TUM School of Life Sciences
Technische Universität München

**TUT**

# Metabolomic analysis to profile thermal food processing emissions

## Leopold Weidner

*Für Gisela und Christian.*

# Acknowledgement

Almost four years of research are condensed in the following pages. Four years of time that were marked for me not only by analytical work but especially by precious friends and colleges; I am very grateful that I got to spend so many memorable moments with you.

My first and very special gratitude goes to my supervisor Prof. Dr. Philippe Schmitt-Kopplin. You welcomed me in your fantastic and open-minded research group, giving me the opportunity to start a scientific journey together, full of exciting discoveries and culinary highlights. By your profound trust, you gave me the possibility to develop my own research interests independently - but never alone - in an inspirational and innovative environment. Your door was always wide open for my ideas, questions, concerns and for inspiring discussions. Thank you for giving me confidence wherever I needed it. I deeply thank you for supporting me during the past four years!

Second, I wish to thank Prof. Dr. Michael Rychlik for hosting me at the Chair of Analytical Food Chemistry. I am very grateful for your help in project administration and for your constant academic support during all steps of my project. You were my guide in the community of analytical food chemistry and always full of inspiring advice.

My very special thanks goes to Dr. Daniel Hemmler. You were the first person ever I spoke to at BGC and from that moment on, you were the most invaluable support, a graduate student could wish for. Your great scientific expertise united with your patience and your steady willingness to share your knowledge helped me to develop novel analytical skills and raised my ability for critical analytical thinking to a new level. Your broad knowledge in analytical chemistry made you a constant source of innovation; always realistic, right in the scope of my project, and at the pulse of current analytical developments. Your critical advice was always for my project's benefit and raised my research to a profounder level. Discussing with you has always been a pleasure; from the lab bench to Taxisgarten!

I especially want to thank Anja Brinckmann and Astrid Bösl. Both of you kept my back clear of administrative and organisational matter enabling me to do my research. Your support saved me from inestimable hours of work and guided me through tangled paths of academic bureaucracy.

My dear colleagues at BGC were one of the biggest pleasures during my time in academia. There is not a single person in the group I did not have great conversations over a cup of coffee or inspiring scientific discussions with. The predominant spirit of helpfulness in the

entire group impressed me from the first day on and made me want to follow your open-minded philosophy. My thank goes to all of you who enriched my daily work.

My special gratitude goes to Jenny Uhl. By knowing every instrument to the last screw and tiniest detail and only thanks to your caring technical support, I was able to fix so many instrumental problems. I highly appreciate all your help on IT-related, instrumental, and technical problems, even in the evening or during the weekend. Thank you further for always having an open ear and giving sound advice.

Likewise, I want to thank Martina Daubmeier and Brigitte Look. Your constant support gave me more time to focus on measurements and computational work.

My gratitude goes to Dr. Norbert Hertkorn for inspiring conversations and for showing me where true beauty can be found in nature.

One of the biggest pleasures was working with you Michelle Berger and Yingfei Yan. Together, as "Team Maillard", we explored glycation signatures under Park7/DJ1 knock-out conditions and always had fun times during our work. I learned a lot from both of you about LC-MS and I highly appreciate all the glycation-related advice you had for my project.

I want to thank Dr. Stefan Pieczonka for fruitful scientific discussions about food analysis and FT-ICR-MS, exciting common student supervision projects, and for the analytical advice I received from you.

Further, I want to express my gratitude towards Philippe Diederich. Thank you very much for many hours of NMR acquisition and spectral analysis. Your deep analytical thirst for knowledge, no matter on what topic or matrix, is enlightening. In times, when I was in doubt about good scientific practice, you always were a guide and a reliable support.

Thank you Nicolas Schmidt for always being an external balance. I want to thank you for your advice and letting me find harmony in our conversations and during our numerous trips to the alps.

Great gratitude goes to Anna-Lena for giving me strong support in every situation. Your always had firm advice for me and cheered me up whenever needed.

My final and greatest appreciation goes to my sister and especially to my parents. Without your irrevocable love and support, this work would not have been possible at all. You tough me curiosity and patience. I want to thank you for your continuous encouragement and wish to dedicate this work to you.

# Abstract

Cooking food is an ancient human habit, substantially differentiating us from other animal species. Over the course of tens of thousands of years, we developed an ensemble of sophisticated cooking methods to increase the nutritional value, deactivate pathogens, and prolong the shelf-life of raw food. This evolutional progression is recently marked by the onset of a chemical understanding of how thermal energy influences the molecular constitution of food by means of cooking. Raw food as a *biological* system is subjected to a cascade and a complex interplay of *chemical* transformations during the cooking process. Even though analytical chemistry thrives on deconvoluting and understanding these controlled transformations, chemical reactions in food resulting from cooking remain heavily understudied up to date. The ultimate objective is to control the chemical aspects of thermal transformations in food, pushing forward the evolution of cooking to a digitalized era marked by nutrition, which is safer and of higher quality due to analytical knowledge about molecular cooking processes.

In this thesis, an analytical focus was set to study thermal transformations of food through emissions that are released from the matrix during the cooking process. The combined application of (ultra)high-resolution mass spectrometric techniques and advanced bioinformatics allowed to construct a holistic overview of processing emissions as a system of small molecules from a metabolomics point of view.

The basis of this work was to sample and analyze processing emissions from cooking processes unbiasedly. This was implemented by developing a cold-trapping apparatus that condenses cooking vapors. An untargeted analysis of the food metabolome from wheat bread roll emissions was conducted in a proof of principle study. Emissions were collected by the developed apparatus attached to a convection oven. Vapor samples from wheat bread baked with varying temperatures (158 to 242 °C) and times (7.8 to 16.2 min) were studied. By metabolomic fingerprinting by DI-FT-ICR-MS, a complex molecular profile in the vapor was revealed, with significantly more features (4,700) detected in vapor from bread baked at 230 °C for 15 min compared to those processed at 170 °C for 15 min (580 features). Various compound classes were identified through visualization techniques, while UHPLC-Q$q$TOF-MS served to identify key metabolites in the mixture. Furthermore, the study successfully monitored and modeled complex chemical reactions reflected in bread processing emissions by RSM, illustrating the potential of vapor sampling as an analytical interface for studying complex transformations in the underlying matrix.

A follow-up study evaluated a comprehensive exploration of the causal correlation between the biological composition of foods and their cooking emissions under varying processing conditions. By collecting emissions from 96 food samples subjected to convection cooking and analyzing them by DI-FT-ICR-MS and UHPLC-Q$q$TOF-MS, untargeted metabolomics analysis was employed to distinguish vapor components unique to plant-based and meat-based foods. Further distinctions were made between meat, fish, plant-based alternatives, vegetables, and bakery products through proposed marker compounds and compound class annotation. Distinct feature patterns linked to the biological properties of the source material were identified and discussed in relation to the emitting matrix. For instance, abundant signatures of the CR and MR reactions were discovered in vapor from bakery products, while meat- and fish-vapor exhibit differences in peptide and fatty acid profiles. Comparative analysis of emissions from bread and steak cooked at different temperatures highlighted temperature-induced changes, and specific structural motives, such as pyridoxine and amino acid derivatives, were identified to play an essential role as thermo-sensitive head space markers. MS/MS fragment ion analysis revealed insights into aromatic N-heterocyclic core structures formed during the MR, particularly under elevated temperatures. Through analysis of mixed Maillard model systems, the study integrates structural and reactional background knowledge to enhance the understanding of processing emissions.

To perform mass spectrometric analysis of food processing emissions during the cooking process, a lab-scaled, highly controllable setup was developed for real-time monitoring of aerosols and volatiles generated during thermal food processing. The application of DBDI-MS permits the transfer of knowledge acquired in the two preceding studies to a time-resolved perspective. The innovative aspect of this methodology lies in its ability to track chemical reactions in real-time in an environment where process parameters can flexibly be designed. As a show-case study, emissions from wheat bread roll dough baked at 210 °C and measured by DBDI-MS were compared with three API sources to define DBDI's selectivity and provide new insights into the molecular composition of baking emissions. Detected emissions were analyzed for molecular composition and correlated with thermal transformations during processing, highlighting changing molecular spaces over time. Time series analysis, enabled by SOM, grouped similar emission profiles into distinct molecular spaces. Based on ambient plasma ionization, this platform successfully characterized over 6,100 food metabolites and reaction products in a time-resolved manner, showing future perspectives for industrial process control and real-time data delivery during food production.

The Python package DBDIpy was introduces as an open-source software solution developed for processing and formally analyzing untargeted, time-sensitive plasma ionization MS datasets. The package aids in detecting adducts and fragments generated during ionization, addressing the challenge of competing ionization mechanisms forming multiple ion species alongside $[M + H]^+$. DBDIpy is available for download from PyPI and GitHub. It integrates

within the matchms-ecosystem for MS data processing in Python. The package employs a two-step open search algorithm to detect in-source fragments and adducts generated during the ionization process, combining temporal correlation and exact mass difference analysis. The user can define custom rules for adduct and fragment detection. A walk-through of its functions, including data import, alignment, adduct detection, plotting, and results export, is provided in the GitHub repository. A demonstration of DBDIpy's application is show-cased using a data set of 4,200 features from DBDI-MS analysis of wheat bread dough baking emissions. A total of 710 potential adducts were detected within the data set.

In conclusion, a methodological tandem for analyzing food processing emissions was developed in this work consisting of at-line and on-line monitoring applications to understand controlled thermal transformations of food. This set is complemented by an accompanying software solution, which gives an enhanced perspective to control food processing in the future based on analytical insights offered by the work presented in this thesis.

# Zusammenfassung

Das Kochen von Lebensmitteln ist eine zehntausende Jahre alte menschliche Praxis, die uns im Wesentlichen von anderen Säugetieren unterscheidet. Im Laufe der Evolution haben wir eine Reihe diverser Kochmethoden entwickelt, um den Nährwert von Lebensmitteln zu erhöhen, Krankheitserreger abzutöten und die Haltbarkeit von Rohkost zu verlängern. Dieser evolutionäre Fortschritt ist in jüngster Zeit durch den Beginn eines chemisch-analytischen Verständnisses gekennzeichnet, wie thermische Energie die molekulare Beschaffenheit von Lebensmitteln während eines Garprozesses beeinflusst. Rohe Lebensmittel sind als *biologisches* System während des Kochvorgangs einer reaktionellen Kaskade sowie einem komplexen Zusammenspiel von *chemischen* Transformationen unterworfen. Obwohl die analytische Chemie bestrebt ist, diese Umwandlungen zu entschlüsseln, zu verstehen und nutzbar zu machen bleiben chemische Transformationen während Kochprozessen weiterhin unzureichend verstanden. Das ultimative Ziel ist es, mittels analytischer Verfahren molekulare Garprozesse zu kontrollieren und die Entwicklung des Kochens hin zu einer digitalisierten Zukunft voranzutreiben. Diese soll durch Lebensmittel gekennzeichnet sein, die aufgrund eines molekularen Wissensschatzes sicherer, individualisierter und von höherer Qualität sind.

In der vorliegenden Arbeit wurden durch Verfahren der instrumentellen Analytik thermische Reaktionsprozesse in Lebensmitteln untersucht, die mittels Prozessemissionen, welche während des Garprozesses aus der Matrix in den Kopfraum freigesetzt werden, analysiert und charakterisiert wurden. Die kombinierte Anwendung von (ultra-)hochauflösenden massenspektrometrischen Verfahren und fortgeschrittener Bioinformatik ermöglichte es, eine ganzheitliche Perspektive von Verarbeitungsemissionen, als Funktion eines komplexen Systems kleiner Moleküle, aus der Sicht der Metabolomik zu konstruieren.

Die Grundlage dieser Arbeit bildete die Entwicklung analytischer Verfahren zur diskriminierungsfreien Entnahme und Analyse von Kopfraum-Proben während Garprozessen. Dies wurde durch die Konzeptualisierung und dem Bau einer Kondensationsapparatur erreicht, welche Kochdämpfe und mitgeschleppte organische Verbindungen auskondensierte. In einer Grundlagenstudie wurde eine ungerichtete Analyse des Lebensmittelmetaboloms in Backemissionen von Weizenbrötchen durchgeführt. Die Emissionen wurden mit einer selbstentwickelten Apparatur aus einem Konvektionsofen heraus gesammelt. Es wurden Emissionsproben von Weizenbrötchen mit unterschiedlichen Backtemperaturen (von 158 bis 242 °C) und Backzeiten (von 7,8 bis 16,2 min) untersucht. Die mittels DI-FT-ICR-MS aufgenommenen metabolomischen Fingerabdrücke zeichneten ein komplexes molekulares Profil im Pro-

zessdampf, wobei in Brötchen, die 15 Minuten lang bei 230 °C gebacken wurden, signifikant mehr Signale (4.700 Stück) festgestellt wurden als in solchen, die 15 Minuten lang bei 170 °C erhitzt wurden (580 annotierte Massensignale). Das Vorkommen verschiedener Verbindungsklassen wurde durch Datenvisualisierungstechniken verdeutlicht und UHPLC-Q$q$TOF-MS diente zur Identifizierung von Schlüsselmetaboliten im Prozessdampf. Darüber hinaus wurden im Rahmen der Studie komplexe chemische Reaktionen, die sich in den Emissionen der Brötchen nachverfolgen lassen, erfolgreich mittels RSM überwacht und mathematisch durch Backtemperatur- und Dauer modelliert, was ihr Potenzial als Medium für die Untersuchung komplexer Reaktionskaskaden in der zugrunde liegenden Matrix verdeutlicht.

In einer Folgestudie wurde eine umfassende Untersuchung des Zusammenhangs zwischen der biologischen Zusammensetzung von Lebensmitteln und ihren Kochemissionen unter verschiedenen Prozessparametern durchgeführt. Durch die Sammlung von Dampfproben aus 96 Garprozessen verschiedener Lebensmittel, die in einem Konvektionsofen zubereitet wurden, gefolgt von deren Analyse mittels DI-FT-ICR-MS und UHPLC-Q$q$TOF-MS wurde eine ungerichtete Metabolomanalyse durchgeführt. Molekulare Signaturen, die charakteristisch für Kochemissionen von pflanzlichen und tierischen Lebensmitteln sind wurden identifiziert. Detaillierte Unterscheidungen der Emissionen von Fleisch, Fisch, pflanzenbasierten Fleischalternativen, Gemüse und Backwaren wurden angestellt und durch vorgeschlagene Markerverbindungen und die Annotation von Verbindungsklassen weiterführend diskriminiert. Es wurden eindeutige Merkmalsmuster identifiziert, die mit den biologischen Eigenschaften oder der Zusammensetzung des rohen Lebensmittels in direkter Verbindung stehen. Diese wurden nachfolgend in Relation zu der emittierenden Matrix diskutiert. Beispielsweise wurden molekulare Bereiche, welche charakteristisch für CR- und MR-Reaktionen in Backwarenkopfräumen sind, identifiziert, während Emissionen von Fleisch und Fisch Unterschiede in ihren Peptid- und Fettsäureprofilen aufweisen. Eine vergleichende Gegenüberstellung der Emissionen von Brot und Steak, die bei unterschiedlichen Temperaturen gegart wurden, zeigte temperaturbedingte Veränderungen auf und es wurden spezifische Verbindungen wie Pyridoxin und Aminosäurederivate identifiziert, welche besonders temperatursensitiv den Prozessverlauf beschreiben. Eine MS/MS-Fragmentionanalyse ermöglichte zudem Einblicke in N-heterozyklische, aromatische Ringstrukturen, die während des Kochprozesses durch die MR gebildet werden, insbesondere bei erhöhten Temperaturen. Durch die Analyse von multi-Edukt-Maillard-Modellsystemen integriert die Studie strukturelles und reaktionstechnisches Hintergrundwissen in molekulare Garprozesse, um das Verständnis von Verarbeitungsemissionen zu verbessern.

Weiterhin wurde für die massenspektrometrische Echtzeitanalyse von Emissionen aus Lebensmittelverarbeitungsprozessen ein exakt kontrollierbarer Laboraufbau entwickelt, der für die Analyse von Aerosolen und Volatilen, die bei Garprozessen entstehen, konzipiert ist. Die auf DBDI-MS basierende Anwendung ermöglicht es nun die zuvor erworbenen

Erkenntnisse über molekulare Reaktionen zeitaufgelöst in analytischer Umgebung zu verfolgen und tiefergreifend auszubauen. Die besondere Innovation der beschriebenen Methodik liegt in ihrer Fähigkeit, chemische Reaktionen in Echtzeit in einer Umgebung zu verfolgen, in der alle Prozessparameter flexibel gestaltet werden können. Es wurden Emissionen aus Weizenbrötchenteig bei 210 ℃ mit Kondensatproben derselben Matrix, die mittels dreier API-Quellen gemessen wurden, verglichen, um die Ionisationsselektivität von DBDI zu bestimmen und neue Erkenntnisse über die molekulare Zusammensetzung von Backemissionen zu gewinnen. Die gemessenen Emissionen wurden auf ihre molekulare Zusammensetzung hin analysiert und mit thermischen Umwandlungen während der Erhitzung korreliert, um die sich im Laufe der Zeit verändernden Substanzklassen aufzuzeigen. Eine mittels SOM durchgeführte Zeitreihenanalyse gruppierte ähnliche Emissionsprofile in systematisch strukturierte, molekulare Räume und deckte molekulare Zusammenhänge zwischen Verbindungen auf, die eine initiale Back- und eine spätere Röstphase charakterisieren. Die entwickelte Plattform, die auf Atmosphärendruck Plasma-Ionisierung basiert, charakterisierte erfolgreich über 6.100 Lebensmittelmetaboliten und Reaktionsprodukte in einer temporalen Dimension und zeigt eine Zukunftsperspektive für die industrielle Prozesskontrolle in Echtzeit während eines Garprozesses auf.

Außerdem wurde im Verlauf dieser Arbeit die Python-Bibliothek DBDIpy entwickelt, die eine Open-Source-Softwarelösung für die Verarbeitung und formale Analyse von ungerichteten, zeitaufgelösten Plasmaionisations-MS-Datensätzen ist. Das Paket hilft bei der Erkennung von Addukten und Fragmenten, die während der Ionisierung entstehen und stellt sich der Herausforderung, konkurrierende Ionenspezies zu identifizieren, die neben der Bildung von $[M+H]^+$ auftreten. DBDIpy kann von PyPI und GitHub heruntergeladen und installiert werden. Es integriert sich in das matchms-Ökosystem, welches für die MS-Datenverarbeitung in Python etabliert ist. Das Paket verwendet einen zweistufigen, ungerichteten Suchalgorithmus, um In-Source-Fragmente und vom Benutzer definierbare Addukte zu erkennen. Dafür wird eine zeitliche Korrelations- und exakte Massendifferenzanalyse kombiniert. Im GitHub-Repository enthalten ist einen Überblick über alle implementierten Funktionen, einschließlich Datenimport, Alignment, Adduktdetektion, Ergebnisvisualisierung und Ergebnisexport. Eine Demonstration der Anwendung von DBDIpy wurde anhand eines Datensatzes mit 4.200 annotierten Massensignalen aus der DBDI-MS-Analyse von Weizenbrotteig-Backemissionen vorgenommen, wobei 710 potenzielle Addukte innerhalb des Datensatzes erkannt wurden.

Zusammenfassend wurde in dieser Arbeit ein methodischer Dualismus für die Analyse von Emissionen aus Lebensmittelgarprozessen entwickelt, welcher at-line und on-line Techniken zum besseren Verständnis von kontrollierten thermischen Umwandlungen in Lebensmitteln kombiniert.

Basierend auf den Ergebnissen der molekularen Untersuchungen, welche durch entwickelte Software komplementiert ist, wird durch die Kombination von instrumenteller Analytik und Bioinformatik eine verbesserte Perspektive für die zukünftige Kontrolle von Kochprozessen geschaffen. Ausgehend vom erhaltenen, detaillierten Verständnis von chemischen Reaktionen, die durch thermische Energie ausgelöst werden, wird es ermöglicht, gesündere und qualitativ hochwertigere Lebensmittel in automatisierten Fertigungsprozessen herzustellen.

# Contents

# Scientific Communications

**Peer-reviewed publications directly addressed in this thesis:**

**Chapter 4.1:** Weidner, L., Hemmler, D., Yan, Y., Rychlik, M. & Schmitt-Kopplin, P. (2022). *Elucidation of the non-volatile fingerprint in oven headspace vapor from bread roll baking by ultra-high resolution mass spectrometry.* Food Chemistry, 2022, Volume 374, Article 131618. doi.org/10.1016/j.foodchem.2021.131618

**Chapter 4.2:** Weidner, L., Cannas, J. V., Rychlik, M. & Schmitt-Kopplin, P. (2023). *Molecular characterization of cooking processes: a metabolomics decoding of vaporous emissions for food markers and thermal reaction indicators.* Journal of Agricultural and Food Chemistry, 2023, Volume 71, Issue 45, Pages 17442–17454. doi.org/10.1021/acs.jafc.3c05383

**Chapter 4.3:** Weidner, L., Hemmler, D., Rychlik, M., & Schmitt-Kopplin, P. (2023). *Real-Time Monitoring of Miniaturized Thermal Food Processing by Advanced Mass Spectrometric Techniques.* Analytical Chemistry, 2023, Volume 95, Issue 2, Pages 1694-1702. doi.org/10.1021/acs.analchem.2c04874

**Chapter 4.4:** Weidner, L., Hemmler, D., Rychlik, M., & Schmitt-Kopplin, P. (2023). *DBDIpy: a Python library for processing of untargeted datasets from real-time plasma ionization mass spectrometry.* Bioinformatics, 2023, Volume 39, Issue 2. doi.org/10.1093/bioinformatics/btad088

**Further peer-reviewed publications besides the framework of the doctoral studies:**

**Appendix C.1:** Prudente De Mello N., Berger M. T., Lehmann K., Yan Y., <u>Weidner L.</u>, Tokarz J., Möller G., Cheng Y., Keipert S., Ciciliot S., Artati A., Wettmarshausen J., Nilsson R., Brandt D, Kutschke M., Leimpek A., Vogt Weisenhorn D., Wurst W., Adamski J., Jain M., Jastroch M., Mandemakers W., Bonifati V., Schmitt-Kopplin P., Perocchi F. & Dyar K. A., *Glycerinyl-modified arginine and lysine are putative biomarkers for PARK7-related early-onset Parkinson's disease.*
This manuscript was submitted to Molecular Neurodegeneration under manuscript number *MOND-D-24-00051* and is under peer review at the time of submitting this thesis.

**Appendix C.2:** Diederich P., Ruf A., Geisberger T., <u>Weidner, L.</u>, Seitz C., Eisenreich W., Huber C. & Schmitt-Kopplin, P. (2023). *C2-addition patterns emerging from acetylene and nickel sulfide in simulated prebiotic hydrothermal conditions.* Communications Chemistry, 2023, Volume 6, Article 220.

## Oral Presentations:

Invited talk at the Plasmion Customer Forum 2021, Augsburg, Germany.
*Real-time monitoring of cereal roasting processes by ultra-high resolution mass spectrometry with Plasmion ionization.*

18th Annual Conference of the Metabolomics Society, Metabolomics 2022, Valencia, Spain.
*A foodomics study on the molecular composition of cooking vapor from the processing of foodstuff.*

Invited talk at Forschungsseminar der Lebensmittelchemie, 2022, Freising, Germany.
*Towards elucidating the chemical entity of cooking vapor from thermal food processing: a foodomics study.*

51st International Symposium on High Performance Liquid Phase Separations and Related Techniques, HPLC 2023, Düsseldorf, Germany.
*Ion Mobility Spectrometry - MS as the next dimension for real-time analysis of food cooking.*

## Poster Presentations:

The 16th Weurman Flavour Research Symposium, 2020, Dijon, France.
*Monitoring the effects of thermal processing on flavour and flavour precursors during cooking processes through oven vapour.*

# Abbreviations

| | |
|---|---|
| AGE | advanced glycation end products |
| AI | artificial inteligence |
| ALE | advanced lipid peroxidation end products |
| APCI | atmospheric pressure chemical ionization |
| API | atmospheric pressure ion source |
| APPI | atmospheric pressure photo ionization |
| ARP | Amadori rearrangement product |
| CR | caramelization reaction |
| CCS | collisional cross section |
| DBDI | dielectric barrier discharge ionization |
| DDA | data-dependant acquisition |
| DI | direct infusion |
| DIA | data-independant acquisition |
| DNA | desoxy ribonucleic acid |
| ESI | electrospray ionization |
| FA | fatty acid |
| FT-ICR | Fourier transform - ion cyclotron resonance |
| GC | gas chromatography |
| GNPS | global natural products social molecular networking |
| HAA | heterocyclic aromatic amines |
| HILIC | hydrophilic interaction liquid chromatography |
| HMF | 5-hydroxymethyl furfural |
| IMS | ion mobility spectrometry |
| LC | liquid chromatography |
| LOOH | lipid hydroperoxyde |
| LPO | lipid peroxidation |
| MetID | metabolite identity |
| ML | machine learning |
| MR | Maillard Reaction |
| MRP | Maillard reaction product |
| MS | mass spectrometry |
| MS/MS | tandem mass spectrometry |
| NMR | nuclear magnetic resonance |

| | |
|---|---|
| PAH | polycyclic aromatic hydrocarbons |
| PTR | proton transfer reaction |
| PM | particulate matter |
| PUFA | polyunsaturated fatty acid |
| Q$q$TOF | quadrupole time of flight |
| RNA | ribonucleic acid |
| RP | reversed phase |
| RSM | response surface modelling |
| SD | steam distillation |
| SIFT | selective ion flow tube |
| SOM | self-organizing maps |
| SPME | solid phase micro extraction |
| TOF | time of flight |
| UHPLC | ultra-high pressure liquid chromatography |
| VOC | volatile organic compound |

# 1 Introduction

Since the earliest onset of human development, the consumption of food has been one of the most fundamental human needs. During our evolutional journey, cooking evolved as a crucial role in shaping our species' society, fostering social bonds, and nourishing both body and culture.

Anthropological researchers take intense action to determine when human predecessors began utilizing heat to treat food. This investigation is of a challenging nature; published findings are often particular and contradicting. In the first place, archaeology currently does not offer a certain answer to the question of when fire was first controlled (*1*). It is commonly accepted that widespread control of fire was established around 250 ka before (*2*). However, some groups report evidence of scattered fireplaces of up to 1 Ma back in time (*3*). In close relation to evolutional biology, a common hypothesis suggests that humans widely adopted the practice of using fire to process their food in a time frame ranging from 400 to 25 ka before (*3*). At this point, the term "cooking" was acknowledged as the use of heat to prepare food (a prospect on modern cooking techniques will be given later on in chapter 1.2).

Less vague, however, is the generally accepted persuasion that in the long term, since the invention of cooking, multiple beneficial aspects for human consumption and society have evolved along. First of all, establishing the habit of cooking among hominids created the necessity to relocate collected or hunted food to a fixed location around a fireplace or early hearth. *In-situ* consumption of collected food or hunted prey, therefore, shifted in order to cook food at a remote location, which can be seen as an early tendency of social gathering for dining.

Second, even if detailed aspects are controversially discussed, cooking food globally contributes to better digestibility, higher nutritional value, and a rise in its net energy value (*4*). On a molecular level, thermal processing triggers various reactions impacting digestibility, which will be discussed in more detail in chapter 1.2.1.

Third, cooking and related techniques such as drying or smoking enable food preservation. The stability of collected goods could significantly be extended by dehydration, deactivation of microorganisms inducing spoilage, and killing of parasites and other pathogens. As a result, it was first possible to store food for longer periods away from the collection site. Today, cooking, pasteurization, and sterilization notably contribute towards higher food safety and extended shelf life, which will also be further elaborated in chapter 1.2.2.

Fourth, it is common knowledge that the cooking of food transforms raw materials by a series of improvements in appearance, texture, and flavor to palatable meals. Thermal energy triggers chemical reactions that lead toward the controlled thermal transformation of precursor compounds to form appealing flavors and tastes via convoluted reactional pathways. As a result of these effects, nutritional studies have demonstrated that even apes prefer the consumption of cooked food when it is offered *ad libitum* next to uncooked (*5*).

During the transformation from ancient towards contemporary cooking techniques, there occurred a second change, which lies in the intentional paradigm behind food processing. In most modern societies, food is ubiquitous and offered in countless forms. Nowadays, cooking has become an art and science at the same time. This thesis aims to contribute to the objective of how chemistry can contribute to augmenting the quality and palatability of our everyday food.

## 1.1 The chemical composition of raw food

The chemical composition of raw food is complex and strikingly dissimilar between product groups. The role of different food constituents in human nutrition differs as much as their diverse chemical properties do. To understand important chemical reactions and physical changes triggered by thermal treatment, a brief overview of food's composition in the raw state is given here within the scope of this thesis.

Food tissue is mainly composed of water, the macronutrient classes carbohydrates, lipids, proteins, further minerals, and a diverse set of small organic compounds such as secondary plant compounds or mammalian cell metabolites, which are often present in low concentration levels. This thesis will uniquely focus on the share of organic food constituents besides water.

The water content of raw food is very variable and directly accounts for its dry matter content. For example, nuts possess up to 95% (*6, 7*) of dry matter, one of the lowest water contents of food and thus have a very high nutrient density. Fresh beef cuts range around approximately 70% of water content (*8*) whereas some vegetables like cucumbers have up to 98% of water or only 2% (*6*) of dry matter. The water content of a tissue is a highly deterministic parameter that influences chemical reactions and pathways during food cooking. Extracellular water is responsible for major biological changes and modifications in tissue structure initiated by thermal treatment (see chapters 1.2.2 - 1.2.3) (*9*).

