEditFieldValueChanged(app,
event)
    app.EDCvalueProton =
app.HchemicalshiftEditField.Value;
%Finde nächsten ppm-Wert, für den ein Datenpunkt in der
%NMR-Matrix (Spectra) vorhanden ist
[I app.ZeileEDCvalueProton] =
min(abs(app.ppmbuckets(:,1) - app.EDCvalueProton));
end

% Value changed function: CchemicalshiftEditField
function CchemicalshiftEditFieldValueChanged(app,
event)
    app.EDCvalueCarbon =
app.CchemicalshiftEditField.Value;
%Finde nächsten ppm-Wert, für den ein Datenpunkt in der
%NMR-Matrix (Spectra) vorhanden ist
[I app.ZeileEDCvalueCarbon] =
min(abs(app.ppmCarbon(:,1) - app.EDCvalueCarbon));
end

% Button pushed function: ShowEDDCButton
function ShowEDDCButtonPushed(app, event)
    EDDC(1, size(app.nmrbuckets, 3)) = zeros;
%Preallocation
for i = 1:size(app.nmrbuckets, 3) %Zähler über alle
Fraktionen, auch wenn in app.fractions weniger ausgewählt sind
    EDDC(1,i) =
app.nmrbuckets(app.ZeileEDCvalueCarbon,
app.ZeileEDCvalueProton, i);
end
%EDDC-Plot:
plot(app.EDDCPlot, app.fractions, EDDC(1,
app.fractions));
end

% Value changed function: mzvalueofEMCEditField
function mzvalueofEMCEditFieldValueChanged(app, event)

```



```

        app.EMCvalue = app.mzvalueofEMCEditField.Value;
%Finde nächsten m/z-Wert, für den ein Datenpunkt in der
%'intensitysub'-Matrix vorhanden ist
[I app.ZeileEMCvalue] = min(abs(app.mzall(:,1) -
app.EMCvalue));
end

% Button pushed function: ShowEMCButton
function ShowEMCButtonPushed(app, event)
%regulärer Plot der MS-Spektren nach Blank-Subtraktion
[X, ~] = meshgrid(app.fractions,
min(app.indexplotms):max(app.indexplotms));
plot3(app.MSSpectra, X,
app.mzall(min(app.indexplotms):max(app.indexplotms)),
app.intensitysub(min(app.indexplotms):max(app.indexplotms),
app.fractions));
xticks(app.MSSpectra, 1:size(app.intensitysub, 2));

%Hinzufügen der EMC-Plots
hold(app.MSSpectra, 'on')
%plot3(app.MSSpectra, app.fractions,
app.mzall(app.ZeileEMCvalue, app.fractions), wholeEMC(1,
app.fractions), 'LineWidth', 2);
plot3(app.MSSpectra, app.fractions,
app.mzall(app.ZeileEMCvalue, app.fractions),
app.intensitysub(app.ZeileEMCvalue, app.fractions),
'LineWidth', 2);
hold(app.MSSpectra, 'off')
end

% Button pushed function: AddEDDCButton
function AddEDDCButtonPushed(app, event)
app.ListBoxEDDCPlotColumn = [];
% Wert des EDDC (in Edit Fields) in ListBox des EDDC/EMC-Plots
% Übernehmen (Items = Bezeichnung, nicht die Daten selbst)
temp = app.ListBoxEDDCPlot.Items;
temp = string(temp);
b = '(' +
string(app.HchemicalshiftEditField.Value) + ';' +
string(app.CchemicalshiftEditField.Value) + ')'; %Anzeige von
1H und 13C Versch. in ListBox (1H;13C)
c = [temp, b];
app.ListBoxEDDCPlot.Items = string(c);
%Carbon-Zeilenummer speichern unter b2, c2, ItemsData der
%ListBox
temp2 = app.ListBoxEDDCPlot.ItemsData;
b2 = app.ZeileEDCvalueCarbon;

%Carbon-Zeilenummern müssen unique sein, damit mehrere gleiche
%13C-Verschiebungen in ListBox richtig den entsprechenden
%1H-Verschiebungen im EDDCPlotProton zugeordnet werden
% --> Wert mit 10000 multiplizieren, dann Spaltenindex addieren
b3 = (b2*10000+(size(temp2,2)+1));
c2 = [temp2, b3];
app.ListBoxEDDCPlot.ItemsData = c2; %unique Werte
in ItemsData Matrix
%Extraktion des Column-Index der spezifischen
%CarbonZeilenummern:
for i = 1:size(app.ListBoxEDDCPlot.ItemsData,2)
app.ListBoxEDDCPlotColumn{1,i} =
num2str(c2(1,i));
app.ListBoxEDDCPlotColumn{1,i} =
str2double(app.ListBoxEDDCPlotColumn{1,i}(end-3:end)); %für
bis zu 3-stellige Zahlen möglich (also bis 999 Einträge in
ListBox)
end
%Proton-Zeilenummer speichern unter b3, c3, in keiner ListBox
temp3 = app.EDDCPlotProton;
b4 = app.ZeileEDCvalueProton;
c3 = [temp3, b4];
app.EDDCPlotProton = c3;
set(app.DeleteEDDCButton, 'Enable', 'on');
set(app.PlotEDDCButton, 'Enable', 'on');
set(app.ClearPlotButton, 'Enable', 'on');
if size(app.ListBoxEDDCPlot.Items) >= 1 &
size(app.ListBoxEMCPlot.Items) >= 1
set(app.NMRMSButton, 'Enable', 'on');
set(app.PearsonTableMS_NMR, 'Enable', 'on');
end
if size(app.ListBoxEDDCPlot.Items) >= 1
set(app.NMRMButton, 'Enable', 'on');
set(app.PearsonTableNMR_NMR, 'Enable', 'on');
end
assignin('base', 'EDDCPlotProton',
app.EDDCPlotProton);
assignin('base', 'ListBoxEDDCPlot',
app.ListBoxEDDCPlot);
assignin('base', 'ListBoxEDDCPlotColumn',
app.ListBoxEDDCPlotColumn);
end

% Button pushed function: AddEMCButton
function AddEMCButtonPushed(app, event)
% Wert des EMC (im Edit Field) in ListBox des EDC/EMC-Plots
% Übernehmen
temp = app.ListBoxEMCPlot.Items;
temp = string(temp);
b = string(app.mzvalueofEMCEditField.Value);
c = [temp, b];
app.ListBoxEMCPlot.Items = string(c);
temp2 = app.ListBoxEMCPlot.ItemsData;
b2 = app.ZeileEMCvalue;
c2 = [temp2, b2];
app.ListBoxEMCPlot.ItemsData = c2;
set(app.DeleteEMCButton, 'Enable', 'on');
set(app.PlotEMCButton, 'Enable', 'on');
set(app.ClearPlotButton, 'Enable', 'on');
if size(app.ListBoxEDDCPlot.Items) >= 1 &
size(app.ListBoxEMCPlot.Items) >= 1
set(app.NMRMSButton, 'Enable', 'on');
set(app.PearsonTableMS_NMR, 'Enable', 'on');
end
if size(app.ListBoxEMCPlot.Items) >= 1
set(app.MSMSButton, 'Enable', 'on');
set(app.PearsonTableMS_MS, 'Enable', 'on');
end

% Button pushed function: DeleteEDDCButton
function DeleteEDDCButtonPushed(app, event)
% Löschen des hervorgehobenen EDC-Wertes in ListBox für 13C-
Zeilenummer (ItemsData) und in
% EDDCPlotProton-Liste für 1H-Zeilenummer. Löschen in Items
% für Bezeichnung in ListBox
[~, idx] = ismember(app.ListBoxEDDCPlot.Value,
app.ListBoxEDDCPlot.ItemsData); %Index ermitteln, der gelöscht
werden soll
%zurückrechnen der regulären Werte (reguläre Zeilennummern der
%13C-Verschiebung):
for i = 1:size(app.ListBoxEDDCPlot.ItemsData,2)
Zeilenummern(1,i) =
(app.ListBoxEDDCPlot.ItemsData(1,i)-i)/10000;
end
%Löschen des ausgewählten Wertes
app.ListBoxEDDCPlot.Items(idx) = [];
Zeilenummern(idx) = [];
app.ListBoxEDDCPlot.ItemsData = []; %komplett
löschen und danach neu berechnen
app.ListBoxEDDCPlotColumn = []; %komplett löschen
und danach neu berechnen
app.EDDCPlotProton(idx) = [];
%Neuberechnung der unique Werte in EDDCListBox.ItemsData
for i = 1:size(Zeilenummern, 2)
app.ListBoxEDDCPlot.ItemsData(1,i) =
(Zeilenummern(1,i)*10000+i);
end
%Extraktion des Column-Index der spezifischen
%CarbonZeilenummern:
for i = 1:size(app.ListBoxEDDCPlot.ItemsData,2)
app.ListBoxEDDCPlotColumn{1,i} =
num2str(app.ListBoxEDDCPlot.ItemsData(1,i));
app.ListBoxEDDCPlotColumn{1,i} =
str2double(app.ListBoxEDDCPlotColumn{1,i}(end));
end
assignin('base', 'EDDCPlotProton',
app.EDDCPlotProton);
assignin('base', 'ListBoxEDDCPlot',
app.ListBoxEDDCPlot);
assignin('base', 'ListBoxEDDCPlotColumn',
app.ListBoxEDDCPlotColumn);
end

% Button pushed function: DeleteEMCButton
function DeleteEMCButtonPushed(app, event)
% Löschen des hervorgehobenen EMC-Wertes in ListBox
[~, idx] = ismember(app.ListBoxEMCPlot.Value,
app.ListBoxEMCPlot.ItemsData);
app.ListBoxEMCPlot.Items(idx) = [];
app.ListBoxEMCPlot.ItemsData(idx) = [];
end

% Button pushed function: PlotEDDCButton
function PlotEDDCButtonPushed(app, event)
%zurückrechnen der regulären Werte (reguläre Zeilennummern der
%13C-Verschiebung):
for i = 1:size(app.ListBoxEDDCPlot.ItemsData,2)
Zeilenummern(1,i) =
(app.ListBoxEDDCPlot.ItemsData(1,i)-i)/10000;
end
Alle Einträge in ListBox in EMC/EDC-Plot plotten
EDDC(size(app.ListBoxEDDCPlot.Items, 2),
size(app.nmrbuckets, 3)) = zeros; %für Preallocation
for i = 1:size(app.nmrbuckets, 3) %Zähler über alle
Fraktionen
for j = 1:size(app.ListBoxEDDCPlot.Items, 2)
%Zähler über Einträge in ListBox
EDDC(j,i) =
app.nmrbuckets(Zeilenummern(j), app.EDDCPlotProton(j), i);
end
end
%EDDC-Plot:
plot(app.EDDCandEMCPlot, app.fractions,
(EDDC(:,app.fractions)));
legend(app.EDDCandEMCPlot,
app.ListBoxEDDCPlot.Items, 'location', 'northeast');

```

```

set(app.NormalizeEDDCIntensitiesinPlotMenu,
'Enable', 'on');
set(app.NormalizeEMCIntensitiesinPlotMenu,
'Enable', 'off');
assignin('base', 'Zeilennummern', Zeilennummern);
assignin('base', 'EDDC', EDDC);
end

% Button pushed function: PlotEMCButton
function PlotEMCButtonPushed(app, event)
    plot(app.EDDCandEMCPlot, app.fractions,
    (app.intensitysub(app.ListBoxEMCPlot.ItemsData,
    app.fractions)))';
    legend(app.EDDCandEMCPlot,
    app.ListBoxEMCPlot.Items, 'location', 'northeast');
    set(app.NormalizeEDDCIntensitiesinPlotMenu,
    'Enable', 'off');
    set(app.NormalizeEMCIntensitiesinPlotMenu,
    'Enable', 'on');
end

% Button pushed function: ClearPlotButton
function ClearPlotButtonPushed(app, event)
    cla(app.EDDCandEMCPlot);
    set(app.NormalizeEDDCIntensitiesinPlotMenu,
    'Enable', 'off');
    set(app.NormalizeEMCIntensitiesinPlotMenu,
    'Enable', 'off');
end

% Button pushed function: NMRMSButton
function NMRMSButtonPushed(app, event)
    %Person-Table etc. löschen (sonst bleiben Werte bestehen, wenn
    %sie aus Listbox gelöscht werden)
    app.PearsonTableMS_NMR.Data = [];
    app.PearsonTableMS_NMR.ColumnName = [];
    app.PearsonTableMS_NMR.RowName = [];
    clear [row, col]
    removeStyle(app.PearsonTableMS_NMR);
    app.p_ValuesTable_MS_NMR.Data = [];
    removeStyle(app.p_ValuesTable_MS_NMR);
    app.p_ValuesTable_MS_NMR.ColumnName = [];
    app.p_ValuesTable_MS_NMR.RowName = [];
    %zurückrechnen der regulären Werte (reguläre Zeilennummern der
    %13C-Verschiebung):
    for i = 1:size(app.ListBoxEDDCPlot.ItemsData,2)
        Zeilennummern(1,i) =
    (app.ListBoxEDDCPlot.ItemsData(1,i)-i)/10000;
    end
    %EDDC-Erstellung
    EDDC(size(app.ListBoxEDDCPlot.Items, 2),
    size(app.nmrbuckets, 3)) = zeros; %Preallocation
    for m = 1:size(app.nmrbuckets, 3) %Zähler über
    Fraktionen
        for n = 1:size(app.ListBoxEDDCPlot.Items, 2)
            %Zähler über Einträge in Listbox
            EDDC(n,m) =
    app.nmrbuckets(Zeilennummern(n), app.EDDCPlotProton(n), m);
        end
    end
    if app.PearsonButton.Value == true
        type = 'Pearson';
    elseif app.KendallButton.Value == true
        type = 'Kendall';
    elseif app.SpearmanButton.Value == true
        type = 'Spearman';
    end
    %Korrelationskoeffizienten zwischen NMR und MS Signalen
    berechnen
    for i = 1:size(app.ListBoxEDDCPlot.Items, 2)
        for j = 1:size(app.ListBoxEMCPlot.Items, 2)
            [tempR, tempP] = corr(EDDC(i,
    app.fractions)',
    app.intensitysub(app.ListBoxEMCPlot.ItemsData(j),
    app.fractions)', 'Type', type);
            app.PearsonTableMS_NMR.Data(i,j) =
    tempR(1,1);
            app.p_ValuesTable_MS_NMR.Data(i,j) =
    tempP(1,1);
        end
    end
    %Tabellen-Beschriftung
    app.PearsonTableMS_NMR.ColumnName =
    app.ListBoxEMCPlot.Items;
    app.PearsonTableMS_NMR.RowName =
    app.ListBoxEDDCPlot.Items;
    app.p_ValuesTable_MS_NMR.ColumnName =
    app.ListBoxEMCPlot.Items;
    app.p_ValuesTable_MS_NMR.RowName =
    app.ListBoxEDDCPlot.Items;
    %Werte>0.95 grün einfärben
    s = uistyle('BackgroundColor', 'green');
    [row, col] = find(app.PearsonTableMS_NMR.Data >
    0.95 | app.PearsonTableMS_NMR.Data < -0.95);
    addStyle(app.PearsonTableMS_NMR, s, 'cell', [row,
    col]);
    addStyle(app.p_ValuesTable_MS_NMR, s, 'cell', [row,
    col]);