Macronutrients prevail in the form of macromolecules in food. Macromolecules are defined as "a molecule of high relative molecular mass, the structure of which essentially comprises the multiple repetition of units derived, actually or conceptually, from molecules of low relative molecular mass" (*10*). For digestion, macromolecules need to be liberated from their ambient cellular structure and are required to be broken down into their building blocks for absorption in the human intestine. The breakdown process is virtually essential for the chemistry of processed food, as the "molecules of low relative molecular mass" are of higher reactivity than their precursors and are deterministic for reactional processes during cooking. A detailed description will be given in chapter 1.2.1.1.

Oligomeric and polymeric carbohydrates consist of linked aldoses and ketoses with the stoichiometric formula $C_n(H_2O)_n$, their respective sugar alcohols, sugar acids, and their derivatives. Single building blocks commonly are pentoses or hexoses, which are variably connected *via* O-glycosidic bonds to acetal. Chemical modifications such as methoxylation or sporadic esterifications can occur in different species to various extents (*6, 9*).

Polymeric carbohydrates are commonly found in all plants, serving as energy storage or structural cellular scaffolds. In animals, contrarily, high molecular weight carbohydrates majorly function as energy storage. Complex carbohydrates need to be reduced in size to smaller structures of mono- to small oligomers to be absorbed in the human intestine (*9*).

On a monomeric level, the most relevant carbohydrates in human nutrition are the hexoses glucose and fructose and, to a smaller extent, pentoses like ribose or xylose. Saccharose ($\alpha$-D-glucopyranosyl-$(1 \rightarrow 2)$-$\beta$-D-fructofuranoside) as a disaccharide is an often added ingredient in processed foods where it serves as a sweetener and is the most common nutritional disaccharide next to maltose (4-O-$\alpha$-D-glucopyranosyl-D-glucose). Important chemical reactions of (reducing) mono- and disaccharides will be summarized in chapter 1.2.1.2-1.2.1.3. Carbohydrates mainly affect two functions in living cells. They serve as reserves for energy storage. The most relevant energy-storage polysaccharide in plants is starch. Starches occur in approximately 15% of all chemically characterized plants (*11*). Depending on the biological origin, starch consists of up to 20-30% of amylose which is linear $\alpha$-1,4-glycosidic glucose with a molecular weight of to $10^5 \, g \, mol^{-1}$ and of 70-80% amylopectin which is $\alpha$-1,4-glycosidic glucose with branched $\alpha$-1,6-glycosidic glucose bonds approximately every 30 monomers with a molecular weight of $10^6 - 10^7 \, g \, mol^{-1}$ (*12*). Shorter starch fragments are called dextrins. Changes occurring to starch during heating are of fundamental importance for the texture of food, for human digestibility, and for chemical reactivity. Polymeric starches are aggregated in in-accessible grains (see chapter 1.2.3 for disintegrating processes of starch) and unreactive compared to their respective monomers.

The primary polysaccharide for energy storage found in animals is glycogen. Glycogen is a multibranched polymer with 2,000 - 60,000 units of glucose per molecule. Linear chains of (glucose-$\alpha$-1,4-glucose) units are its main glycosidic bond type, followed by (glucose-$\alpha$-1,6-glucose)-branches every 8-12 glucose units on the main stem (*13*). Glycogen's 3D structure resembles a sphere of branches of glucose stands centered around a glycogenic protein. In comparison to starch, glycogen is highly symmetric and significantly more branched than amylopectin (*14*). Analogously, its physiological function is to store glucose; glycogen can mainly be found in liver and muscle tissue. The most important implication of glycogen on cooking is meat quality. Degradation of glycogen releases free glucose and it is *post mortem* successively degraded to lactic acid, lowering the pH of respective tissue (*15*).

Structural polysaccharides, which predominantly cannot be digested by human enzymes, are predominantly found in plants and are referred to as dietary fiber. Their biological function is to make up cell walls, and they play an important role in the structure and texture of food. Commonly found representatives are pectin or (hemi-) celluloses (*9*). They surpass the small intestine and are fermented by bacteria in the colon, having an important contribution to intestinal health (*16*).

Lipids can be found in all plant- and animal-derived foods. As lipidomics research has shown during recent years, lipids are an extensive superclass of compounds with fundamental biological functions for energy storage, cell signaling, and cell structure by being essential and functional components of cell membranes (*17*, *18*). The most important class of dietary lipids are glycerolipids and free fatty acids (FA) (*19*). Glycerolipids are neutral lipids and composed of a glycerol backbone with one to three fatty acids bond *via* esters, respec-

tively named mono-, di- or triacylglycerides. Due to their important biological roles, dietary intake patterns and lipid quality are of fundamental importance for human health (*19*). The physicochemical and nutritious properties of glycerolipids are direct functions of their FA spectrum, in particular of chain length and saturation level, and of their processing history. Generally, edible oils from plants are rich in unsaturated FA with relatively longer chain lengths. In contrast, animal-derived fats show substantial amounts of shorter chained, saturated FA and, therefore, tend to be solid, whereas most plant lipids are liquid at room temperature (*20*). Notably, all lipids found in food are highly complex mixtures and contain next to glycerolipids, many further accompanying lipid classes such as sterol lipids, phospholipids, and solubilized lipophilic vitamins which influence nutritious qualities and physicochemical properties of the lipid mixture. Their chemical degradation reactions and interactions with other food constituents during processing have strong impacts on the sensory properties and nutritious values of cooked food (*18*). Thus, an overview of important reactions of lipids induced by heat will be given in chapter 1.2.1.4.

Proteins are polymeric chains of amino acid residues and can be found in plant-based and animal-derived food. Their building blocks are so-called proteinogenic amino acids. There are 21 proteinogenic amino acids encoded by human deoxyribonucleic acid (DNA), of which eight are classified as essential and, therefore, have to be supplied by food intake. Protein can only be absorbed in the human intestine through free amino acids or small peptides (*21*). Moreover, several hundreds of non-proteinogenic amino acids have been described in fungi, bacteria, or plants to play an important biochemical role in their organisms (*22*). Large amounts of free amino acids and peptides are formed by proteolysis from proteins *via* numerous pathways during the storage and processing of food (*23*). The 3D structure of proteins is mainly determined by their amino acid sequence, leading to characteristic folds *via* inter- and intra-molecular interaction of the side chain residues. Proteins exhibit manifold roles in living organisms, which are relevant in terms of cooking research. Muscle myofibrils, which are composed of actin and myosin, or collagen as the main structural protein in connective tissue, have an important impact on the texture of meat even if collagen only constitutes to 2% of most skeletal muscle (*24*). Collagen is a macromolecule of approximately 300 kDa molecular weight composed of a coiled $\alpha$−triple-helix structure (*25*). Collagen fibrils are interconnected by hydrogen bonds and covalent intermolecular crosslinks. The tenderness of meat is associated with the extent of covalent collagen crosslinks, which increase in living organisms with age (*26, 27*). The behavior of collagen under the influence of heat will be emphasized in chapters 1.2.1.1 and 1.2.3. Some enzymes are important in terms of the durability of food (see chapter 1.2.3) and in a nutritious background (see chapter 1.2.2).

Raw food items are heterogeneous, complex biological systems, next to macromolecules consisting of myriads of small organic compounds that apparently do not belong to the classes of carbohydrates, lipids, or proteins. Small molecules is a commonly utilized term

for organic compound with a molecular weight less than 1,500 Da. They can roughly be introduced based on their biological background:

Next to foreign xenobiotics, first there are monomers and building blocks of the macronutrients named above, which can be found in plant- and animal-based foods: free amino acids, monosaccharides, and FA. Second, there are secondary plant metabolites, including wide varieties of alkaloids or polyphenols, glycosides, and many others. Third, there are cellular metabolites of various functions. These involve cellular signal transmitters, metabolites of central energy metabolism, vitamins, co-factors, and others. Fourth, in every organism, there are precursors, degradation products, conjugates, and spontaneous chemical modifications occurring under physiological conditions of all of the above, leading to a further increase in molecular complexity. All members of these four groups serve as important precursors in reactional pools of small organic compounds during cooking processes. The molecular complexity of food is hard to access in a comprehensive manner. However, modern analytical approaches aim to deconvolute the continuum of small organic compounds found in food matrices (*28–31*). Their untargeted description in a holistic way is an ongoing challenge in modern food chemistry, which will further be discussed in chapter 2.1.1 from an analytical perspective.

The challenge behind the analysis of cooking processes lies in the coalescence of two worlds: this chapter was a mere introduction to raw food whose chemical composition results from living, *biological*, cellular systems. Once food is subjected to cooking, thermal energy induces *chemical* reactions, triggering reactional pathways and opening up a diverse space of generated compounds, which shall be introduced in the following sections.

## 1.2 Thermal processing of food

The term "cooking" shall initially be defined for further discussion as "the application of thermal energy to food during its preparation". For further simplification, the term "heat" is taken as synonymous with "thermal energy".

Cooking strategies of various flavors evolved depending on the properties of the raw food being processed and on a selection of the desired processing effects made by the cook. Most modern cooking techniques can be classified according to the thermodynamic proliferation of thermal energy.

First, there is conductive heat transfer $\dot{Q}_{cond}$. Thermal energy thereby is transferred through matter from the more energetic to less energetic particles along a temperature gradient $dT$ without bulk motion of matter (*32*). Given the case, the thickness of the transient layer $\Delta x \rightarrow 0$, Fourier's law of conduction applies, and the heat transfer can be determined by the thermal conductivity $k$ of the material and the cross-sectional area $A$, see equation 1.1. The derived cooking method is called conductive cooking, a common example of which is pan frying.

$$\dot{Q}_{cond} = -kA\frac{dT}{dx} \tag{1.1}$$

Second, heat can be transferred in the form of radiation, $\dot{Q}_{rad}$. Matter can emit thermal energy through photons or electromagnetic waves (*32*). The Stefan-Boltzmann Law describes radiative heat emission by the material property emittance $\epsilon$, the Stefan-Boltzmann constant $\sigma$ and the absolute temperature $T$, see equation 1.2.

$$\dot{Q}_{rad} = \epsilon\sigma T^4 \tag{1.2}$$

Radiation energy can be absorbed (absorptivity $\alpha$), reflected (reflectivity $\rho$), or transmitted (transmissivity $\tau$) when it reaches a surface, see equation 1.3. Upon absorption of radiation, a body is heated. Radiation is an important factor in cooking methods such as rotisserie grilling.

$$\alpha + \rho + \tau = 1 \tag{1.3}$$

Third comes convective heat transfer $\dot{Q}_{conv}$, which is defined as the transfer of thermal energy between a surface and a fluid that moves over the surface. It consists of two modalities, which are diffusion of energy (conduction) and bulk fluid motion (advection) (*32*). Convective heat transfer is defined by Newton's Law of Cooling with the convective heat transfer coefficient $h$, $T_W$ as the surface temperature, and $T_\infty$ as the temperature of the fluid, see equation 1.4. Examples are oil-frying, water-boiling, or steaming.

$$\dot{Q}_{conv} = hA(T_W - T_\infty) \tag{1.4}$$

In many cases, cooking techniques combine more than one heating principle. E.g., an air convection oven is an important heating setup to prepare food samples which was also employed during this thesis. During convection cooking, the convective transfer of thermal energy plays a primary role. Besides, there are also minor influences of radiative heat transfer from hot surfaces inside the cooking cabinet and conductive effects from hot trays, so the entire transferred heat can be expressed for convection cooking as in equation 1.5.

$$\dot{Q}_{total} = \dot{Q}_{cond} + \dot{Q}_{rad} + \dot{Q}_{conv} \tag{1.5}$$

Cooking undoubtedly manipulates the chemistry, flavor, and molecular composition of food. An overview of changes induced by thermal processing with a focus on chemical aspects shall be given in the following chapter as a base for a later discussion of the findings of this thesis. Major degradative and transformative processes shall be outlined. Further aspects like physical, biological, and nutritional effects likewise exist, are often closely related to or co-triggered by chemical reactions, and are crucial to understanding the complex transformations associated with cooking to their entire extent. However, as chemical transformations are the main focus of this thesis, supporting topics shall only be briefly mentioned to the extent that serves as a base to comprehensively discuss the results presented in chapter 4.

### 1.2.1 Thermally induced chemical changes in the matrix

Food chemistry thrives on understanding the exact course of molecular rearrangements, bond breakages, and condensation reactions driven by thermal energy to obtain better knowledge about cooked food. Cellular processes in plants and animals provide a defined set of precursors, which were briefly sketched in chapter 1.1. This set of compounds composes an initial reactional pool of raw food metabolites, ready to be altered by heat to human desire. Even though extensive efforts are made to chemically characterize plant material and human cells in large screening programs, their molecular profile is still incompletely defined (see chapter 2.1.1 for further aspects of uncharacterized metabolites). Even less systematic knowledge is present about the reactome of these biological precursors, which evolves as a result of thermal processing. To set a mechanistic base for chapters 4 and 5, the following pages will outline important reactional frameworks occurring during the cooking process.

Self-evidently, none of the following reactional systems takes place in an isolated way during cooking. The complexity of food-born thermal reaction products lies in the beauty of the interplay between different reaction pools, branching, and merging downstream intermediates to generate a heterogeneous hybrid of the here presented reactions whose entity was studied in this thesis.

### 1.2.1.1 Hydrolysis of macronutrients: feeding a reactional pool of small compounds from biopolymers

High-molecular macronutrients are, in comparison to their monomeric building blocks, spatially organized in an efficient way and of comparable low reactivity in order to fulfill their biological function without extensive unintended chemical modification. Upon dissipation to units of relatively lower molecular weight, the number of possible reaction partners and the number of reactive sites offered increases steadily. Consequently, hydrolysis processes directly contribute to a complex landscape of small molecules and reactive precursors in a food system.

Further, biopolymers cannot be absorbed by the human intestine; they need to be reduced in size to fragments of lower molecular weight before they can actively be transported out of the intestine lumen or before they can passively diffuse through epithelial gut tissues [21]. This breakdown is achieved by a set of enzymes, namely lipases, amylases, or proteases, which digest their respective substrates endogenously in the organic tissue or after consumption in the human digestive system [33]. Additionally, cooking initiates extra somatic chemical reactions leading to the breakdown of macromolecules. These reactions are collectively called hydrolysis [9].

Hydrolysis is the nucleophilic addition of water to an electropositive site, resulting in the cleavage of a covalent bond, releasing two component parts of the original compound. Hydrolysis can be catalyzed by $H^+$- or $OH^-$-ions or enzymes, see figure 1.1 for a summary. Heat thereby supplies the required activation energy, which varies for different substrate classes [9].

Therefore, hydrolysis extensively impacts the digestibility of biopolymers, which will further be elaborated in chapter 1.2.2. More important from a reactional perspective, however, is that the lysis of larger molecules feeds the pool of reactive precursors for further downstream reactions.

Hydrolysis of carbohydrates is characterized by dissociation of the O-glycosidic bond. In the case of saccharose, an equimolar mixture of glucose and fructose called invert sugar is generated. In contrast to both generated monosaccharides, saccharose is referred to as a non-reducing carbohydrate. Both of its anomeric carbon atoms are connected *via* a 1,2-glycosidic bond, resulting in the fact that saccharose cannot enter acyclic configurations and thus cannot form a reactive aldehyde/enolate form, which is attributed to the high reactivity of reducing sugars. Therefore, as shown in the example of saccharose, hydrolysis yields two reactive monosaccharides from one relatively unreactive precursor. Of even higher reactional importance for cooking processes is the hydrolysis of starch. Starch, as an energy-storing biopolymer, is built up of several thousands of glucose units. Hydrolysis of starch is called dextrinization and favorably happens during heating at reduced moisture as it happens, for example, during roasting [34, 35]. Dextrinization unites the processes of hydrolysis,
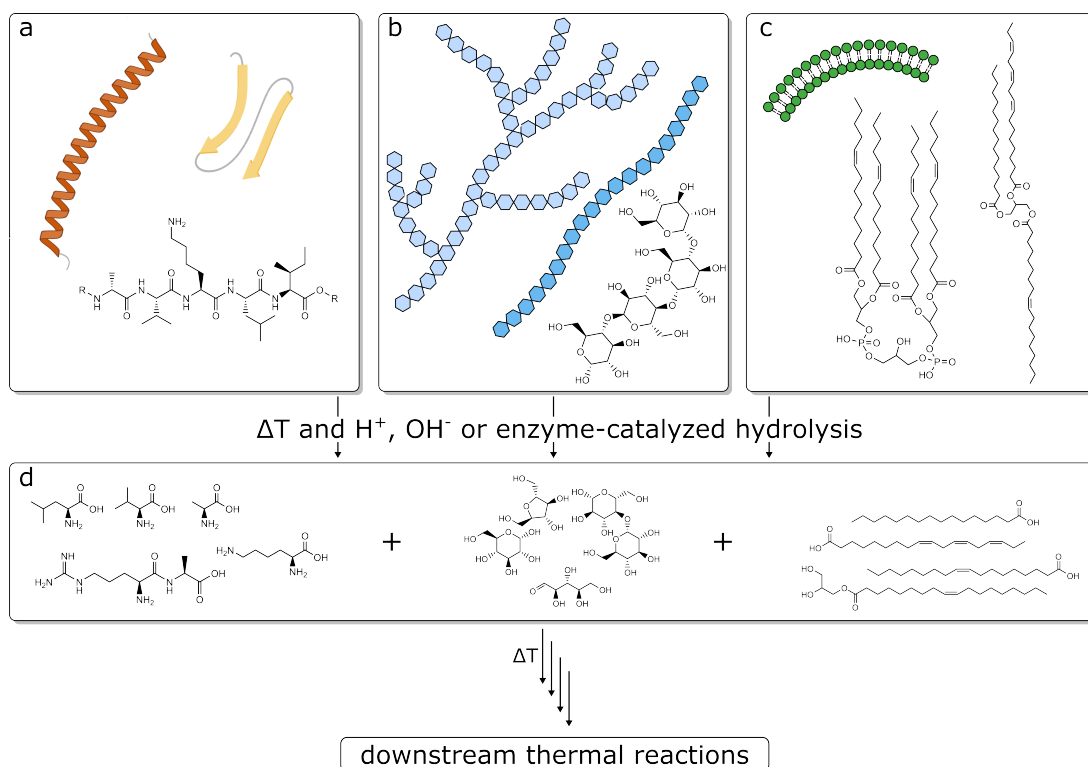
**Figure 1.1** Different pathways of hydrolysis of macronutrients occurring during the cooking of food: peptides (**a**), polysaccharides (**b**), and lipids (**c**) are enzymatic or non-enzymatic hydrolyzed to reactive monomers feeding a reactive pool of small molecules (**d**).

transglycosidation, and repolymerization (*36*). Hydrolysis generates low molecular weight reducing sugars such as glucose, maltose, maltotriose, and further homologues. During the baking of starch-rich cereal products, for example, quantitative data shows that reducing sugars accumulate in the matrix during baking, even if they are further consumed in downstream reactions in parallel (*37*). Next to chemical hydrolysis, the activity of $\alpha$-amylases from the grain also contribute to the efficient generation of small sugars during enzymatic baking phases (*37*). Reducing sugars have a sweet taste (*38*) and serve as precursors for flavor-generating reactional systems in food like the caramelization reaction (CR) and the Maillard reaction (MR), which will be outlined in chapters 1.2.1.2 and 1.2.1.3.

Hydrolysis of glycerolipids requires catalyzation, which can be performed by lipases or in an acidic as well as in an alkaline milieu. In all cases, the ester bonds are broken by the addition of water, resulting in the release of free FA, less substituted glycerol, and free glycerol; transesterifications can occur as side reactions (*9*, *33*). In a traditional sense, alkaline hydrolysis of glycerolipids with sodium or potassium hydroxide is called saponification (*6*). The reaction yields a mixture of glycerol and sodium or potassium salts of FA, which was historically used as soap. Derived from the degree of hydrolyzed ester bonds, the saponification value indicates the amount of free, unbound FA in a mixture of lipids. The release of free FA

influences the texture of food, e.g., by disintegrating lipid droplets in aqueous solutions, the mixture becomes more dispersed by increasing the lipid's surface. Solubilized constituents such as lipophilic vitamins and sterols hence become easier to digest (*9*). At the same time, lipid hydrolysis is accompanied by increasing acidity through free FA in a product (*39*). Lipid hydrolysis is associated with spoilage during storage induced by microorganisms. The development of rancid notes is attributed to releasing small, volatile FA such as butyric acid as a well-known phenomenon in expired butter (*39*). Free FA are more readily available for further chemical reactions, and their liberation marks the launch of a complex oxidative reaction cascade called lipid peroxidation, which will be further elaborated in chapter 1.2.1.4.

Hydrolysis of proteins is characterized by cleavage of peptide bonds, yielding peptides of variable length down to free amino acids. Generation of high-quality protein hydrolysates succeeds best by enzymatic digestions (*40*). Thereby, one differentiates between the cleavage of interior peptide bonds as performed by endopeptidases and the hydrolysis of terminal peptide bonds as catalyzed by exopeptidases. Enzymatic processes conserve the original amino acid profile of the hydrolyzed protein best and induce little to no chemical adulteration to the amino acid side chains. In contrast stands the chemical hydrolysis of protein mixtures, which runs under very harsh conditions. E.g., boiling matrices in 6 N hydrochloric acid for about 24 h is common practice in amino acid analysis and industrial preparation of protein hydrolysates (*41*). However, this leads to a loss of labile amino acids like tryptophan, cysteine, glutamine, and asparagine (*41*). Protein hydrolysates of countless physicochemical properties are commercially available on the market to serve as flavor enhancers, emulsifiers, stabilizers, and nutritional enhancers in highly processed, composed foods (*40, 42–44*).

Hydrolysis of structural proteins has major effects on the texture of food. As peptide bonds are cleaved, the natural structure of a protein is lost - it denatures. It should be pointed out that protein denaturation during the cooking process, however, happens for several reasons, foremost because of thermal energy impact, changes in pH, or the influence of other technological procedures. The best-studied structural protein that is subject to major changes during cooking is collagen. *Post-mortem* adulteration of collagen by proteases makes meat more tender by hydrolyzing crosslinks between fibrils (*26*), see chapter 1.2.3 for details about collagen and how its denaturation has an important effect on processing emissions.

From a nutritional point of view, proteolysis has beneficial effects by making food more palatable as the amino acid glutamate acts in its free form as a flavor enhancer and tastes savory and umami. $\gamma-$Glutamyl dipeptides were shown to boost kokumi flavor (*45*).

Protein hydrolysis also plays a crucial role in enhancing protein digestibility and bioavailability as only amino acids and small peptides can be absorbed in the intestine (*1, 9, 21*).

Further, hydrolysis is an important tool to deactivate harmful peptides such as enzyme inhibitors like trypsin inhibitors or extra somatic peptidases to prevent damage to the gastrointestinal system (*46*).

Of fundamental importance for the chemistry of cooking processes, however, is the fact that the hydrolysis of polypeptides deploys a multitude of free amino- and carboxylic acid groups to the reactive pool of food. Hence, the cleavage of peptide bonds releases functionalities that become available for downstream amino acid modification reactions. The cleavage process feeds additional N-termini to the mixture, offering more available glycation sides. Further, the breakdown of the original 3D structure reveals sterically hindered reactive side chains of amino acids. Consequently, chemical reactions found to be of fundamental importance for the generation of flavor, such as the MR, can occur more efficiently.

### 1.2.1.2 The Maillard reaction

When it comes to studying thermal processes in food, countless studies and reviews found the MR has a tremendous influence on flavor (*47–50*) and color (*50–53*) development during cooking and to impact nutritional (*54–56*) properties of heated food.

Commonly, the term "Maillard reaction" is understood as an ensemble of reactional cascades rather than a single, linear chemical reaction. From a historical perspective, Louis Camille Maillard introduced 1912 the topic as the reactions of reducing sugars and amino acids (*57*), which nowadays is still acknowledged as the foundation of the entire reactional complex. Although a wide selection of carbonyl compounds and amino functionality-bearing molecules react in food matrices in *Maillard-analogues* ways, the core concept of the MR is restricted to the reactional system initiated by the spontaneous condensation of the carbonyl group of a reducing carbohydrate with an amino group of an $\alpha$-amino acid.

Over the last 110 years, the MR was extensively studied and reviewed (*58–64*). Up to date, a reactional scheme of the MR proposed by Hodge et al. in 1953 is still valid to a large extent (*65–67*) and only got particularly refined. Accordingly, the MR is conceptualized into three main reactional phases (*68*):

    i. Initial stage:

        (a) Sugar - amine condensation

        (b) Amadori rearrangement

    ii. Intermediate stage:

        (c) Sugar dehydration

        (d) Sugar fragmentation

        (e) Amino acid degradation

    iii. Final stage:

        (f) Aldol condensation

        (g) Aldehyde - amine polymerization; formation of N-heterocyclic compounds

Under a reactional aspect, the sugar - amine condensation forms a Schiff base (*57*), see figure 1.2, structure (**a**). The intermediate condensation product readily undergoes enolization. The resulting structure is commonly referred to as Amadori rearrangement product (ARP, structure (**b**)) in case the carbohydrate was an aldose or as Heyns rearrangement product if the carbohydrate was a ketose (*69*). The ARP is a relatively stable intermediate of the early MR, sometimes used as an indicator for thermal treatment of food (*6*), and can react *via* enolization and subsequent dehydration to form highly reactive $\alpha$-dicarbonyl compounds, the deoxyglucosones, see figure 1.2 structures **c**, **d** and **e**. The deoxyglucosones are of pivotal importance for the Maillard chemistry. Emerging from them, the literature describes several reactional cascades shaping phase iii. of the Hodge scheme. Late-phase Maillard reaction products (MRP) and advanced glycation end products (AGE) are commonly attributed to a deoxyglucosone-specific pathway as a function of their early dehydration site. For example, downstream products of 3-deoxyglucosone often possess $\beta$-pyranon- or furanoide structures, e.g., like 5-hydroxymethyl furfural (HMF). Products emerging from 1-deoxyglucosone, on the other hand, can be recognized by a characteristic methyl- or acetyl-group at position 2 of heterocyclic structures, which are formed by formal reduction of the C1 atom of the carbohydrate (*6*).

Subsequent reactions enrolling in phase iii. of the MR include further dehydration reactions leading towards forming aromatic structures, incorporating N-bearing entities, dicarbonyl cleavages, retro aldolization, and aldolization reactions, among others. Which reactions are favored depends on factors like the pH value of the matrix, temperature, and available reactional partners.

Under a modern western diet, it was estimated that consumers have a dietary uptake of approximately 1 g/d of Amadori products and 25-25 mg/d of AGE (the latter formed during phase iii. **g**) (*37*). Therefore, products and intermediates of all three phases must be considered during food analysis.

An enriching modification to the Hodge scheme was proposed by Varoujan A. Yaylayan in 1997. It is attributed to the complex matrix of different interacting pathways (*70*). The MR cannot be understood as a linear reaction in a temporal course as most amino acids and carbohydrates were found to undergo independent degradation and modification reactions by themselves (*70*). Hodge did not consider in his initial work that other precursors than Amadori- and Heyns products commit towards the MR. Yaylayan defined different pools of intermediates, where "self-interacting fragmentation pools" of amino acids and carbohydrates may very well serve as substrates during all phases (*70*). It is an outstanding fact that these pools can behave recursively, which breaks with the general assumption of defined, linear reactional pathways. The consideration of these self-interacting pools became good practice in the untargeted analysis of MRP (*71, 72*) to not overestimate the impact of the MR in complex mixtures which likewise contain carbohydrate and amino acid modification products (*73*).
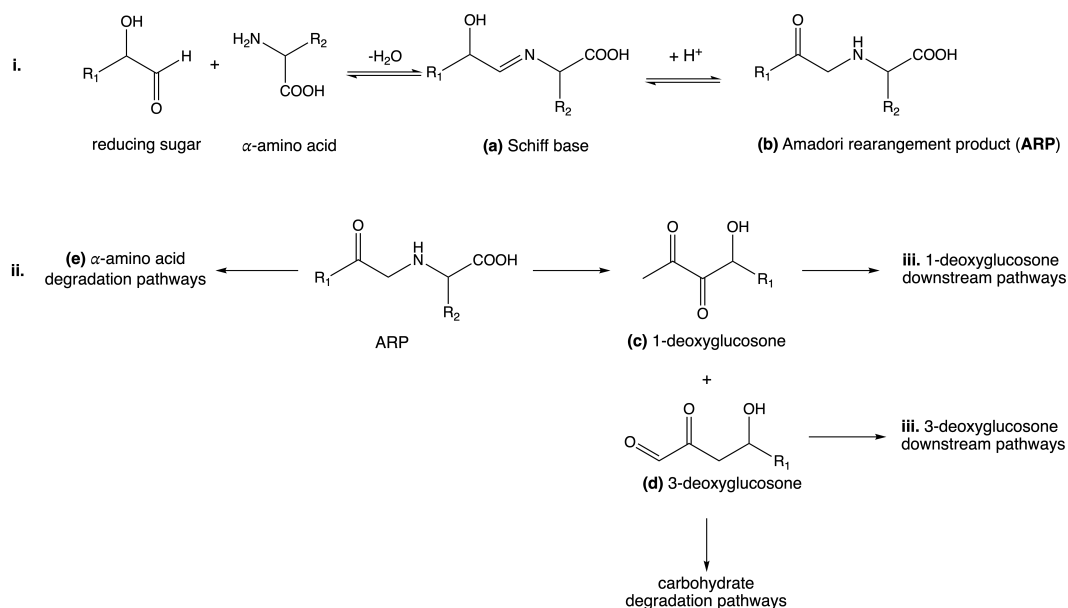
**Figure 1.2** Scheme of the Maillard reaction related to Hodge. In the initial phase **i.** a reducing sugar and an amino-functionality condense and form *via* a Schiff base **(a)** an Amadori rearrangement product (ARP, **b**). In phase **ii.** subsequent sugar dehydration reactions lead to the formation of deoxyglucosones **(c)** and amino acid dehydration products **(e)**. Deoxyglucosones are reactive key intermediates and compromise the base for subsequent reactional cascades in **iii.**.