    app.TabGroup.SelectedTab = app.NMRMSTab;
end

% Button pushed function: ShowallcorrelatedEDDCsButton
function ShowallcorrelatedEDDCsButtonPushed(app, event)
    %Progress-Dialog
    d1 =
    uiprogressdlg(app.SCOREmetaboliteIDSquaredUIFigure, 'Title',
    'Please Wait', 'Message', ...
    'Calculating...', 'Indeterminate', 'on');
    drawnow
    %Ziel: Suche nach passenden EDDCs (mit Korrelationskoeffizient
    %> 0.95 oder < -0.95) für einen ausgewählten MS-Peak;
    %Button ist nur aktiv, wenn genau ein Wert in EMC-ListBox
    %ausgewählt ist;
    %Bestehende Daten löschen:
    AllEMCCorrelations = [];
    HighEMCCorrelations = [];
    ind = [];
    %Finde Index von ausgewähltem Wert in EMC-ListBox:
    [~, idx] = ismember(app.ListBoxEMCPlot.Value,
    app.ListBoxEMCPlot.ItemsData); %idx = Index des in der Listbox
    ausgewählten Eintrags
    if app.PearsonButton.Value == true
        type = 'Pearson';
    elseif app.KendallButton.Value == true
        type = 'Kendall';
    elseif app.SpearmanButton.Value == true
        type = 'Spearman';
    end
    %Berechnung aller möglichen Korrelationskoeffizienten zwischen
    %EMC und allen EDDCs:
    AllEMCCorrelations(size(app.nmrbuckets, 1),
    size(app.nmrbuckets, 2)) = zeros; %Preallocation
    for j = 1:size(app.nmrbuckets,1) %Zähler über 13C-
    Versch.
        for i = 1:size(app.nmrbuckets, 2) %Zähler über
    1H-Versch.
            [tempR, tempP] =
    corr(permute(app.nmrbuckets(j,1,app.fractions),[3,2,1]),
    app.intensitysub(app.ListBoxEMCPlot.ItemsData(idx),
    app.fractions)', 'type', type);
            AllEMCCorrelations(j,i) = tempR(1,1);
            AllEMCCorrelations_pValues(j,i) =
    tempP(1,1);
        end
    end
    %Abfrage, ab welchem Wert die Korrelationskoeffizienten
    %angezeigt werden sollen
    edit_corrcoef = inputdlg('Please enter the minimum
    value of correlation coefficients to be shown','Limit for
    correlation coefficient',[1 50], {'0.95'});
    corrcoef_limit = str2double(edit_corrcoef);
    %Zeige nur Korrelationen an, bei denen der
    %Korrelationskoeffizient größer ist als der Wert, der im Input-
    Dlg angegeben wurde
    for i = 1:size(AllEMCCorrelations, 1) %Zähler über
    13C-Versch.
        for j = 1:size(AllEMCCorrelations, 2) %Zähler
    über 1H-Versch.
            if AllEMCCorrelations(i,j) >=
    corrcoef_limit || AllEMCCorrelations(i,j) <= -corrcoef_limit
                HighEMCCorrelations(i*j,3) =
    AllEMCCorrelations(i,j); %Spalte 3: Korrelationskoeffizient
                HighEMCCorrelations(i*j,1) =
    app.ppmbuckets(j,1); %Spalte 1: 1H-Verschiebung
                HighEMCCorrelations(i*j,2) =
    app.ppmCarbon(i,1); %Spalte 2: 13C-Verschiebung
                HighEMCCorrelations(i*j,4) =
    max(app.nmrbuckets(i, j, app.fractions)); %Spalte 4: maximale
    Intensität
                HighEMCCorrelations(i*j,5) =
    min(app.nmrbuckets(i, j, app.fractions)); %Spalte 5: minimale
    Intensität (wegen multiplizitäts-editiertem HSQC-Spektrum)
                HighEMCCorrelations(i*j,6) =
    AllEMCCorrelations_pValues(i,j); %Spalte 6: p-Werte
            end
        end
    end
    %Erstellung eines neuen Fensters mit neuer Tabelle
    fig = uifigure('Position', [100 300 800 700]);
    fig.Name = ('Correlating EDDCs to m/z ' +
    string(app.ListBoxEMCPlot.Items(idx)) + ' in fractions ' +
    string(min(app.fractions)) + ' - ' +
    string(max(app.fractions)));
    s1 = uistyle('BackgroundColor', 'green');
    s2 = uistyle('BackgroundColor', [0.65 1 0.65]);
    %mittleres grün
    s3 = uistyle('BackgroundColor', [0.85 1 0.85]);
    %sehr abgeschwächtes grün
    s4 = uistyle('BackgroundColor', 'red');
    s5 = uistyle('BackgroundColor', [1 0.65 0.65]);
    s6 = uistyle('BackgroundColor', [1 0.85 0.85]);
    %Lösche die übrigen Zeilen (bestehend aus Nullen)
    if size(HighEMCCorrelations) ~= 0 %um Fehlermeldung
    zu vermeiden, wenn keine Korrelationen gefunden werden können
        ind = HighEMCCorrelations(:,1) ~= 0;
    end
end

```

```

HighEMCCorrelations(ind,:);
uit = uitable(fig, 'Data', HighEMCCorrelations,
'ColumnName', ...
    {'1H chemical shift [ppm]', '13C chemical
shift [ppm]', ...
    'correlation coefficient', 'maximal
intensity', 'minimal intensity', 'p-value'}, 'Position', [50 50
700 630], 'ColumnSortable', true, 'ColumnFormat', {'shortG'});
    if size(HighEMCCorrelations,1) >= 30
15); %finde die 15 NMR-Signale, die maximale Intensität haben
    addStyle(uit, s1, 'row', I(1:5));
    addStyle(uit, s2, 'row', I(6:10));
    addStyle(uit, s3, 'row', I(11:15));
    [~, J] = mink(HighEMCCorrelations(:,5),
15); %finde die 15 NMR-Signale, die minimale Intensität haben
    addStyle(uit, s4, 'row', J(1:5));
    addStyle(uit, s5, 'row', J(6:10));
    addStyle(uit, s6, 'row', J(11:15));
    elseif size(HighEMCCorrelations,1) >= 10
    [~, I] = maxk(HighEMCCorrelations(:,4), 5);
%finde die 5 NMR-Signale, die maximale Intensität haben
    addStyle(uit, s1, 'row', I(1:5));
    [~, J] = mink(HighEMCCorrelations(:,5), 5);
%finde die 5 NMR-Signale, die minimale Intensität haben
    addStyle(uit, s4, 'row', J(1:5));
    end
else
    uialert(fig, 'No correlating EDDCs found!',
'EDDC Correlations');
    end
%Erstellung der CorrelationPlot-Matrix, die nur dort Werte
%enthält, wo der Korrelationskoeffizient >0.95/<-0.95 ist
    CorrelationPlot(size(AlLEMCCorrelations, 1),
size(AlLEMCCorrelations, 2)) = zeros; %Preallocation
    for i = 1:size(AlLEMCCorrelations, 1)
        for j = 1:size(AlLEMCCorrelations, 2)
            if AlLEMCCorrelations(i,j) >=
corrcoef_limit || AlLEMCCorrelations(i,j) <= -corrcoef_limit
%Plot der Korrelations-Koeffizienten:
%CorrelationPlot(i,j) = AlLEMCCorrelations(i,j);
%Plot der Intensitätswerte:
                CorrelationPlot(i,j) =
app.SpectraSumBuckets(i,j);
            end
        end
    end
%Erstellung eines neuen Fensters für 2D-Plot der
%hoch-korrelierenden EDDCs zu EMC:
    fig2 = figure('Position', [950 300 800 700]);
    fig2.Name = ('Plot of EDDCs with correlation
coefficient >= ' + string(edit_corrcoef) + ' ...
' and <= -' + string(edit_corrcoef) + ' to m/z
' + string(app.ListBoxEMCPlot.Items(idx)) + ' ...
' in fractions ' + string(min(app.fractions)) +
' - ' + string(max(app.fractions)));
    ax = uiaxes(fig2, 'Position', [50 50 700 600]);
%2DPlot:
    contour(ax, app.ppmbuckets(:,1),
app.ppmCarbon(:,1), CorrelationPlot, 'LevelStep', 0.5);
    set(ax, 'colormap', winter, 'YAxisLocation',
'right');
    ax.YLabel.String = '13C chemical shift [ppm]';
    ax.XLabel.String = '1H chemical shift [ppm]';
    ax.XDir = 'reverse';
    ax.YDir = 'reverse';
    close(d1);
    %assignin('base', 'CorrelationCoefficientstoEMC',
AlLEMCCorrelations);
    %assignin('base', 'TableofCorrelationCoefficients',
HighEMCCorrelations);
    assignin('base', 'CorrelationPlot',
CorrelationPlot);
    %Context-Menü im öffnenden Fenstern mit Tabelle
    cm = uicontextmenu(fig);
    m1 = uimenu(cm, 'Text', 'Copy data to workspace');
    fig.ContextMenu = cm;
    m1.MenuSelectedFcn = @MenuSelectedFcn;
%Funktion, um Daten aus Tabelle in MATLAB Workspace zu
%übertragen
    function MenuSelectedFcn(src, event)
        assignin('base', 'CorrelatingEDDCstoEMC',
HighEMCCorrelations);
    end
end

% Button pushed function: ShowallcorrelatedEMCstoEMCButton
function ShowallcorrelatedEMCstoEMCButtonPushed(app,
event)
%Zeige alle hoch-korrelierenden MS-Signale zu einem
%ausgewählten MS-Signal
%Button ist nur aktiv, wenn genau ein Wert in EMC-ListBox steht
    AlLEMCCorrelations = [];
    HighEMCCorrelations = [];
    ind = [];
%Berechnung aller möglichen Korrelationskoeffizienten zwischen
%EMC und allen NMR-Signalen in nmrbuckets:
    [~, idx] = ismember(app.ListBoxEMCPlot.Value,
app.ListBoxEMCPlot.ItemsData); %idx = Index des in der ListBox
ausgewählten Eintrags
    if app.PearsonButton.Value == true
        type = 'Pearson';
    elseif app.KendallButton.Value == true
        type = 'Kendall';
    elseif app.SpearmanButton.Value == true
        type = 'Spearman';
    end
    for j = 1:size(app.intensitysub,1)
        [tempR, tempP] = corr(app.intensitysub(j,
app.fractions),
app.intensitysub(app.ListBoxEMCPlot.ItemsData(idx),
app.fractions), 'type', type);
        AlLEMCCorrelations(j,1) = tempR(1,1);
        AlLEMCCorrelations(j,2) = tempP(1,1);
    end
%Abfrage, ab welchem Wert die Korrelationskoeffizienten
%angezeigt werden sollen
    edit_corrcoef = inputdlg({'Please enter the minimum
value of correlation coefficients to be shown'}, 'Limit for
correlation coefficient', [1 50], {'0.95'});
    corrcoef_limit = str2double(edit_corrcoef);
%Zeige nur Korrelationen an, bei denen der
%Korrelationskoeffizient größer ist als der Wert, der im Input-
Dlg angegeben wurde
    for i = 1:size(AlLEMCCorrelations, 1)
        if AlLEMCCorrelations(i,1) >= corrcoef_limit
            HighEMCCorrelations(i,2) =
AlLEMCCorrelations(i,1);
            HighEMCCorrelations(i,1) = app.mzall(i,1);
            HighEMCCorrelations(i,3) =
max(app.intensitysub(i, app.fractions));
            HighEMCCorrelations(i,4) =
AlLEMCCorrelations(i,2);
        end
    end
%Erstellung eines neuen Fensters mit neuer Tabelle
    fig = uifigure;
    fig.Name = ('Correlating EMCs to EMC at m/z ' +
string(app.ListBoxEMCPlot.Items(idx)) + ' in fractions ' +
string(min(app.fractions)) + ' - ' +
string(max(app.fractions)));
    s1 = uistyle('BackgroundColor', 'green');
    s2 = uistyle('BackgroundColor', [0.65 1 0.65]);
%mittleres grün
    s3 = uistyle('BackgroundColor', [0.85 1 0.85]);
%sehr abgeschwächtes grün
%lösche die übrigen Zeilen (bestehend aus Nullen)
    if size(HighEMCCorrelations) ~= 0 %um Fehlermeldung
zu vermeiden, wenn keine Korrelationen gefunden werden können
        ind = HighEMCCorrelations(:,1) ~= 0;
        HighEMCCorrelations =
HighEMCCorrelations(ind,:);
        uit = uitable(fig, 'Data', HighEMCCorrelations,
'ColumnName', {'m/z', 'correlation coefficient', 'maximal
intensity', 'p-value'}, 'Position', [50 50 450 350],
'ColumnSortable', true, 'ColumnFormat', {'shortG'});
        if size(HighEMCCorrelations,1) >= 15
            [~, I] = maxk(HighEMCCorrelations(:,3),
15); %finde die 15 NMR-Signale, die maximale Intensität haben
            addStyle(uit, s1, 'row', I(1:5));
            addStyle(uit, s2, 'row', I(6:10));
            addStyle(uit, s3, 'row', I(11:15));
            elseif size(HighEMCCorrelations,1) >= 5
            [~, I] = maxk(HighEMCCorrelations(:,3), 5);
%finde die 5 NMR-Signale, die maximale Intensität haben
            addStyle(uit, s1, 'row', I(1:5));
        end
    else
        uialert(fig, 'No correlating EMCs found!', 'EMC
Correlations');
    end
%Context-Menü im öffnenden Fenstern mit Tabelle
    cm = uicontextmenu(fig);
    m1 = uimenu(cm, 'Text', 'Copy data to workspace');
    m2 = uimenu(cm, 'Text', 'Create plot of
correlations');
    fig.ContextMenu = cm;
    m1.MenuSelectedFcn = @MenuSelectedFcn;
    m2.MenuSelectedFcn = @MenuSelectedFcn2;
%Funktion, um Daten aus Tabelle in MATLAB Workspace zu
%übertragen
    function MenuSelectedFcn(src, event)
        assignin('base', 'CorrelatingEMCstoEMC',
HighEMCCorrelations);
    end
%Funktion, um STOCSSY/SHY-like Plot zu erstellen
    function MenuSelectedFcn2(src, event)
        VisLoadings1D_MS(app.intensitysub(:,
app.fractions), AlLEMCCorrelations(:,1), app.mzall(:,1));
    end
end