In the analysis of processed food, an important aspect of Yaylayan's prospect must be picked up again. During recent years, it became clear that intermediates and products of different molecular families and reactional pathways interact with endogenous food constituents (*74–76*). For example, extensive research show lipid - Maillard interaction processes were long time underestimated (*77–79*), see chapter 1.2.1.4. In the same sense, Maillard-analogues processes contribute important aspects to the chemistry of heated food as their pathways are woven with the traditional MR. Thiamine (formerly Vitamin $B_1$) and its degradation product 5-(aminomethyl)-2-methylpyridin-4-amine have free amino functionalities that react with carbonyl compounds in an analogous way to amino acids. Thiamine degradation during cooking is associated with vitamin loss, browning (*80*), and is involved in flavor generation (*81, 82*). On the carbonyl side of the MR, ascorbic acid plays an important role (*83*): after it readily oxidizes to dehydroascorbic acid, various degradation pathways, which have been excellently reviewed in depth (*84, 85*), lead to the proliferation of mono- and dicarbonyl compounds which again react with amino acids to form colored compounds (*83*) and aroma active compounds (*86*).

Higher processing temperatures were shown to accelerate the MR in food (*87*). Further, water activity was found to be an important modality to control the MR. Extensive studies showed that maximum browning rates occur at water activity rates of $a_w$ 0.65-0.75 (*88*) and that reactions come to a halt below $a_w$ 0.30 (*89*). Analytes were found to be diluted at high $a_w$-values, whereas lower water availability decreases the mobility of the analytes in the ma-

trix (*90*). The influence of the pH value is more complex as it can promote different reactional pathways within the Maillard system. On the one hand, pH values < 7 favor the formation of furfurals from ARP, whereas pathways leading to fission reactions and reductone formation are boosted at pH >7 (*87*). However, at pH ≈ 3.5, most amino groups of amino acids are protonated, and the proportion of precursors available for the MR, therefore, is reduced; other reactional systems, such as the caramelization reaction are favored under this condition.

### 1.2.1.3 The caramelization reaction

Caramelization is the general term for an ensemble of carbohydrate degradation reactions that occur when sugars are heated at high temperatures and in contrast to the MR in the absence of reactive amino groups (*91*, *92*).

Heat-induced decomposition of carbohydrates leads to a complex compound mixture that gives caramel its characteristic color, aroma, and flavor. The CR can be catalyzed by acids or bases, i.e., ammonia caramel is a coloring food additive generated in the alkaline milieu (*93*). For all carbohydrates, except monosaccharides, prior to the CR comes fission of glycosidic linkages releasing free reducing sugar units (*94*), see also chapter 1.2.1.1. Following, initial phases of the CR are characterized by intramolecular rearrangements and dehydration of carbohydrates (*95*): the transformation of reducing sugars begins with an enolization known as the Lobry-de-Bruijn-van-Eckenstein-rearrangement which is attributed pivotal importance (*96*). Successive reactions from the key osulose-intermediates on create a diverse spectrum of precursors and involve dehydration or $\beta-$elimination, dicarboxylic cleaving, and retro-aldol reactions (*95*). Later reactional phases lead to the formation of a complex mixture of stable compounds. Volatile constituents of caramel, O-heterocyclic- and carbocyclic compounds have extensively been studied (*92*, *95*), simultaneously enrolling transglycosidation reactions, generation of inter- and intra-molecular sugar anhydrides and the *de-novo* formation of oligosaccharides by polymerization reactions create a complex mixture of compounds which still remains partially uncharacterized (*95*, *97*). On a low-molecular level, furans from monosaccharides such as HMF and other downstream degradation products like furfural and 5-methylfurfural, or pyranes from disaccharides such as maltol are important aroma-active products of the CR and are generated in large molar proportions (*95*). HMF, however, is a good example that various reactions lead to the same compounds as HMF is also well known to arise from the MR (*98*).

The CR is likewise responsible for the formation of coloring agents. From reactive intermediates on, brown polymers are formed, which give caramel its characteristic color. Their structure is complex and up to date, poorly understood. However, reactive aldehydes like HMF, furfural, or hydroxyacetyl furan are postulated to build regularly structured macromolecules (*95*).

During the heating process, the CR is limited to regions in foods where carbohydrate-rich compartments are exposed to intense heat. Practical and desired applications occur,

for example, during the roasting of coffee (*99*) or toasting of bread (*100*), whereas browning during the production of milk powders is undesired (*101*).

### 1.2.1.4 Lipid peroxidation processes

Unheated lipids primarily contribute to the flavor of food by their taste and mouthfeel. Whereas the textural perceptions of lipids has been elaborated for a longer time, evidence of a gustatory pathway to perceive dietary fat as a sixth human taste modality consolidated just recently (*102*, *103*). Orosensory detection of lipids is presumed to be mediated by free FA (*103*), see also chapter 1.2.1.1 about the hydrolysis of glycerolipids to free FA. The olfactory properties of lipids are further mediated by volatile compounds, which are generated by lipid degradation processes induced by cooking. The main responsible pathway for the generation of aroma-active compounds from lipids is the lipid peroxidation (LPO). The exact biological and chemical conditions of LPO are as complex as the diverse substance classes of lipids are and, therefore, not exhaustively understood (*104*, *105*). There are several LPO mechanisms that can be classified as enzymatic, non-enzymatic, non-radical peroxidation, and non-enzymatic free-radical mediated peroxidation (*106*). Of highest relevance for the transformation of nutritional lipids during cooking processes is the free-radical non-enzymatic peroxidation of polyunsaturated fatty acids (PUFA). The primary products of this pathway are highly reactive lipid hydroperoxides (see figure 1.3 structure **d**), which further decay to more stable secondary products. The reaction can be divided into three stages, summarized in the following scheme, originally proposed by Guéraud et al. (*105*). Central reactions of the free-radical mechanism are accordingly visualized in figure 1.3.

i. Initiation: abstraction of an H· radical from a lipid (LH) to form an L· radical; an initiator is required (L'H in figure 1.3). L· radicals of PUFA are resonance stabilized and therefore favored. The initiation can, e.g., be metal-catalyzed.

ii. Propagation: the initial L· radical reacts with oxygen to form a lipoperoxyl LOO· radical, which in turn reacts with a second lipid to form an L· radical and the primary product of LPO, a lipid hydroperoxyde (LOOH). LOOH are unstable and decay into secondary products.

iii. Termination: a combination of radical species to yield non-radical or non-propagating species.

For the initial formation of a radical starter, a trigger is needed: of the highest importance in food is the photo-induced formation of singlet oxygen, which is supported by sensitizers contained in food, e.g., riboflavin (*107*). The subsequent radical formation is formally catalyzed by the presence of $M^{n+}$ metal ions like $Fe^{2+}$, $Cu^+$, and $Co^{2+}$ (*107*). Especially heme-rich tissues like muscles, which contain large amounts of myoglobins, are known to cause strong induction of radical-mediated LPO.

The secondary products of PUFA peroxidation are of great interest for cooking research. LOOH can further react to alkoxyl radicals (LO·), which can undergo $\beta$−scission to yield products of shorter alkyl chain length. Dependent on further oxidation reactions, these can have variable chemical functionalities such as aldehydes or carboxylic acids, which preferably evolve from $\omega6$- and $\omega3$-FA (*108*). Further, LOOH are known to rearrange to hydroxyl- and epoxy-acids, to dimerize or to polymerize (*105*).
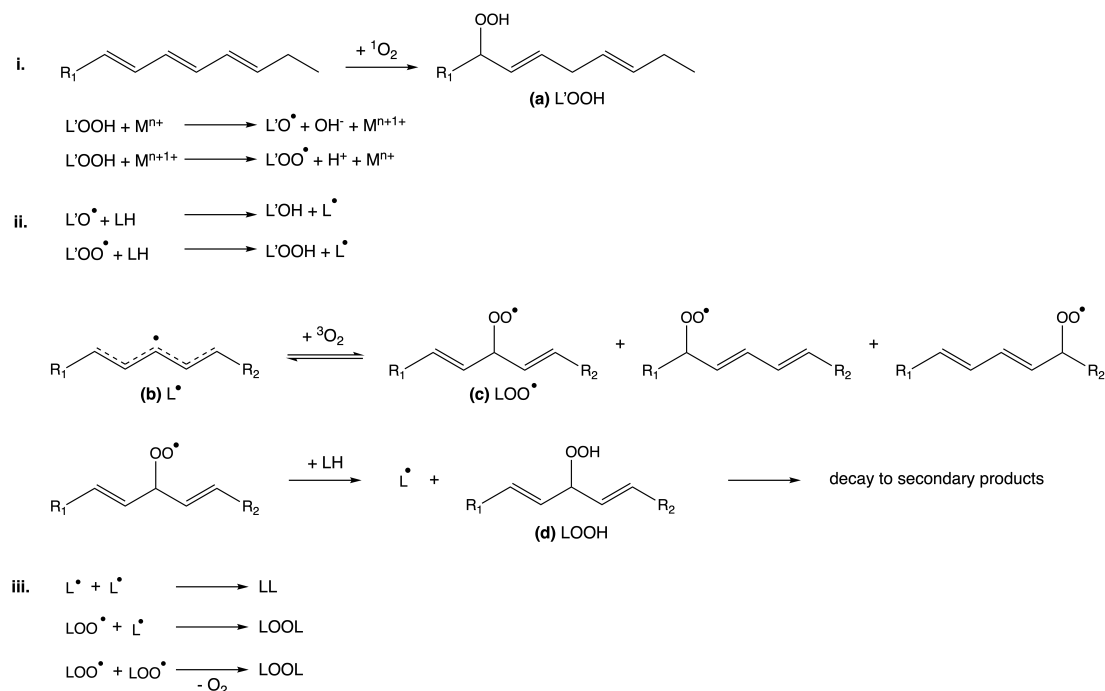


**Figure 1.3** Lipid autoxidation after Guéraud (*105*): In the initial phase **i.**, singlet oxygen forms a first lipoperoxid **(a)** by an Ene-reaction which decays under metal catalysis to lipoperoxyl- and lipalcoxy radicals. In phase **ii.**, hydrogen abstraction from another lipid L starts a chain reaction. The alkyl radical **(b)** reacts with triplet oxygen to form a lipoperoxyl radical **(c)**. Subsequent hydrogen abstraction creates a new alkyl radical and a lipid hydroperoxyde LOOH **(d)**, which undergoes manifold downstream reactions. In phase **iii.**, the reaction of two radicals stops the cascade.

Likewise, the decomposition of secondary products is known to form other derived compound classes such as ketones, alcohols, esters, hydrocarbons, furans, and lactones, all of which being highly relevant for the aroma of cooked food (*108*, *109*). Of substantial interest regarding dietary health are small unsaturated aldehydes such as acrolein and 1,2-dicarbonyls as glyoxal or methylglyoxal. Named compounds are arising in significant levels during the LPO that were found to be of higher concentration compared to those generated by the MR from carbohydrates (*17*). Furthermore, dietary advanced lipid peroxidation end products (ALE) were proven to activate inflammation response upon resorption, and many known ALE are associated with negative implications on human health (*17*, *108*).

The interaction of lipid hydroperoxyde radicals naturally is not limited to other FA. Other lipid classes, such as sterols and lipophilic vitamins, can suffer from their oxidational dam-

age, resulting in a loss of nutritional value (*108*). Additionally, lipid-Maillard-interactions, as already primed in chapter 1.2.1.2, is a topic on the intersect of carbohydrate and lipid research (*77*, *110*). As both the MR and LPO happen simultaneously during the cooking process, e.g. reactive aldehydes like 4-hydroxynonenal, which is a common product from $\omega$6-FA degradation, can react with cysteine or lysine residues to form stable thiophene, thiapyran or pyridine adducts (*77*). On the positive side, it is noteworthy that lipid-Maillard products were shown to contribute beneficially towards the aroma of cooked meat (*111*).

### 1.2.1.5 Formation of thermal reaction markers

In the course of studying the reactional effects of cooking, many scientists came to the point of defining thermal reaction markers. The concept of thermal reaction markers refers to specific compounds that possess an indicative character that undergo characteristic concentration changes or reactions under variable thermal conditions. These markers are utilized to monitor and evaluate the extent of thermal load submitted to food products. Compounds with high sensitivity towards temperature and water activity are generally regarded as promising heating markers (*112*). Generally, the introduction of novel thermal processing indicators can be divided into three phases:

i. Untargeted discovery of a thermo-sensitive feature. Reactional understanding of its formation pathway.

ii. Quantitative studies on its formation kinetics and behavior under different processing conditions.

iii. Establishment of the marker in routine analytical protocols.

Many contemporary processing markers that are routinely employed in food monitoring programs were discovered during untargeted studies of reactional frameworks like the CR, MR, or LPO. Embedding a potential marker compound into a reactional context and understanding its generative background helps to strengthen its reliability for further routine analysis. Thermal processing markers can be classified depending on their reactional background and functionality.

First, one of the oldest applications of thermal processing indicators is to identify heat adulteration of food. Possibly the most popular heating marker is HMF, see section 1.2.1.3 for its background and formation. HMF is widely used in honey analytics to detect the usage of fraudulent heat treatment during honey harvest to increase extraction yields from the honeycomb or to determine whether honey might be illegally blended with high fructose corn syrup (*98*, *113*, *114*).

Second, there are processing indicators imparting nutritional information. Various MRP and AGE take a prominent part of this section, as the MR is a highly popular research topic

in analytical food chemistry. Some authors evaluate the extent of thermal damage ([113]) and the quality of food based on the MR ([115]), e.g. Hellwig et al. propose to assess the degree of heat treatment of pasta products based on some AGE ([116]). Most notably, the MR goes along with the loss of essential amino acids, decreasing the nutritional value of a protein source ([56]). Nowadays, a wide variety of AGE such as furosine, N-$\epsilon$-carboxymethyl lysine, pyrraline, pentosidine, and pronyl-lysine serve as markers for thermal treatment of different foods ([56]). Further, intense research is carried out about the effects of dietary AGE uptake. The discovery of a receptor for AGE led to advances in understanding their role in inflammatory processes ([117]). Further, there is a close connection between AGE, aging, and other metabolic processes which are under ongoing medical investigation ([55, 117–119]).

Another intensely studied marker compound class indicating harsh heat exposure is the family of heterocyclic aromatic amines (HAA). HAA arise from Maillard reaction-like interactions of reducing sugars, amino acids, and creatine under high-temperature pyrolysis, especially in cheese, meat, and fish ([120, 121]). HAA are aimed to be reduced in processed food due to their mutagenic and carcinogenic properties ([122–124]).

In addition, melanoidines are complex condensed, high-molecular MRP responsible for pleasant coloration of heated food ([125]). Although their structure is barely understood up to date ([126]), recent studies suggest health-beneficial properties such as antioxidative activity ([125, 127, 128]).

Third, thermal processing markers can be used to increase the palatability of food by precisely boosting reactional pathways responsible for generating desired flavor compounds ([129]). Many pleasant volatile aroma compounds such as furans, N- or S-heterocyclic compounds, or LPO are handled as thermal reaction markers ([112, 130–132]). For example, as demonstrated in studies of bakery products, targeted adulteration of formulations ([133]) and careful selection of appropriate processing conditions ([134]) can lead to higher consumer acception.

By transporting structural, reactional or compositional aspects, processing indicators can be used as infochemicals which grant particular insight into cooking processes. They can be used as valuable tools for ensuring food safety and quality by providing chemical indications of whether a product has been subjected to the appropriate temperature and time during processing. These indicators play a crucial role in verifying the effectiveness and quality of thermal treatments, reducing the risk of foodborne illnesses, and maintaining product consistency.

### 1.2.2 Biological and nutritional aspects of cooking

The previously discussed chemical reactions focus on transformations of food on a molecular level. From a life sciences-oriented perspective, these reactions cannot stand isolated. The human intention in cooking includes further nutritional purposes than chemical transformations, leading to a complex integration of biological aspects. These shall shortly be outlined to support the later discussion.

It was extensively shown that extra somatic hydrolysis of starch, proteins, and lipids facilitates gastrointestinal digestion. Better ileal digestibility of nutrients attains that less nutrient-rich food remnants reach the colon. Thus, the net content of energy and biological substrates the human body can extract from nutrition is increased by cooking (*1, 9, 135*).

Since clear evidence was gathered that human health is directly related to a healthy gut microbiome, several studies indicate that a person's microbiome is strongly dependent on personal food uptake and the way it is processed (*16, 136*). It was evident that the microbiomes of rodents fed with raw versus cooked tuber were fundamentally distinct, with positive aspects for the host triggered by a healthier gut microbiota under cooked diet (*16*). These findings reinforce the hypothesis that humans and the human gut microbiome co-evolved under cooking-related pressure (*16*). Next to structural changes in macromolecules, cooking was likewise shown to have additional beneficial effects as heat can denature antimicrobial and antinutritional compounds like protease inhibitors found in food (*137*). Nevertheless, cooking is not the only processing technique found to influence the microbiome. Likewise, fermentation processes can also pre-digest nutrients and affect the gut microbiota (*138*). Artificial food additives are also known to impact gut microbial composition (*136*).

An original intention of cooking lies in the inactivation of food-borne pathogens, spoilage-inducing microorganisms, and the killing of parasites. Spoilage leads to a loss of qualitative properties in food, which often results from the metabolism of bacteria, yeast, and molds (*39*). Even if the consumer's evaluation of spoilage relies on individual and objective organoleptic assessment (*139*), excessive microbial growth can likewise result in the production of food-borne toxins. Fungi can produce undesired mycotoxins (*140–142*) and bacterial genera like *Bacillus spp., Clostridium spp., Staphylococcus spp.*, and various others can produce endo- and exotoxins which are a substantial threat to consumer safety (*143–146*). Thus, microbial safety assessment, including good hygienic practices and reduction strategies, are indispensable elements of modern food processing strategies. Preservation protocols nowadays have become multi-step interdisciplinary approaches (*147*): traditional techniques such as drying, chilling, or chemical preservation present the first step for microbial control. Here, heating techniques such as pasteurization or sterilization play a central role in reducing the number of colony-forming units. The harsher the bacterial heat-inactivation is carried out, the stronger the chemical side reactions that enroll during the preservation process are. Furthermore, emerging alternatives for the inactivation of microorganisms supplement preservative methods. These technologies can, for example, involve irradiation (*148*), Ohmic heating (*149*), high pressure processing (*150*), pulsed electric field processing (*151*), microwaving (*152*), ultrasound processing (*153*) and many others. The effects on pathogens and side effects on the matrix are as diverse as the physical principles of these technologies, and thus, the interested reader shall be referred to the cited literature.

Lothar Leistner carried out extensive work around the concept of "hurdle technology," which thrives for the proficient combination of various bactericidal and bacteriostatic treat-

ments. This involves the manipulation of parameters like $a_w$- and pH-values, redox potential, heat treatment, and other modern technological solutions (*154–156*).

Common to modern preservation techniques is that most approaches try to find complementary mechanisms different than heat to inactivate microorganisms. On the one hand, this can create synergistic effects in the spirit of a hurdle concept. On the other hand, these techniques do not stand in the direct focus of this thesis as their approaches trigger no or different chemical reactions in the matrix compared to traditional heating.

In conclusion, cooking has profound biological effects on food. It not only enhances the digestibility of nutrients, making them more readily available for absorption in the intestine but also destroys harmful pathogens and antinutritional compounds, thus ensuring food safety. Cooking can also lead to the loss of certain heat-sensitive vitamins and antioxidants, emphasizing the importance of balanced thermal processing conditions to preserve the nutritional value of food and to avoid generating thermal process contaminants.

### 1.2.3 Thermally induced physical changes in the matrix

Molecular transformations during the cooking process are accompanied by principal physical changes occurring to food. The process of convection cooking simultaneously triggers heat and mass transfer in the processed matrix. While heat is transferred from the hot air to the food, an equilibrium between absorption of moisture under high-humidity dry-food cooking conditions and evaporation under low-humidity moist-food conditions is reached during the process (*157, 158*). Accurate modeling of simultaneous heat and mass transfer processes is a complex problem: most theoretical models lack real-life applicability as they assume simplified geometries or constant processing conditions (*159*). However, heat transfer coefficients vary through the cooking process as density, thermal conductivity, and heat capacity are functions of food temperature, moisture, and fat contents and therefore alter under thermal influence (*157, 159*). Heat flux models help to determine accurate processing conditions in relation to core temperature, energy costs, and help optimizing product quality and yield by finite elements or finite differences calculation methods (*159–161*). However recent advanced were made in the modelling of transport phenomena, texture and color changes during thermal treatment of poultry meat (*162, 163*).

However, the most relevant physical aspect for the discussion of this thesis is the mobility of moisture under thermal influence. The studies carried out in this work deal with the analysis of thermal processing emissions in the head space of cooked food. Next to evaporated volatiles, small organic compounds are transported to the head space by relocation processes of water from the matrix. Water evaporation takes place in direct relation to the temperature of food. The surfacing area of food is the interface between the heated air and deeper-lying matrix compartments. Thus, it is heated most intensely and most rapidly. As a

result, an everyday phenomenon takes place: the formation of the crust by initial drying of the surface (*158*, *164*).

Case-specific moisture diffusion patterns vary as a function of matrix morphology. The literature generally agrees that there is a water evaporation front that moves towards the inside of the processed item during convection cooking. This initially results in the differentiation of two zones. The crust is a low-moisture compartment, and the crumb - crumb is a terminus borrowed from bakery sciences and shall be used illustratively for interior regions of food below the differentiated crust - as a high-moisture compartment. The evaporation front progressively advances towards the center of the product, causing the core temperature to asymptotically approach 100 °C. The surface, by contrast, tends to approximate the oven temperature (*165*, *166*). Accordingly, flux models found that water vapor movement divides into two directions: once parts of the contained water diffuse to the surface, whereas another share moves deeper towards the center of the product, where it condenses in colder regions (*158*). During cooking, crust thickness progressively increases towards the inside of the product, leading to a restriction of water diffusion from internal parts towards the outside by forming a barrier.

Accordingly, the surface of food offers different reactional conditions compared to its interior. In the low-moisture crust, the $a_w$-values significantly lie below the interior, the mobility of small organic compounds is reduced, and the higher thermal impact enables reactions which require higher activation energy (*112*, *167*). As a result, most browning and flavor generation takes place on the surface of food (*112*, *164*).

Again, many moisture transport models presume ideal and homogeneous matrix structures, which often are not found in real-life food and neglect other important physical processes. Portanguen et al., for instance, found that the principal share of water that is lost from convection-cooked meat is juice expelled by protein denaturation and fiber contraction caused by denaturation (*164*).

Thermal energy has direct physical effects on the conformation of macromolecules and, thereby, on the texture of food. Under the influence of heat, denatured proteins lose their native secondary, tertiary, and quaternary structure. Enzymes thereby lose their functionality, which commonly is desired during cooking as the durability of food increases because degenerative enzymes like hydrolases (*135*) or harmful compounds like protease inhibitors are deactivated (*46*).

Heating of muscle meat induces shrinkage of the cut for two reasons. First, denaturation of myofibrillar actin and myosin causes contraction both in diameter and longitudinal axis (*168*). Second, collagen is quite resistant to thermo-induced breakdown. However, collagen likewise shrinks during denaturation, leading to increased fluid loss by the pressure exerted by this thermo-induced contraction (*24*, *26*, *169*). The degree of tension caused depends on the number of mature crosslinks in the fibrils, which is attributed directly to tenderness (*169*). Due

to its thermo-resistant nature, *post-mortem* hydrolysis of collagen is technologically favored, e.g., by supplemented collagenases (*26*).

In bakery products, the temperature-induced coagulation of glutenin and gliadin, which are the main constituents of gluten, helps to maintain the shape and structure of the baked goods. During dough production, a network of interconnected gluten strands forms which, as it denatures, retains evaporating gases and thereby supports a light texture (*6, 170, 171*).

A further highly desired physical process is the gelatinization of starch, which is of fundamental importance for bakery goods. In fact, anthropologists suspect that human cooking found its origin in heating carbohydrate-rich tubers to modify contained starches (*135*). Cooking of these significantly boosts gelatinization and dextrinization of the root starches, increasing their digestibility (*1*). The gelatinization of starch is a process that occurs when starch is heated in the presence of water. It is a fundamental transformation that alters its physical and chemical properties by a transition from an ordered to a disordered state (*172, 173*). Gelatinization turns unsolvable starch granules into a gel-like solution of its individual constituents, amylose and amylopectin (*174*). According to Donovan, gelatinization involves three major characteristics (*175*):

  i. Loss of crystallinity of the granules, formation of an amorphous state.

 ii. Heat absorption of water to amorphous space of amylopectin.

iii. Hydration of the starch accompanied by swelling of the granule. Amylose diffuses to the surrounding water, and the granule disintegrates.

The baking process is characterized by partial gelatinization: amylose absorbs water to form a gel with swollen amylopectin (*6*). The partial gel has several beneficial properties for the final bakery product, including moisture retention, stabilization of other constituents, and texture enhancement. Additionally, solubilized dextran chains are better accessible for enzymatic and chemical hydrolysis, see chapter 1.2.1.1.

The opposing process of retrogradation,l where starch re-enters more crystalline structures, shall not be explained here as it only plays a role *posterior* to cooking during food storage.

Conclusively, these few examples of the behavior of macromolecules emphasize the complexity of physical responses in food triggered by thermal energy. Associated effects do not only influence the texture of cooked food, they likewise directly influence chemical processes by break-downs of macromolecules or deactivation of enzymes. Most importantly, physical aspects help to explain the transport of water and conveyed organic compounds to the head space.

## 1.3 The concept of studying thermal cooking emissions

Extensive literature research suggests that the practice of analyzing emissions of cooking processes dates back to the year 1952 when Ernest Albert Pence introduced the principle of condensing vapor from bread baking processes (*176*). Pence used cold-trapped baking emissions from a convective oven to gain insights into the chemical characteristics of bread flavor. By application of fractional distillation techniques, solvent extraction, crystallization, chromatographic separation, and a combination of these techniques, Pence identifies several molecules and compound classes to be present in bread vapor. The most important constitutions were found to be alcohols, acids, esters, aldehydes, and ketones. Among the identified compounds were ethanol, formic acid, acetic acid, and several others (*176*). By analyzing its thermal processing emissions, Pence drew conclusions about the aroma profile of baked bread.

Pence likewise conducted extensive theoretical work about how to collect vapor samples most efficiently. He compared different means of adsorption sampling and optimized cooling procedures for vapor collection. As Pence concluded that cold-trapping of baking fumes is best suited for his purposes, he already identified the central difficulties of the vapor sampling methodology. The main analytical challenge lies in the chemical and physical nature of gas collection (*176*). Evaporating water from bread dilutes analyte molecules significantly. Further, convection ovens are not hermetically closed and, therefore, open systems, which leads to analyte loss. Analyte molecules might not evaporate quantitatively, and efficient cold-trapping, especially of very volatile compounds, is challenging. Thus, quantitative calculations of exact concentrations in the matrix based on processing emissions were found to be of a rather approximate nature (*176*). Nevertheless, this study can be regarded as the first introduction to oven vapor research. Pence paved the road for efficient qualitative screening of organic constituents released from cooked food to its head space. By the means of modern instrumental analytical techniques, a complex system of molecular spaces awaited to be discovered.

In the past 70 years since Pence's reflections, the applications for the analysis of thermal processing emissions have partially shifted. This chapter will give a short bibliographic overlook about current applications and centers of interest that modern analytical chemistry offers on cooking emissions. Nowadays, the considerations about cooking fumes have become more diverse and more specialized. Particular interests arose with advancing analytical instrumentation. Therefore, the field is split currently into two major disciplines. On the one hand side, cooking emissions are studied behind the background of environmental sciences in the contexts of urban and indoor air quality. On the other hand, food scientists still use cooking head spaces to draw conclusions about the underlying matrix, as Pence did.

Environmental scientists carry out extensive research to identify sources of man-made air pollution. A huge community aims to identify substances harmful to the human respiratory

system arising from cooking processes and how to reduce them. Especially metropolitan areas in the United States, China, and other countries of Southeast Asia are at the center of research interest due to a dense concentration of many high-temperature cooking sites [177]. In the environmental context, cooking methods such as barbecuing, meat grilling, or wok frying are investigated [177, 178]. Cooking fumes are unanimously regarded as one of the most important sources for the emission of volatile organic compounds (VOC) and organic aerosols to urban atmospheres next to traffic [177, 178]. Studies on the constitution of urban aerosols found several compound classes to be major constituents of cooking aerosols. The nature of emitted compounds heavily depends on the cooking method employed. So far, no holistic analytical investigation on the composition of cooking-induced air pollution in direct comparison to the processed matrix is available in the literature [179].

The environmental focus of cooking emission analysis primarily lies on air-borne (ultra-) fine particulate matter (PM) consisting of solid particles and liquid droplets that have an aerodynamic diameter smaller than 2.5 $\mu$m (PM2.5). PM is termed *primary* if it is directly emitted to the atmosphere, and *secondary* if it arises from chemical reactions of primary PM. PM carries along primary air pollutants such as nitric and sulfuric acid, respective salts, ammonium, $NO_x$, elemental carbon, and VOC, which often bear reactive carbonyl groups [180]. It can be transformed in photochemical smog to ozone and cause nitrogen transfer by photolysis of $NO_2$.

Most important and quantitatively relevant constituents of indoor cooking emissions are non-methane gases such as methanol, dimethyl sulfide, CO [181], $NO_x$ [182], diverse VOC [183], alkanes and polycyclic aromatic hydrocarbons (PAH) [184, 185], and volatile reaction products of the MR and LPO [177], see also chapters 1.2.1.2 and 1.2.1.4 for the background of their emergence. The most severe implications on personal health were found to arise from PAH, carcinogenic and mutagenic VOC such as formaldehyde, acrolein, and acrylamide, and of OH· radicals from LPO processes [182]. Further, the contribution of cooking emissions to photochemical smog and ozone production was proven [186].

As a matter of fact, indoor air pollution is significantly more concentrated and severe than cooking fumes exhausted to the urban atmosphere, e.g., resulting in elevated cancer risks for chefs in poorly vented, traditional Chinese kitchens [187, 188].

Although the head space-centered sampling approach offers various advantages, the vast majority of studies investigating the effects of the thermal processing of food focuses on direct sampling protocols of the processed matrix. In contrast to environmental-focused studies, a handful of food researchers give a closer look at processing emissions right during the cooking process under in-oven conditions. The advantages of sampling directly from the head space during convection cooking involves several points. First, there is no need for time-consuming extraction protocols specified for single compounds. No bulk-matrix constituents need to be separated from the matrix, making the extraction process greener with

less solvent being consumed. The only bulk constituent of thermal processing emissions is water. Second, organic compounds over a wide range of volatility and polarity can be assessed with a single method (*130*). Third, sensitive reaction products like furanes have no time to degenerate when directly measured from the head space (*112*). Fourth, on-line sampling and measurement techniques are indispensable concepts to best understand chemical reactions that enroll *during* cooking. As discussed earlier in chapter 1.2.3, food's thermal flux is not necessarily linear and uniform. Therefore, thorough kinetic calculations require online data acquisition, which is best practiced out of the head space (*130*).

As a result, several food products were already investigated during their processing in convection ovens from a head space's perspective. Many of the previously named environmental studies deal with aerosols from meat products. Rochat and Chaintreau, e.g., set up a noteworthy study to investigate meat cooking emissions from an in-oven perspective (*189*). Their work focuses on carbonyl odorants arising during the cooking process of roast beef and emphasizes that aromatic notes, which are measured in the thermal processing vapors, contribute to the first sensory impression a consumer has of the final product (*189*). In contrast to environmental analysis, this study enlightens valued sensory aspects of meat processing emissions and does not merely treat them as harmful pollutants.

Rosati et al. carried out an interesting study on the emissions from microwave cooking of popcorn (*190*). By means of head space sampling, the authors were able to record time-resolved VOC and particle emission profiles of popcorn during the popping process and in the successive 40 min after bag opening. Next to several butter-flavoring aroma notes, the profiling study revealed different organic compounds from the popcorn bag components (*190*).

Klein et al. characterized non-methane organic gases from the cooking of different vegetables and meats (*181*). Different VOC like acrolein, hexanal, or 2,4-decadienal from the decomposition of frying oils were found to dominate the emissions. Dimethyl sulfide was described to be emitted by the cooking of vegetables from the *Cruciferae spp.* in significant amounts (*181*). Different aldehydes were proposed as tracers for the identification of cooking emissions indoor (*181*).

Most studies published about thermal processing emissions in convection ovens deal with the analysis of bakery products. Joana Pico et al. for example studied the loss of valued aroma compound from gluten-free bread to the oven head space during the baking process (*170, 171*). Gluten-free breads possess poor aromatic properties, which is why extensive research is carried out to create formulations with appealing aromatic notes, increasing consumer acceptance. Pico found head space sampling as a straightforward method to understand loss processes in real-time and was able to model release patterns with differently shaped curves (*170*). Gluten-free bread showed a higher loss of volatiles than opposed wheat bread, which was offered as a perspective for future research to improve the aroma of gluten-free bread (*170*).

Excellent work about volatiles emitted to the head space during the baking process of pastries was carried out around Barbara Rega. She developed non-disruptive in-oven sampling devices to investigate the baking processes of sponge cake and biscuits (*191*, *192*). In several works, Rega managed to compare the head space concentration of volatiles to the matrix itself (*112*, *192*). Positively evaluated aroma compounds, especially from the MR, but also from CR and LPO reactions like furans, N-heterocyclic compounds, and aldehydes, were regarded in these studies, also in the context of reaction markers. An extensive overview of kinetics and release dynamics was compiled on them to follow the evolution of compounds of interest (*129*, *130*).

Most authors studying thermal processing emissions pursue matrix-specific questions and do not aim for a holistic description or for a classification of the head space *per se*. In combination with computational data processing approaches, modern analytical methods offer the possibility of untargeted screening of thermal processing emissions for a profound and holistic understanding of cooking emissions. In the following chapter, section 2.3 will refer to the technological procedures used during the studies presented above. Furthermore, methodological and computational approaches for characterizing thermal processing emissions will be outlined.

# 2 Methods

Analytical food chemistry had its beginnings in the early to mid $20^{th}$ century with the reliance on predominantly "wet chemistry" laboratory methods. Soon, these methods gradually gave way to modern instrumental techniques, which were rapidly developed and got established in routine laboratories (*193, 194*). The unrivaled speed and precision of modern instrumental procedures had a meaningful impact on food analysis in multiple fields, such as assessment of food quality and authenticity, to ensure food safety and compliance with legislation and regulations (*193*). Gaining an increasing molecular understanding of chemical reactions in food allows to control nutritional and sensorial aspects during food production.

The following chapter gives an introduction to contemporary branches of modern food analysis. Special emphasis is given to methods that were used in this thesis, including mass spectrometry-based techniques, hyphenated techniques, modern chromatographic developments, and bioinformatic tools for data processing and evaluation.

## 2.1 Analytical strategies to decipher the molecular complexity of food

Most traditional methods used in food analysis primarily focus on detecting selected compounds or a predefined set of analytes. Along with the development of untargeted methods evolved particular challenges. This is mainly due to the inherent nature and variety of food samples, as well as the wide concentration range of analytes (*195*). Chapter 1.1 drew a synopsis of the chemical substance classes found in food. Depicting their entire molecular complexity, from original matrix components on to thermal reaction products, requires untargeted screening methods that enable the unbiased profiling of samples, allowing for the detection and characterization of unexpected compounds without *à priori* knowledge of suspects or preceding targeted analysis.

Historic examples like the melamine scandal of 2008, where milk was thinned with water and blended with the nitrogen-rich compound melamine to mimic the protein content of unadulterated milk (*196*), or increasing fraud in the production of honey (*197*) as well as in "extra virgin" olive oil (*198*) teach us two lessons. First, analytical chemists need to expect the unexpected. Second, one should not focus on a limited set of marker compounds but rather pursue untargeted investigations to reveal adulterated molecular profiles.

Different technological platforms with corresponding data processing and evaluation pipelines were developed and are readily available to tackle present analytical challenges (*195*). The most powerful platforms for molecular fingerprinting include nuclear magnetic resonance (NMR) spectroscopy and mass spectrometry (MS). The following chapters will discuss MS-based methodologies suited for discovering reactional changes in food.

### 2.1.1 Metabolomics: molecular characterization of a complex system

The concept of "metabolomics" has been consolidated around the turn of the millennium in the endeavors of system sciences to holistically describe cellular metabolic processes. Compared to other *-omics* approaches like genomics (*199, 200*) and proteomics (*201, 202*), which evolved in the 1980s and 1990s respectively, metabolomics is a relatively young discipline targeting on the discovery of small organic compounds involved in metabolic pathways, see figure 2.1. The evolution from genomics to metabolomics *via* proteomics reflects a change in scientific paradigm and corresponds to a progression from gene to function to consideration of the entire phenotype. The genome of an organism is static for its entire existence. Rapidly arose the cognition that even if next-generation sequencing techniques allow batch analysis of entire genomes, more dynamic approaches are needed to understand fast-paced cellular processes. Transcriptomics was developed to analyze ribonucleic acid (RNA) and to consider regulatory transcriptional effects. However, the end product of transcription and translation are proteins that finally carry out the true function encoded by a gene, which led to the establishment of proteomics. The evolution of *-omics* goes along with an increase in analyte complexity: one exon encodes one protein, whereas a number of different gene loci can encode one protein with different isoforms and post-translational modifications. A metabolite, which is commonly regarded as an organic compound with a molecular mass below 1,500 Da, however, can be subject to multiple interacting reactional pathways by being a substrate of multiple enzymes. Whereas the human genome was found to contain less than 27,000 protein-encoding genes (*203*), the number of metabolites is estimated to range around 100,000 for humans (*204*) and 200,000 for plants (*205*). There is an estimated number of millions of xenobiotics that need to be taken into account in addition to endogenous metabolites as they are part of a system's exposome and are estimated to exert a strong impact on human health (*206*). For this reason, the metabolome of a cell grants the most direct insight into the exposome of a system of genetic and environmental regulations. It is therefore non-neglectable to completely describe a phenotype (*207, 208*).

In contrast to genomics and proteomics, the physicochemical properties of the analyte compounds are much more diverse in metabolomics. Proteins, as well as DNA, are biopolymers and built by a defined set of precursor compounds (20 proteinogenic amino acids or, respectively, four deoxyribonucleotides), which are repetitively connected on a known backbone. Small organic compounds, by contrast, are very heterogeneous in their atomic compo-

sition, size, structure, and stereochemical orientation. These variations together represent a fundamental challenge for the *de novo* characterization of small metabolites (*208*).

This vast complexity led to communal efforts to define the frame of metabolomics research. Many research units contributed to the discussion, including pioneering groups around Oliver Fiehn and Jeremy Nicholson (*208, 209*). In 2002, Fiehn initially defined the term *metabolomics* as the comprehensive analysis of all the metabolites of a biological system, including their quantitation and identification (*208*). By his definition, Fiehn drew a concrete line to closely related practices of *metabolic profiling* which identifies and quantitates a predefined number of metabolites in a restriction to a specific pathway, and to *metabolomic fingerprinting*, which can be understood as a rapid characterization of "the net result of metabolic regulation of a biological sample" (*208*).

Therefore, a true metabolomics approach must eliminate all possible biases against certain compound classes by chemical structure or by abundance from sample preparation to data acquisition and has to guarantee accurate metabolite quantitation, which comes along with great challenges (*204, 208*). To keep up with rigid quality standards, the Metabolomics Society evolved from the community. It hosts different scientific task groups to develop best-practice guidelines for key challenges such as metabolite identification, data standardization, and data analysis (*210*). Important landmark publications arose from the metabolomics standards initiative impacting the community by creating consensual minimum standards for the reporting of metabolomics experiments and other universal tasks (*211, 212*).

### 2.1.2 Metabolomics in special respect to food research

As recognized by Oliver Fiehn, the profile of small organic molecules represents "the ultimate and most immediate descriptor of a system in relation to its phenotype" (*208*). The exposome is an important aspect of the phenome, considering the sum of exogenous environmental influences. Notably, the human exposome is heavily impacted by individual nutrition and is directly linked to personal health (*213*).

Meeting the core of its original definition, the concept of metabolomics can be applied to a food item as food likewise is a biological system *per se*. In that sense, Alejandro Cifuentes defined in 2009 the term *foodomics* as the "application and integration of *-omics* technologies to improve consumer's well-being, health, and knowledge" (*214*). Therefore, metabolomics is notably only one analytical track included in foodomics next to an ensemble of other *-omics* technologies, see figure 2.1. Further, Cifuentes stresses the importance of bioinformatic approaches to integrate and evaluate cross-discipline foodomics datasets (*215*). Thus, chapter 2.4 will give special attention to computational aspects of contemporary food research.

From a traditional perspective, food and consumer safety was the primary interest of food analysis. In recent years, however, technological advances have promoted powerful applica-
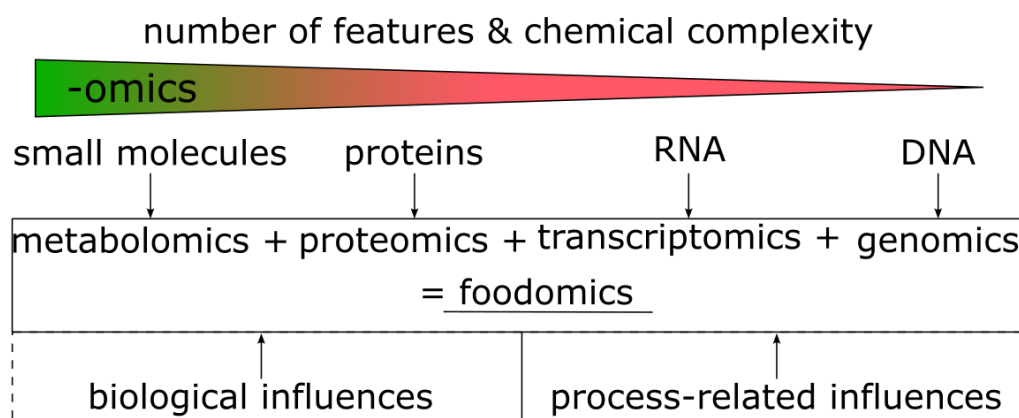
number of features & chemical complexity



-omics

small molecules     proteins          RNA          DNA

metabolomics + proteomics + transcriptomics + genomics
= foodomics

biological influences        process-related influences

**Figure 2.1** Relation and evolution of different *-omics* disciplines relevant to food sciences

tions in the fields of food traceability, quality control, health, and processing research. An extensive review by Valdez et al. gives an excellent overview of contemporary challenges and opportunities in analytical food chemistry (*216*). In special respect to cooking research, metabolomic studies can, for example, contribute to improved flavor development (*217*), to enforce food integrity and traceability (*218*), and can be applied to identify thermal processing markers (*219*), see also chapter 1.2.1.5.

Along with the evolution of foodomics, the term *foodome* was defined as "the pool of all compounds present in a food sample and/or in a biological system interacting with the investigated food at a given time" (*220*). The metabolomics branch of foodomics stands in the central focus of this thesis as a direct consequence of this definition. As the metabolome is regarded to be the most immediate link to the phenome of a living system, the metabolome of food accordingly represents an explanatory function to reflect the sum of influences a food item is subjected to. A set of intrinsic, biological properties of the raw food product is joined with external, chemical aspects triggered by human processing techniques like cooking, see figure 2.1. The true power of applied small molecule research lies in the ability to extract information related to both aspects from one methodological source. Most chemical reactions that have direct implications on the quality, nutritional value, and flavor of food affect compounds with a molecular weight below 1,500 Da. The introduction portrayed how the hydrolysis of macromolecules feeds exactly this reactive pool of organic compounds, which proceedingly flows to recursive and (self-)interacting thermal modification pathways. Therefore, metabolomic strategies are a well-suited tool for the unbiased investigation of cooking processes holistically, when special attention is given to heat-induced chemical transformations on a molecular level.

### 2.1.3 Current and critical progressions in metabolomics

Rapidly ongoing instrumental and computational developments constantly push the technological limits of food analysis. These advancement, however, require the community to keep pace with co-evolving challenges. Current estimations presume that only less than 2% of metabolomics data can be properly annotated (*221*). Even if computational tools thrive on reducing the proportion of the unknown, the share of "unilluminated dark metabolomic matter" is predominant (*221*). For this reason, current trends in the analytical community suggest abandoning stubborn feature collection practices in favor of thorough metabolite annotation. The quality of feature annotation and the comprehensive description of pathways is prioritized in favor of reporting unreliably identified features (*206*). It is highly beneficial for the resilience of computational models and conclusions drawn from experiments to restrict datasets to reliable metabolite identities (MetID) compared to the sheer collection of chemical noise. Certainly, reporting correct MetIDs is one of the prevailing challenges in metabolomics experiments at the moment. Most importantly, the annotation of a metabolite must not be mistaken for its identification. Mass spectrometry, in the first place, is a technology suited for the *characterization* of a compound (or sample). Additional compound information such as retention time, collisional cross-section, and other available knowledge must be integrated for unambiguous compound *identification* (*206*).

Additionally, computational strategies aim to constantly improve the quality of metabolite annotation by various knowledge-driven approaches. Network-based calculations, deep learning algorithms, and improved MS/MS deconvolution strategies can lead to a better understanding of metabolomics datasets. They will further be discussed in chapter 2.4.

Modern instrumentation offers further benefits such as data-independent acquisition (DIA): truly comprehensive and unbiased metabolomic datasets of former experiments can retrospectively be re-analyzed in the context of new-arising biological questions. Further, higher instrument sensitivity and greater dynamic ranges are urgently needed as xenobiotics often are four to five orders of magnitude lower concentrated in biological systems than endogenous metabolites (*222*).

By increasing specialization and particularization of experimental protocols, generalization and comparability of metabolomics studies are hindered. The metabolomics community responds to this challenge with an increased effort in standardizing reporting guidelines and harmonization of data repositories (*223*, *224*). In general, the opinion prevails that raw data from metabolomic experiments needs to be reported in a harmonized way to the public. The overarching goal of national biobanks and extensive cohort studies must necessarily be to contribute to a unified and integrated synthesis of all comprehensive metabolomic data for the greater good than a single, isolated study (*225*, *226*). This emphasizes the need to expand the usage of public repositories across the community. Additionally, neighboring *-omics* tech-

nologies' findings need to be better integrated into computationally demanding *multi-omics* approaches (*227*, *228*).

Transparent disposal of metabolomics data from food analysis can also lead to increased food authenticity in the future. Detection of food fraught and permitting better traceability of food in a progressively globalized food market is a core interest of foodomics (*31*). Blockchain technology recently found a growing interest in tracking and authenticating the food supply chain (*229*, *230*). Proportionate integration of metabolomics data to the blockchain of food items would be a promising perspective to prevent food fraud on a molecular level.

Altogether, metabolomics approaches in food analysis offer great possibilities for present analytical challenges already. With the co-evolution of computational methods, future perspectives for analyzing the food metabolome will greatly advance the understanding of our nutrition.

## 2.2 Instrumental analytical methods for the study of cooking emissions

Up to date, there is not a single analytical method to holistically describe the full metabolome of food or any other biological system. Therefore, a combination of several (ultra-) high-resolution analytical methods is best suited to estimate the ensemble of small organic compounds contained in food. A summary of MS-based technologies used in the experimental work of this thesis will be given in the following, elaborating on state-of-the-art instrumentation and discussing central aspects of data analysis.

### 2.2.1 Ambient plasma ionization techniques

Prior to mass spectrometric analysis, an analyte compound has to be ionized. Different technological solutions for the ionization process exist, which are selected by the operator as a function of physicochemical analyte properties and desired instrumental setup. The most popular ion sources in routine analysis are atmospheric pressure ionization (API) sources such as electrospray ionization (ESI), atmospheric pressure chemical ionization (APCI), and atmospheric pressure photo ionization (APPI). Especially ESI setups are commonly found in metabolomics studies as this technology offers highly efficient ionization for analytes of a wide range of polarity and molecular mass and generates intact pseudo-molecular ions with a low extent of in-source fragmentation (*231*).

The implementation of on-line MS analysis excludes the possibility of chromatographic separation. It requires an analyte stream to be continuously delivered to the MS. For this reason, most API solutions, which need to be supplied with a constant solvent flow either from a liquid chromatography (LC) system or from a syringe pump, cannot be used in on-line monitoring procedures. Therefore, a variety of alternative tools was developed, of which special attention is given to dielectric barrier discharge ionization (DBDI) as a plasma-based technique in the following paragraphs. An extensive review of other ionization technologies (used in real-time analysis) is given elsewhere (*231, 232*).

Plasma is a form of ordinary matter that is characterized by being composed of a mixture of free electrons, ions, and neutral species, which underlay constant interaction and are present in arbitrary proportions (*233*). Molecules and atoms can be ionized upon interaction with a plasma. Whereas high-temperature plasmas have relatively long been used for the ionization of metals in inductively coupled plasma MS, the application of room-temperature plasmas at ambient pressure is a relatively new field in MS (*233, 234*). Different plasma-based ambient ionization techniques are commercially available, among them is DBDI, a popular option (*235*): DBDI sources typically consist of two parallel electrodes separated by a dielectric material. Application of a high voltage, alternating current of high frequency creates a discharge forming a non-thermal atmospheric pressure plasma between the electrodes. The fundamental difference to thermal plasmas lies in the distribution of electron energies. In

DBDI, the electron distribution is characterized by a high population of low-energy electrons, whereas thermal plasmas possess a high fraction of electrons with high kinetic energy. Analytes introduced into a non-thermal plasma are subjected to a variety of ionization mechanisms, which are highly dependent on atmospheric conditions, analyte properties, and the energetic state of the plasma (*233*). These mechanisms are non-trivial and partially remain ununderstood, which represents a severe challenge in the application of DBDI in untargeted metabolomics experiments (*233*). Evidence has been gathered that ionization mechanisms include atom bombardment, electron ionization, electron transfer, and Penning ionization (*235*). Generally, DBDI, however, is a soft ionization technique generating primarily $[M + H]^+$ pseudo-molecular ions. Several factors make DBDI an advantageous technology for real-time MS. These include its ability to operate at ambient atmospheric conditions, minimal sample pre-treatment requirements, low sample consumption, and compatibility with a wide range of analyte types (*235–237*). Especially its native compatibility with a wide range of analyte polarities offers a high potential for the *in-situ* analysis of emissions directly from a cooking process without prior sample treatment. Limitations of the technology include plasma instabilities, high dependency of the ionization process on surrounding atmospheric gases such as $H_2O$ or $CO_2$, and uncertain side reactions during the ionization process for some analytes (*238*).

### 2.2.2 Ultra-high resolution Fourier transform - ion cyclotron - MS

Fourier transform - ion cyclotron (FT-ICR) - MS is a powerful profiling technique in metabolomics. FT-ICR is a mass analyzer that measures *m/z*-values based on an induced cyclotron motion. A schematic drawing of an FT-ICR mass spectrometer can be seen in figure 2.2. Positively or negatively charged ions are generated in the source of the MS **(2)**. The ion beam is guided through the ion path of the instrument **(3 - 8)**, and finally introduced on stable trajectories into an ICR-cell **(9)** which is located in a homogeneous high magnetic field $B$ **(10)**. An ion of mass $m$ and electric charge $q = n \cdot e$ with a velocity $v > 0$ is subject to a constant, inward force in the plane perpendicular to the magnetic field, called the Lorentz force $F_L$, as expressed in equation 2.1.

$$F_L = q \cdot v \cdot B \tag{2.1}$$

The angular velocity $\omega_c$, also called the cyclotron frequency, was defined by Lawrence as the equilibrium of the Lorentz force and the centrifugal force and is forcing an ion on a stable circular trajectory in the ICR-cell, see equation 2.2.

$$\omega_c = \frac{q \cdot B}{m} \quad f = \frac{\omega}{2 \cdot \pi} \tag{2.2}$$

The cyclotron motion is directly related to the mass $m$ and charge $q$ of an ion and linearly increases with the strength of the magnetic field $B$. Therefore, increased field strength also

increases the resolution power of the MS and enhances several other instrumental parameters (*239*).

The ICR cell hosts a pair of opposed excitation plates, which create an alternating electrical field orthogonal to the magnetic field. If the cyclotron motion of an ion has the same frequency as the electromagnetic field, resonance increases the ion's kinetic energy and raises it to a higher orbit in the ICR cell. A second orthogonal pair of detection plates registers the excited oscillation as the ions introduce an alternating current as a function of their cyclotron frequency on their trajectory. The signal intensity corresponds to the number of ions in the respective ion package. The resulting current, which is a superimposition of signals from all ions in the cell, is measured as a frequency *versus* time spectrum, called free-induced decay. By application of Fourier transformation, the final *m/z*-values of the ions can be calculated.

One of the most remarkable characteristics of FT-ICR-MS is the instrument's high mass resolution, which Alan Marshall defined as in equation 2.3 (*240*).

$$\frac{\omega_c}{d\omega_c} = -\frac{m}{dm} \tag{2.3}$$

The resolution *R* is defined as the full width of a peak at half-maximum peak height and is closely related to the accuracy of a mass measurement. The mass accuracy is reported in *ppm* or *ppb* and usually ranges below 200 *ppb* for routine measurements carried out on instruments with a field strength of $B = 12\,T$ (*240*).

$$\Delta m/z = \frac{measured\ mass - theoretic\ mass}{theoretic\ mass} \cdot 10^6 ppm \tag{2.4}$$

Mass accuracy is a crucial parameter because, during data evaluation, molecular formulae are computationally annotated to acquisitioned mass signals to initiate the process of chemical interpretation of the measured sample. Even the mass of an electron has to be taken into account during the annotation process, as at a molecular mass of 200 Da a mismatching electron configuration would yield an error of ±2.74 *ppm*. Every known molecule or molecular complex is composed of a linear combination of elements. In the scope of this thesis, compounds made of carbon, hydrogen, oxygen, nitrogen, and sulfur were studied. Based on numeric combinations of these five elements and a set of chemical rules (*241*), it is possible to putatively assign potential candidates of molecular formulae to detected mass signals. These annotations can further be refined by detected isotopic fine structures or by a network-based annotation that associates detected masses by exact mass differences of known chemical and biological building blocks (*242*).

FT-ICR-MS's highest potential unfolds when operated under direct-infusion (DI) conditions. DI-FT-ICR-MS allows rapid metabolic profiling in an automated and simplified process for high sample throughput (*243*). The sacrifice of chromatographic separation saves time and avoids discrimination of single compound classes by selection of a stationary col-
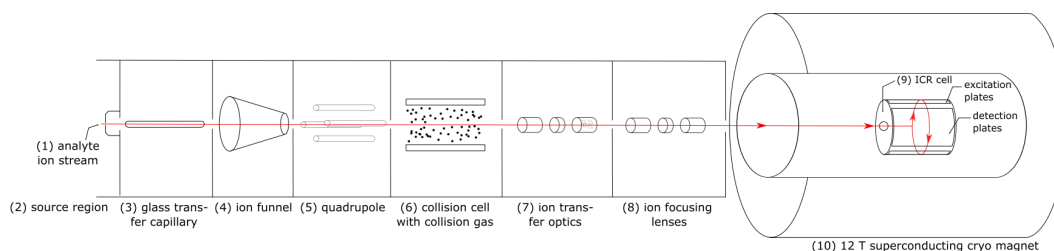
**Figure 2.2** Schematic drawing of the architecture of an FT-ICR-MS instrument: the analyte ion stream (1) is supplied by an ion source (2), e.g., by ESI or DBDI at atmospheric pressure. The ions are guided through the ion path (3 - 8) to reach the ICR detector cell (9) located within a 12 T superconducting cryo magnet. In the ICR, the ion packages are analyzed at a pressure of $10^{-11} mbar$.

umn phase. By contrast, no isomeric compounds can be distinguished by DI-MS. Clusters of analytes form in the ion sour or ion path, ion suppression effects can occur, and routine applications are restricted to MS1 profiling studies (*244*). Due to its universal nature, ESI is used in most FT-ICR-MS applications to detect a maximum of metabolites with wide ranges of polarity. A remarkable challenge in DI-MS is the formation of adducts and in-source fragments during the ionization process. Small cations and anions can attach to neutral molecules during the electrospray process to form non-covalent complexes. The mass of the intact complex is detected and cannot be distinguished from covalently bound molecules. In positive ion mode, $[M + NH_4]^+$, $[M + Na]^+$, and $[M + K]^+$ are formed to the disadvantage of $[M + H]^+$. In negative ion mode, the prevailing adduct species next to $[M - H]^-$ is $[M + Cl]^-$ (*231*). Adduct formation is an intensively studied field, and detailed knowledge has been collected about how steric and constitutional effects influence the ion affinity of neutral compounds. Exemplary reference to the attachment of chlorine to saccharides shall be given (*245*, *246*). In LC-MS applications, sophisticated data processing algorithms were established to account for the adduct formation problems. Algorithms such as CAMERA annotate isotope and adduct peaks and propose the correct compound mass (*247*). In DI-MS, the temporal dimension of the data is missing due to the lack of chromatographic separation; therefore, less confident approaches are applied to handle adducts. For example, chlorine-containing annotations can easily be recognized because of the characteristic isotopic ratio of $^{35}Cl$:$^{37}Cl$ and are converted *in silico* to the corresponding pseudo-molecular ion. However, one cannot be certain if the chlorine atom is covalently bound to the molecule or just associated in the form of a complex. Hence, differentiation between both species is not possible. The isotopic proof is not always possible as in the case of sodium which is a mononuclidic element and can lead to miss-annotations if overlooked. $[M + NH_4]^+$-adducts are even more difficult to handle as there is virtually no possibility to distinguish between two potential species $[C_cH_hN_nO_o + NH_4]^+$ and $[C_cH_{h+3}N_{n+1}O_o + H]^+$. Adduct peaks are often of high abundance and may dominate the spectrum. They can lead to space charge effects in the ICR-cell, which has adverse effects on spectral quality and mass accuracy (*248*). Therefore, reduction of adducts by sample preparation or adequate dilution is desired in FT-ICR-MS. Likewise, in-

source fragmentation processes occur during electrospraying. This phenomenon describes the neutral loss of small inorganic species, often $H_2O$, $CO$, $CO_2$ or $NH_3$ are lost prior to entering the mass analyzer. In-source fragmentation takes place when the energetic conditions in the source are not adapted to the analyte by various mechanisms that are reviewed elsewhere [231]. Especially unstable compounds are prone to lose the species mentioned earlier. Therefore, the operator of a mass spectrometer needs to make a proper selection of source parameters as a function of solvent and analyte properties. The loss of $H_2O$, for example, can lead to misinterpretation in carbohydrate research as many central mechanisms of CR and MR chemistry involve dehydration reactions, which can likewise be artificially induced in the source.

FT-ICR-MS characterization of small organic compounds gives valuable insights into the metabolomic profile of the foodome. The first metabolomic profiling study by FT-ICR-MS was remarkably carried out on food: Asaph Aharoni et al. investigated the ripening process of strawberries in 2002 and found diverging metabolic profiles between green and red strawberries [249]. In the past two decades, numerous other profiling studies have been performed on food. Among them was extensive research carried out about the decomposition of the molecular complexity of beer by FT-ICR-MS. Pieczonka et al. constructed molecular profiles of different beer types, elucidated the role of the MR in beer, made reflections about the German purity law, and characterized an ancient beer sample [28, 250–252]. The spectrum of applications solely related to brewing research emphasized the high content of chemical information that can be derived from ultra-high resolution MS profiling studies. Further, untargeted FT-ICR-MS studies contributed valuable insights about the MR of different amino acids and ribose [71, 253], was successfully applied in foodomics studies of wine [30] and whiskey [254], was used to study molecular processes during the fermentation of cocoa [255], and in a variety of targeted and untargeted studies on the complex metabolome of plants [256].