% Value changed function: lowermzEditField
function lowermzEditFieldValueChanged(app, event)

```

```

        app.mzlimit(1,1) = app.lowermzEditField.Value;
        app.indexplots(1,1) = find(app.mzall(:,1) >
min(app.mzlimit), 1, 'first');
        app.indexplots(1,2) = find(app.mzall(:,1) <
max(app.mzlimit), 1, 'last');
        [X, ~] = meshgrid(app.fractions,
min(app.indexplots):max(app.indexplots));
        plot3(app.MSSpectra, X,
app.mzall(min(app.indexplots):max(app.indexplots)),
app.intensitysub(min(app.indexplots):max(app.indexplots),
app.fractions));
        xticks(app.MSSpectra, 1:size(app.intensitysub, 2));
    end

% Value changed function: uppermzEditField
function uppermzEditFieldValueChanged(app, event)
    app.mzlimit(1,2) = app.uppermzEditField.Value;
    app.indexplots(1,2) = find(app.mzall(:,1) >
max(app.mzlimit), 1, 'first');
    app.indexplots(1,1) = find(app.mzall(:,1) <
min(app.mzlimit), 1, 'last');
    [X, ~] = meshgrid(app.fractions,
min(app.indexplots):max(app.indexplots));
    plot3(app.MSSpectra, X,
app.mzall(min(app.indexplots):max(app.indexplots)),
app.intensitysub(min(app.indexplots):max(app.indexplots),
app.fractions));
    xticks(app.MSSpectra, 1:size(app.intensitysub, 2));
end

% Button pushed function: ShowallcorrelatedEMCsButton
function ShowallcorrelatedEMCsButtonPushed(app, event)
%Ziel: Suche nach passenden EMCs (mit Korrelationskoeffizient >
%0.95 oder < -0.95) für einen ausgewählten NMR-Peak (= EDDC);
%Button ist nur aktiv, wenn genau ein Wert in EDDC-ListBox
%ausgewählt ist
%Tabellen löschen:
    AllEDCCorrelations = [];
    HighEDCCorrelations = [];
    ind = [];
    %Index des ausgewählten Eintrags in ListBox finden:
    [~, idx] = ismember(app.ListBoxEDDCPlot.Value,
app.ListBoxEDDCPlot.ItemsData); %idx = Index des in der ListBox
ausgewählten Eintrags
%zurückrechnen der regulären Werte (reguläre Zeilennummern der
%13C-Verschiebung):
    for i = 1:size(app.ListBoxEDDCPlot.ItemsData,2)
        Zeilennummern(1,i) =
(app.ListBoxEDDCPlot.ItemsData(1,i)-i)/10000;
    end
%Erstellung des EDDCs des ausgewählten Eintrags (Index idx)
%über alle Fraktionen:
    EDDC(1, size(app.nmrbuckets, 3)) = zeros;
%Preallocation
    for m = 1:size(app.nmrbuckets, 3) %Zähler über
Fraktionen
        EDDC(1,m) = app.nmrbuckets(Zeilennummern(idx),
app.EDDCPlotProton(idx), m);
    end
    if app.PearsonButton.Value == true
        type = 'Pearson';
    elseif app.KendallButton.Value == true
        type = 'Kendall';
    elseif app.SpearmanButton.Value == true
        type = 'Spearman';
    end
%Berechnung aller möglichen Korrelationskoeffizienten zwischen
%EDC und allen MS-Signalen in intensitysub
    AllEDCCorrelations(size(app.intensitysub, 1), 1) =
zeros; %Preallocation
    for j = 1:size(app.intensitysub,1)
        [tempR, tempP] = corr(app.intensitysub(j,
app.fractions), EDDC(1, app.fractions), 'type', type);
        AllEDCCorrelations(j,1) = tempR(1,1);
        AllEDCCorrelations(j,2) = tempP(1,1);
    end
%Abfrage, ab welchem Wert die Korrelationskoeffizienten
%angezeigt werden sollen
    edit_corrcoef = inputdlg('Please enter the minimum
value of correlation coefficients to be shown','Limit for
correlation coefficient',[1 50], {'0.95'});
    corrcoef_limit = str2double(edit_corrcoef);
%Zeige nur Korrelationen an, bei denen der
%Korrelationskoeffizient größer ist als der Wert, der im
%Input-Dlg angegeben wurde
    for i = 1:size(AllEDCCorrelations, 1)
        if AllEDCCorrelations(i,1) >= corrcoef_limit ||
AllEDCCorrelations(i,1) <= -corrcoef_limit
            HighEDCCorrelations(i,2) =
AllEDCCorrelations(i,1);
            HighEDCCorrelations(i,1) = app.mzall(i,1);
            HighEDCCorrelations(i,3) =
max(app.intensitysub(i, app.fractions));
            HighEDCCorrelations(i,4) =
AllEDCCorrelations(i,2);
        end
    end
%Erstellung eines neuen Fensters mit neuer Tabelle
        fig = uifigure;
        fig.Name = ('Correlating EMCs to EDDC at ' +
string(app.ListBoxEDDCPlot.Items(idx)) + ' ppm in fractions ' +
string(min(app.fractions)) + ' to ' +
string(max(app.fractions)));
        s1 = uistyle('BackgroundColor', 'green');
        s2 = uistyle('BackgroundColor', [0.65 1 0.65]);
        %mittleres grün
        s3 = uistyle('BackgroundColor', [0.85 1 0.85]);
        %sehr abgeschwächtes grün
        %lösche die übrigen Zeilen (bestehend aus Nullen)
        if size(HighEDCCorrelations) ~= 0 %um
Fehlermeldung zu vermeiden, wenn keine Korrelationen gefunden
werden können
            ind = HighEDCCorrelations(:,1) ~= 0;
            HighEDCCorrelations =
HighEDCCorrelations(ind,:);
            uit = uitable(fig, 'Data',
HighEDCCorrelations, 'ColumnName', {'m/z', 'correlation
coefficient', 'maximal intensity', 'p-value'}, 'Position', [50
50 450 350], 'ColumnSortable', true, 'ColumnFormat',
{'shortG'});
            if size(HighEDCCorrelations,1) >= 15
                [~, I] = max(HighEDCCorrelations(:,3),
15); %finde die 15 m/z-Werte, die maximale Intensität haben
                addStyle(uit, s1, 'row', I(1:5));
                addStyle(uit, s2, 'row', I(6:10));
                addStyle(uit, s3, 'row', I(11:15));
            elseif size(HighEDCCorrelations,1) >= 5
                [~, I] = max(HighEDCCorrelations(:,3),
5); %finde die 15 m/z-Werte, die maximale Intensität haben
                addStyle(uit, s1, 'row', I(1:5));
            end
        else
            uialert(fig, 'No correlating EMCs found!', 'EMC
Correlations');
        end
        %Context-Menü im öffnenden Fenster mit Tabelle
        cm = uicontextmenu(fig);
        m1 = uimenu(cm, 'Text', 'Copy data to workspace');
        m2 = uimenu(cm, 'Text', 'Create plot of
correlations');
        fig.ContextMenu = cm;
        m1.MenuSelectedFcn = @MenuSelectedFcn;
        m2.MenuSelectedFcn = @MenuSelectedFcn2;
        %Funktion, um Daten aus Tabelle in MATLAB Workspace zu
%übertragen
        function MenuSelectedFcn(src, event)
            assignin('base', 'CorrelatingEMCstoEDDC',
HighEDCCorrelations);
        end
        assignin('base', 'EDDCPlotProton',
app.EDDCPlotProton);
        assignin('base', 'ListBoxEDDCPlot',
app.ListBoxEDDCPlot);
        %Funktion, um STOCYSV-SHY-like Plot zu erstellen
        function MenuSelectedFcn2(src,event)
            VisLoadings1D_MS(app.intensitysub(:,
app.fractions), AllEDCCorrelations(:,1), app.mzall(:,1));
        end
    end

% Button pushed function: ShowallcorrelatedEDDCstoEDDCButton
function ShowallcorrelatedEDDCstoEDDCButtonPushed(app,
event)
    %Progress-Dialog
    d1 =
uiprogressdlg(app.SCOREmetaboliteIDSquaredUIFigure, 'Title',
'Please Wait', 'Message', ...
    'Calculating...', 'Indeterminate', 'on');
    drawnow
    %Zeige alle hoch-korrelierenden EDDCs zu einem ausgewählten
%EDDC
    %Button ist nur aktiv, wenn genau ein Wert in EMC-ListBox steht
    AllEDCCorrelations = [];
    HighEDCCorrelations = [];
    ind = [];
    %Berechnung aller möglichen Korrelationskoeffizienten zwischen
%EMC und allen NMR-Signalen in nmrbuckets:
    [~, idx] = ismember(app.ListBoxEDDCPlot.Value,
app.ListBoxEDDCPlot.ItemsData); %idx = Index des in der ListBox
ausgewählten Eintrags
%zurückrechnen der regulären Werte (reguläre Zeilennummern der
%13C-Verschiebung):
    for i = 1:size(app.ListBoxEDDCPlot.ItemsData,2)
        Zeilennummern(1,i) =
(app.ListBoxEDDCPlot.ItemsData(1,i)-i)/10000;
    end
%Erstellung des EDDCs des ausgewählten Eintrags (Index idx)
%über alle Fraktionen:
    EDDC(1, size(app.nmrbuckets, 3)) = zeros;
%Preallocation
    for m = 1:size(app.nmrbuckets, 3) %Zähler über
Fraktionen
        EDDC(1,m) = app.nmrbuckets(Zeilennummern(idx),
app.EDDCPlotProton(idx), m);
    end
    if app.PearsonButton.Value == true

```