### 2.2.3 Hyphenated mass spectrometric techniques

Hyphenation refers to the practice where a mass spectrometer is coupled to a separation technique to serve as a detector unit. Separation techniques routinely employed in metabolomics research need to be fast enough for online coupling to scan over chromatographic peaks with satisfactory cycle time. This criterion is met by capillary electrophoresis, gas chromatography (GC), high-performance liquid chromatography (HPLC), and supercritical fluid chromatography in routine applications. As stated in chapter 2.2, single analytical methods are insufficient to access the entire spectrum of small molecules in complex systems such as food. Whereas DI-MS has been established for rapid metabolomic profiling studies, especially GC for the analysis of volatiles with its high separation efficiency, and HPLC are the most employed techniques in comprehensive metabolomics experiments [257].

During chromatographic analysis, an analyte compound is separated from a mixture based on their differential interactions with a stationary phase, which is contained in a column, and a mobile phase. The resolution $Rs$ of a chromatographic system, which is defined as the completeness of separation from one analyte to another, is defined by the capacity factor $k'$, the selectivity of the stationary phase $\alpha$, and the theoretical plate number $N$, as explained in equation 2.5.

$$Rs = \frac{1}{4} \cdot \frac{\alpha - 1}{\alpha} \cdot \sqrt{N} \cdot \frac{k'}{1 + k'} \tag{2.5}$$

The retention factor $k'$ of an analyte with retention time $t_R$, the system's void time $t_0$, $\alpha$ as the ratio of two adjacent peaks $i$ and $j$, $N$ as a function of retention time $t_R$, peak with $W$, and a constant factor $A$ together define fundamental retention relationships as seen in equation 2.6.

$$k' = \frac{t_R - t_0}{t_0} \qquad \alpha = \frac{k_J}{k_i} \qquad N = A \cdot \frac{t_R}{W} \tag{2.6}$$

The van Deemter equation relates the basic physical, kinetic, and thermodynamic effects of a separation process to evaluate the performance of a chromatographic system, see equation 2.7. This theorem includes the consideration of eddy diffusion $A$, axial molecular diffusion $B$, finiteness of transfer coefficient $C$, and mobile phase velocity $v$ (*258*).

$$H = A + \frac{B}{v} + C \cdot v \tag{2.7}$$

To achieve a maximum of chromatographic resolution power by meeting the optimum mobile phase velocity as a given function of a column's plate height $H$, which is defined as the fraction of column length $L$ and theoretical plate number $N$, a chromatographic separation process can be optimized.

$$H = \frac{L}{N} \tag{2.8}$$

The significance of optimum chromatographic resolution for in-depth exploration of food samples is reasoned as higher peak capacities allow better feature detection, enable higher MS sensitivity, and superior MS spectral quality (*257*). For that reason, HPLC has successively been replaced by ultra-high pressure liquid chromatography (UHPLC) instrumentation in small molecule research during the last few years. Technological advancements supply analytical columns with constantly smaller particle sizes of sub 2 *μm* particles, resulting in increased separation power as measured by the number of theoretical plates. Additionally, alternative stationary phase designs like capillary or monolithic LC columns and fully porous particles mature into routine analysis (*259*). These developments in higher separation capacity are accompanied by shorter analysis time, which is beneficial for high-throughput applications (*257*).

The surface of the stationary phase carrier, which is traditionally made of porous silica particles, is covalently modified to offer specific interaction sites for analyte retention. The stationary phase bonding determines the interactive effects with an analyte compound during the separation process and its retention on the column. Therefore, column materials must be adapted to the polarity of respective analytes, which can be estimated based on the analytes $log_P$-value. Myriads of different column materials are commercially available, relying on various retention mechanisms such as analyte polarity, charge state, or molecular size. However, the selection of a bond phase implies an *à priori* introduction of a bias in the analytical process. Therefore, most metabolomics studies thrive for comprehensiveness by performing a dualism of reversed phase chromatography (RP) for the separation of non- to semipolar analytes and hydrophilic interaction chromatography (HILIC) for the analysis of water-soluble compounds, resulting in a broad metabolic coverage. Historically, RP is the gold standard of LC and yields the best results in terms of peak shape, reproducibility, and robustness (*260*). RP C8 and C18 bond phases prevail in metabolomics experiments (*261*). HILIC separation relies on complimentary retention mechanisms for the separation of polar compounds, which are eluted in the void volume of RP columns. Even if HILIC phases are relatively young compared to RP, various bond materials specifically tailored to retain different compound classes exist and have become more robust over the last decade (*262*). Therefore, a combination of two orthogonal separation principals represents a significant increase in peak capacity. Analysis of a sample can be performed on RP and HILIC by two successive injections or by 2D-LC (*263*). Certainly, the separation of isomeric and isobaric compounds adds a valuable dimension for the characterization of complex systems (*205*). Further, retention behavior contains physicochemical information about the retained compounds. Using retention time standards like homologous series of hydrocarbons in RP-LC can contribute further knowledge about an analyte (*264*).

Several mass analyzers are routinely employed in LC-MS hyphenation; however, in the field of untargeted metabolomics, quadrupole time of flight (Q$q$TOF) and orbitrap systems prevail (*265*, *266*). The first-mentioned is shortly presented and schematically drawn in figure 2.3.

Inside a Q$q$TOF instrument, ions are usually generated by an API source **(2)**. Analyte ions are guided through the ion optics **(3-8)** and finally pushed by the electric field of an orthogonal accelerator **(9)** into an evacuated flight tube **(10)**. The reflectron **(11)** focuses ions of the same mass with slightly different kinetic energies and returns the ion beam into a detector, e.g., a multichannel plate **(12)**. The linear velocity (v) of an ion depends on its mass-to-charge ratio at a given length (L) of the flight path, the pulse potential (V), and the elementary charge e. The final relationship is described by equation 2.9.

$$TOF = \frac{L}{v} = L \cdot \sqrt{\frac{m}{2zeV}} \tag{2.9}$$

In practice, Q*q*TOF instruments can operate at high scan rates of up to 100 Hz and a resolution of up to 70,000 (*265*). According to equation 2.9, higher mass resolution is obtained for higher masses. The stability over time in modern instruments is dramatically improved by, e.g., on-line reference mass delivery or a heated flight tube. As a result, extensive chromatographic batches can be measured without significant drifts in mass accuracy over time.
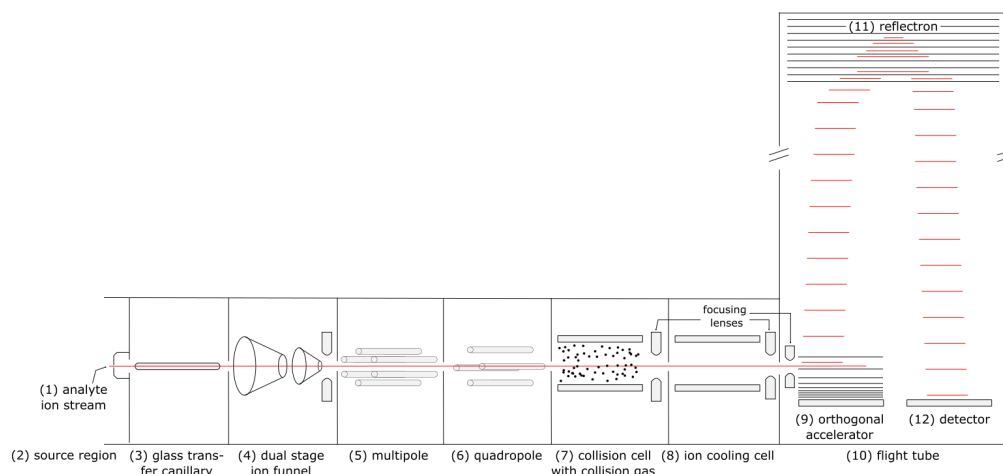
**Figure 2.3** Schematic drawing of the architecture of a Q*q*TOF-MS instrument: the analyte ion stream (1) is supplied by an ion source (2), e.g. by ESI at atmospheric pressure. The ions are guided through the ion path (3 - 8) to the flight tube (10). An orthogonal accelerator pushes the ions towards a reflection (11), which focuses ions of the same species with different kinetic energy and directs them toward the detector (12). In data-dependent acquisition mode, an ion species of a defined mass is isolated by the quadrupole (6) and successively fragmented to its product ions by collision-induced decay in the collision cell (7). The resulting fragment ions are analyzed and detected by a full MS/MS scan.

Compared to FT-ICR-MS, the mass accuracy of Q*q*TOF instruments is relatively low, which leads to poorer acquisition of isotopic fine structures of analytes, resulting in less confident metabolite annotation. However, chromatographic separation offers, compared to DI-MS, superior reproducibility rates, reduced ion suppression and matrix effects, and facilitates the correct detection and attribution of isobaric and isomeric compounds and ion adducts (*266*). Narrow peak intervals in UHPLC require high scan rates of the detector system. Otherwise, insufficient data points are acquired. Due to their relatively slow acquisition times, FT-ICR-MS instruments are not well-suited for LC-MS coupling.

A further advantage of TOF instruments is that they are generally available as hybrid mass spectrometers, where a second mass analyzer, most often a quadrupole **(6)**, permits the performance of $MS/MS$ or $MS^E$ experiments. In data-dependant acquisition (DDA) mode, the *n* most abundant peaks of a full-scan precursor ion spectrum are successively isolated in the first mass analyzer and fragmented in a collision cell **(7)**. Resulting fragment ions are then detected in the second MS dimension and serve to characterize or identify the metabolite. In

modern mass spectrometers, DIA is an alternative to DDA, where all analytes in a given mass range are fragmented, and resulting fragment ions are assigned to the associated precursor ions by computational approaches (*267*). The characterization of unknowns by acquisition of tandem mass spectra has direct implications on the understanding of a system's metabolome. Spectral quality and comprehensive metabolite fragmentation are important in terms of reliable annotation of small molecules, which together allow the integrative discussion of a studied sample in a biological or chemical context (*267*, *268*). The metabolomics community has created various computational tools for curating, treating, and interpreting mass spectrometric datasets, which will further be discussed in chapter 2.4.

In metabolomics research, capillary electrophoresis, related chromatographic techniques, and other approaches are equally successfully applied as hyphenated separation techniques. At this place, the concept of ion mobility spectrometry (IMS) shall be highlighted for better deconvolution of complex samples. IMS as a post-ionization separation technique in combination with mass spectrometric detection can serve as a fourth data dimension (next to retention time, accurate precursor mass, and tandem mass spectra) in untargeted metabolomics (*257*). IMS separates gas-phase ions based on their mobility in a carrier gas under the influence of an electric field. Analyte ions are introduced into an ion mobility spectrometer, where they enter a drift tube filled with a neutral gas, often nitrogen or helium. As the ions traverse through the drift tube, they experience collisions with the neutral gas molecules. The ions' mobility, which is their speed of migration under the influence of the electric field, depends on their size, shape, charge, and interactions with the gas molecules. Larger and more complex ions tend to have slower mobility, while smaller and simpler ions are dragged faster through the drift tube. Different isomers travel through the drift tube at disparate speeds due to their characteristic collisional cross section (CCS). Feature detection is enhanced by removing spectral overlaps and matching acquired CCS values with public databases, giving further confidence in metabolite identification (*257*, *269*, *270*).

The application of hyphenated technologies in food analysis allows deeper insights into the metabolic signatures of food. Increasing peak capacities, higher sensitivity due to lower matrix effects, and separation of isomeric compounds dramatically improve the informative content that can be extracted from complex samples. By chromatographic separation, the comprehensive acquisition of high-quality MS/MS fragmentation spectra becomes feasible, which allows the characterization of samples and identification of key metabolites beyond the possibilities offered by MS1 spectra. The combination of RP and HILIC protocols compromises a fundamentally longer analysis time compared to DI-MS analysis. However, attention is paid to polar and non-polar compounds, giving access to the entire metabolome with respect to Fiehn's definition.

Current trends are especially focusing on miniaturization techniques that enable high throughput of samples and allow shorter analysis times. Enhanced capillary LC and nano-LC

columns in combination with nano-ESI interfaces allow $\mu$L to nL flow rates and are routinely applied in proteomics, dramatically boosting the sensitivity of LC-MS experiments (*271*). The transformation of nano-LC into food analysis is on its onset, equal to the integration of "lab on a chip" solutions that are designed for highly automated, high throughput miniaturized measurements (*259*). Future directions are estimated to further decrease analysis time to accelerate the analysis of large cohort studies, to grant near-real-time data access on the chromatographic side, and to further improve spectral quality and sensitivity on the mass spectrometric side.

Notably, modern computational approaches are increasingly applied in evaluating undirected metabolomic experiments to handle extensive data sets from UHPLC-MS/MS analysis. Li et al. recently reviewed the usage of molecular networking approaches to evaluate the function, safety, and quality of foods (*272*). They concluded that molecular networking is an emerging field in food metabolomics that offers perspectives for the characterization and identification of novel metabolites, biomarkers, and quality markers but needs further establishment in the field (*272*). Advanced instrumentation with increasing resolution and scan rates continuously increases the data size of metabolomics datasets. Whereas the application of IMS-MS in foodomics indisputably offers increased sensitivity and selectivity, acquired drift maps are large in size and require high computational processing capacities and expertise in the analysis of obtained drift data (*273*). Developing advanced algorithms for data analysis is a central research point to take full advantage of data supplied by modern hyphenated analytical setups.

## 2.3 Methodology of cooking research interfaced by the head space

A central point in the concept of metabolomics prerequisites completely unbiased experiments. One problem in the analysis of food is that, on the one hand, compounds of substantially different polarities and molecular sizes are present. On the other hand, excess matrix constituents decrease extraction yields for trace analytes. The concentration range of small molecules in food is widespread and can range over several orders of magnitude. These challenges can impede the unbiased analysis of a sample. Therefore, nondiscriminatory liquid extraction of food is difficult to achieve and, if at all, can only be performed by cumbersome fractional extraction techniques using solvents of multiple polarities. This process is labor-intensive and time- and resource-consuming. By contrast, head space sampling permits non-discriminating, direct access to small organic compounds from food matrices of various polarities and concentration ranges without oversaturation of extraction media.

### 2.3.1 Holistic head space sampling approaches

For further methodological conceptualization, shall the head space be defined as the empty space above processed food filled with air. In the ideal case for collecting samples from cooking processes, the head space is a closed or alternatively isolated system. As closed systems are barely implementable in real-live cooking processes, interference and contamination of cooking fumes by external influences must be controlled by collecting and measuring surrounding atmospheric control samples.

Transport of molecules to the head space from food items during the cooking processes can be attributed to different mechanisms, as later discussed in chapter 5.3. The sum of mass transport leads to the presence of vapor, volatiles, aerosols, and particulate matter in the head space. The following chapter will introduce existing methodological approaches to sample fractions of these emissions from cooking head spaces. A methodological differentiation will be drawn between direct, on-line measurement, where processing emissions are transferred from the cooking site directly to a mass spectrometer through a sampling interface, and at-line sampling, where sample collection is performed during a cooking process, and respective samples are manually delivered to a mass spectrometer.

### 2.3.2 On-line analysis of food processing emissions

As discussed in chapter 1.3, most work on processing emissions is carried out on the analysis of small VOC in the scope of (indoor) air quality analysis and on volatile aroma-active compounds in the context of flavor analysis.

In terms of process analysis, on-line measurement is the most direct and immediate option to monitor cooking processes. On-line measurements permit real-time process control and

supply continuous data streams. As there is no need for offline sample transfer, analyte loss can be minimized. The bottleneck in real-time MS, however, is often located at the interface between the cooking site and the mass spectrometer as the loss-free delivery of an analyte stream is challenging. High atmospheric water loads arising from the processed food can dilute analytes in the head space and cause suppressive problems during the ionization process. In principle, there are three techniques that are widely used in on-line MS and consequently suited for the monitoring of cooking processes: APCI-MS, proton transfer reaction (PTR) MS, and selective ion flow tube (SIFT) MS (274). APCI, PTR, and SIFT all together utilize a soft ionization mechanism to create $[M + H]^+$ pseudo-molecular ions. All three technological platforms come without prior chromatographic separation of analytes and are, therefore, DI-MS methods. Even though ultra-fast GC, (275) and LC (276) methods with few seconds of separation time exist, coupling of chromatographic separation techniques almost always excludes the possibility of real-time data acquisition. The need for complex injector modules compared to simple transfer lines presents a further delay in the analytical process so that detector signals lag behind the actual process.

PTR-MS is usually employed together with a TOF mass analyzer, giving access to high scan rates, which is commonly regarded as an important parameter in real-time mass spectrometry (277). TOF-MS counts to high-resolution techniques, which permits identifying isobaric compounds (274). PTR-TOF-MS is the most popular platform used for on-line monitoring of cooking processes. It was successfully applied in real-time monitoring studies to assess the loss of volatile aroma compounds to oven head spaces (170), where numerous volatiles were tracked in a time-resolved way. Food processing-related applications of PTR-TOF-MS are diverse and include the measurements of emissions from the roasting of coffee (278, 279), evaluating volatiles from the deep-frying of tubers (280) or quality assessment of processed milk (281).

Although SIFT-MS instruments are currently only commercially available with quadrupole mass analyzers, which provide unit mass resolution, their strength lies in detecting trace gases and in a quantitative measurement principle if the analyte is known. Multiple reagent ions are successively reacted with an analyte so that conclusions about the analyte concentration can be drawn from compound-specific reaction constants (282). Due to long cycle times in full scan mode and low mass resolution, the application of SIFT-MS in untargeted analysis comes with notable drawbacks. SIFT-MS is routinely used for the detection of trace gases (282) and contributed to a facilitated adjustment of the roasting conditions of cocoa for chocolate production by detecting, for example, pyrazines which positively contribute to the pleasant aroma of cocoa (283).

APCI-MS found its early applications in flavor research (274). Unlike PTR-MS and SIFT-MS, which generate intermediate primary reagent ions for the protonation of analyte molecules, APCI-MS directly ionizes the analyte in a corona discharge. Commercial APCI sources can

be installed on mass spectrometers of many manufacturers and present a valuable ionization tool with higher efficiency towards hydrophobic analytes compared to ESI (*284*). By its universal availability, APCI gives access to mass spectrometers with high resolution and tandem mass analyzers for the recording of MS/MS spectra. APCI-MS finds most of its applications in environmental and flavor or breath analysis (*232*, *274*) but it was also successfully applied for the real-time authentication of honey (*285*).

### 2.3.3 At-line analysis of food processing

Usually, the benefits of immediate on-line measurements are sacrificed to access more powerful analytical instrumentation. At-line sampling is often the better-suited strategy to apply special ionization techniques and chromatographic separation to access instruments with higher mass resolution or sensitivity. It can include a brief sample preparation step. To transfer the analyte from the cooking site to the mass spectrometer, a trapping technique is usually employed instead of a sample transfer line. At-line sampling is also useful if a sample of a cooking process shall be analyzed by multiple techniques or conserved for a later purpose. The sample workup can include a concentration step, dilution, solvent exchange, or addition of an internal standard.

The earliest example of a trapping technique in process vapor analysis was the setup described by Ernest Pence to study emissions from bread-baking processes (*176*). The setup was based on the simple physical principle of condensing vapor by actively pumping it through a frigorific mixture. The resulting fluid was collected in a cold trap and analyzed later. Years after, Barbara Rega et al. similarly used a membrane pump to "continuously carry [baking vapors] through a glass inlet hood to a refrigerated extraction chamber" (*130*). In the extraction chamber, the vapors passed by a solid phase microextraction (SPME) fiber, which absorbed volatile compounds of interest, which were later analyzed by thermodesorption GC-MS (*130*). The condensate itself was discharged, however. Whilst Pence applied different wet chemical analysis methods, Rega et al. followed a completely different analytical approach by using a thermodesorption GC-MS setup to study volatiles from the baking process. This demonstrates the high flexibility of cold-trapping approaches.

Various other SPME approaches are described to directly sample volatiles from the head space of food cooking systems. Despite SPME generally suffers from low reproducibility and the selection of a fiber, respectively its sorbent properties introduces a first bias in terms of analyte polarity to the analysis. SPME is a popular tool to analyze molecules with relatively low molecular masses (*286*, *287*). SPME sorbents are commercially available for various applications and target different compound classes. However, they are generally unsuited to bond non-volatiles, making them an ideal tool for separating analytes from bulk matrix components. Absorbed volatiles are usually desorbed from the fiber by thermodesorption in an injector module of a GC. Notable contributions to the understanding of odorants from bread

baking processes (*171, 288*) and to the roasting of pine nuts (*289*) were accomplished by SPME setups.

Tenax tubes are sorbent-filled cartridges that are often used for the analysis of volatiles. Processing emissions are pumped through the tube and later desorbed for chromatographic analysis. Likewise to SPME, the polarity of the sorbent determines which analytes are retained. The major advantage of Tenax is that samples can be taken at any place with low instrumental effort. A simple pump suffices to draw air or processing vapor through a sample tube, which is sealed on-site and analyzed in an instrumented laboratory. For this reason, Tenax tubes are popular in environmental studies (*179, 290*). Tenax tubes were, e.g., used to measure odor emission sources from exhaust air outlets of food production plants (*291*) or to characterize airborne aldehydes from oil-frying (*292*).

Further methodologies for the analysis of volatiles from food exist. However, their application is often not established in process analysis or limited to particular cases (*293*). In summary, at-line sampling allows great instrumental possibilities for the characterization of food processing emissions. However, great attention has to be paid to not introduce analytical bias in the spirit of a truly untargeted metabolomics study. The final choice between on-line and at-line sampling has to be made on a case-dependent basis in consideration of the respective analytical question.

## 2.4 Bioinformatic approaches in untargeted metabolomics and food research

Applying hyphenated techniques in metabolomics research results in a constant generation of big, untargeted datasets. As Cifuentes defined foodomics as an interdisciplinary and integrative framework, bioinformatics methods are needed to process and assemble data from various sources to converge an uniform picture of an analyzed matrix (*216*). The processing of this data presents a substantial challenge. The community develops computational methods to solve these objections, which focus on three central topics: a) post-processing, removal of noise, tidying the data, and extraction of features, b) identification of analytes of interest followed by their correct annotation, c) streamlining raw data to uni- and multivariate statistics, application of machine learning (ML) and artificial intelligence (AI) algorithms, creating interfaces to relevant reporting systems. Detailed abstracts of these topics will be given in the following chapter, with special respect to cooking research.

### 2.4.1 Data handling: post-acquisition processing and curation of metabolomics data

The extraction of human-interpretable information is the main goal during the work-up of mass spectrometric data. An array of commercial and open-source software solutions exists for the processing of mass spectrometric datasets, which ultimately all create data matrices of samples, chromatographic features, and the matching abundance expressed as peak height or peak area.

Elementary processing steps which are implemented in nearly all processing software suits include the following: after centroidation of profile mode data from high-resolution instruments, raw data is imported, and mass detection is initially performed, where each detected signal under a user-defined threshold is discharged as noise. Detected mass signals are subjected to feature detection algorithms, where a feature is characterized by a retention time and *m/z*-value. Therefore, extracted ion chromatograms (XIC) are built along the temporal axis for each detected mass. Following, co-eluting and overlapping chromatography peaks are resolved by local minimum detection. Resulting feature tables are successively refined by isotopic pattern and adduct detection (*247*). Afterward, feature alignment algorithms align corresponding features across all chromatographic runs and correct possible retention time shifts. Ultimately, gap filling helps to avoid missing values in the data set by picking original intensity values from raw data (*294*).

The final MS1 data matrix is suitable for further analysis by uni- and multivariate statistics to train models or can be submitted to pre-defined data evaluation pipelines. Next to the processing of MS1 spectra, fragmentation data from MS/MS or $MS^E$ experiments is simultaneously picked, cleaned, consolidated, and assigned to an MS1 precursor feature. MS/MS data can be used for the characterization of the sample and for improved metabolite iden-

tification (*294*). In the case of DI-MS, several processing steps like adduct detection or consolidation of tandem mass spectra are hindered due to missing chromatographic resolution, revealing the need for tailored processing solutions.

This general workflow can be extended in dependence on the applied instrumental setup, data quality, and analytical question. Additional separation techniques like IMS raise the need for an additional 2D feature detection of drift data, which needs to be merged with MS1 feature tables. Optional detector systems like photo diode array or fluorescence detectors might supply additional data which need to be consolidated.

Fundamental challenges during data processing demand the selection of adequate processing parameters by the operator to correctly define noise levels, refine features, and to correctly align features across samples. Good practice during planning of the measurement batch (*295*) includes the analysis of (known) quality control samples and of standards mixes to assure optimal instrumental performance (*296*) and correct processing parameters (*297*, *298*).

### 2.4.2 Metabolite identity: annotation of metabolites

Precise assignment of exact chemical identities to chromatographic features is an important bottleneck and crucial for correctly interpreting untargeted metabolomics studies (*299*). To assure community-wide working standards, the Metabolomics Standard Initiative around Lloyd W. Sumner proposed minimum reporting standards for metabolomics experiments to harmonize the reporting of experimental protocols, quality control, and metabolite annotation (*212*). Sumner most importantly drew a line between reliably *identified* compounds (by an authentic reference material or confirmation by NMR experiments) and putatively *annotated* compounds (e.g., by fragmentation data from a public repository) (*212*). As standard material is expensive, often unavailable, or needs to be synthesized, many researchers cope with the lack of available standards with two alternative strategies for metabolite annotation. On the one hand, acquired spectra can be matched against fragmentation spectra from public repositories. On the other hand, advanced computational tools for the *in-silico* prediction of compound IDs are becoming increasingly popular. However, one must remember that the term *annotation* is defined as "a note of explanation or comment" and, therefore, must not be mistaken for an experimental proof of a compound's identity.

Multiple public data repositories have been built over the last few years. The most popular ones include the human metabolome database (*300*), the global natural products social molecular networking (GNPS) repository (*301*), the mass bank of North America (*302*) or various commercial libraries, e.g., from the national institute of standards and technology. Recorded spectra can be matched to the library by many spectral comparison methods. Traditionally, mass fragmentation spectra are transformed into vectors and mathematically compared, e.g., by cosine scoring. Modifications of the scoring method can account for neutral losses, shared fragment ions, consider the entropy of two spectra, and can improve matching results case-dependently (*303*). Other sophisticated similarity metrics are based on transformative

comparisons and machine learning algorithms (*303*). The choice of similarity metric should depend on the analyte class and the content of the selected database and must be made with great caution. Even if an annotation is incorrect, it might propose valid hints for common structural motives. Additionally, it should also be noted that spectral similarity does not necessarily go hand in hand with structural similarity, which can be determined by means of the Tanimoto score (*304*). False-positive annotations might prevail and should be checked for plausibility with great care. Additional secureness to annotations can be given by comparison of retention times and published CCS values.

Despite all combined efforts, only about 1.8% of spectra in an untargeted metabolomics experiment can be annotated by reference spectra (*221*). To overcome this "dark matter of metabolomics" and further increase the content of information that can be extracted from metabolomics experiments, several computational solutions are developed across the community to obtain information about spectra that are not contained in public repositories (*221*).

First to mention are network-based approaches, of which GNPS is the most popular (*301*). Molecular networks can be understood as graphs, where each MS/MS-spectrum represents a node and edges represent spectrum-to-spectrum alignments like similarity measures or mass difference building blocks. Sets of spectra from compounds of the same molecular family group together and can form sub-networks (*301*). Seeds of structural knowledge can be introduced into the population of unknowns by library comparison, helping in the search for analogous compounds within a sub-population of MS/MS spectra. Different modifications of the GNPS networking like molecular network enhancing (*305*) exist and are only one facet of a family of algorithms thriving to improve informative relationships in networks (*306*).

Besides networking approaches, deep learning methods based on database similarity searching are among the most reliable strategies for the illumination of dark metabolomic matter (*303*). One of the most popular tools in the field is the compound structure identification (CSI) FingerID tool developed by Dührkop et al. (*307*). Based on tandem MS spectra, fragmentation trees are computed to explain the fragmentation spectrum of an unknown molecule. The algorithm tries to connect the different fragments and compares the predicted structures with pre-learned molecular fingerprints. Successively, a support vector machine classifier is used to separate proposed molecular fingerprints, which finally will be scored (*307*). Accompanying functions in the software suite propose compound classes by a pre-trained model (*308*) or offer network-based refinements of the predictions (*309*). Related software solutions like ClassyFire developed by Wishart et al. provide hierarchical classification of chemical entities (*310*), or Spec2Vec developed by Huber et al. leans on the mature machine learning discipline of natural language processing and is based on spectral embeddings learned from the fragmental relationships within a large set of spectral data (*311*).

Huge drawbacks of machine learning models come with the limited availability and goodness of training data. As roughly 98% of metabolomics data remains uncharacterized up to date, wide-spanning chemical spaces are not covered in the public libraries used to train the

models. This leads towards a strong bias of a model's output in regard to unexplored molecular entities (*303*) and stresses the need for every scientist to contribute reliably identified compounds to public repositories.

In conclusion, the transformation of spectral data into structural knowledge remains a challenge and a fundamental bottleneck in the interpretation of untargeted data. To further promote open science and report high-quality metabolomics data to the public, communal efforts in data sharing and maintaining of data infrastructures as, e.g., proposed by Janna Neumann in the FAIR data principles need to be made (*312*). Nevertheless, *in-silico* tools are becoming more sophisticated. Mass spectrometry is a technology for characterizing compounds *per se* and needs support by retention times and CCS data for secure identification. Supporting experimental coverage can, for example, be made by the usage of stable-isotope labeling approaches or NMR for further structural elucidations (*313, 314*).

### 2.4.3 Applied data sciences in food research

In 2001, Frank Desiere et al. postulated that continued progression in bioinformatics will have a profound impact on many aspects of modern food and nutritional sciences in the future. The proposed impacts range from the development of new crops, food production, and processing under the background of personal health to improving food quality and safety (*315*). Now, more than 20 years later, analytical chemistry has become a data-rich discipline due to advancing *-omics* technologies. This section will outline how ML and AI algorithms are successfully applied in food and nutrition sciences.

The prevailing application of ML and AI in food sciences lies in the quality assessment of food. Currently, there is an economic transformation going on termed "the fourth industrial revolution", or "industry 4.0", which is driven by growing digitalization and automation (*316*). The associated branch of "food quality 4.0" is enabled by big data-providing methods like the ones from the *-omics*-family, non-destructive fingerprinting techniques and bioinformatic tools (*317*). Accurate and fault-free data is indispensable for training computational models, stressing the fundamental significance of meticulous data-reporting pipelines for metabolomics data as described in chapter 2.4.1. Industry 4.0 technologies include smart sensing solutions, internet of things technologies, and real-time process control solutions. Corresponding data pipelines are constructed to extract information from food analysis to achieve rapid, reliable, and objective assessment of food quality (*318*). Especially deep and reinforced learning algorithms are well-suited for feature selection processes. These molecular targets can be applied to probe quality dynamics, predict food safety (*319*), and monitor shelf lives (*320*). Process parameters can be adapted case-dependently, and real-time data pipelines offer a strategy to respond to the fluctuating quality of raw goods (*320*). So far, the identified target features were often provided by genomics (*321*) and proteomics (*322*) approaches. A reason for this seems to be that both technologies are longer established than,

e.g., metabolomics or lipidomics and have matured into industrial applications. The goal of emerging processing 4.0 strategies is to increase overall sustainability and efficiency of food production by simultaneously decreasing production costs by dynamic responses to the natural variability of living organisms (*315*).

Further applications of ML and AI lie in food authenticity control, where classification algorithms accomplish to detect subtle adulteration of a food matrix (*323*, *324*). Besides processing-related applications, bioinformatic approaches likewise provide knowledge-based solutions to food sciences: integrated results from foodomics studies can contribute to the development of food personalized for consumers' health and individual preference, safer products, and a global understanding of nutrition on a molecular level (*215*, *315*). The supply of metabolomic data thereby contributes to an unbiased, molecular description of thermal food treatment. Especially under the increasing consideration of information gained from the analysis of small molecules - as they represent the most immediate indicator of the exposome - in combination with bioinformatic advances, food production is expected to perform a digitalized transformation in the near future (*315*).

# 3  Motivation and aim of the thesis

Food consumption is an imperative part of every human's life. The way, our daily food supply is processed -on an individual or industrial level- is highly determining for the quality, flavor, and nutritional value of the final meal. It is becoming increasingly clear that individual health is strongly dependent on personal nutrition and that there is a close connection between dietary effects and the conditions a meal was prepared under (*55*).

The preparation of food often involves application of thermal energy in the frame of a cooking process. In an evolutionary context, cooking increases the nutritional value of food, deactivates harmful substances and microorganisms and prolongs shelf life. Most importantly however, a cooking process transforms dull raw material into highly palatable meals with pleasant olfactory properties. This transformation of raw biological material can be attributed to an interplay of physical, biological and especially chemical processes.

Whereas analytical chemists started to study the molecular composition of food in the early $19^{th}$ century (*325*), mechanistical reflections about browning reactions, flavor generation, and physicochemical changes were prospering from the early $20^{th}$ century on, when also Louis Camille Maillard published his pivotal work about non-enzymatic browning reactions (*57*). From then on, multiple aspects of cooking processes were analyzed. Advancing analytical instrumentation soon permitted to make reflection on transformative processes of most substance classes commonly found in food.

Although, major chemical knowledge about raw food, which has its origin as a biological system composed of plant material or animal-based origin, was collected, the share of small organic compounds modified as a result of controlled thermal exposure remained relatively unstudied. Finding a method that achieves to depict the entity of chemical modifications happening during cooking processes is near to be an impossible task due to the heterogeneity and complex composition of food. Nevertheless, there is a lack of holistic analytical approaches that study thermal food transformation *as a system*. The metabolome of food unites extensive chemical information about its biological origin and imposed processing steps. Although during the last years numerous studies have proven that these small molecules are emitted to the head space of cooking sites and comprehend a plurality of food- and process-related information (as closer elaborated in chapter 1.3), no comprehensive analytical method has been developed yet to fully exploit the potential of the head space metabolome from food processing.

It is of high importance to develop a comprehensive understanding of chemical transformations occurring during cooking processes to guarantee the production of healthy and nutritious food of high quality. In account of the pressing need for a resource-saving, sustainable and health-promoting food supply chain, metabolomic analysis of food processing emissions provides a green chemistry approach for drawing direct conclusions about reactive processes during cooking and controlling food processing in real time.

This thesis aims to contribute to a better understanding of the molecular spaces contained in food processing emissions, to acquire knowledge about how we can recognize and control thermal processing *via* the head space as an interface to the underlying matrix, and how the analysis of the head space metabolome can contribute to healthier, more palatable and sustainable food production.

A graphical outline of this thesis is given in figure 3.1. In **Chapter 4.1** (Weidner et al. Food Chemistry 2022, 374, 131618), we present an innovative and holistic strategy to access molecular information contained in processing emissions from wheat bread rolls sampled from the head space of a convective oven. We showed that even complex chemical reactions such as the formation of AGE can be retrieved and validly modelled through the head space.

This prove-of-concept was refined in a second study presented in **Chapter 4.2** (Weidner et al. Journal of Agricultural and Food Chemistry 2023, 71, 45) by applying our approach to a selection of 96 different food items. We defined food-type characteristic, molecular spaces, investigated thermal processing indicators and extracted structural information from oven vapor metabolites based on mixed Maillard model systems.

In **Chapter 4.3** (Weidner et al. Analytical Chemistry 2023, 95, 2), we then introduced a miniaturized, laboratory-scaled, highly controllable setup for the real-time measurement of thermal food processing emissions by FT-ICR-MS. Knowledge gained from the at-line protocols presented in chapters 4.1 and 4.2 was converted to an on-line experimental setup which permits to follow thermal food processing in a time resolved way by advanced mass spectrometric techniques.

Further, we developed the software package DBDIpy written in Python which is closely presented in **Chapter 4.4** (Weidner et al. Bioinformatics 2023, 39, 2). The relatively young technology of dielectric barrier discharge ionization, which was used as the principal ionization setup in the study presented in chapter 4.3, is known to introduce spectral artifacts which need be cleaned after data acquisition. DBDIpy offers open source solutions to the community for the interpretation and curation of untargeted metabolomics datasets from time-resolved ambient plasma ionization techniques.
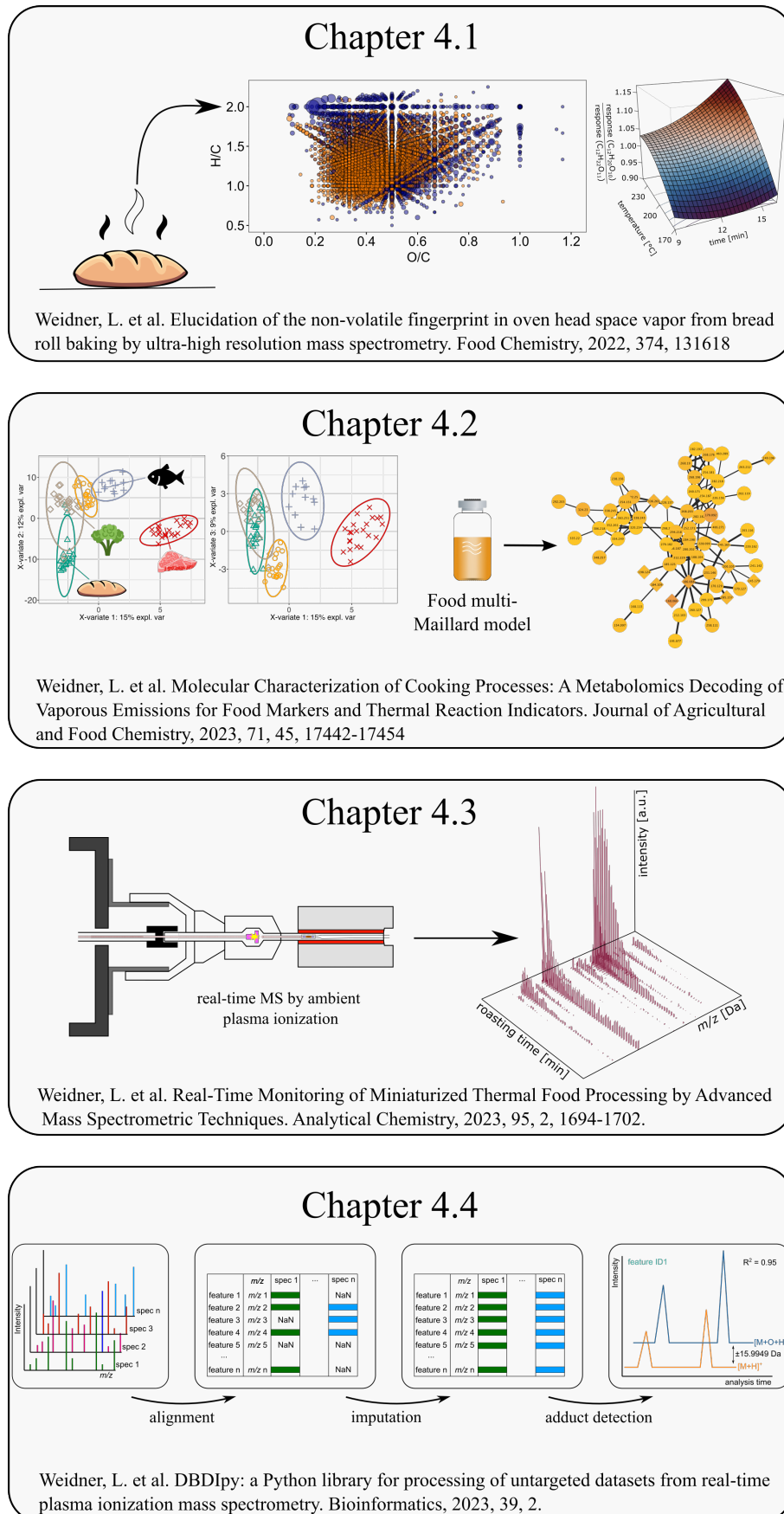
## Chapter 4.1

Weidner, L. et al. Elucidation of the non-volatile fingerprint in oven head space vapor from bread roll baking by ultra-high resolution mass spectrometry. Food Chemistry, 2022, 374, 131618

## Chapter 4.2

Weidner, L. et al. Molecular Characterization of Cooking Processes: A Metabolomics Decoding of Vaporous Emissions for Food Markers and Thermal Reaction Indicators. Journal of Agricultural and Food Chemistry, 2023, 71, 45, 17442-17454

## Chapter 4.3

real-time MS by ambient plasma ionization

Weidner, L. et al. Real-Time Monitoring of Miniaturized Thermal Food Processing by Advanced Mass Spectrometric Techniques. Analytical Chemistry, 2023, 95, 2, 1694-1702.

## Chapter 4.4

alignment · imputation · adduct detection

Weidner, L. et al. DBDIpy: a Python library for processing of untargeted datasets from real-time plasma ionization mass spectrometry. Bioinformatics, 2023, 39, 2.

**Figure 3.1** Structure of the thesis

# 4 Results

In this dissertation, I have investigated the chemical changes in food triggered by controlled thermal transformations. I contributed to four research articles publishes in peer reviewed, international journals. I am the first author of all four articles, where three of them were original research articles and one was an application note.

1. **Chapter 4.1:** Weidner, L., Hemmler, D., Yan, Y., Rychlik, M. & Schmitt-Kopplin, P. (2022). *Elucidation of the non-volatile fingerprint in oven headspace vapor from bread roll baking by ultra-high resolution mass spectrometry.* Food Chemistry, 2022, Volume 374, Article 131618.
   doi.org/10.1016/j.foodchem.2021.131618

2. **Chapter 4.2:** Weidner, L., Cannas, J. V., Rychlik, M. & Schmitt-Kopplin, P. (2023). *Molecular Characterization of Cooking Processes: A Metabolomics Decoding of Vaporous Emissions for Food Markers and Thermal Reaction Indicators.* Journal of Agricultural and Food Chemistry, 2023, Volume 71, Issue 45, Pages 17442–17454.
   doi.org/10.1021/acs.jafc.3c05383

3. **Chapter 4.3:** Weidner, L., Hemmler, D., Rychlik, M., & Schmitt-Kopplin, P. (2023). *Real-Time Monitoring of Miniaturized Thermal Food Processing by Advanced Mass Spectrometric Techniques.* Analytical Chemistry, 2023, Volume 95, Issue 2, Pages 1694-1702.
   doi.org/10.1021/acs.analchem.2c04874

4. **Chapter 4.4:** Weidner, L., Hemmler, D., Rychlik, M., & Schmitt-Kopplin, P. (2023). *DBDIpy: a Python library for processing of untargeted datasets from real-time plasma ionization mass spectrometry.* Bioinformatics, 2023, Volume 39, Issue 2.
   doi.org/10.1093/bioinformatics/btad088

In the following chapter, I provide a one-page summary for each of these articles and indicate my scientific contribution. The full articles are included in Appendix A. Respective collections of Supplementary Materials are included in Appendix B.

Subsequently, I will set these articles into a conclusive framework. I will discuss their relevance and connection and give a final outlook for future research perspectives.

## 4.1 Elucidation of the non-volatile fingerprint in oven headspace vapor from bread roll baking by ultra-high resolution mass spectrometry.

*Summary: During this work, we performed an untargeted fingerprinting study of the food metabolome from wheat bread rolls contained in their processing emissions. To collect these emissions, we initially designed a sampling apparatus relying on the physical principal of cold-trapping vapor carrying organic atmospheric constituents pumped out of the baking chamber of a convection oven. Wheat bread rolls were baked according to a design of experiments plan with baking temperature and time as factors to study the influence of variable processing parameters on the head space signature. The factors ranged between 158 to 242 °C and from 7.8 to 16.2 min, respectively. With minimal sample preparation, these vapor condensates were analyzed by DI-FT-ICR-MS. This revealed a molecular complexity of more than 4,700 annotated features in vapor from bread rolls baked at 230 °C for 15 min opposed to 580 features in vapor from bread rolls processed at 170 °for 15 min. Visualization techniques suited for the illustration of untargeted datasets from ultra-high resolution MS showed abundant traces of members of various compound classes of which selected representatives were identified by UHPLC-QqTOF-MS. Distinct signatures of the CR and the MR were identified in the vapor, especially under higher thermal load. Response surface calculations showed, that the relative mass spectrometric intensity of a feature in the vapor sample can validly be modelled by baking temperature and time. In total, 990 valid models for small molecules present in the head space were calculated. Thermo-induced early dehydration reactions of the mono- and disaccharide feature were compared to the literature and found to be in excellent accordance with quantitative data published about saccharide concentration during baking of cereal products. This finally led to the demonstration, that even complex chemical transformations, like the formation of the AGE maltosine from saccharose and lysine, can be monitored and modelled through processing emissions as an interface.*

---

*Candidate's contributions:* L.W. developed the sampling protocol and experimental plan, performed the experiments and the validation, prepared and measured the samples, analysed the data, wrote the programming code, prepared the figures, and wrote and revised this research paper.

## 4.2 Molecular characterization of cooking processes: a metabolomics decoding of vaporous emissions for food markers and thermal reaction indicators.

*Summary:* *The research presented in this work presents an inclusive description of the interrelationship between individual biological sources and technological compositions of food and their cooking emissions under variable processing conditions. For this reason, we collected processing emissions from convection cooking of 96 food samples, which were analyzed after minimal sample treatment by DI-FT-ICR-MS and UHPLC-QqTOF-MS. Multivariate analysis by sPLS-DA discriminated vapor constituents specific for plant-based and meat-based food. Further, meat and fish on the one hand side as well as plant-based meat alternatives, vegetables, and bakery products on the other hand could be further discriminated by proposed markers compounds. Metabolomic fingerprinting performed by ultra-high resolution MS revealed distinctive feature spaces which could be attributed to biological properties of the emitting matrix. Given the example of bakery and pastry products, traces of the CR and MR reaction can be used for their recognition from the head space whereas meat and fish differ by their peptideous and fatty acid fingerprints. The defined molecular spaces were annotated with compounds classes and compared for their communal population. Consecutively, features detected in emissions from bread and steak collected under different processing temperatures were compared for their sensitivity to temperature changes. Pyridoxine and several amino acid derivatives could be annotated in the steak experiment. The same analysis was performed with detected MS/MS fragment ions, which revealed structural motives of aromatic N-heterocyclic core structures, as they evolve during the MR, to be most sensitive variable processing temperatures. Analysis of mixed Maillard model systems allowed to integrate features in the rump steak vapor sample into a Maillard reactional context for a better reactional understanding of processing emissions.*

---

*Candidate's contributions:* L.W. developed the analytical approach, performed the experiments, prepared and measured the samples, developed algorithms, performed data processing and performed statistical analyzes, evaluated the data, prepared the figures, wrote and revised this research paper.

## 4.3 Real-Time Monitoring of Miniaturized Thermal Food Processing by Advanced Mass Spectrometric Techniques.

*Summary:* *In this work, we designed a laboratory-scaled, miniaturized, and highly controllable setup for the online monitoring of aerosols and volatiles from thermal food processing based on DBDI-MS. The setup is easily transferable between MS systems of different vendors to offer full flexibility in selection of the available mass analyzers. The most unique point of the presented methodology is the capability to trace chemical reactions during cooking in real time, as the heating setup is closely attached to the MS orifice to direct the entity of* in-situ *generated thermal reaction products of volatile and aerosol processing emissions to the mass analyzer. A show-case was performed by monitoring small organic compounds emerging during the thermal processing of wheat bread roll dough at 210 ℃ by FT-ICR-MS. After an initial illustration and following discussion of the developed setup, we drew a comparison to three API sources to define the selectivity of DBDI, to show communal intersecting molecular spaces, and to highlight new analytical perspectives given by DBDI. The detected processing emissions were characterized for their molecular composition and discussed in the background of thermal transformations during food processing. A closer perspective was thereby given to the temporal aspects, where it was found that distinctive molecular spaces are populated during different stages of the experimental run. Time series analysis by the AI tool self-organizing maps grouped different temporal metabolite profiles together and revealed that similarly-shaped emission profiles belong to distinct spaces of molecular entities. The developed analytical platform based on ambient plasma ionization was all together able to characterize more than 6,100 food metabolites and thermal reactions products in a time-resolved way. The experiment was divided in an initial baking-followed by a successive roasting phase after prolonged thermal exposure. The developed setup offers direct applications for industrial on-line process control and enables the delivery of time-resolved extensive data during the production of food.*

---

*Candidate's contributions:* L.W. developed the analytical approach, performed the experiments, measured the samples, developed algorithms, analysed the data, prepared the figures, and wrote and revised this research paper.

## 4.4 DBDIpy: a Python library for processing of untargeted datasets from real-time plasma ionization mass spectrometry.

**Summary:** *In this application note, we presented the description of the Python package DBDIpy which we developed for the processing and formal analysis of untargeted, time-sensitive plasma ionization MS datasets. DBDIpy is open-source and implemented in the Python programming language (Version $\geq$ 3.7). It can be downloaded and installed from the Python package repository PyPI or from the official GitHub release channel. DBDIpy's main functionality is the detection of adducts and fragments generated in the ion source during time-sensitive experiments. Ambient plasma ionization techniques are relatively young compared to established API solutions like ESI but show increasing application during the last years. However, different artifact arise during the ionization process resulting in miscellaneous ion species next to $[M + H]^+$. This fact makes acquired data especially in untargeted metabolomic datasets ambiguous. DBDIpy offers a straightforward solution for the user to overcome this convolution. The developed functionalities are integrated into the matchms-ecosystem, a popular package for processing of MS data in Python. A walk-through of the implemented functions including data-import, data organization by alignment and zero-filling, via adduct detection, to plotting and exporting of results is given on GitHub. The kernel algorithm consists of a two-step open search strategy to detect in-source fragments and adducts which are pre-set and can be individually extended by the user. It consists of a combination of temporal correlation analysis of detected XIC combined with a refinement based on exact mass difference analysis. The application of DBDIpy was showcased on a demonstrational data set of 4,200 features from DBDI-MS analysis of thermal processing emissions from wheat bread dough baking. It detected 710 potential adducts with a Pearson correlation threshold of $r > 0.95$ and mass offset of $2\,ppm$.*

---

*Candidate's contributions:* L.W. developed the analytical and computational approaches, performed the experiments, measured the samples, supplied the demonstrational data, developed algorithms, developed algorithm tests, wrote the documentation, built and published the Python package, prepared the figures, and wrote and revised this research paper.

# 5 Concluding Discussion and Outlook

The presented analytical work reports on the molecular complexity of processing emissions arising during the controlled thermal transformation of food. Modern instrumental analytical approaches were developed, bioinformatic data analyses implemented, and data visualization strategies for the illustration of complex, untargeted metabolomics data were deployed to gain a holistic chemical insight into food processing through the head space as an analytical interface. The research results are compiled and thoroughly discussed in the respective publications presented in chapters 4.1 to 4.4. The following chapter will serve as a global discussion: first, the published research articles will be contextualized and discussed for their logical progression. During a critical integration of the thesis into the contemporary scientific discourse, chemical changes of food during the cooking process, release dynamics to the head space, and developed bioinformatic tools will be debated. A connection will be drawn, how the findings presented in the previous chapters offer perspectives for an enhanced understanding of cooking applications. To conclude, an outlook will be given, and future research perspectives will be presented.

## 5.1 Contextualisation of single research papers

To build molecular understanding about the controlled thermal transformation of food, the core concept of monitoring chemical reactions in processed food through the head space as an interface was applied in this thesis. To reach this goal, two analytical setups were conceptualized, built, refined, and successively applied in the analytical routine.

The origin of the research concept lies, as elaborated in chapter 3, in a lack of systematical chemical knowledge about the thermal reactome of food contained in processing emissions. Therefore, constructing an apparatus capable of the unbiased sampling of processing emissions stood in the immediate focus of this work. The first developed setup, closely described in chapter 4.1, relies on simple physical principles. It consists of a negative pressure system actively conveying the atmosphere from a convective oven and condensing all constituents in a cold trap. It is important to stress that the described setup follows an at-line sampling strategy. This allows to concentrate and analyze the collected sample using powerful analytical tools as it was done with FT-ICR-MS and UHPLC-Q$q$TOF-MS. The acquired data revealed diverse molecular spaces to be present in the collected condensate and permitted to validly

model chemical reactions in the matrix through the head space by a response surface analysis. Even complex reactions, such as the formation of the AGE maltosine, could be modeled.

However, the initial study was solely conducted with wheat bread rolls, a relatively uncomplex food product mainly consisting of cereal material. From this proof-of-concept arose the potential to disclose food-specific molecular spaces and to investigate the structural consequences of variable cooking temperatures on chemical head space profiles. The resulting comprehensive characterization of processing emissions from 96 food items under different thermal influences is incorporated in chapter 4.2. Profound relations between matrix constituents and formulations and resulting emitted molecular profiles were determined and resolved in molecular spaces. Further, structural motives of molecules that can serve as thermal processing indicators were identified by fragment ion analysis.

Despite the extensive molecular knowledge gathered by the developed at-line sampling protocol, it was targeted to observe chemical transformations in the head space in a time-resolved way. Therefore, chapter 4.3 describes the construction of a miniaturized thermal food processing apparatus that empowers the on-line measurement of processing emissions. The setup permitted to monitor arising processing emissions again from baking wheat dough by FT-ICR-MS as a detector system. Different temporal patterns of detected metabolites were determined over the course of experimental baking events and could be attributed to distinct molecular families, mostly in the CHNO space.

The principal ionization technology employed in this second setup is an ambient plasma ionization source. The used DBDI technology is relatively young compared to other API sources and is known to unite variable, competing ionization pathways, which impede the processing and interpretation of untargeted metabolomics data. Different ion species, including in-source adducts and fragments, might distort data analysis if not correctly taken into account. A lack of chromatographic separation in real-time mode complicates the post-acquisition recognition of, e.g., a series of in-source oxygen adducts. Consequently, algorithms developed for handling, processing, and curating time-resolved DBDI-MS datasets were made available to the public in an open-source Python package format. The accompanying article is included in chapter 4.4 and presents a bioinformatic approach to group multiple ion-species of the same analyte by a two-step filter algorithm.

Both developed sampling setups work together on the core research question of this thesis, although they focus on different aspects of processing emissions. Both methods contribute to revealing various clarifying aspects of processing emissions.

A detailed comparison of both approaches reveals the following characteristics: An at-line collected vapor sample is of steady character. It is available for sample preparation tasks, e.g., concentration or dilution of the collected condensate. The sample can be subjected to multiple, complementary measurements over an extended period of time, when stored in a deep-frozen state. Analysis can be performed at remote locations with powerful or spe-

cialized instrumentation. This enabled us to perform metabolomic profiling experiments by FT-ICR-MS, on the one hand side, combined with in-depth structural characterizations by UHPLC-Q$q$TOF-MS, on the other hand side, with the same collected condensate sample. This advantage in terms of instrumental flexibility is sacrificed in the DBDI setup for the benefit of acquiring real-time data of cooking processes. The sample, however, is consumed after the experiment, and a measurement needs to be repeated by a physical replicate of the original sample. In response to this drawback, the DBDI setup was conceptualized in a way that it can easily be operated at different vendors' mass spectrometers to increase the method's flexibility. This instrumental flexibility offers opportunities to implement the developed method in industrial food production processes.

Vapor collection techniques can approximate a time-resolved mode by decreasing the sample interval lengths to successive sampling progressions per cooking process. Still, it can never reach the temporal resolution of an on-line coupled detector. Defined by the sampling interval length, collected vapors are a cumulative sample representing the time frame of a specific baking phase. In future experiments, vapor sampling of multiple time intervals during an experimental run can contribute valuable information about the chemical composition of the head space in early and advanced baking phases. Insight on release dynamics as a function of surface temperature could thereby be generated. In time-resolved mass spectrometric methods, the scan speed is commonly regarded as a crucial factor and needs to keep the pace of the analyzed process. Long-cycling methods bear the risk of missing important events and of lacking the possibility to implement real-time process control (*274*).

Both developed methods focus on divergent targeted analyte classes. DBDI is best suited to analyze volatiles and organic aerosols. The collected condensate samples were analyzed by API sources, mainly ESI, to detect non- and semi-volatile compounds. As API sources and DBDI primarily generate pseudo molecular ions, 56% of all annotated analytes found in ESI(+)-mode were likewise detected by the DBDI setup. This indicates a significant overlap in ionization selectivity of both techniques and is a mutual confirmation, that the developed DBDI setup successfully down-scaled the baking process to a miniaturized level. As expected, did DBDI specifically reveal aromatic heterocyclic compounds in CHNO spaces from volatile emissions that were not accessible by vapor collection due to their high vapor pressure.

Next to adding a mass spectrometric real-time prospect on food processing, the work presented in chapter 4.3, further unites perspectives on food processing emissions by multiple ionization techniques. All applied ion sources contribute different ionization mechanisms and specificities to illustrate a comprehensive picture of molecular spaces contained in thermal emissions from wheat bread baking in an unbiased metabolomics perspective. Especially the DBDI setup strongly follows this paradigm as no sample preparation is implicitly needed, and the unadulterated entity of processing emissions is directly made available for mass spectrometric analysis. As proposed in the original definition of metabolomics (see chapter 2.1.1), does the collective of small organic molecules represent the ultimate and most

immediate descriptor of a system concerning its intrinsic constituents and external influences (*208*). Cooking adulterates the biochemical system of a food item by the external application of heat; which is reflected in more than 16,000 heating-related annotated molecular formulae observed during the studies reported in this thesis.

A fundamental difference between both cooking setups concerns the proliferation of thermal energy to the processed food item. The vapor collection studies were conducted in a convection oven. A fan circulates the oven atmosphere to create an evenly heated environment. From a thermodynamic perspective, the oven atmosphere, primarily constituted of air and steam, acts as fluid to mediate convective heat transfer. As shown in equation 1.5, radiative heat transfers thermal energy from hot surfaces of the cooking cabinet, and conductive heat transfers thermal energy from the baking tray to the food at comparably low rates. In the SPME module of the DBDI setup prevails by contrast, radiative heat transfer. In the first place, the sample containing borosilicate glass tube is heated by a heating coil. The heated tube again transfers heat in a conductive way to the sample in a second step.

The differences in applied heating strategies are likely to cause diverging structural properties in the studied matrix: typically, a stationary thermal boundary layer of cooler air forms on the surface of a baked food item. Turbulent airflow caused by the fan breaks this insulating layer and increases the convective heat transfer coefficient $h$ to efficiently heat the matrix. Nevertheless applies:

$$h_{convective} \ll h_{conductive} \qquad (5.1)$$

The real-time setup was explicitly designed to handle miniaturized amounts of sample to exploit the relation given in equation 5.1. An instantaneous, homogeneous, and reproducible heat flux profile in the studied sample was desired to avoid differentiation of temperature zones in the matrix during the experiment. As no thermal compartments are formed, water contained in the sample evaporates rapidly and uniformly to the exterior, so no $a_w$-value gradient can emerge. As a result, chemical reactions can be assumed to occur in the entire matrix homogeneously. However, this is not the case during bulk experiments performed in the oven setup where distinct heat and material flux gradients exist, as elaborated in chapter 1.2.3.