```

        type = 'Pearson';
    elseif app.KendallButton.Value == true
        type = 'Kendall';
    elseif app.SpearmanButton.Value == true
        type = 'Spearman';
    end
%Berechnung aller möglichen Korrelationskoeffizienten zwischen
%EDDC und allen EDDCs:
    AllEDCCorrelations(size(app.nmrbuckets, 1),
    size(app.nmrbuckets, 2)) = zeros; %Preallocation
    for j = 1:size(app.nmrbuckets,1) %Zähler über 13C-
Versch.
        for i = 1:size(app.nmrbuckets, 2) %Zähler über
1H-Versch.
            [tempR, tempP] =
corr(permute(app.nmrbuckets(j, i, app.fractions), [3,2,1]),
EDDC(1, app.fractions), 'type', type);
            AllEDCCorrelations(j,i) = tempR(1,1);
            AllEDCCorrelations_pValues(j,i) =
tempP(1,1);
        end
    end
%Abfrage, ab welchem Wert die Korrelationskoeffizienten
%angezeigt werden sollen
    edit_corrcoef = inputdlg('Please enter the minimum
value of correlation coefficients to be shown','Limit for
correlation coefficient',[1 50], {'0.95'});
    corrcoef_limit = str2double(edit_corrcoef);
%Zeige nur Korrelationen an, bei denen der
%Korrelationskoeffizient größer ist als der Wert, der im Input-
Dlg angegeben wurde
    for i = 1:size(AllEDCCorrelations, 1) %Zähler über
13C-Versch.
        for j = 1:size(AllEDCCorrelations, 2) %Zähler
über 1H-Versch.
            if AllEDCCorrelations(i,j) >=
corrcoef_limit || AllEDCCorrelations(i,j) <= -corrcoef_limit
                HighEDCCorrelations(i*j,3) =
AllEDCCorrelations(i,j); %Spalte 3: Korrelationskoeffizient
                HighEDCCorrelations(i*j,1) =
app.ppmbuckets(j,1); %Spalte 1: 1H-Verschiebung
                HighEDCCorrelations(i*j,2) =
app.ppmCarbon(1,i); %Spalte 2: 13C-Verschiebung
                HighEDCCorrelations(i*j,4) =
max(app.nmrbuckets(i, j, app.fractions)); %Spalte 4: maximale
Intensität
                HighEDCCorrelations(i*j,5) =
min(app.nmrbuckets(i, j, app.fractions)); %Spalte 5: minimale
Intensität (wegen multiplizitäts-editiertem HSQC-Spektrum)
                HighEDCCorrelations(i*j,6) =
AllEDCCorrelations_pValues(i,j); %Spalte 6: p-Werte
            end
        end
    end
%Erstellung eines neuen Fensters mit neuer Tabelle
    fig = uifigure('Position', [100 300 800 700]);
    fig.Name = ('Correlating EDDCs to EDDC at ' +
string(app.ListBoxEDDCPlot.Items(idx)) + ' ppm in fractions ' +
string(min(app.fractions)) + ' - ' +
string(max(app.fractions))));
    s1 = uistyle('BackgroundColor', 'green');
    s2 = uistyle('BackgroundColor', [0.65 1 0.65]);
%mittleres grün
    s3 = uistyle('BackgroundColor', [0.85 1 0.85]);
%sehr abgeschwächtes grün
    s4 = uistyle('BackgroundColor', 'red');
    s5 = uistyle('BackgroundColor', [1 0.65 0.65]);
    s6 = uistyle('BackgroundColor', [1 0.85 0.85]);
%Lösche die übrigen Zeilen (bestehend aus Nullen)
    if size(HighEDCCorrelations) ~= 0 %um Fehlermeldung
zu vermeiden, wenn keine Korrelationen gefunden werden können
        ind = HighEDCCorrelations(:,1) ~= 0;
        HighEDCCorrelations =
HighEDCCorrelations(ind,:);
        uit = uitable(fig, 'Data', HighEDCCorrelations,
'ColumnName', {'1H chemical shift [ppm]', '13C chemical shift
[ppm]', 'correlation coefficient', 'maximal intensity',
'minimal intensity', 'p-value'}, 'Position', [50 50 700 630],
'ColumnSortable', true, 'ColumnFormat', {'shortG'});
        if size(HighEDCCorrelations,1) >= 30
            [~, I] = maxk(HighEDCCorrelations(:,4),
15); %finde die 15 NMR-Signale, die maximale Intensität haben
            addStyle(uit, s1, 'row', I(1:5));
            addStyle(uit, s2, 'row', I(6:10));
            addStyle(uit, s3, 'row', I(11:15));
            [~, J] = mink(HighEDCCorrelations(:,5),
15); %finde die 15 NMR-Signale, die minimale Intensität haben
            addStyle(uit, s4, 'row', J(1:5));
            addStyle(uit, s5, 'row', J(6:10));
            addStyle(uit, s6, 'row', J(11:15));
        elseif size(HighEDCCorrelations,1) >= 10
            [~, I] = maxk(HighEDCCorrelations(:,4), 5);
            addStyle(uit, s1, 'row', I(1:5));
            [~, J] = mink(HighEDCCorrelations(:,5), 5);
            addStyle(uit, s4, 'row', J(1:5));
        end
    end
    else
        uialert(fig, 'No correlating EDDCs found!',
'EDDC Correlations');
    end
%Erstellung der CorrelationPlot-Matrix, die nur dort Werte
%enthält, wo der Korrelationskoeffizient >0.95/<-0.95 ist
    CorrelationPlot(size(AllEDCCorrelations, 1),
size(AllEDCCorrelations, 2)) = zeros; %Preallocation
    for i = 1:size(AllEDCCorrelations, 1)
        for j = 1:size(AllEDCCorrelations, 2)
            if AllEDCCorrelations(i,j) >=
corrcoef_limit || AllEDCCorrelations(i,j) <= -corrcoef_limit
                %Plot der Korrelations-Koeffizienten:
                %CorrelationPlot(i,j) =
AllEMCCorrelations(i,j);
                %Plot der Intensitätswerte:
                CorrelationPlot(i,j) =
app.SpectraSumBuckets(i,j);
            end
        end
    end
%Erstellung eines neuen Fensters für 2D-Plot der
%hoch-korrelierenden EDDCs zu EDDC:
    fig2 = uifigure('Position', [950 300 800 700]);
    fig2.Name = ('Plot of EDDCs with correlation
coefficient >= ' + string(edit_corrcoef) + ' ...
' and <= -' + string(edit_corrcoef) + ' to EDDC
at ' + string(app.ListBoxEDDCPlot.Items(idx)) + ' ...
' ppm in fractions ' +
string(min(app.fractions)) + ' - ' +
string(max(app.fractions))));
    ax = uiaxes(fig2, 'Position', [50 50 700 600]);
    %2DPlot:
    contour(ax, app.ppmbuckets(:,1),
app.ppmCarbon(:,1), CorrelationPlot, 'LevelStep', 0.5);
    set(ax, 'colormap', winter, 'YAxisLocation',
'right');
    ax.YLabel.String = '13C chemical shift [ppm]';
    ax.XLabel.String = '1H chemical shift [ppm]';
    ax.XDir = 'reverse';
    ax.YDir = 'reverse';
    close(d1);
    %assignin('base', 'CorrelationCoefficientsstoEMC',
AllEMCCorrelations);
    %assignin('base', 'TableofCorrelationCoefficients',
HighEMCCorrelations);
    assignin('base', 'CorrelationPlot',
CorrelationPlot);
    %Context-Menü im öffnenden Fenster mit Tabelle
    cm = uicontextmenu(fig);
    m1 = uimenu(cm, 'Text', 'Copy data to workspace');
    fig.ContextMenu = cm;
    m1.MenuSelectedFcn = @MenuSelectedFcn;
%Funktion, um Daten aus Tabelle in MATLAB Workspace zu
%übertragen
    function MenuSelectedFcn(src, event)
        assignin('base', 'CorrelatingEDDCstoEDDC',
HighEDCCorrelations);
    end
    assignin('base', 'EDDCPlotProton',
app.EDDCPlotProton);
    assignin('base', 'ListBoxEDDCPlot',
app.ListBoxEDDCPlot);
    end
% Menu selected function: Copydatatoworkspace_NMRMS
    function Copydatatoworkspace_NMRMSSelected(app, event)
        assignin('base', 'fractions_min',
min(app.fractions));
        assignin('base', 'fractions_max',
max(app.fractions));
        for i = 1:size(app.PearsonTableMS_NMR.Data,1)
            PearsonTableMS_NMR((i+1),1) =
app.PearsonTableMS_NMR.RowName(i);
            PValuesMS_NMR((i+1),1) =
app.PearsonTableMS_NMR.RowName(i);
        end
        for i = 1:size(app.PearsonTableMS_NMR.Data,2)
            PearsonTableMS_NMR(1,(i+1)) =
app.PearsonTableMS_NMR.ColumnName(i);
            PValuesMS_NMR(1,(i+1)) =
app.PearsonTableMS_NMR.ColumnName(i);
        end
        for i = 1:size(app.PearsonTableMS_NMR.Data,1)
            for j = 1:size(app.PearsonTableMS_NMR.Data,2)
                PearsonTableMS_NMR{(i+1), (j+1)} =
app.PearsonTableMS_NMR.Data(i,j);
                PValuesMS_NMR{(i+1), (j+1)} =
app.p_ValuesTable_MS_NMR.Data(i,j);
            end
        end
        assignin('base', 'PearsonTableMS_NMR',
PearsonTableMS_NMR);
        assignin('base', 'PValuesMS_NMR', PValuesMS_NMR);
    end
% Menu selected function: Copydatatoworkspace_NMRNMR

```

```

function Copydatatoworkspace_NMRNMRMenuSelected(app,
event)
    assignin('base', 'fractions_min',
min(app.fractions));
    assignin('base', 'fractions_max',
max(app.fractions));
    for i = 1:size(app.PearsonTableNMR_NMR.Data,1)
        PearsonTableNMR_NMR{(i+1),1} =
app.PearsonTableNMR_NMR.RowName(i); %ok<*ADPROPLC>
        PValuesNMR_NMR{(i+1),1} =
app.PearsonTableNMR_NMR.RowName(i);
    end
    for i = 1:size(app.PearsonTableNMR_NMR.Data,2)
        PearsonTableNMR_NMR{(i+1),1} =
app.PearsonTableNMR_NMR.ColumnName(i);
        PValuesNMR_NMR{(i+1),1} =
app.PearsonTableNMR_NMR.ColumnName(i);
    end
    for i = 1:size(app.PearsonTableNMR_NMR.Data,1)
        for j = 1:size(app.PearsonTableNMR_NMR.Data,2)
            PearsonTableNMR_NMR{(i+1), (j+1)} =
app.PearsonTableNMR_NMR.Data(i,j);
            PValuesNMR_NMR{(i+1), (j+1)} =
app.p_valuesTable_NMR_NMR.Data(i,j);
            if (i+1) == (j+1)
                PearsonTableNMR_NMR{(i+1), (j+1)} = [];
                PValuesNMR_NMR{(i+1), (j+1)} = [];
            end
        end
    end
    assignin('base', 'PearsonTableNMR_NMR',
PearsonTableNMR_NMR);
    assignin('base', 'PValuesNMR_NMR', PValuesNMR_NMR);
end

% Value changed function: ListBoxEMCPlot
function ListBoxEMCPlotValueChanged(app, event)
    if size(app.ListBoxEMCPlot.Value) == 1 %wenn ein
Wert in der EMC-Listbox ausgewählt wurde,
        set(app.ShowallcorrelatedEDDCsButton, 'Enable',
'on') %dann aktiviere den Button, damit alle Korrelationen mit
EDCs berechnet werden können
        set(app.ShowallcorrelatedEMCsToEMCButton,
'Enable', 'on');
    else
        set(app.ShowallcorrelatedEDDCsButton, 'Enable',
'off') %wenn kein oder mehr Werte ausgewählt wurden, können
keine automatische Korrelationen berechnet werden
        set(app.ShowallcorrelatedEMCsToEMCButton,
'Enable', 'off');
    end
    assignin('base', 'ListBoxEMCPlotValue',
app.ListBoxEMCPlot.Value);
end

% Value changed function: ListBoxEDDCPlot
function ListBoxEDDCPlotValueChanged(app, event)
    if size(app.ListBoxEDDCPlot.Value) == 1 %wenn ein
Wert in der EDC-Listbox ausgewählt wurde,
        set(app.ShowallcorrelatedEMCsButton, 'Enable',
'on') %dann aktiviere den Button, damit alle Korrelationen mit
EMCs berechnet werden können
        set(app.ShowallcorrelatedEDDCsToEDDCButton,
'Enable', 'on');
    else
        set(app.ShowallcorrelatedEMCsButton, 'Enable',
'off') %wenn kein oder mehr Werte ausgewählt wurden, können
keine automatische Korrelationen berechnet werden
        set(app.ShowallcorrelatedEDDCsToEDDCButton,
'Enable', 'off');
    end
    assignin('base', 'ListBoxEDDCPlot',
app.ListBoxEDDCPlot.Value);
end

% Menu selected function: CopyPlotofNMRSpectraMenu
function CopyPlotofNMRSpectraMenuSelected(app, event)
    fig1 = figure;
    ax = app.NMRSpectra;
    copyobj(ax, fig1);
end

% Menu selected function: CopyPlotofMSSpectraMenu
function CopyPlotofMSSpectraMenuSelected(app, event)
    fig1 = figure;
    ax = app.MSSpectra;
    copyobj(ax, fig1);
end

% Menu selected function: CopyEDDCorEMCPlotMenu
function CopyEDDCorEMCPlotMenuSelected(app, event)
    fig1 = figure;
    ax = app.EDDCandEMCPlot;
    copyobj(ax, fig1);
end

% Menu selected function: CreatePlotofEDDCandEMCMenu
function CreatePlotofEDDCandEMCMenuSelected(app, event)

    fig2 = figure('Position', [950 300 800 700]);
    ax = uiaxes(fig2, 'Position', [50 50 700 600]);
    yyaxis(ax, 'left');
    for i = 1:size(app.ListBoxEDDCPlot.ItemsData,2)
        Zeilennummern(1,i) =
(app.ListBoxEDDCPlot.ItemsData(1,i)-i)/10000;
    end
    %Alle Einträge in Listbox in EMC/EDC-Plot plotten
    EDDC(size(app.ListBoxEDDCPlot.Items, 2),
size(app.nmrbuckets, 3)) = zeros; %für Preallocation
    for i = 1:size(app.nmrbuckets, 3) %Zähler über alle
Fraktionen
        for j = 1:size(app.ListBoxEDDCPlot.Items, 2)
            %Zähler über Einträge in Listbox
            EDDC(j,i) =
app.nmrbuckets(Zeilennummern(j), app.EDDCPlotProton(j), i);
        end
    end
    %EDDC-Plot
    plot(ax, app.fractions, (EDDC(:,app.fractions)));
    yyaxis(ax, 'right')
    plot(ax, app.fractions,
(app.intensitysub(app.ListBoxEMCPlot.ItemsData,
app.fractions)));
    %Legend(ax, app.ListBoxEMCPlot.Items, 'location', 'east');
end

% Menu selected function: CopyEDDCPlotMenu
function CopyEDDCPlotMenuSelected(app, event)
    fig1 = figure;
    ax = app.EDDCPlot;
    copyobj(ax, fig1);
end

% Button pushed function: NMRNMRButton
function NMRNMRButtonPushed(app, event)
    %Pearson-Table etc. löschen (sonst bleiben Werte bestehen, wenn
%sie aus Listbox gelöscht werden)
    app.PearsonTableNMR_NMR.Data = [];
    app.PearsonTableNMR_NMR.ColumnName = [];
    app.PearsonTableNMR_NMR.RowName = [];
    clear [row, col]
    removeStyle(app.PearsonTableNMR_NMR);
    app.p_valuesTable_NMR_NMR.Data = [];
    app.p_valuesTable_NMR_NMR.ColumnName = [];
    app.p_valuesTable_NMR_NMR.RowName = [];
    app.p_valuesTable_NMR_NMR.Data = [];
    removeStyle(app.p_valuesTable_NMR_NMR);
    %zurückrechnen der regulären Werte (reguläre Zeilennummern der
%13C-Verschiebung):
    for i = 1:size(app.ListBoxEDDCPlot.ItemsData,2)
        Zeilennummern(1,i) =
(app.ListBoxEDDCPlot.ItemsData(1,i)-i)/10000;
    end
    %EDDC-Erstellung
    EDDC(size(app.ListBoxEDDCPlot.Items, 2),
size(app.nmrbuckets, 3)) = zeros; %Preallocation
    for m = 1:size(app.nmrbuckets, 3) %Zähler über
Fraktionen
        for n = 1:size(app.ListBoxEDDCPlot.Items, 2)
            %Zähler über Einträge in Listbox
            EDDC(n,m) =
app.nmrbuckets(Zeilennummern(n), app.EDDCPlotProton(n), m);
        end
    end
    if app.PearsonButton.Value == true
        type = 'Pearson';
    elseif app.KendallButton.Value == true
        type = 'Kendall';
    elseif app.SpearmanButton.Value == true
        type = 'Spearman';
    end
    %Berechnung der Korrelationskoeffizienten zwischen NMR-Signalen
    for i = 1:size(app.ListBoxEDDCPlot.Items, 2)
        for j = 1:size(app.ListBoxEDDCPlot.Items, 2)
            [tempR, tempP] = corr(EDDC(i,
app.fractions), EDDC(j, app.fractions), 'Type', type);
            app.PearsonTableNMR_NMR.Data(i,j) =
tempR(1,1);
            app.p_valuesTable_NMR_NMR.Data(i,j) =
tempP(1,1);
        end
    end
    %Tabellen-Beschriftung
    app.PearsonTableNMR_NMR.ColumnName =
app.ListBoxEDDCPlot.Items;
    app.PearsonTableNMR_NMR.RowName =
app.ListBoxEDDCPlot.Items;
    app.p_valuesTable_NMR_NMR.ColumnName =
app.ListBoxEDDCPlot.Items;
    app.p_valuesTable_NMR_NMR.RowName =
app.ListBoxEDDCPlot.Items;
    %Einfärbung der sehr hohen Korrelationskoeffizienten
    s = uistyle('BackgroundColor', 'green');
    s1 = uistyle('FontColor', 'white',
'BackgroundColor', 'white');
    %Style für Korrelationskoeffizienten > 0.95
    [row, col] = find(app.PearsonTableNMR_NMR.Data >
0.95 | app.PearsonTableNMR_NMR.Data < -0.95);
end

```

```

        addStyle(app.PearsonTableNMR_NMR, s, 'cell', [row,
col]);
        addStyle(app.p_valuesTable_NMR_NMR, s, 'cell',
[ row, col]);
%Tabelleneinträge in Diagonale weiß einfärben (Style s1)
for i = 1:size(app.PearsonTableNMR_NMR.ColumnName)
    for j = 1:size(app.PearsonTableNMR_NMR.RowName)
        if i == j
            addStyle(app.PearsonTableNMR_NMR, s1,
'cell', [i, j]);
            addStyle(app.p_valuesTable_NMR_NMR, s1,
'cell', [i, j]);
        end
    end
end
app.TabGroup.SelectedTab = app.NMRNMRTab;
end

% Button pushed function: MSMSButton
function MSMSButtonPushed(app, event)
    app.PearsonTableMS_MS.Data = [];
    app.PearsonTableMS_MS.ColumnName = [];
    app.PearsonTableMS_MS.RowName = [];
    clear [row, col]
    removeStyle(app.PearsonTableMS_MS);
    app.p_valuesTable_MS_MS.Data = [];
    removeStyle(app.p_valuesTable_MS_MS);
    app.p_valuesTable_MS_MS.ColumnName = [];
    app.p_valuesTable_MS_MS.RowName = [];
    if app.PearsonButton.Value == true
        type = 'Pearson';
    elseif app.KendallButton.Value == true
        type = 'Kendall';
    elseif app.SpearmanButton.Value == true
        type = 'Spearman';
    end
    for i = 1:size(app.ListBoxEMCPlot.Items, 2)
        for j = 1:size(app.ListBoxEMCPlot.Items, 2)
            [tempR, tempP] =
corr(app.intensitysub(app.ListBoxEMCPlot.ItemsData(i),
app.fractions),
app.intensitysub(app.ListBoxEMCPlot.ItemsData(j),
app.fractions), 'Type', type);
            app.PearsonTableMS_MS.Data(i,j) =
tempR(1,1);
            app.p_valuesTable_MS_MS.Data(i,j) =
tempP(1,1);
        end
    end
    for i = 1:size(app.p_valuesTable_MS_MS.Data,2)
        app.p_valuesTable_MS_MS.ColumnFormat{1,i} =
{'shortG'};
    end
    app.PearsonTableMS_MS.ColumnName =
app.ListBoxEMCPlot.Items;
    app.PearsonTableMS_MS.RowName =
app.ListBoxEMCPlot.Items;
    app.p_valuesTable_MS_MS.ColumnName =
app.ListBoxEMCPlot.Items;
    app.p_valuesTable_MS_MS.RowName =
app.ListBoxEMCPlot.Items;
    s = uistyle('BackgroundColor', 'green');
    s1 = uistyle('FontColor', 'white',
'BackgroundColor', 'white');
    [row, col] = find(app.PearsonTableMS_MS.Data >
0.95);
    addStyle(app.PearsonTableMS_MS, s, 'cell', [row,
col]);
    addStyle(app.p_valuesTable_MS_MS, s, 'cell', [row,
col]);
%Tabelleneinträge in Diagonale weiß einfärben (Style s1)
for i = 1:size(app.PearsonTableMS_MS.ColumnName)
    for j = 1:size(app.PearsonTableMS_MS.RowName)
        if i == j
            addStyle(app.PearsonTableMS_MS, s1,
'cell', [i, j]);
            addStyle(app.p_valuesTable_MS_MS, s1,
'cell', [i, j]);
        end
    end
end
app.p_valuesTable_MS_MS.ColumnFormat =
{'shortE'};
app.TabGroup.SelectedTab = app.MSMSButton;
end

% Menu selected function: CopydatatoworkspaceMenu_MS
function CopydatatoworkspaceMenu_MS_Selected(app,
event)
    assignin('base', 'fractions_min',
min(app.fractions));
    assignin('base', 'fractions_max',
max(app.fractions));
    for i = 1:size(app.PearsonTableMS_MS.Data,1)
        PearsonTableMS_MS((i+1),1) =
app.PearsonTableMS_MS.RowName(i);
        PValuesMS_MS((i+1),1) =
app.PearsonTableMS_MS.RowName(i);
    end
    for i = 1:size(app.PearsonTableMS_MS.Data,2)
        PearsonTableMS_MS(1,(i+1)) =
app.PearsonTableMS_MS.ColumnName(i);
        PValuesMS_MS(1,(i+1)) =
app.PearsonTableMS_MS.ColumnName(i);
    end
    assignin('base', 'PearsonTableMS_MS',
PearsonTableMS_MS);
    assignin('base', 'PValuesMS_MS', PValuesMS_MS);
end

% Menu selected function: NormalizeEDDCIntensitiesinPlotMenu
function NormalizeEDDCIntensitiesinPlotMenuSelected(app, event)
%zurückrechnen der regulären Werte (reguläre Zeilennummern der
%13C-Verschiebung):
for i = 1:size(app.ListBoxEDDCPlot.ItemsData,2)
    Zeilennummern(1,i) =
(app.ListBoxEDDCPlot.ItemsData(1,i)-i)/10000;
end
%Alle Einträge in ListBox in EMC/EDC-Plot plotten
EDDC(size(app.ListBoxEDDCPlot.Items, 2),
size(app.nmrbuckets, 3)) = zeros; %Für Preallocation
for i = 1:size(app.nmrbuckets, 3) %Zähler über alle
Fraktionen
    for j = 1:size(app.ListBoxEDDCPlot.Items, 2)
        EDDC(j,i) =
app.nmrbuckets(Zeilennummern(j), app.EDDCPlotProton(j), i);
    end
end
for i = 1:size(EDDC, 1)
    maxValue = max(EDDC(:,app.fractions), [], 2);
    EDDC_norm(i,:) = (EDDC(i,:)/maxValue(i));
end
%EDDC-Plot:
cla(app.EDDCandEMCPlot);
plot(app.EDDCandEMCPlot, app.fractions,
(EDDC_norm(:,app.fractions)));
legend(app.EDDCandEMCPlot,
app.ListBoxEDDCPlot.Items, 'location', 'northeast');
end

% Menu selected function: NormalizeEMCIntensitiesinPlotMenu
function NormalizeEMCIntensitiesinPlotMenuSelected(app,
event)
    for i = 1:size(app.ListBoxEMCPlot.Items, 2)
        maxValue =
max(app.intensitysub(app.ListBoxEMCPlot.ItemsData,
app.fractions), [], 2);
        EMC_norm =
(app.intensitysub(app.ListBoxEMCPlot.ItemsData,
app.fractions))/maxValue;
    end
    cla(app.EDDCandEMCPlot);
    plot(app.EDDCandEMCPlot, app.fractions,
(EMC_norm));
    legend(app.EDDCandEMCPlot,
app.ListBoxEMCPlot.Items, 'location', 'northeast');
end

% Changes arrangement of the app based on UIFigure width
function updateAppLayout(app, event)
    currentFigureWidth =
app.SCOREmetaboliteIDSquaredUIFigure.Position(3);
    if(currentFigureWidth <= app.onePanelWidth)
        % Change to a 3x1 grid
        app.GridLayout.RowHeight = {1029, 1029, 1029};
        app.GridLayout.ColumnWidth = {'1x'};
        app.CenterPanel.Layout.Row = 1;
        app.CenterPanel.Layout.Column = 1;
        app.LeftPanel.Layout.Row = 2;
        app.LeftPanel.Layout.Column = 1;
        app.RightPanel.Layout.Row = 3;
        app.RightPanel.Layout.Column = 1;
    elseif (currentFigureWidth > app.onePanelWidth &&
currentFigureWidth <= app.twoPanelWidth)
        % Change to a 2x2 grid
        app.GridLayout.RowHeight = {1029, 1029};
        app.GridLayout.ColumnWidth = {'1x', '1x'};
        app.CenterPanel.Layout.Row = 1;
        app.CenterPanel.Layout.Column = [1,2];
        app.LeftPanel.Layout.Row = 2;
        app.LeftPanel.Layout.Column = 1;
        app.RightPanel.Layout.Row = 2;
        app.RightPanel.Layout.Column = 2;
    else
        % Change to a 1x3 grid
    end
end

```

```

        app.GridLayout.RowHeight = {'1x'};
        app.GridLayout.ColumnWidth = {245, '1x', 586};
        app.LeftPanel.Layout.Row = 1;
        app.LeftPanel.Layout.Column = 1;
        app.CenterPanel.Layout.Row = 1;
        app.CenterPanel.Layout.Column = 2;
        app.RightPanel.Layout.Row = 1;
        app.RightPanel.Layout.Column = 3;
    end
end
% Component initialization
methods (Access = private)

% Create UIFigure and components
function createComponents(app)
% Create SCOREmetaboliteIDSquaredUIFigure and hide until all
components are created
    app.SCOREmetaboliteIDSquaredUIFigure =
uifigure('Visible', 'off');
    app.SCOREmetaboliteIDSquaredUIFigure.AutoResizeChildren =
'off';
    app.SCOREmetaboliteIDSquaredUIFigure.Color =
[0.9412 0.9412 0.9412];
    app.SCOREmetaboliteIDSquaredUIFigure.Position =
[100 100 1736 1029];
    app.SCOREmetaboliteIDSquaredUIFigure.Name = 'SCORE-
metabolite-ID Squared';
    app.SCOREmetaboliteIDSquaredUIFigure.Icon =
'Lupe_Icon.png';
    app.SCOREmetaboliteIDSquaredUIFigure.SizeChangedFcn =
createCallbackFcn(app, @updateAppLayout, true);
    % Create GridLayout
    app.GridLayout =
uigridlayout(app.SCOREmetaboliteIDSquaredUIFigure);
    app.GridLayout.ColumnWidth = {245, '1x', 586};
    app.GridLayout.RowHeight = {'1x'};
    app.GridLayout.ColumnSpacing = 0;
    app.GridLayout.RowSpacing = 0;
    app.GridLayout.Padding = [0 0 0 0];
    app.GridLayout.Scrollable = 'on';
    % Create LeftPanel
    app.LeftPanel = uipanel(app.GridLayout);
    app.LeftPanel.BackgroundColor = [0.9412 0.9412
0.9412];
    app.LeftPanel.Layout.Row = 1;
    app.LeftPanel.Layout.Column = 1;
    % Create EDDCPlot
    app.EDDCPlot = uiaxes(app.LeftPanel);
    xlabel(app.EDDCPlot, 'fraction');
    ylabel(app.EDDCPlot, 'intensity');
    app.EDDCPlot.Position = [7 366 220 197];
    % Create ShowEMCButton
    app.ShowEMCButton = uibutton(app.LeftPanel,
'push');
    app.ShowEMCButton.ButtonPushedFcn =
createCallbackFcn(app, @ShowEMCButtonPushed, true);
    app.ShowEMCButton.Enable = 'off';
    app.ShowEMCButton.Position = [25 82 187 48];
    app.ShowEMCButton.Text = 'Show EMC';
    % Create ShowEDDCButton
    app.ShowEDDCButton = uibutton(app.LeftPanel,
'push');
    app.ShowEDDCButton.ButtonPushedFcn =
createCallbackFcn(app, @ShowEDDCButtonPushed, true);
    app.ShowEDDCButton.Enable = 'off';
    app.ShowEDDCButton.Position = [26 572 186 49];
    app.ShowEDDCButton.Text = 'Show EDDC';
    % Create SubtractBlankMSSpectrumButton
    app.SubtractBlankMSSpectrumButton =
uibutton(app.LeftPanel, 'push');
    app.SubtractBlankMSSpectrumButton.ButtonPushedFcn =
createCallbackFcn(app, @SubtractBlankMSSpectrumButtonPushed,
true);
    app.SubtractBlankMSSpectrumButton.Enable = 'off';
    app.SubtractBlankMSSpectrumButton.Position = [24
206 187 49];
    app.SubtractBlankMSSpectrumButton.Text = 'Subtract
Blank MS Spectrum';
    % Create ImportMSSSpectraButton
    app.ImportMSSSpectraButton = uibutton(app.LeftPanel,
'push');
    app.ImportMSSSpectraButton.ButtonPushedFcn =
createCallbackFcn(app, @ImportMSSSpectraButtonPushed, true);
    app.ImportMSSSpectraButton.Enable = 'off';
    app.ImportMSSSpectraButton.Position = [24 267 187
49];
    app.ImportMSSSpectraButton.Text = 'Import MS
Spectra';
    % Create ImportHSQCSpectraButton
    app.ImportHSQCSpectraButton =
uibutton(app.LeftPanel, 'push');
    app.ImportHSQCSpectraButton.ButtonPushedFcn =
createCallbackFcn(app, @ImportHSQCSpectraButtonPushed, true);
    app.ImportHSQCSpectraButton.Position = [24 927 187
49];
    app.ImportHSQCSpectraButton.Text = 'Import HSQC
Spectra';