Accordingly, the two created methods for monitoring processing emissions do not only target two complementary analyte classes but also serve different analytical objectives.

At-line analysis of cooking vapor focuses on a set of non- to semi-volatile compounds, whereas the on-line protocol targets on the description of semi-volatiles to volatiles. Detailed description about the complementary nature of both methods is given in the discussion section of the publication presented in chapter 4.3. In-oven vapor collection allows for probing real-life cooking events of bulk food material as in industrial, gastronomic, and personal food preparation. The DBDI-based real-time setup, however, offers opportunities to simulate

cooking processes under highly controllable, miniaturized conditions to evaluate the effect of different formulations or processing techniques on a laboratory scale.

Ultimately, the differences in heating strategies point to the fact that starting from the vapor sampling setup, which is suited to analyze real-life food processing, a scientific implementation was created to model cooking processes and to investigate chemical reaction mechanisms on a laboratory-scale level. The study's results presented in chapter 4.3 demonstrated that the presented methodological dualism achieved to construct a functional bridge between bulk food production and a miniaturized research tool. The communal similarities of both analytical tools were proven and represent the linkage of an innovative strategic tandem of methods to extract chemical information offered by head space analysis. The study of further cooking techniques like grilling, roasting, or steaming will deliver further valuable insights about the effects the cooking technique has on the emitted compounds.

## 5.2 Embedding of the thesis into contemporary literature about processing emission analysis

In section 1.3, different aspects of cooking emission analysis were summarized. Altogether, the available literature about the analysis of thermal food processing emissions shows a lack of comprehensive investigations and molecular profiling studies. Hardly any study investigated so far the exact relation between food composition and resulting processing emissions; thereby, the high biological variability of raw food and diverse formulations were neglected. Most important, however, is that nearly all methods used to study food processing emissions, reviewed in chapter 2.3, focus on volatile compounds. The share of non- and semi-volatile constituents remains unconsidered in most assessments of cooking head spaces. Untargeted approaches were only sparsely pursued; the true complexity of cooking emission remained uncharacterized as often selected, targeted compounds were studied. Hence, the shortcoming of integrative methods prevented to draw a comprehensive, molecular picture of food processing emissions so far.

Currently, the environmental research interest of cooking emission analysis primarily lies on PM2.5 and its implication on health. The absolute focus of compound classes studied in this field lies on VOC, including alkanes and aldehydes studied by GC-MS and especially on airborne PAH, which are analyzed by GC-MS and by LC with fluorescence detection (*179*). The presence of these substances in urban atmospheres due to cooking processes is remarkable and demonstrates the outreach of organic compounds originating from food preparation. Compared to urban sampling sites, the work presented in this thesis deals with analyzing relatively concentrated cooking emissions at the site of their formation. This reduces the influence of external contaminants and secondary atmospheric reactions with other environmental influences, e.g., from automotive exhaust gases.

In the scope of environmental air analysis, Mcdonald et al. performed an outstanding study: the authors sampled emissions from charbroiling and grilling of chicken and beef and identified by GC-MS complex organic structures as $\gamma$-lactones from oxidation of $\beta$-hydroxy FA, and PAH of predominantly two to three unsubstituted, annealed rings like naphthalene (*326*). In in the work presented in this thesis, both poultry and bovine meat vapors were sampled and analyzed by UHPLC-Q$q$TOF-MS. PAH could not be detected in any of our analyzed samples, which is reasonable as PAH are formed between 550 to 950 °C (*327*), a temperature the convection oven did not reach.

Volatile compounds such as VOC, which possess high vapor pressures, are hardly retainable by cold trapping; most protocols suggest the usage of trapping sorbents for their analysis (*186*). Cryo-trapping of volatiles requires tailored instrumental setups and the usage of liquid nitrogen. Respective protocols were pioneered around Peter Schieberle and prove especially useful for the analysis of very volatile or thermo-labile compounds (*328*). Further, the molecular weight of these species is below the mass range accessible by FT-ICR-MS.

Therefore, the molecular intersection of compounds studied in this thesis with VOC from the literature is limited.

Further, McDonald suggested sterole (among other species) as a marker compound to distinguish meat cooking emissions from other carbonaceous sources (*326*). We were not able to identify sterole in meat-related processing vapors. However, we suggest other marker compounds for the processing of meat, such as N-methyl-L-histidin-derivatves or dicarboxylic acids, which were also identified by Rogge et al. in fine organic aerosols (*185*). Additional cooking marker substances like monosaccharide anhydrides were proposed earlier (*329*). We likewise identified them, especially in emissions from cereal products. Further, Zhao et al. suggested to use combinations of molecular biomarkers to distinguish cooking source types (*330*), which gives a perspective towards multi-feature classification approaches as we, e.g., employed for the definition of food-related molecular spaces. Together, this demonstrates that our untargeted approaches confirm formerly suggested marker compounds and complement them with an additional landscape of food-specific molecular spaces.

The analytical approaches presented in this work likewise offer a perspective to advance the field of environmental air analysis. It was extensively shown that cooking processes liberate compounds with wide ranges of polarity and molecular mass. The transfer of metabolomics or fingerprinting protocols could offer holistic perspectives for the analysis of indoor air and urban atmospheres to overcome the bottlenecks of established methods towards highly polar compounds.

Besides the context of environmental PM analysis, flavor chemists especially studied aroma-active compounds in food processing emissions. Rega et al. constructed a refrigerated extraction chamber, where baking vapor from several sources of cereal products were sampled by SPME fibers and later analyzed by GC-MS. Besides aliphatic aldehydes and ketones, compounds of interest were mainly furanic compounds like furfural or HMF and several pyrazines. These compounds were identified in vapor from cookies and sponge cake (models) (*112*, *130*, *191*). We detected abundant series of pyrazines by online DBDI-MS analysis of wheat bread processing and discussed their suitability as thermal reaction markers contained in rump steak vapor under the background of the MR. Further, HMF was identified by UHPLC-Q*q*TOF-MS in vapor samples from bread roll emissions as a central marker compound for the processing of cereal products.

Several other studies, e.g., about the odorous emissions from popcorn (*190*) or gluten-free bread (*170*, *171*) were carried out but again were restricted to volatile compounds. Volatiles with low molecular masses ($\leq 200\,\mathrm{Da}$) stood not in the direct focus of the work presented in the preceding chapters. Their retention in a cold trap is challenging because of their relatively low boiling point; significant amounts can be lost during lyophilization, and ESI is not the most-efficient ion source for analyzing hydrophobic volatiles. The DBDI-setup, which is efficient in the analysis of volatiles, was only operated on an FT-ICR-MS platform, which is best-suited for the analysis of compounds with a molecular mass above $200\,\mathrm{Da}$, which is

due to instrumental characteristics of the ion optics. Future experiments should perform DBDI-MS monitoring of cooking processes on MS platforms appropriate for the analysis of low-mass compounds like tandem quadrupole or Q$q$TOF mass detectors, which would additionally offer the benefit of tandem MS data acquisition.

In the here-presented work, a methodological dualism was used: volatiles and aerosols were monitored by DBDI-MS, and non- and semi-volatile compounds were assessed by vapor-trapping, bringing together molecular knowledge from two perspectives. Semi-volatile compounds are distributed between the gaseous and particle form as a function of compound-specific vapor pressure, particle nature, and environmental conditions, e.g., cooking temperature (*326*). Therefore, our partially overlapping and complementary set of methods proves to be versatile as it confirms the presence of several cooking markers known in the literature and extends the knowledge to new limits of analyte polarity and molecular mass. Other head space sampling techniques did rarely account for heavier compounds, nor could they depict entire reactional systems, as these usually involve polar compounds or precursor compounds not present in the gas phase. However, only unbiased metabolomics datasets can enable the understanding of branched reactional systems or give a comprehensive base for the definition of food-related molecular fingerprints.

Compared to extraction-based metabolomics studies of the food reactome, the vapor-based approach offers several advantages: first, the omitted extraction step enables the real-time analysis of food processing from the head space. Second, selecting an extraction solvent represents an *à priori* bias towards analyte polarity. Cooking emissions on the opposite contain substances of wide polarity ranges: the presence of hydrophilic carbohydrates to reaching lipophilic substances such as FA was reported in the study presented in chapter 4.1. Third, the incorporation of processing emissions as an interface to the matrix avoids the release of cell material or bulk matrix components, causing analytical problems. Extraction-based protocols often require purification steps to remove excessive matrix constituents like lipids or carbohydratous starch remnants or insoluble components. These do hinder chromatographic analysis and decrease analyte signals due to suppression effects. Especially regarding fingerprinting experiments, the presence of bulk matrix components can conceal trace-level metabolites. So far, no sampled food class was found to release excessive or highly suppressive matrix constituents in their processing emissions; likewise, no literature is known that needed to perform purification of cooking emissions; only explicit concentration steps regarding *targeted* analysis of single compound classes like PAH are described (*179*).

Also, in comparison to emission studies based on SIFT-MS, PTR-MS, or SPME thermodesorption GC-MS, do both presented methodological approaches offer closer insights into cooking vapor due to higher instrumental sensitivity and better resolution. However, there were no quantitative measurements performed on the developed setups so far. Especially SIFT-MS is an excellent tool to record quantitative in-oven ratios of organic compounds in

processing emissions. Extracting quantitative data from vapor collection would, therefore, require the usage of internal standards, which is expensive and labor-intensive.

In the context of a sustainable society, analytical chemistry needs to implement green chemistry principles. Considerations of sustainability need to be taken into account for all steps of manufacturing processes (*331*), hence especially on-line analysis of processing emissions, which requires no sample preparation, offers a forward-looking perspective for the analysis of food production.

## 5.3 Release dynamics of semi- and non-volatile compounds to the head space

A fundamental aspect of processing emission analysis deals with considering the extent to which the head space is suitable as an analytical interface for the processed food. In other terms: do processing emissions reflect the profile of small molecules in the emitting matrix unbiased and comprehensively? Do molecular ratios in the head space reproduce the situation in the actual food with respect to its complexity and external triggers? To answer this question, the release dynamics of organic compounds to the head space shall be discussed in the following.

A precondition for a compound to be detected in processing emissions is to become airborne. In the case of volatiles, evaporation is a sufficient explanation for their presence in the head space. Volatility is a function of the ambient parameters temperature and pressure and intrinsic parameters like vapor pressure, and retention in the matrix (based on compound and matrix $log_P$ and pH values). Further, diffusion effects along a concentration gradient play a role. Under the influence of applied thermal energy, like within the baking chamber of a convection oven or in the miniaturized baking setup, at least 200 °C prevail, under which efficient evaporation of volatile compounds can be assumed. In the work presented in this thesis, only the miniaturized baking setup based on DBDI-MS, is suited for analyzing highly volatile compounds, which are hardly retainable by the cold-trapping apparatus.

As discussed in chapter 1.2.3, the predominant share of volatiles is generated on the surface, respectively the outer crust, of a food item due to high thermal impact. Therefore, the prevailing share of odor-active compounds detected in the head space is attributed to origin from the surface of the food. The proportion of volatiles evaporated during a baking process can be significant: Pico et al. constructed a PTR-MS setup for the in-oven monitoring of volatiles from gluten-free bread during baking, as these often find low-consumer acceptance due to poor sensory qualities ([170]). The study permitted to follow release patterns of qualitatively relevant compounds to the head space in significant amounts. Regardless, residues of volatile compounds always remain in the matrix, enclosed in the crust and pores, and are responsible for the sensorial experience of the consumer ([332]).

The presence of volatile compounds in processing emissions is evident from the olfactory notes perceived during daily cooking processes. Less clear, however, is the presence of a magnitude of non- and semi-volatiles being present in processing emissions. Their initial characterization is a central finding of this thesis and raises the question of how these compounds can reach the head space. The atmospheric composition of cooking fumes is complex, and two main explanations for mechanisms contributing to the release of non-volatiles will be given to rationalize their presence in the head space.

First comes the particulate matter model: fumes from cooking processes consist of (oil) droplets, evaporated water, combustion products, as well as condensed and coagulated organic substances (*179*). It was shown that supersaturated fumes can, depending on ambient conditions, cool and nucleate to form larger aerosols (*179*). There is strong consensus that ions, water vapor, and small inorganic trace gases present in very low concentration in the head space facilitate and contribute to the nucleation process (*333*). These growing particles possess great surface areas. Depending on their polarity and ambient conditions, adsorption processes of organic molecules to their surface are described. There is a steady state of diffusion between the gas and particle phase of compounds depending on their vapor pressure (*179*, *326*). Aerosols are a highly dynamic and interacting system, especially under elevated processing temperatures. Additionally, primary matrix constituents can be modified by secondary chemical reactions in the atmosphere to form species with relatively lower vapor pressure that condense under given reactional conditions (*334*). These can form ultra-fine dispersed droplets acting as a nucleus for agglomeration processes. The fan of the convective oven causes a constant atmospheric motion, which leads to ongoing collisions of the particles with the matrix, allowing the incorporation of non-volatile matrix constituents like amino acids or carbohydrates during their interaction. This stands in logical consequence to the observation that head space profiles largely depend on the surface morphology of the processed food. Open structures, e.g., found in plant-based meat alternatives, yield very abundant emission patterns.

Besides that, aqueous juice, e.g. as it is expelled from meat, can drip and splash in the cooking cabinet as a result to thermal exposure. This is due to the contraction of functional protein fibrils, see chapter 1.2.3, and a loss of water-binding capacity as a reason of *post-mortem* production of lactic acid. Consequently, washed-out non-volatile organic compounds from the matrix can reach the head space as expelled fluids quickly evaporate (*166*).

The presented theorem of PM transport mechanisms needs to be probed in future research. Up to careful literature research are the formation of PM and associated adsorption processes only known from atmospheric chemistry but were not validated in the context of cooking emissions. However, the share of fine aqueous droplets offers a perspective to substantiate the presence of highly polar compounds in processing emissions.

The second transport model comes especially into action under steam-rich processing conditions and exploits the physical principles of steam distillation (SD). SD is a traditional purification technique used to extract natural products at a temperature below the components' normal boiling point (*335*). Usually, SD is employed for extracting non-polar mixtures as essential oils (*335*) or recently also for heavy oil recovery (*336*). The SD mechanism is especially relevant for semi-volatile compounds under humid cooking conditions, e.g., during the preparation of water-rich vegetables or steaming processes.

Recently, there are reports on the extraction of polar compounds like phenols by SD (*337*, *338*). Therefore, SD is not only limited to highly hydrophobic compounds and steam-mediated

liberation of polar compounds to the head space will likely contribute to the complex chemical constitution of food cooking emissions. Especially in porous matrices like, e.g., puff pastry or bread, water evaporates through pores to the surface of the processed food item and finally to the head space. The vapor is hypothesized to transport a share of hydrophilic organic compounds to the cooking cabinet.

Constructing a reliable synopsis of the different transport mechanisms is challenging, as so far, no experimental investigation of the interplay of these transport functions has been made. Their effect size indeed depends on the studied matrix, the employed cooking setup, and the respective analyte family. The revealing perspective given by the molecular atlas of food processing emissions drawn in chapter 4.2 solely presents a metabolomic end-point. The combinations of the PM and SD models thrive to explain the presence of thousands of polar compounds in processing emissions. However, experimental proof is necessary to elucidate the exact proportions of involved transport mechanisms. This would also contribute clarifying information to the question to which extent single compound classes are represented in the head space. As the evaporation of volatiles, PM transport, and SD carriage all depend on molecular and matrix properties, the final mechanism of how a specific compound reaches the head space is difficult to postulate at this point. Potential experiments could include tests under low-moisture and high-moisture conditions to assess the role of steam. Visual or spectral perception under variable fan settings could determine the motion of matter within baking chambers to give information about release patterns.

In chapter 1.2.3, the topic of moisture mobility during a cooking process was addressed. Computational moisture models are very specific to particular scenarios and do not consider the transport of small organic compounds within a processed matrix. Spiking experiments with tracer compounds could give insight into spatial mobility in certain food types, e.g., caused by evaporating water. These experiments could also clarify if semi- and non-volatiles from the core of a food item do reach the oven atmosphere in significant amounts.

The combined mode of action of the discussed transport mechanisms is illustrated in figure 5.1. Notably, all three transport processes occur on the food surface, where the highest temperatures are found during cooking, see chapter 1.2.1. Thus, the head space profile is most likely linked to the surface chemistry of the emitting food item. Consequently, this would lead to a distorted chemical mapping of the molecular head space profile for heterogeneous food items with different core constituents. E.g., the cooking emission of pâté might not correctly reflect the entire original matrix composition. In chapter 4.2, the influence of complex formulations of composed food items in relation to emitted molecular profiles were determined. However, predominantly homogeneous food items were studied. Consequently, careful evaluation of the influence of different coating materials and the effect of their thickness should be assessed in the future.

Even if certain analytes are relatively over- or under-represented in processing emissions, it allows to cover a broad range of analyte polarity and molecular mass in the sense of

metabolomics screening due to the combination of several transport mechanisms. Clarifying experiments require quantitative measurements to determine the exact concentration ratios between a compound and its liberated proportion. This could be achieved after matrix extraction by targeted MS methods or comprehensive NMR studies. Initial approaches with $^{14}C$-labeled saccharose were made in the early phases of process emission analysis to follow the propagation of the radioactive label through the baking process (*339*). Nowadays, an approach with stable isotope labeled $^{13}C$ or $^{15}N$ precursors seems more reasonable and could also contribute to the understanding of new reaction mechanisms.
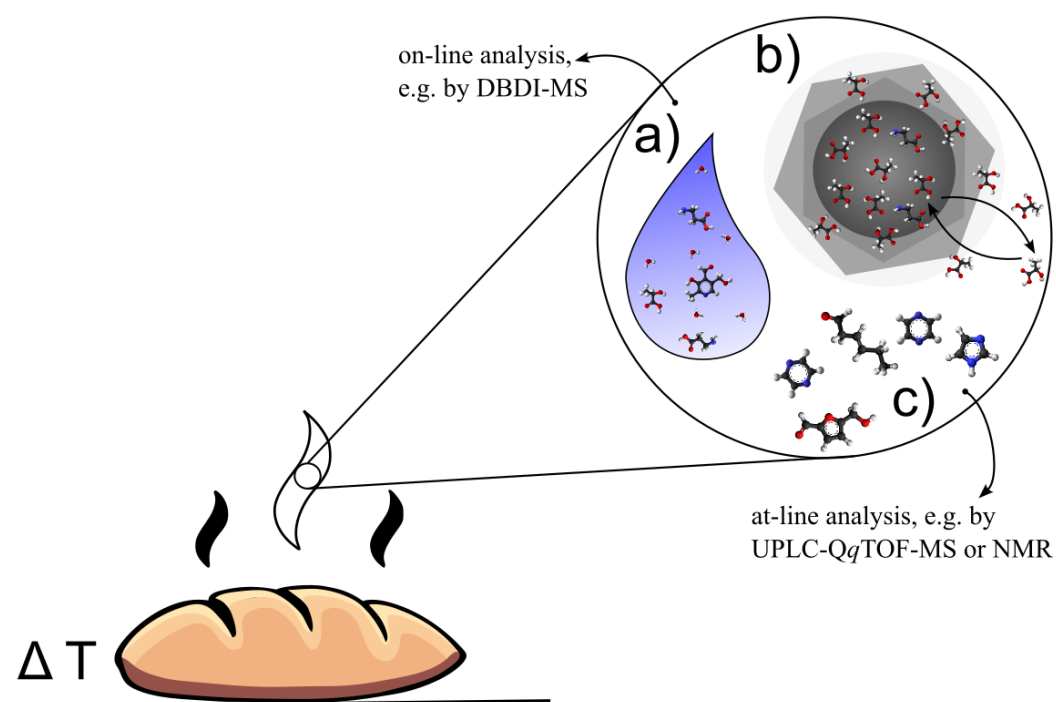


**Figure 5.1** Schematic drawing of the constituents of processing emissions emerging from heated food: non- and semi-volatile compounds contained in a) liquid droplets, b) particulate matter with continuous adsorption and desorption processes, and c) evaporated volatile compounds.

At this point, the main benefit of head space sampling, the potential to obtain non-destructive and extrication-free insights into a processed matrix, shall be highlighted. Current technological developments provide increasing instrumental sensitivity and throughput, enabling metabolomics research on a single cell level (*340*). In every analytical field, the development of high-throughput methods yields a wealth of biological and chemical data, which, in combination with computational approaches and enriched by mechanistic insights, grands data-driven insights to molecular systems (*257*). Accordingly, head space sampling offers an option to implement the collection of big data to the production of food and to follow this global analytical progression.

The characterization of food processing emissions presented in this work, combined with real-time analysis, offers perspectives for data-driven process control of food production. In

the context of food production 4.0, where automated processes steadily become more critical (*317*), chemical on-line monitoring of production plants can contribute towards immediate action and digitalized control on the one hand and to quality assurance protocols by seamless process surveillance on the other hand. In combination with advanced data processing pipelines, AI guided process control will become an essential aspect of future of food production.

## 5.4 Controlled chemical transformations in the matrix seen from the head space

The results of this thesis include descriptive food-dependant molecular profiles, which were acquired and discussed in the background of metabolomic fingerprinting analysis and the emitting matrix. However, a closer look at the chemical reactions enrolling during cooking needs to be given to understand the complex transformations happening to food during the heating process. By monitoring thermal reactions, ultimate control over the qualitative and nutritional properties of the final food product can be reached. To approach this goal, the studies presented in chapters 4.1 and 4.2 investigated the effects of modulating cooking parameters to assess reactional outcomes.

Overall, processing emissions were shown to be more dense and more abundantly populated by food metabolites under the influence of increased heat. It cannot be determined with certainty whether this is due to the increased formation of reaction products in the matrix or due to improved transport and release efficiency under harsh processing conditions; however, the ultimately resulting molecular fingerprint in the head space remains unaffected by this question and was found to be characteristic for food type and processing style. Comparative matrix extraction studies would contribute answers to this question.

Section 1.2.1 drew a consecutive picture of thermal reactions enrolling in food during cooking. Critical reconsideration of the results of chapters 4.1 and 4.2 leads to the conclusion that processing emissions contain a mixture of unadulterated compounds from the raw matrix on the one hand side and thermal reaction products on the other hand side. Heavy biopolymers and compounds exceeding a molecular mass of 850 Da were not detected. Such heavy molecules are too firmly integrated into the matrix to be released and do not meet the physicochemical parameters to enter the head space efficiently. Under this consideration, initial thermal breakdown reactions of larger precursors are crucial for organic compounds to populate processing emissions.

For the compound class of lipids, only free FA and monoglycerides were found to play a significant role in processing emissions. Especially the latter could significantly contribute to the formation of amphiphilic droplets in the head space. Free glycerol from the hydrolysis of lipids and various other biochemical sources was ubiquitous in all studied emission types. A series of oxidized FA from LPO processes were found, especially under harsh cooking conditions. Vapor from fish showed, in comparison to vapor from mammalian meat, abundant traces of PUFAs, as it was expected from the original matrix constitution. Annotations fitting to 3-chloro-1,2-propanediol and associated fatty acid esters were observed, however they were non-distinguishable from isobaric non-covalent chloride adducts under given analytical conditions. Exact molecular identities were not assigned on a large scale, as no lipidomics analysis was performed.

The largest carbohydrate species that could be detected was a trisaccharide; the exact identity could not be determined. The relative intensity maximum of this substance class lies on the contrary clearly on mono- and disaccharides species. The study of baking emissions from cereal products showed consecutive series of dehydration reactions typical for the CR. Many highly unsaturated, aromatic caramelization products were identified, especially by APCI(+) and APPI(+) analysis, indicating that the CR can efficiently be traced in the head space.

Free amino acids and mono- to tripeptides contribute strongly to the nitrogenous domain in the peptide space. Initiated by thermal degradation processes and interacting reactional pathways like the MR, it unfolds an abundant continuum of thousands of amino acid-related thermal transformation products. The vapor interface proves in chapter 4.1 that the formation of advanced thermal reaction products like maltosine can be traced from initial carbohydrate degradation via condensation with the amino acid to final AGE formation. Further, fragment ion studies of aromatic N-heterocyclic core structures of MRP permitted to estimate the amount of thermal energy a food item was exposed to during cooking. Results acquired by DBDI-MS supportively also showed an abundant population of these feature spaces. When considering the Hodge scheme, molecular spaces following characteristics for the late and intermediate phase of the MR are populated most densely. This indicates that processing emissions are not ideal for following initial Maillard processes because early intermediates like ARP are not stable enough in the head space under thermal stress and react to downstream products.

Conclusively, head space sampling allows following complex thermal reaction cascades from biological precursors to advanced end-products for all major nutrient classes. The wide availability of diverse molecular spaces in the cooking vapors makes this matrix an ideal tool for studying thermal processing markers. The analysis of processing indicators from an easy-to-sample and widely available vaporous system offers a straightforward possibility for on-line process control by means of chemical marker compounds used as infochemicals.
Once released to the head space, the spatial distribution of metabolites by tissue or compartment no longer exists. Secondary chemical reactions of atmospheric pollutants are a well-described phenomena (*179*). In contrast to the global atmosphere, cooking cabinets restrict processing emissions to a narrow space, making secondary reactions of food metabolites more probable based on frequent particle collisions. Consequently, cross-overs of linear carbohydrate, peptide, or lipid degradation pathways are suspected of contributing to the molecular complexity of processing emissions.

It would be helpful to construct an experimental convective oven, where the cooked food could be separated from the surrounding emissions after a cooking experiment. The remaining atmosphere could sequentially be resampled to determine the degree of secondary reactions taking place in cooking emissions independently of the original matrix.

In addition to thermal reaction products, endogenous matrix-specific metabolites were detected in all vapor samples to a wide extent. Creatine, purines, and vitamins were detected in meat emissions. Plant material emitted cinnamic acid derivatives or amino acid sulfoxides, and fish vapor was rich in abundant series of PUFA, to only name a few examples..

Endogenous matrix constituents were found to be useful to distinguish cooking emissions from various food sources. On the contrary, thermal reaction products can be fed to classification models to allow conclusions about the progress of thermal processing.

It became apparent during the evaluation of UHPLC-Q*q*TOF-MS data that only minor parts of the food-related metabolites and their thermal reaction products could be fully annotated and identified by standards. Besides the tracks of human metabolism and the characterization of natural products, MS/MS databases show a profound lack of compounds that arise from thermal reaction mechanisms (*341*). This lack of reference spectra, evidently accompanied by a systematic shortage of authentic standard material, makes it complicated to perform an ultimate comprehension of processing emissions at the given time.

Utilizing model systems to illuminate undisclosed parts of reactional networks is a wet-chemical approach to approximate thermal reaction products from the base of known precursors. In direct connection to food analysis, model systems prove to be helpful, e.g., for understanding the formation of HAA (*342*) in meat and fish or in determining the contribution of the MR to the beer metabolome (*250*). The more precursor compounds are added to a reaction model, the more realistic it becomes. However, its interpretation also becomes more difficult. To reconstruct reactional pathways of the MR during the convection cooking of meat, we designed a mixed Maillard reaction model of five amino acids and glucose, the predominant carbohydrate in muscle tissue. Thermal reaction products from the incubated mix enriched other features from vapor analysis with a reactional context. Successive spectral similarity networking helped to contextualize and characterize proximate features. Even a high spectral similarity score does not necessarily imply structural similarity nor chronological reactional order. Nevertheless did the constructed network contribute to a better reactional understanding of the widely uncharacterized process metabolome.

A promising approach for future research would be to spike stable isotope labeled precursors to raw food matrices. The miniaturized DBDI-MS cooking setup is the ideal tool to perform these investigations as the smallest amounts of $^{15}$N-labeled amino acids and short peptides or $^{13}$C-labeled carbonyl compounds would be sufficient to study the label's propagation and incorporation to the matrix. Q*q*TOF-MS or IMS-MS were valuable tools to study the thermal propagation of the label in real-time "food-flux" experiments and to contribute valuable structural information about the evolving products. First, experiments could be performed with bread dough, where labeled precursors could be incorporated homogeneously into the matrix. The limitations inevitably associated with a model system would be overcome with this approach and yield more realistic results as it accounts for the complexity

of food. Matrix effects could be considered with this approach. Reactional pathways could be studied in a time-resolved perspective. Additionally, cross-over connections of traditional linear reaction cascades, as mixed Maillard-lipid-oxidation-processes, could systematically be studied to investigate the interplay of the intermediate food metabolome flowing together to common reactive pools. Such experiments would help to overcome isolated research, which is restricted to single reactional systems. Amongst others, the discovery and characterization of thermal reaction markers could directly profit from this approach.

## 5.5 Bioinformatics perspectives on process emission analysis

The rapid development of computational capacities, along with applied computer sciences, offers revolutionary perspectives for food production through advanced data handling, analysis, and interpretation models. With the on-line access to the chemical information contained in processing emissions, big data sets can be delivered during food production. Treated correctly, this data can be used to analyze and control processes.

Continuous digitalization of quality control increasingly enables the integration of the "quality by design" principle, which the Federal Drug Administration originally introduced for the production of pharmaceuticals. It states that the quality of a product should be assured by process design and not by post-production quality testing as it is common in traditional production lines (*343*). A gradual transition from decoupling of monitoring and control towards continuous data acquisition of core quality parameters during production is the central pillar of this transformation. Bioinformatics methods then help to evaluate the acquired multivariate data by pre-trained and intelligent models to enable active process control and to manufacture in-spec products (*343*). In contrast to pharmaceutical production, raw food is a variable system with numerous degrees of freedom. Due to biological variability, a natural product's composition, size, and shape fluctuate regularly. Local consumer preferences regarding browning state, point of doneness, or formulation pose further challenges and require a high-level automation strategy to adapt process parameters under a data-driven impact for future-directed and sustainable food production.