    % Create NMRSpectraLabel
    app.NMRSpectraLabel = uilabel(app.LeftPanel);
    app.NMRSpectraLabel.FontSize = 16;
    app.NMRSpectraLabel.FontWeight = 'bold';
    app.NMRSpectraLabel.Position = [25 990 105 22];
    app.NMRSpectraLabel.Text = 'NMR Spectra';
    % Create MSSSpectraLabel
    app.MSSSpectraLabel = uilabel(app.LeftPanel);
    app.MSSSpectraLabel.FontSize = 16;
    app.MSSSpectraLabel.FontWeight = 'bold';
    app.MSSSpectraLabel.Position = [25 334 93 22];
    app.MSSSpectraLabel.Text = 'MS Spectra';
    % Create HchemicalshiftEditField
    app.HchemicalshiftEditField =
uieditfield(app.LeftPanel, 'numeric');
    app.HchemicalshiftEditField.ValueChangedFcn =
createCallbackFcn(app, @HchemicalshiftEditFieldValueChanged,
true);
    app.HchemicalshiftEditField.Enable = 'off';
    app.HchemicalshiftEditField.Position = [133 692 80
46];
    % Create HchemicalshiftEditFieldLabel
    app.HchemicalshiftEditFieldLabel =
uilabel(app.LeftPanel);
    app.HchemicalshiftEditFieldLabel.HorizontalAlignment = 'right';
    app.HchemicalshiftEditFieldLabel.Position = [17 691
97 47];
    app.HchemicalshiftEditFieldLabel.Text = '1H
chemical shift';
    % Create CchemicalshiftEditField
    app.CchemicalshiftEditField =
uieditfield(app.LeftPanel, 'numeric');
    app.CchemicalshiftEditField.ValueChangedFcn =
createCallbackFcn(app, @CchemicalshiftEditFieldValueChanged,
true);
    app.CchemicalshiftEditField.Enable = 'off';
    app.CchemicalshiftEditField.Position = [134 632 80
46];
    % Create CchemicalshiftEditFieldLabel
    app.CchemicalshiftEditFieldLabel =
uilabel(app.LeftPanel);
    app.CchemicalshiftEditFieldLabel.HorizontalAlignment = 'right';
    app.CchemicalshiftEditFieldLabel.Position = [12 631
103 47];
    app.CchemicalshiftEditFieldLabel.Text = '13C
chemical shift';
    % Create mzvalueofEMCEditField
    app.mzvalueofEMCEditField =
uieditfield(app.LeftPanel, 'numeric');
    app.mzvalueofEMCEditField.ValueChangedFcn =
createCallbackFcn(app, @mzvalueofEMCEditFieldValueChanged,
true);
    app.mzvalueofEMCEditField.Enable = 'off';
    app.mzvalueofEMCEditField.Position = [132 146 80
46];
    % Create mzvalueofEMCEditFieldLabel
    app.mzvalueofEMCEditFieldLabel =
uilabel(app.LeftPanel);
    app.mzvalueofEMCEditFieldLabel.HorizontalAlignment =
'right';
    app.mzvalueofEMCEditFieldLabel.Position = [22 145
100 47];
    app.mzvalueofEMCEditFieldLabel.Text = 'm/z value of
EMC';
    % Create BucketingButton
    app.BucketingButton = uibutton(app.LeftPanel,
'push');
    app.BucketingButton.ButtonPushedFcn =
createCallbackFcn(app, @BucketingButtonPushed, true);
    app.BucketingButton.Enable = 'off';
    app.BucketingButton.Position = [24 748 109 49];
    app.BucketingButton.Text = 'Bucketing';
    % Create CalibrationButton
    app.CalibrationButton = uibutton(app.LeftPanel,
'push');
    app.CalibrationButton.ButtonPushedFcn =
createCallbackFcn(app, @CalibrationButtonPushed, true);
    app.CalibrationButton.Enable = 'off';
    app.CalibrationButton.Position = [24 866 187 49];
    app.CalibrationButton.Text = 'Calibration';
    % Create NormalizationButton
    app.NormalizationButton = uibutton(app.LeftPanel,
'push');
    app.NormalizationButton.ButtonPushedFcn =
createCallbackFcn(app, @NormalizationButtonPushed, true);
    app.NormalizationButton.Enable = 'off';
    app.NormalizationButton.Position = [24 807 187 49];
    app.NormalizationButton.Text = 'Normalization';
    % Create BucketsizeLabel
    app.BucketsizeLabel = uilabel(app.LeftPanel);
    app.BucketsizeLabel.HorizontalAlignment = 'center';
    app.BucketsizeLabel.Enable = 'off';
    app.BucketsizeLabel.Position = [146 748 67 21];
    app.BucketsizeLabel.Text = 'Bucket size';
    % Create BucketsizeEditField

```

```

        app.BucketsizeEditField =
uieditfield(app.LeftPanel, 'text');
        app.BucketsizeEditField.Enable = 'off';
        app.BucketsizeEditField.Position = [146 772 65 24];
        % Create ESIionizationmodeButtonGroup
        app.ESIionizationmodeButtonGroup =
uibuttongroup(app.LeftPanel);
        app.ESIionizationmodeButtonGroup.Enable = 'off';
        app.ESIionizationmodeButtonGroup.Title = 'ESI
ionization mode';
        app.ESIionizationmodeButtonGroup.BackgroundColor =
[0.9608 0.9608 0.9412];
        app.ESIionizationmodeButtonGroup.Position = [25 23
187 48];
        % Create positiveButton
        app.positiveButton =
uiradiobutton(app.ESIionizationmodeButtonGroup);
        app.positiveButton.Text = 'positive';
        app.positiveButton.Position = [15 2 62 22];
        app.positiveButton.Value = true;
        % Create negativeButton
        app.negativeButton =
uiradiobutton(app.ESIionizationmodeButtonGroup);
        app.negativeButton.Text = 'negative';
        app.negativeButton.Position = [96 2 67 22];
        % Create ChangeTheBucketsizeButton
        app.ChangeTheBucketsizeButton =
uibutton(app.LeftPanel, 'push');
        app.ChangeTheBucketsizeButton.ButtonPushedFcn =
createCallbackFcn(app, @ChangeTheBucketsizeButtonPushed, true);
        app.ChangeTheBucketsizeButton.Enable = 'off';
        app.ChangeTheBucketsizeButton.Visible = 'off';
        app.ChangeTheBucketsizeButton.Position = [24 748
109 49];
        app.ChangeTheBucketsizeButton.Text = {'Change the
'; 'Bucket size'};
        % Create CenterPanel
        app.CenterPanel = uipanel(app.GridLayout);
        app.CenterPanel.Layout.Row = 1;
        app.CenterPanel.Layout.Column = 2;
        % Create MSSpectra
        app.MSSpectra = uiaxes(app.CenterPanel);
        title(app.MSSpectra, 'MS Spectra');
        xlabel(app.MSSpectra, 'fraction');
        ylabel(app.MSSpectra, 'm/z');
        zlabel(app.MSSpectra, 'intensity');
        app.MSSpectra.YDir = 'reverse';
        app.MSSpectra.SortMethod = 'depth';
        app.MSSpectra.Position = [31 38 692 460];
        % Create NMRspectra
        app.NMRspectra = uiaxes(app.CenterPanel);
        title(app.NMRspectra, 'HSQC Spectra');
        xlabel(app.NMRspectra, '1H chemical shift [ppm]');
        ylabel(app.NMRspectra, '13C chemical shift [ppm]');
        app.NMRspectra.XDir = 'reverse';
        app.NMRspectra.YDir = 'reverse';
        app.NMRspectra.YAxisLocation = 'right';
        app.NMRspectra.SortMethod = 'depth';
        app.NMRspectra.Position = [31 38 692 460];
        % Create EditField_2Label
        app.EditField_2Label = uilabel(app.CenterPanel);
        app.EditField_2Label.HorizontalAlignment = 'right';
        app.EditField_2Label.Position = [781 82 25 22];
        app.EditField_2Label.Text = '-';
        % Create FractionstoplotListBox
        app.FractionstoplotListBox =
uilibox(app.CenterPanel);
        app.FractionstoplotListBox.Items = {};
        app.FractionstoplotListBox.Multiselect = 'on';
        app.FractionstoplotListBox.ValueChangedFcn =
createCallbackFcn(app, @FractionstoplotListBoxValueChanged,
true);
        app.FractionstoplotListBox.Enable = 'off';
        app.FractionstoplotListBox.Position = [749 178 113
689];
        app.FractionstoplotListBox.Value = {};
        % Create FractionstoplotListBoxLabel
        app.FractionstoplotListBoxLabel =
uilabel(app.CenterPanel);
        app.FractionstoplotListBoxLabel.HorizontalAlignment
= 'right';
        app.FractionstoplotListBoxLabel.FontWeight =
'bold';
        app.FractionstoplotListBoxLabel.Position = [756 866
99 22];
        app.FractionstoplotListBoxLabel.Text = 'Fractions
to plot';
        % Create mzLabel
        app.mzLabel = uilabel(app.CenterPanel);
        app.mzLabel.Position = [868 82 25 22];
        app.mzLabel.Text = 'm/z';
        % Create mzregionLabel
        app.mzregionLabel = uilabel(app.CenterPanel);
        app.mzregionLabel.Position = [743 118 65 22];
        app.mzregionLabel.Text = 'm/z region';
        % Create lowermzEditField
        app.lowermzEditField = uieditfield(app.CenterPanel,
'numeric');
        app.lowermzEditField.ValueChangedFcn =
createCallbackFcn(app, @lowermzEditFieldValueChanged, true);
        app.lowermzEditField.Position = [744 76 52 34];
        % Create uppermzEditField
        app.uppermzEditField = uieditfield(app.CenterPanel,
'numeric');
        app.uppermzEditField.ValueChangedFcn =
createCallbackFcn(app, @uppermzEditFieldValueChanged, true);
        app.uppermzEditField.Position = [810 76 52 34];
        % Create ContourLevelsLabel
        app.ContourLevelsLabel = uilabel(app.CenterPanel);
        app.ContourLevelsLabel.HorizontalAlignment =
'right';
        app.ContourLevelsLabel.Position = [750 964 86 22];
        app.ContourLevelsLabel.Text = 'Contour levels: ';
        % Create ContourLevelsSlider
        app.ContourLevelsSlider =
uislider(app.CenterPanel);
        app.ContourLevelsSlider.ValueChangedFcn =
createCallbackFcn(app, @ContourLevelsSliderValueChanged, true);
        app.ContourLevelsSlider.Enable = 'off';
        app.ContourLevelsSlider.Position = [753 945 107 7];
        app.ContourLevelsSlider.Value = 50;
        % Create PearsonTableMS_NMR_alt
        app.PearsonTableMS_NMR_alt =
uitable(app.CenterPanel);
        app.PearsonTableMS_NMR_alt.ColumnName = '';
        app.PearsonTableMS_NMR_alt.RowName = {};
        app.PearsonTableMS_NMR_alt.Enable = 'off';
        app.PearsonTableMS_NMR_alt.Visible = 'off';
        app.PearsonTableMS_NMR_alt.Position = [193 641 447
262];
        % Create PearsonTableNMR_NMR_alt
        app.PearsonTableNMR_NMR_alt =
uitable(app.CenterPanel);
        app.PearsonTableNMR_NMR_alt.ColumnName = '';
        app.PearsonTableNMR_NMR_alt.RowName = {};
        app.PearsonTableNMR_NMR_alt.Enable = 'off';
        app.PearsonTableNMR_NMR_alt.Visible = 'off';
        app.PearsonTableNMR_NMR_alt.Position = [193 356 447
262];
        % Create RightPanel
        app.RightPanel = uipanel(app.GridLayout);
        app.RightPanel.BackgroundColor = [0.9412 0.9412
0.9412];
        app.RightPanel.Layout.Row = 1;
        app.RightPanel.Layout.Column = 3;
        % Create EDDCandEMCPlot
        app.EDDCandEMCPlot = uiaxes(app.RightPanel);
        title(app.EDDCandEMCPlot, 'EDDC & EMC');
        xlabel(app.EDDCandEMCPlot, 'fraction');
        ylabel(app.EDDCandEMCPlot, 'intensity');
        zlabel(app.EDDCandEMCPlot, 'Z');
        app.EDDCandEMCPlot.YColor = [0.149 0.149 0.149];
        app.EDDCandEMCPlot.NextPlot = 'add';
        app.EDDCandEMCPlot.Position = [11 726 283 261];
        % Create ListBoxEDDCPlot
        app.ListBoxEDDCPlot = uilibox(app.RightPanel);
        app.ListBoxEDDCPlot.Items = {};
        app.ListBoxEDDCPlot.Multiselect = 'on';
        app.ListBoxEDDCPlot.ValueChangedFcn =
createCallbackFcn(app, @ListBoxEDDCPlotValueChanged, true);
        app.ListBoxEDDCPlot.Enable = 'off';
        app.ListBoxEDDCPlot.Position = [318 777 117 141];
        app.ListBoxEDDCPlot.Value = {};
        % Create DeleteEDDCButton
        app.DeleteEDDCButton = uibutton(app.RightPanel,
'push');
        app.DeleteEDDCButton.ButtonPushedFcn =
createCallbackFcn(app, @DeleteEDDCButtonPushed, true);
        app.DeleteEDDCButton.Enable = 'off';
        app.DeleteEDDCButton.Position = [317 726 118 44];
        app.DeleteEDDCButton.Text = 'Delete EDDC';
        % Create ListBoxEMCPlot
        app.ListBoxEMCPlot = uilibox(app.RightPanel);
        app.ListBoxEMCPlot.Items = {};
        app.ListBoxEMCPlot.Multiselect = 'on';
        app.ListBoxEMCPlot.ValueChangedFcn =
createCallbackFcn(app, @ListBoxEMCPlotValueChanged, true);
        app.ListBoxEMCPlot.Enable = 'off';
        app.ListBoxEMCPlot.Position = [447 777 117 141];
        app.ListBoxEMCPlot.Value = {};
        % Create DeleteEMCButton
        app.DeleteEMCButton = uibutton(app.RightPanel,
'push');
        app.DeleteEMCButton.ButtonPushedFcn =
createCallbackFcn(app, @DeleteEMCButtonPushed, true);
        app.DeleteEMCButton.Enable = 'off';
        app.DeleteEMCButton.Position = [447 726 118 44];
        app.DeleteEMCButton.Text = 'Delete EMC';
        % Create AddEDDCButton
        app.AddEDDCButton = uibutton(app.RightPanel,
'push');
        app.AddEDDCButton.ButtonPushedFcn =
createCallbackFcn(app, @AddEDDCButtonPushed, true);
        app.AddEDDCButton.Enable = 'off';

```



```

app.AddEDDCButton.Position = [318 926 117 44];
app.AddEDDCButton.Text = 'Add EDDC';
% Create AddEMCButton
app.AddEMCButton = uibutton(app.RightPanel,
'push');
app.AddEMCButton.ButtonPushedFcn =
createCallbackFcn(app, @AddEMCButtonPushed, true);
app.AddEMCButton.Enable = 'off';
app.AddEMCButton.Position = [447 926 117 44];
app.AddEMCButton.Text = 'Add EMC';
% Create PlotEDDCButton
app.PlotEDDCButton = uibutton(app.RightPanel,
'push');
app.PlotEDDCButton.ButtonPushedFcn =
createCallbackFcn(app, @PlotEDDCButtonPushed, true);
app.PlotEDDCButton.Enable = 'off';
app.PlotEDDCButton.Position = [31 666 105 45];
app.PlotEDDCButton.Text = 'Plot EDDC';
% Create PlotEMCButton
app.PlotEMCButton = uibutton(app.RightPanel,
'push');
app.PlotEMCButton.ButtonPushedFcn =
createCallbackFcn(app, @PlotEMCButtonPushed, true);
app.PlotEMCButton.Enable = 'off';
app.PlotEMCButton.Position = [178 665 106 46];
app.PlotEMCButton.Text = 'Plot EMC';
% Create ClearPlotButton
app.ClearPlotButton = uibutton(app.RightPanel,
'push');
app.ClearPlotButton.ButtonPushedFcn =
createCallbackFcn(app, @ClearPlotButtonPushed, true);
app.ClearPlotButton.Enable = 'off';
app.ClearPlotButton.Position = [104 606 106 46];
app.ClearPlotButton.Text = 'Clear Plot';
% Create ShowallcorrelatedEDDCsButton
app.ShowallcorrelatedEDDCsButton =
uibutton(app.RightPanel, 'push');
app.ShowallcorrelatedEDDCsButton.ButtonPushedFcn =
createCallbackFcn(app, @ShowallcorrelatedEDDCsButtonPushed,
true);
app.ShowallcorrelatedEDDCsButton.WordWrap = 'on';
app.ShowallcorrelatedEDDCsButton.Enable = 'off';
app.ShowallcorrelatedEDDCsButton.Position = [318 651 117 58];
app.ShowallcorrelatedEDDCsButton.Text = 'Show all
correlated EDDCs';
% Create ShowallcorrelatedEMCsButton
app.ShowallcorrelatedEMCsButton =
uibutton(app.RightPanel, 'push');
app.ShowallcorrelatedEMCsButton.ButtonPushedFcn =
createCallbackFcn(app, @ShowallcorrelatedEMCsButtonPushed,
true);
app.ShowallcorrelatedEMCsButton.WordWrap = 'on';
app.ShowallcorrelatedEMCsButton.Enable = 'off';
app.ShowallcorrelatedEMCsButton.Position = [318 651 117 58];
app.ShowallcorrelatedEMCsButton.Text = 'Show all
correlated EMCs';
% Create ShowallcorrelatedEDDCstoEDDCButton
app.ShowallcorrelatedEDDCstoEDDCButton =
uibutton(app.RightPanel, 'push');
app.ShowallcorrelatedEDDCstoEDDCButton.ButtonPushedFcn =
createCallbackFcn(app,
@ShowallcorrelatedEDDCstoEDDCButtonPushed, true);
app.ShowallcorrelatedEDDCstoEDDCButton.WordWrap =
'on';
app.ShowallcorrelatedEDDCstoEDDCButton.Enable =
'off';
app.ShowallcorrelatedEDDCstoEDDCButton.Position =
[319 585 116 58];
app.ShowallcorrelatedEDDCstoEDDCButton.Text = 'Show
all correlated EDDCs';
% Create ShowallcorrelatedEMCstoEMCButton
app.ShowallcorrelatedEMCstoEMCButton =
uibutton(app.RightPanel, 'push');
app.ShowallcorrelatedEMCstoEMCButton.ButtonPushedFcn =
createCallbackFcn(app, @ShowallcorrelatedEMCstoEMCButtonPushed,
true);
app.ShowallcorrelatedEMCstoEMCButton.WordWrap =
'on';
app.ShowallcorrelatedEMCstoEMCButton.Enable =
'off';
app.ShowallcorrelatedEMCstoEMCButton.Position =
[447 585 116 58];
app.ShowallcorrelatedEMCstoEMCButton.Text = 'Show
all correlated EMCs';
% Create EDDCsLabel
app.EDDCsLabel = uilabel(app.RightPanel);
app.EDDCsLabel.FontSize = 16;
app.EDDCsLabel.FontWeight = 'bold';
app.EDDCsLabel.Position = [346 983 60 22];
app.EDDCsLabel.Text = 'EDDCs';
% Create EMCsLabel
app.EMCsLabel = uilabel(app.RightPanel);
app.EMCsLabel.FontSize = 16;
app.EMCsLabel.FontWeight = 'bold';
app.EMCsLabel.Position = [480 983 50 22];

app.EMCsLabel.Text = 'EMCs';
% Create Label
app.Label = uilabel(app.RightPanel);
app.Label.HorizontalAlignment = 'center';
app.Label.Position = [9 564 557 22];
app.Label.Text =
';
% Create NMRMSButton
app.NMRMSButton = uibutton(app.RightPanel, 'push');
createCallbackFcn(app, @NMRMSButtonPushed, true);
app.NMRMSButton.WordWrap = 'on';
app.NMRMSButton.FontWeight = 'bold';
app.NMRMSButton.Enable = 'off';
app.NMRMSButton.Position = [135 517 126 46];
app.NMRMSButton.Text = 'NMR + MS';
% Create TabGroup
app.TabGroup = uitabgroup(app.RightPanel);
app.TabGroup.Position = [20 19 542 446];
% Create NMRMSTab
app.NMRMSTab = uitab(app.TabGroup);
app.NMRMSTab.Title = 'NMR + MS';
% Create TabGroup_MS_NMR
app.TabGroup_MS_NMR = uitabgroup(app.NMRMSTab);
app.TabGroup_MS_NMR.Position = [0 -2 541 424];
% Create Coefficients
app.Coefficients = uitab(app.TabGroup_MS_NMR);
app.Coefficients.Title = 'Coefficients';
% Create PearsonTableMS_NMR
app.PearsonTableMS_NMR = uitable(app.Coefficients);
app.PearsonTableMS_NMR.ColumnName = '';
app.PearsonTableMS_NMR.RowName = {};
app.PearsonTableMS_NMR.Position = [1 -1 539 400];
% Create p_values
app.p_values = uitab(app.TabGroup_MS_NMR);
app.p_values.Title = 'p-Values';
% Create p_ValuesTable_MS_NMR
app.p_ValuesTable_MS_NMR = uitable(app.p_values);
app.p_ValuesTable_MS_NMR.ColumnName = '';
app.p_ValuesTable_MS_NMR.RowName = {};
app.p_ValuesTable_MS_NMR.Position = [1 0 538 399];
% Create NMRNMRTab
app.NMRNMRTab = uitab(app.TabGroup);
app.NMRNMRTab.Title = 'NMR + NMR';
% Create TabGroup_NMR_NMR
app.TabGroup_NMR_NMR = uitabgroup(app.NMRNMRTab);
app.TabGroup_NMR_NMR.Position = [0 -2 542 424];
% Create CoefficientsTab
app.CoefficientsTab = uitab(app.TabGroup_NMR_NMR);
app.CoefficientsTab.Title = 'Coefficients';
% Create PearsonTableNMR_NMR
app.PearsonTableNMR_NMR =
uitable(app.CoefficientsTab);
app.PearsonTableNMR_NMR.ColumnName = '';
app.PearsonTableNMR_NMR.RowName = {};
app.PearsonTableNMR_NMR.Position = [1 0 541 398];
% Create pValuesTab
app.pValuesTab = uitab(app.TabGroup_NMR_NMR);
app.pValuesTab.Title = 'p-Values';
% Create p_valuesTable_NMR_NMR
app.p_valuesTable_NMR_NMR =
uitable(app.pValuesTab);
app.p_valuesTable_NMR_NMR.ColumnName = '';
app.p_valuesTable_NMR_NMR.RowName = {};
app.p_valuesTable_NMR_NMR.Position = [1 0 541 398];
% Create MSMSTab
app.MSMSTab = uitab(app.TabGroup);
app.MSMSTab.Title = 'MS + MS';
% Create TabGroup_MS_MS
app.TabGroup_MS_MS = uitabgroup(app.MSMSTab);
app.TabGroup_MS_MS.Position = [1 -2 540 425];
% Create CoefficientsTab_MS_MS
app.CoefficientsTab_MS_MS =
uitab(app.TabGroup_MS_MS);
app.CoefficientsTab_MS_MS.Title = 'Coefficients';
% Create PearsonTableMS_MS
app.PearsonTableMS_MS =
uitable(app.CoefficientsTab_MS_MS);
app.PearsonTableMS_MS.ColumnName = '';
app.PearsonTableMS_MS.RowName = {};
app.PearsonTableMS_MS.Position = [1 1 536 398];
% Create pValuesTab_MSMS
app.pValuesTab_MSMS = uitab(app.TabGroup_MS_MS);
app.pValuesTab_MSMS.Title = 'p-Values';
% Create p_valuesTable_MS_MS
app.p_valuesTable_MS_MS =
uitable(app.pValuesTab_MSMS);
app.p_valuesTable_MS_MS.ColumnName = '';
app.p_valuesTable_MS_MS.RowName = {};
app.p_valuesTable_MS_MS.Position = [1 0 538 399];
% Create NMRNMRButton
app.NMRNMRButton = uibutton(app.RightPanel,
'push');
app.NMRNMRButton.ButtonPushedFcn =
createCallbackFcn(app, @NMRNMRButtonPushed, true);
app.NMRNMRButton.WordWrap = 'on';
app.NMRNMRButton.FontWeight = 'bold';

```

```

        app.NMRNMRButton.Enable = 'off';
        app.NMRNMRButton.Position = [284 517 126 46];
        app.NMRNMRButton.Text = 'NMR + NMR';
        % Create CorrelationCoefficientsLabel
        app.CorrelationCoefficientsLabel =
        uilabel(app.RightPanel);
        app.CorrelationCoefficientsLabel.HorizontalAlignment =
        'center';
        app.CorrelationCoefficientsLabel.FontSize = 16;
        app.CorrelationCoefficientsLabel.FontWeight =
        'bold';
        app.CorrelationCoefficientsLabel.Position = [20 519
67 39];
        app.CorrelationCoefficientsLabel.Text =
        {'Correlation'; 'Coefficients'};
        % Create MSMSButton
        app.MSMSButton = uibutton(app.RightPanel, 'push');
        app.MSMSButton.ButtonPushedFcn =
        createCallbackFcn(app, @MSMSButtonPushed, true);
        app.MSMSButton.WordWrap = 'on';
        app.MSMSButton.FontWeight = 'bold';
        app.MSMSButton.Enable = 'off';
        app.MSMSButton.Position = [432 517 126 46];
        app.MSMSButton.Text = 'MS + MS';
        % Create ButtonGroup
        app.ButtonGroup = uibuttongroup(app.RightPanel);
        app.ButtonGroup.Position = [136 473 422 36];
        % Create PearsonButton
        app.PearsonButton = uiradiobutton(app.ButtonGroup);
        app.PearsonButton.Enable = 'off';
        app.PearsonButton.Text = 'Pearson';
        app.PearsonButton.Position = [29 7 67 22];
        app.PearsonButton.Value = true;
        % Create KendallButton
        app.KendallButton = uiradiobutton(app.ButtonGroup);
        app.KendallButton.Enable = 'off';
        app.KendallButton.Text = 'Kendall';
        app.KendallButton.Position = [177 7 65 22];
        % Create SpearmanButton
        app.SpearmanButton =
        uiradiobutton(app.ButtonGroup);
        app.SpearmanButton.Enable = 'off';
        app.SpearmanButton.Text = 'Spearman';
        app.SpearmanButton.Position = [322 6 77 22];
        % Create TypeofcorrelationLabel
        app.TypeofcorrelationLabel =
        uilabel(app.RightPanel);
        app.TypeofcorrelationLabel.HorizontalAlignment =
        'center';
        app.TypeofcorrelationLabel.Position = [20 480 104
22];
        app.TypeofcorrelationLabel.Text = 'Type of
correlation';
        % Create ContextMenu_NMR_MS
        app.ContextMenu_NMR_MS =
        uicontextmenu(app.SCOREmetaboliteIDSquaredUIFigure);
        % Create Copydatatoworkspace_NMRMS
        app.Copydatatoworkspace_NMRMS =
        uimenu(app.ContextMenu_NMR_MS);
        app.Copydatatoworkspace_NMRMS.MenuSelectedFcn =
        createCallbackFcn(app, @Copydatatoworkspace_NMRMSSelected,
true);
        app.Copydatatoworkspace_NMRMS.Text = 'Copy data to
workspace';
        % Assign app.ContextMenu_NMR_MS
        app.PearsonTableMS_NMR_alt.ContextMenu =
        app.ContextMenu_NMR_MS;
        app.TabGroup_MS_NMR.ContextMenu =
        app.ContextMenu_NMR_MS;
        app.PearsonTableMS_NMR.ContextMenu =
        app.ContextMenu_NMR_MS;
        app.p_valuesTable_MS_NMR.ContextMenu =
        app.ContextMenu_NMR_MS;
        % Create ContextMenu_NMR_NMR
        app.ContextMenu_NMR_NMR =
        uicontextmenu(app.SCOREmetaboliteIDSquaredUIFigure);
        % Create Copydatatoworkspace_NMRNMR
        app.Copydatatoworkspace_NMRNMR =
        uimenu(app.ContextMenu_NMR_NMR);
        app.Copydatatoworkspace_NMRNMR.MenuSelectedFcn =
        createCallbackFcn(app, @Copydatatoworkspace_NMRNMRMenuSelected,
true);
        app.Copydatatoworkspace_NMRNMR.Text = 'Copy data to
workspace';
        % Assign app.ContextMenu_NMR_NMR
        app.PearsonTableNMR_NMR_alt.ContextMenu =
        app.ContextMenu_NMR_NMR;
        app.TabGroup_NMR_NMR.ContextMenu =
        app.ContextMenu_NMR_NMR;
        app.PearsonTableNMR_NMR.ContextMenu =
        app.ContextMenu_NMR_NMR;
        app.p_valuesTable_NMR_NMR.ContextMenu =
        app.ContextMenu_NMR_NMR;
        % Create ContextMenu_CopyNMRsSpectra
        app.ContextMenu_CopyNMRsSpectra =
        uicontextmenu(app.SCOREmetaboliteIDSquaredUIFigure);
        % Create CopyPlotofNMRsSpectraMenu
        app.CopyPlotofNMRsSpectraMenu =
        uimenu(app.ContextMenu_CopyNMRsSpectra);
        app.CopyPlotofNMRsSpectraMenu.MenuSelectedFcn =
        createCallbackFcn(app, @CopyPlotofNMRsSpectraMenuSelected,
true);
        app.CopyPlotofNMRsSpectraMenu.Text = 'Copy Plot of
NMR Spectra';
        % Assign app.ContextMenu_CopyNMRsSpectra
        app.NMRsSpectra.ContextMenu =
        app.ContextMenu_CopyNMRsSpectra;
        % Create ContextMenu_CopyMSSSpectra
        app.ContextMenu_CopyMSSSpectra =
        uicontextmenu(app.SCOREmetaboliteIDSquaredUIFigure);
        % Create CopyPlotofMSSSpectraMenu
        app.CopyPlotofMSSSpectraMenu =
        uimenu(app.ContextMenu_CopyMSSSpectra);
        app.CopyPlotofMSSSpectraMenu.MenuSelectedFcn =
        createCallbackFcn(app, @CopyPlotofMSSSpectraMenuSelected, true);
        app.CopyPlotofMSSSpectraMenu.Text = 'Copy Plot of MS
Spectra';
        % Assign app.ContextMenu_CopyMSSSpectra
        app.MSSSpectra.ContextMenu =
        app.ContextMenu_CopyMSSSpectra;
        % Create ContextMenu_CopyEDDCandEMCPlot
        app.ContextMenu_CopyEDDCandEMCPlot =
        uicontextmenu(app.SCOREmetaboliteIDSquaredUIFigure);
        % Create CopyEDDCorEMCPlotMenu
        app.CopyEDDCorEMCPlotMenu =
        uimenu(app.ContextMenu_CopyEDDCandEMCPlot);
        app.CopyEDDCorEMCPlotMenu.MenuSelectedFcn =
        createCallbackFcn(app, @CopyEDDCorEMCPlotMenuSelected, true);
        app.CopyEDDCorEMCPlotMenu.Text = 'Copy EDDC or EMC
Plot';
        % Create CreatePlotofEDDCandEMCMenu
        app.CreatePlotofEDDCandEMCMenu =
        uimenu(app.ContextMenu_CopyEDDCandEMCPlot);
        app.CreatePlotofEDDCandEMCMenu.MenuSelectedFcn =
        createCallbackFcn(app, @CreatePlotofEDDCandEMCMenuSelected,
true);
        app.CreatePlotofEDDCandEMCMenu.Text = 'Create Plot
of EDDC and EMC';
        % Create NormalizeEDDCIntensitiesinPlotMenu
        app.NormalizeEDDCIntensitiesinPlotMenu =
        uimenu(app.ContextMenu_CopyEDDCandEMCPlot);
        app.NormalizeEDDCIntensitiesinPlotMenu.MenuSelectedFcn =
        createCallbackFcn(app, @NormalizeEDDCIntensitiesinPlotMenuSelected, true);
        app.NormalizeEDDCIntensitiesinPlotMenu.Enable =
        'off';
        app.NormalizeEDDCIntensitiesinPlotMenu.Text =
        'Normalize EDDC Intensities in Plot';
        % Create NormalizeEMCIntensitiesinPlotMenu
        app.NormalizeEMCIntensitiesinPlotMenu =
        uimenu(app.ContextMenu_CopyEDDCandEMCPlot);
        app.NormalizeEMCIntensitiesinPlotMenu.MenuSelectedFcn =
        createCallbackFcn(app, @NormalizeEMCIntensitiesinPlotMenuSelected, true);
        app.NormalizeEMCIntensitiesinPlotMenu.Enable =
        'off';
        app.NormalizeEMCIntensitiesinPlotMenu.Text =
        'Normalize EMC Intensities in Plot';
        % Assign app.ContextMenu_CopyEDDCandEMCPlot
        app.EDDCandEMCPlot.ContextMenu =
        app.ContextMenu_CopyEDDCandEMCPlot;
        % Create ContextMenu_CopyEDDCPlot
        app.ContextMenu_CopyEDDCPlot =
        uicontextmenu(app.SCOREmetaboliteIDSquaredUIFigure);
        % Create CopyEDDCPlotMenu
        app.CopyEDDCPlotMenu =
        uimenu(app.ContextMenu_CopyEDDCPlot);
        app.CopyEDDCPlotMenu.MenuSelectedFcn =
        createCallbackFcn(app, @CopyEDDCPlotMenuSelected, true);
        app.CopyEDDCPlotMenu.Text = 'Copy EDDC Plot';
        % Assign app.ContextMenu_CopyEDDCPlot
        app.EDDCPlot.ContextMenu =
        app.ContextMenu_CopyEDDCPlot;
        % Create ContextMenu_MS_MS
        app.ContextMenu_MS_MS =
        uicontextmenu(app.SCOREmetaboliteIDSquaredUIFigure);
        % Create CopydatatoworkspaceMenu_MS_MS
        app.CopydatatoworkspaceMenu_MS_MS =
        uimenu(app.ContextMenu_MS_MS);
        app.CopydatatoworkspaceMenu_MS_MS.MenuSelectedFcn =
        createCallbackFcn(app, @CopydatatoworkspaceMenu_MS_MSSelected,
true);
        app.CopydatatoworkspaceMenu_MS_MS.Text = 'Copy data
to workspace';
        % Assign app.ContextMenu_MS_MS
        app.TabGroup_MS_MS.ContextMenu =
        app.ContextMenu_MS_MS;
        app.PearsonTableMS_MS.ContextMenu =
        app.ContextMenu_MS_MS;
        app.p_valuesTable_MS_MS.ContextMenu =
        app.ContextMenu_MS_MS;
        % Show the figure after all components are created
        app.SCOREmetaboliteIDSquaredUIFigure.Visible =
        'on';

```

```

end
end
% App creation and deletion
methods (Access = public)
% Construct app
function app = SCORE_Metabolite_ID_Squared
% Create UIFigure and components
createComponents(app)
% Register the app with App Designer
registerApp(app,
app.SCOREmetaboliteIDSquaredUIFigure)

if nargin == 0
clear app
end
end
% Code that executes before app deletion
function delete(app)
% Delete UIFigure when app is deleted
delete(app.SCOREmetaboliteIDSquaredUIFigure)
end
end
end

```

8.6.3 Skript zur Prozessierung der DI-Massenspektren

Das folgende Skript wurde in Compass Data Analysis 4.2 (BRUKER DALTONICS GmbH) zur automatischen Prozessierung der Massenspektren inkl. interner Kalibrierung gegen den Kalibrierstandard genutzt. Die gewählten Retentionszeitfenster wurden wie in Abschnitt 6.11.1 beschrieben angepasst.

```

Dim TIC
Set TIC = CreateObject(„DataAnalysis.TICChromatogramDefinition“)

Analysis.Chromatograms.AddChromatogram TIC
Analysis.Chromatograms.DeleteChromatogram 1

Analysis.AddChromatogramRangeSelection 2.3, 2.3, 0, 0

Dim MyChromatogram
For each MyChromatogram in Analysis.Chromatograms
    MyChromatogram.AverageMassSpectrum true, false
Next

Analysis.ClearChromatogramRangeSelections
Analysis.AddChromatogramRangeSelection 0.2, 0.6, 0, 0

Dim MyChromatogram2
For each MyChromatogram in Analysis.Chromatograms
    MyChromatogram.AverageMassSpectrum true, false
Next

Analysis.RecalibrateInternal

```

8.6.4 Skript „DataInput SMARTNMR“

Das folgende MATLAB-Skript „DataInput_SMARTNMR“ wurde zur einfachen Verknüpfung der *SCORE-metabolite-ID*² Methode mit SMART 2.0 (b) entwickelt. Darüber können die durch die *SCORE-metabolite-ID*² Methode ermittelten HSQC-Korrelationssignale in das entsprechende

Format zur direkten Eingabe auf der Website des SMART 2.0 Webtools überführt werden.^[200] Die im Skript verwendete Funktion *findClusters* wurde aus der *“2D NMR Spectra Fitter”* Toolbox übernommen.^[199]

```
clear clusters
clear clusters_cell
clear fig1
clear fig2

%Input für SMART-NMR aus EDDCs to EMC Correlation Matrix generieren
%genutzte Funktion: findclusters.m aus "2D NMR spectra fitter" Toolbox

clusters = findClusters(CorrelatingEDDCstoEMC(:,1:2), 0.02, 0.5);

%Umwandlung in Matrix
for i = 1:size(clusters,1)
    clusters_cell(i,1) = cell2mat(clusters{i,1}(1,1));
    clusters_cell(i,2) = cell2mat(clusters{i,1}(1,2));
end

%ursprüngliche Peaks im Scatter-Plot:
fig1 = figure('Position', [100 300 800 700]);
fig1.Name = ('All Peaks');
ax = uiaxes(fig1, 'Position', [50 50 700 600]);
ax.YLabel.String = '13C chemical shift [ppm]';
ax.XLabel.String = '1H chemical shift [ppm]';
ax.XDir = 'reverse';
ax.YDir = 'reverse';
scatter(CorrelatingEDDCstoEMC(:,1), CorrelatingEDDCstoEMC(:,2), 'LineWidth', 1.5);

%Scatter-Plot der Cluster
fig2 = figure('Position', [950 300 800 700]);
fig2.Name = ('Calculated Cluster');
ax = uiaxes(fig2, 'Position', [50 50 700 600]);
ax.YLabel.String = '13C chemical shift [ppm]';
ax.XLabel.String = '1H chemical shift [ppm]';
ax.XDir = 'reverse';
ax.YDir = 'reverse';
for i = 1:size(clusters_cell,1)
    scatter(clusters_cell(i,1), clusters_cell(i,2), 'LineWidth', 1.5);
    hold on
end

%Ausgabe der Daten in passendem Format für SMART-NMR (HSQC):
formatSpec = '%4.2f,%4.2f\n';
fprintf('1H,13C\n')
for i = 1:size(clusters_cell,1)
    fprintf(formatSpec, clusters_cell(i,1), clusters_cell(i,2));
end
```


9 Danksagung

An dieser Stelle der Arbeit möchte ich mich bei all denen bedanken, die mich auf diesem Weg begleitet haben und ohne deren Hilfe und Unterstützung ich nicht diese Zeilen schreiben könnte.

Zunächst möchte ich Dr. Thomas Hackl für die Bereitstellung des interessanten Themas, die Betreuung der Arbeit und die umfassende Einführung in die Welt der NMR-Spektroskopie und analytischen Chemie danken. Ebenso danke ich Prof. Dr. Christian Stark für die Co-Betreuung und der Ermöglichung meiner Promotion. Weiterhin danke ich Prof. Dr. Maria Buchweitz für die wertvollen fachlichen Diskussionen und die Übernahme der Zweitkorrektur sowie Prof. Dr. Stephan Schwarzingen für die Übernahme des Drittgutachtens. Ebenso danke ich Prof. Dr. Ralph Holl und Prof. Dr. Wolfgang Maison als Mitglieder der Prüfungskommission für ihre Teilnahme an der Disputation.

Ein herzliches Dankeschön geht an den gesamten Arbeitskreis Hackl. Danke an Caro, Navid, Thorsten, Adu, Marie, Johanna und Jasmin für die schöne Zeit und die angenehme Atmosphäre! Ein ganz besonderer Dank gilt insbesondere auch meinen Praktikanten und Bacheloranden Marie, Tobias, Julia K., Julia E. und Sinah, die mich tatkräftig unterstützt haben und mit denen ich immer gut zusammenarbeiten konnte. Danke an den besten Arbeitskreis, dass durch euch die Arbeitstage so viel angenehmer waren!

Vielen Dank auch an René für den intensiven Austausch und die wertvollen Impulse. Ein großes Dankeschön auch an dich und das Landeslabor Schleswig-Holstein für die Bereitstellung der Apfelproben.

Ich bedanke mich außerdem herzlich bei dem gesamten AK Meier. Ich habe mich bei euch immer sehr wohl gefühlt und danke euch für eure Offenheit, die entspannten gemeinsamen Mittagspausen sowie die vielen fachlichen und nicht-fachlichen Gespräche, die mir so viel Spaß bereitet haben. Ebenso danke ich dem AK Fischer für den wertvollen Austausch im Rahmen der Doktorandenseminare und die gute Zusammenarbeit.

Ein großer Dank gilt außerdem der NMR-Abteilung. Danke an Vera, Eva und Gaby und insbesondere an Claudia für die tatkräftige Unterstützung bei der Messung der vielen NMR-Spektren und den netten Austausch! Prof. Dr. Arwen Pearson und Marty Rogers danke ich

außerdem für die Unterstützung und die Möglichkeit, weitere NMR-Messungen am HARBOR-Institut durchführen zu können.

Vielen Dank auch an Dr. Maria Riedner, Dr. Jennifer Menzel, Gaby, Dany, Christine und Erik für die Unterstützung bei den MS-Messungen sowie die angenehme Zusammenarbeit und netten Gespräche im 4. Stock!

Weiterhin möchte ich mich bei Dr. Werner Pauer und Dr. Daniel Griebe für die spannende Zusammenarbeit mit dem Institut für Technische und Makromolekulare Chemie bedanken.

Herzlichsten Dank an Dr. Brita Werner für die langjährige und schöne Zeit miteinander. Die gemeinsame Zeit im und um das OCP hat mir immer großen Spaß bereitet. Danke für das entgegengebrachte Vertrauen und dafür, dass du immer ein offenes Ohr für mich hattest und mir mit Rat und Tat zur Seite standest.

Die gesamte Zeit an der Universität Hamburg war für mich weitaus mehr als nur Studium und Promotion – viele Kommilitonen und Kollegen sind zu engen Freunden geworden und haben die Stadt Hamburg zu einem Teil meiner Heimat gemacht. Ich bedanke mich herzlich bei Andi, Caro, Tobi und Nils für eure Freundschaft, eure Unterstützung und dafür, dass ihr diesen Lebensabschnitt zu so viel mehr gemacht habt, als nur das Studium.

Ein ganz besonderer Dank gilt Michelle: Ohne dich, ohne unsere unzähligen (und überraschend produktiven) Kaffeepausen, ohne das ein oder andere Gläschen Wein und vor allem ohne unsere tiefe Freundschaft, wäre vieles schwieriger gewesen. Deine Unterstützung und dein Humor haben diese Zeit für mich unvergesslich gemacht. Danke, dass du immer da bist!

Ich danke ganz besonders meinen Eltern, meiner Familie und allen, die in Hamburg zu einem Teil meiner Familie geworden sind. Tausend Dank an Mama und Papa für eure uneingeschränkte Unterstützung, euren Rückhalt und dafür, dass ihr immer an mich geglaubt habt! Ohne euch wäre dieser Weg nicht möglich gewesen.

Zuletzt gilt mein allergrößter Dank Chris. Ich danke dir von tiefstem Herzen für deine bedingungslose Unterstützung, deine Zuversicht und deine Liebe. Ich bin unendlich dankbar, dich an meiner Seite zu haben.

10 Eidesstattliche Erklärung

Hiermit versichere ich an Eides statt, die vorliegende Dissertationsschrift selbst verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt zu haben. Sofern im Zuge der Erstellung der vorliegenden Dissertationsschrift generative Künstliche Intelligenz (gKI) basierte elektronische Hilfsmittel verwendet wurden, versichere ich, dass meine eigene Leistung im Vordergrund stand und dass eine vollständige Dokumentation aller verwendeten Hilfsmittel gemäß der Guten wissenschaftlichen Praxis vorliegt. Ich trage die Verantwortung für eventuell durch die gKI generierte fehlerhafte oder verzerrte Inhalte, fehlerhafte Referenzen, Verstöße gegen das Datenschutz- und Urheberrecht oder Plagiate.

_____

Hamburg, den 11.06.2025