By spotlighting a molecular decoding of the food metabolome in processing emissions, we presented in chapter 4.3 a method to deliver a wealth of matrix- and process-related biochemical data along with a strategy for on-line mass spectrometric data acquisition. Together with indicative marker compounds proposed in chapter 4.2, these targets can serve as a data basis for "processing 4.0 food production". Even if most published AI-based models rely on proteomics data so far, extending process-related knowledge to small molecules, applications proved to be the most immediate link to a matrix by drawing a comprehensive picture of the food item and its history.

Raw mass spectrometric data has various pitfalls that need to be taken into account by application of *in-silico* processing pipelines and data cleaning. Next to instrumental variability and noise, spectral artifacts are common to mass spectrometric data. By implementing DB-DIpy, presented in chapter 4.4, we showed how the development of new analytical tools goes hand in hand with the co-implementation of associated data-processing tools. Computational solutions for the curation of raw data are an imperative element in providing clean and reliable datasets from complex computational models. There is broad consensus that poor data quality was a primary contributor to the reproducibility crisis in biomedical research (*344*). Evidently, the performance of a computational model is upper-bounded by the goodness of the supplied data (*345*), and thereby, the benefits it can offer to an intelligent production pro-

cess are limited in the same way. Consequently, the development of data curation algorithms, which stand early in the data pipelines of smart food production systems, is a critical aspect to supply valid metabolomics data for the successful training of high-quality models.

Ongoing developments on DBDIpy at the time of writing of this thesis include run-time optimization to enable adduct and in-source detection during data acquisition by an improved data interface. Further, the existing adduct detection algorithm currently relying on MS1 data only will be redesigned to improve adduct detection by incorporating tandem MS data. Consideration of spectral similarity metrics that attribute to defined mass shift of fragment ions as they are caused by, e.g., in-source oxygenation, will provide higher accuracy in adduct detection in future versions. These improvements follow a necessary evolution to reduce the number of false-positive hits and to deliver better insights into the ionization behavior of DBDI in complex mixtures.

The molecular characterizations that were carried out across different food products and processing modalities contribute, next to their descriptive knowledge component, to the perspective to model food-distinctive molecular spaces by bioinformatics methods. So far, AI in cooking surveillance primarily relies on computer vision and optical image recognition. However, these algorithms cannot consider *intrinsic* molecular characteristics of the product (*346*). In-oven molecular sensing could make a valuable contribution to food processing 4.0 models by representing an independent data source or by contributing to robust ensemble models together with computer vision in innovative processing systems.

Especially reactional aspects are poorly covered in public food-related spectral repositories. FooDB currently is the most extensive database gathering information about molecular food constituents and focuses mainly on nutrients in their unprocessed state (*347*). Enriching public repositories by the aspect of processed food would be a great extension of general knowledge permitting to associate detected features with a concrete thermal background.

## 5.6 Outlook and future research directions

The overall analytical concept presented in this thesis offers a knowledge-based contribution to an in-depth molecular description of processing emissions to describe controlled thermal transformations of food. By consolidating innovative sampling strategies, characterization of small molecules, and applied bioinformatic data treatment, a solid base to better comprehend vaporous cooking emissions was set. Nevertheless, open points for future research arise from these findings:

One of the most elemental points of cooking emission research is the question of where the compounds present in the head space were formed. As already stated in section 5.4, the extent of secondary atmospheric reactions is unclear. This raises the question of to what extent a head space profile is modified *a posteriori* to its liberation from the matrix. Clarifying experiments are non-trivial and would require thoughtful spatial and temporal separation of matrix and emissions under ongoing thermal stimulus and progressive analysis of the latter. The classification of food metabolites in the head space proposed in chapter 4.2 could accordingly be extended to three categories: *a) endogenous metabolites* related to the raw matrix, *b) primary* matrix-related thermal reaction products, and *c) secondary* reaction products of both of the previously-named.

Closely related is the question of which of the discussed transport mechanisms, see section 5.3, apply to a compound as a function of its physicochemical properties. Necessary experiments for this open point include matrix-extraction studies with quantitative measurements of compounds or compound families to better understand release dynamics and ratios of organic compounds during cooking processes. Additionally, material-flux modeling would contribute supporting insights to better understand occurring release dynamics.

Ideal physicochemical parameters for a compound to be liberated to the head space could be described based on this data, and a physicochemical cutoff could be defined for which compounds cannot be set free to the atmosphere.

A potential use case for processing emission analysis that should be probed is the determination of reaction kinetics. Detailed knowledge from the previously proposed experiments provided, could on-line monitoring of cooking processing emissions be applied to calculate reaction constants under various processing conditions. These, again, could feed prognostic models which postulate ideal cooking parameters. It was shown that kinetics calculated in liquid reaction models are barely transferable to solid matrices (*129*). Provided the necessary conditions are given, the emissions of cooking processes could be a straightforward medium to reflect the matrix and explore reaction kinetics.

Agriculture and food manufacturing is one of the biggest industries in the world. As a result of a steadily growing world population and demanding climatic developments, food production needs to be optimized to keep up with these challenges.

Since humans began to use fire for the controlled thermal transformation of food, the chemical knowledge about cooking processes and the interfering technological solutions have never been as advanced as today.

Or as Desiere predicted 20 years ago: "The great future for food processing, however, is not in simply processing for greater safety, but in merging biological knowledge of living organisms with the biomaterial knowledge necessary to convert them to foods (*315*)".

Molecular comprehension of thermal processing emissions presented in this thesis brings us one step closer to the point where analytical chemistry will directly affect the quotidian supply of food. Individually tailored, innovative processing solutions will join an evolutionary development of cooking that goes back tens of thousands of years and is now on the onset of a new, modern era.

# A Appendix: Original research articles as a first author

In the following chapter, I present the officially published full-text versions of the research articles I contributed to as a first author for my dissertation. Explicit permission for the reproduction of the articles within this thesis has been granted by all journals or were given under the Creative Commons Attribution License.

Page numbers, layout and citations are conserved from the official publications and do not follow the style of this thesis.

## A.1 Article I: Elucidation of the non-volatile fingerprint in oven headspace vapor from bread roll baking by ultra-high resolution mass spectrometry.

Weidner, L., Hemmler, D., Yan, Y., Rychlik, M. & Schmitt-Kopplin, P. (2022). *Elucidation of the non-volatile fingerprint in oven headspace vapor from bread roll baking by ultra-high resolution mass spectrometry.* Food Chemistry, 2022, Volume 374, Article 131618. doi.org/10.1016/j.foodchem.2021.131618

# Elucidation of the non-volatile fingerprint in oven headspace vapor from bread roll baking by ultra-high resolution mass spectrometry

Leopold Weidner [a,b], Yingfei Yan [b], Daniel Hemmler [a,b], Michael Rychlik [a], Philippe Schmitt-Kopplin [a,b,*]

[a] Comprehensive Foodomics Platform, Chair of Analytical Food Chemistry, TUM School of Life Sciences, Technical University of Munich, Maximus-von-Imhof-Forum 2, 85354 Freising, Germany

[b] Helmholtz Zentrum Muenchen, Analytical BioGeoChemistry, Ingolstaedter Landstr. 1, 85764 Neuherberg, Germany

## ABSTRACT

Untargeted research on vapor arising during the thermal processing of food has so far focused on volatile aroma compounds. In this study, we present an oven atmosphere sampling strategy to trap headspace aerosols along with semi- and non-volatile molecules liberated during the baking of wheat bread rolls. The collected vapor condensate was analyzed for its molecular fingerprinting using direct infusion ultra-high resolution mass spectrometry. We detected up to 4,700 molecular species in a vapor sample from bread rolls baked at 230 °C for 15 min. Beyond the global profiling of the underlying matrix, our method can follow complex reaction cascades during the baking process, such as the formation of advanced glycation end-products like maltosine through the interface of trapped vapor. Further, process parameters such as baking temperature and duration were used to model the dynamic liberation of molecules to the oven atmosphere by a response surface methodology approach.

## 1. Introduction

During production of cereal products, baking of the dough represents the final process step before human consumption. By the application of heat, a multitude of chemical reactions and transformations is triggered inducing huge changes to the dough. The exact elucidation of chemical transformations induced by thermal processing of bakery products thereby is of fundamental interest in analytical food chemistry. During every heating process water rises from the matrix to the headspace of the processing vessel. In 1952, Pence introduced the novel concept of studying this vapor by the principle of physically condensing vapor from bread baking as an additional and easily accessible source of information about processes occurring during cooking. Collected condensate was analyzed by thin-layer chromatography, fractional distillation and identification of functional groups (Pence, 1952). Later, Lee et al. fortified bread dough with ^{14}C labeled sucrose and studied its dispersion in oven vapor by the same sampling approach (Lee & Chen, 1966). These studies were the first untargeted investigations of oven vapor. As analytical food chemistry advanced in the following years, the focus on research of processed food has been on the matrix itself. In the few cases when its headspace was studied during processing, most protocols focused on the analysis of volatile molecules contained in the vapor. In particular, solid-phase microextraction (SPME) coupled to gas chromatography with mass spectrometric detection (GC–MS) (Rega, Guerard, Delarue, Maire, & Giampaoli, 2009; Rochat & Chaintreau, 2005) and proton transfer reaction (PTR) MS (Pico, Khomenko, Capozzi, Navarini, & Biasioli, 2020; Wieland et al., 2012) were used to study aroma active compounds evolving during thermal processing. Both approaches focus on the analysis of volatile compounds. PTR-MS is a direct infusion technique and as such capable of analyzing the oven headspace in real time during baking. In contrast, analytes can be absorbed on a SPME fiber online during baking, but desorption and analysis of these samples by GC–MS is performed after the baking experiment. Nevertheless, it benefits from chromatographical separation before mass spectrometric detection. As Pence already demonstrated that oven vapor is composed of various substance classes beyond volatiles, there currently seems to be a methodological lack of holistic investigations about semi- and non-volatile molecules within this matrix. Based on the original study from Pence, we intend to develop an advanced vapor sampling protocol and to apply it by sampling vapor released from commercial wheat bread

rolls. The collected condensate sample shall be analyzed by direct injection analysis Fourier-transform ion cyclotron resonance mass spectrometry (DIA-FT-ICR-MS), the up-to-date most powerful analytical platform for in-depth metabolomic profiling of complex products (Hertkorn et al., 2008; Marshall, Hendrickson, & Jackson, 1998; Marshall & Rodgers, 2004). Ultra-high resolution mass spectrometry datasets from FT-ICR-MS studies already gave insights into the molecular complexity of foodstuffs like wine, whiskey and beer (Pieczonka, Lucio, Rychlik, & Schmitt-Kopplin, 2020; Roullier-Gall et al., 2014; 2018). Similar to the SPME approach developed by Rega et al. our approach comprises an atline-measurement strategy to fully take advantage of the benefits offered by FT-ICR-MS. In contrast to SPME and PTR techniques both focusing on volatiles, we aim to extend the knowledge about compounds contained in oven vapor to a broader range of analyte polarity and molecular weight by our research. We want to understand the release dynamics of compounds liberated to the oven headspace triggered by thermal causes and to further investigate the information about the underlying food reactome mediated by oven vapor. Based on the profiling, we developed a response surface methodology (RSM) in the second part of this work which permits us to monitor the presence of molecules in the oven headspace as a function of processing parameters such as baking temperature and duration.

## 2. Materials and methods

### 2.1. Materials

Ultra-pure water was supplied by a Milli-Q integral system (Merck, Darmstadt, Germany). Arginine (reagent grade, 98%) was purchased from Sigma Aldrich (Taufkirchen, Germany). Methanol (LC–MS grade) was obtained from Fisher Scientific (Schwerte, Germany); Ethanol (99%) was purchased from Brenntag (Essen, Germany) and dry ice pellets were purchased from Polar (Neustadt an der Donau, Germany).

### 2.2. Food samples

Wheat bread roll samples (approximately 65 g each) were purchased from an industrial bakery (Froneri, Nürnberg, Germany). Formulation as well as nutritional composition of the samples are documented in Supplementary Table 1. The doughs were risen, partially pre-baked and deep-frozen. Only the final baking phase, which is usually performed by the end-customer, was investigated during this study. Doughs were stored at –21 °C immediately until baking.

### 2.3. Baking of bread rolls:

All doughs were baked in a professional oven. Doughs were baked in the "hot-air" instrument-mode. Temperature and baking time ranged around a center point of 200 °C and 12 min, which was optimized for the utilized oven based on the manufacturer's recommendations during pretests. The studied temperature and time values reached from 158 to 242 °C and from 7.8 to 16.2 min (see Supplementary Table 2). Humidity was set to 100%. A total of six doughs per experiment were baked on an ordinary metal sheet at the middle slot of the baking chamber and baked under a medium level of convection. Bread rolls were placed inside the oven after pre-heating and equilibration of the baking chamber. All baking events were performed in triplicate and in randomized order.

### 2.4. Oven vapor sampling:

The oven was daily cleaned and baked out at 300 °C prior to sampling. Void-runs were sampled on a daily basis to monitor possible contaminants in the background. A small glass tube connected to a hose made of polytetrafluoroethylene (PTFE) was inserted into the baking chamber through the sealing of the door. A negative pressure system was used to draw vapor and gases from the inside of the baking chamber towards a chilled condenser-apparatus. The condensate was collected in a round-bottomed flask (see Fig. 1 for a schematic drawing of the entire apparatus). Sample collection was performed cumulatively over the entire baking duration. The pump was operated at a constant speed of 1.3 L min$^{-1}$. Hose and condenser were rinsed with 5 mL of ultra-pure water after each baking experiment to collect remaining analytes. Samples were stored at –80 °C and consecutively lyophilized until complete dryness. Subsequently, residues were reconstituted in 200 μL of ultra-pure water. After incubating in an ultrasonic bath for 10 min, samples were centrifuged at 15,000 rpm for 15 min. Finally, the clear supernatant was diluted with methanol by a factor of 10 ($v/v$) prior to mass spectrometric analysis.

### 2.5. Direct infusion FT-ICR mass spectrometry data acquisition:

Ultra-high resolution mass spectra were acquired in direct flow injection mode on a solariX FT-ICR mass spectrometer equipped with a 12 T superconducting magnet (Bruker Daltonics, Bremen, Germany). The diluted samples were automatically injected by a PAL RTC system autosampler (CTC Analytics, Zwingen, Switzerland) at a constant flow rate of 2 μL min$^{-1}$. Electrospray ionization (ESI) was performed in negative ion mode using an APOLLO II ion source. Source parameters were set as follows: dry gas flow rate 4 L min$^{-1}$ at 180 °C, nebulizer gas flow 2.2 bar (both nitrogen), capillary voltage 3600 V and spray shield voltage –500 V. Spectra were acquired with a time-domain of 4 megawords. Ions were accumulated for 0.4 s per scan and a total number of 200 scans from $m/z$ 122 to $m/z$ 1000 were accumulated per sample. External calibration of spectra was performed with ion clusters of arginine (5 ppm in methanol). The injection order of samples was completely randomized.

### 2.6. Processing of FT-ICR-MS data:

Raw spectra were preprocessed by Compass DataAnalysis 5.0 (Bruker Daltonics, Bremen, Germany). In addition to calibration of the instrument's mass analyzer, internal calibration of spectra with a reference mass list was performed. The reference list was composed of ubiquitous masses like fatty acids on the one hand and of reliably identified masses specific for oven vapor on the other. Internal calibration and further data analysis were performed over a mass range from 122 Da to 600 Da. Ion signals with a signal-to-noise ratio of at least six were exported into mass lists. All further data post-processing and analysis steps were performed in R programming language (R Foundation for Statistical Computing, Vienna, Austria). Only singly charged ions were considered for analysis. Exported mass lists were consecutively aligned into a matrix consisting of averaged $m/z$ values for every detected peak and corresponding intensity values for all samples. The mass list was cleaned up for side bands, heavy isotope peaks and ion-adducts (Kanawati, Bader, Wanczek, Li, & Schmitt-Kopplin, 2017).

Chloride adducts were converted to the corresponding [M−H]$^-$ ions. On average only one $^{34}$S isotope peak was identified per analyzed sample; thus sulfur was assumed to play a minor role in the matrix. Assignment of molecular formulas was therefore restricted to the CHNO space. Based on a mass difference network analysis (Tziotis, Hertkorn, & Schmitt-Kopplin, 2011), molecular formulas for the remaining exact mass values were calculated and assigned to the matrix. Thereby, isobaric compounds cannot be distinguished, as direct flow injection techniques do not permit the separation of isomers. Only features which were present in at least two out of three sample replicates were considered for annotation and further analysis.

### 2.7. Response surface methodology:

To investigate the relative abundance of compounds released into the headspace during bread baking at different processing conditions, a fractional factorial experimental plan with a central composite compo-
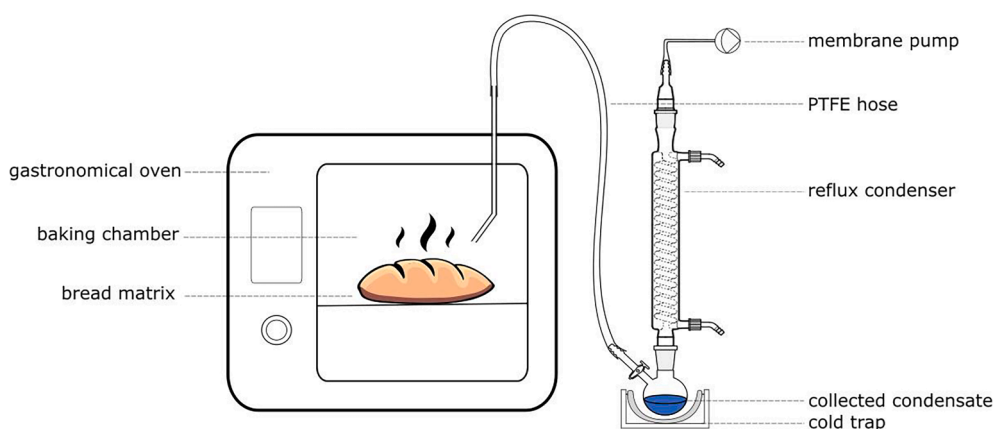
**Fig. 1.** Schematic drawing of the sampling apparatus. Oven vapor is drawn by a negative pressure system out of a baking chamber and condensed in a round bottom flask for later analysis.

nent was set up. Baking temperature and baking time were chosen as experimental factors to investigate. Their uncoded levels ranged from 158 to 242 °C and from 7.8 to 16.2 min, respectively. A total number of 21 factor combinations within these ranges were investigated. Information about coding of factor levels for all baking experiments is given in Supplementary Table 2. Prior to RSM analysis, missing values within the matrix were filled by a Random Forest (RF) imputation strategy (Stekhoven & Bühlmann, 2012). Subsequently, we performed response surface modelling for every feature in the matrix to describe its relative concentration in the headspace as a function of variable processing parameters. Initially, second order polynomial models (see equation (1)) were fitted to the feature's intensities acquired at different temperature (T) and time (t) pairings.

$$\widehat{Y} = \widehat{\beta_0} + \widehat{\beta_1}T + \widehat{\beta_2}t + \widehat{\beta_{12}}Tt + \widehat{\beta_{11}}T^2 + \widehat{\beta_{22}}t^2 \qquad (1)$$

A full second order model consists of first order (FO), two way interaction (TWI) and pure quadratic (PQ) term parts. Variable selection and optimization of models were carried out based on comparative likelihood-ratio tests using the Akaike information criterion (AIC) (Akaike, 1974). The AIC was used in an automated script to estimate the prediction errors of every possible composition of a model by stepwise leave-one-out validation of single term parts. The model-version yielding the lowest prediction error, and thus the lowest AIC-score, was regarded as most accurate. Furthermore, only those models were accepted which had an adjusted $R^2 > 0.6$.

### 2.8. Validation of models

Baking processes at two further, randomly selected processing conditions were used to assess the prediction capacity of the calculated models. Precisely, a random number generator was used to create two temperature–time pairings within the covered factor ranges which are subsequently addressed as validation points (VP). Bread rolls were baked, sampled and analyzed at these points in triplicates similar to the 21 sample points. Intensities of features measured at the VP were not included in the calculations of the models. For each feature-specific model, the measured intensities at these VP were compared to the model's predicted values. Therefore, the standard deviation of a model's residues was calculated. Then, it was checked if the prediction of a model deviated less than two standard deviations from the actual intensity recorded in both VP samples. Only models which validly predicted both VP were regarded as resilient.

### 3. Results and discussion

#### 3.1. Cold trapping steam from baking processes

We constructed a simple and robust sampling apparatus targeted for trapping oven vapor. A schematic drawing is shown in Fig. 1. The glass parts of the apparatus are easily accessible and permit careful cleaning between sampling runs to avoid carry-over. The glass tube introduced into the baking chamber as well as the attached PTFE hose were subject to relatively heavy pollution as they conducted the steam-assisted sample transfer and thus were exchanged on a regular basis. Although we did focus on one professional oven, the apparatus could be used to sample any other baking system with similar door sealings without the need of installing permanent modifications on the system. By operating the membrane pump at a constant speed of 1.3 L min⁻¹ we conveyed between approximately 2 and 8 mL of condensate per baking experiment at a constant load of six bread rolls per experiment. With increasing baking temperature from 170 to 242 °C and baking duration from 9.0 to 16.2 min, the amount of collected liquid increased likewise.

Sample collection was performed cumulatively over the entire baking run and the full amount of collected liquid was used for analysis. The cooling bath and the reflux condenser were filled with ice water, at a temperature of 4 °C, to support condensation of the conducted steam and to preserve sample constitution prior to storing the sample at –80 °C. We additionally tested the option to cool the round-bottomed flask with a cold mixture consisting of ethanol and dry ice at –72 °C. Under this condition, the volume of the trapped condensate did not increase compared to cooling with wet ice in water, so we considered the latter as a sufficient cooling medium for our experiments. As the pumped steam warms up the cooling bath during sampling, colder trapping temperatures might nevertheless be beneficial and should be considered in the case of (i) very hot processing conditions, (ii) sampling large volumes of vapor, (iii) extensive sampling runs, (iv) targeted sampling of labile analytes or (v) targeted sampling of semi-volatile analytes. Recently, there have been advances in investigations about the volatile components of oven vapor (Ait Ameur, Rega, Giampaoli, Trystram, & Birlouez-Aragon, 2008; Rega et al., 2009). The referred authors applied a similar sampling procedure: they likewise pumped vapor out of the baking chamber of an oven towards a refrigerated extraction chamber. Analytes were therein adsorbed on solid-phase microextraction (SPME) fibers and subjected to gas chromatography analysis with various detection systems. However, during these studies the condensate which was still containing semi- and non-volatiles was discharged after SPME. The presented online SPME analysis yields excellent results in the monitoring of volatiles such as furfurals, short-chained aliphatic acids and aldehydes or pyrazines that evolve during the baking process. Further,

related analytical work on oven vapor was executed by proton transfer reaction PTR-MS (Pico et al., 2020; Wieland et al., 2012). In PTR-MS volatile organic compounds (VOC) are ionized by transferring a proton from $H_3O^+$ ions to analytes having a greater proton affinity than water. Whereas PTR-MS analysis of baking processes can be performed online and in real-time, it operates under the limitations of ionizing uniquely molecules with proton-accepting functionalities. FT-ICR-MS on the contrary is favored by using universal electro spray ionization which can be operated in positive and negative ion mode and performs versatile in the analysis of complex organic mixtures (Hertkorn et al., 2008). The findings based on PTR-MS and SPME-analysis of oven vapor strongly emphasize the opportunities offered by oven vapor sampling and encourage us to extend the field of research toward semi- and non-volatile compounds.

### 3.2. Molecular complexity of oven vapor during bread roll baking:

We obtained molecular fingerprints of oven vapor by electrospray (–) FT-ICR-MS with a primary interest on semi- and non-volatile molecules. After aligning mass lists and applying data pre-treatment filters, the matrix for the entire dataset of 63 collected condensate samples consisted of >6,700 features, each representing a monoisotopic mass signal. An original spectrum, as well as an example crop showing the distinct fingerprinting pattern of six bread rolls baked at 230 °C for 15 min are shown in Fig. 2a/**b**. In the 0.16 Da section, 33 ion signals (S/N > 6) were detected, already illustrating the high density of information contained in oven vapor. For this spectrum, a resolution of 740,000 at *m/z* 200 and of 410,000 at *m/z* 400 was achieved. The assignment of molecular formulas to the matrix was restricted to a ± 100 ppb error window. Of the assigned matrix, 77% of all annotated formulas could be found within -50 ppb < error < +50 ppb (see Fig. 1c). As opposed to the entire matrix, we detected a total of 4,738 annotated mass signals in the condensate

collected from six bread rolls baked at 230 °C for 15 min. This refers to the final number of features after application of pre-filtering algorithms and replicate filtering. A second sample arising from six bread rolls baked for the same time but at 170 °C was determined to contain only 580 detectable features. Comparing all annotated ion signals in both samples, we observed a median molecular mass of 319 Da (170 °C, 15 min) opposed to 366 Da (230 °C, 15 min) under the analytical conditions applied (see points 2.4–2.6). Thus, increasing baking temperature by 60 °C triggers an eightfold change in feature number with significantly higher median molecular mass in the oven headspace (Wilcoxon Rank test, W = 1,817,740, $p \leq 0.0001$, $n = 5,318$ features).

We further observed a change in the chemical composition of the samples: whilst the 170 °C condensate showed a CHO:CHNO ratio of 0.7, the rate decreasing to 0.5 in the 230 °C sample. The inter-sample ratios of feature numbers assigned to the CHO, $CHN_1O$ and $CHN_3O$ roughly followed the eight-fold change of total feature numbers, whereas there are 17 times more compounds assigned to the $CHN_2O$ space in oven vapor from bread rolls baked at 230 °C for 15 min. From both trends we conclude a favored entry of nitrogen into the CHO space by rising baking temperature (see Fig. 3e for visualization). Likewise, the evaluation of characteristic bulk descriptors such as double bond equivalents, aromaticity index and average H/C-rate altogether confirm a higher degree of unsaturation among features detected at higher baking temperatures. Variation of baking time was found to have less impact on the number of features, molecular mass and chemical composition than baking temperature. Next to these two samples, which we exemplarily discussed here, we conducted experiments at 19 further baking temperature-duration combinations for our RSM calculations. These vapor samples follow the same trends as discussed above (see Supplementary Fig. 1). The compositional differences between both samples are cross-plotted by Van-Krevelen and respective H/C- to *m/z*-plots (see Fig. 3a–d): Van-Krevelen plots are a versatile, graphical tool for the
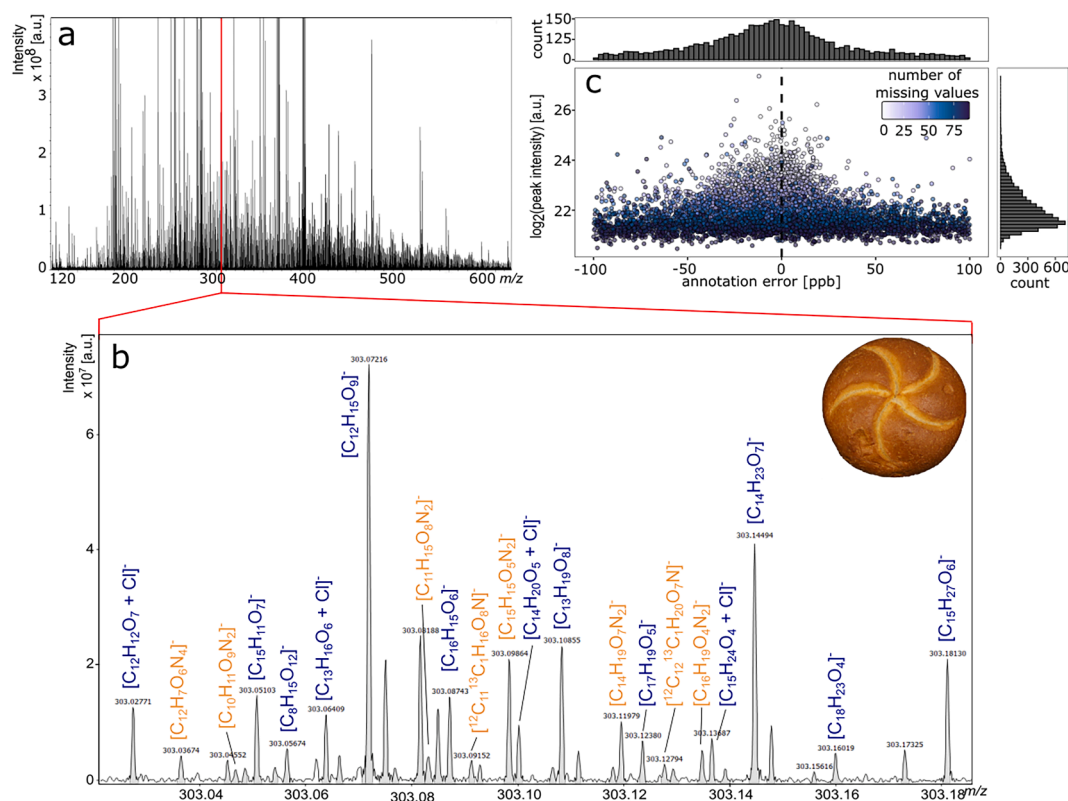


**Fig. 2. a** Complete FT-ICR-MS spectrum of a vapor sample from six bread rolls baked at 230 °C for 15 min. **b** We detected a total of 33 mass signals, of which two are ^{13}C heavy isotope peaks, in a 160 mDa crop of the spectrum at *m/z* 303. **c** Annotation error plot of the aligned matrix with histograms showing their distributions. A total of 77% of all annotated formulas range within the sub 50 ppb range. Point color refers to the detection frequency of the feature.
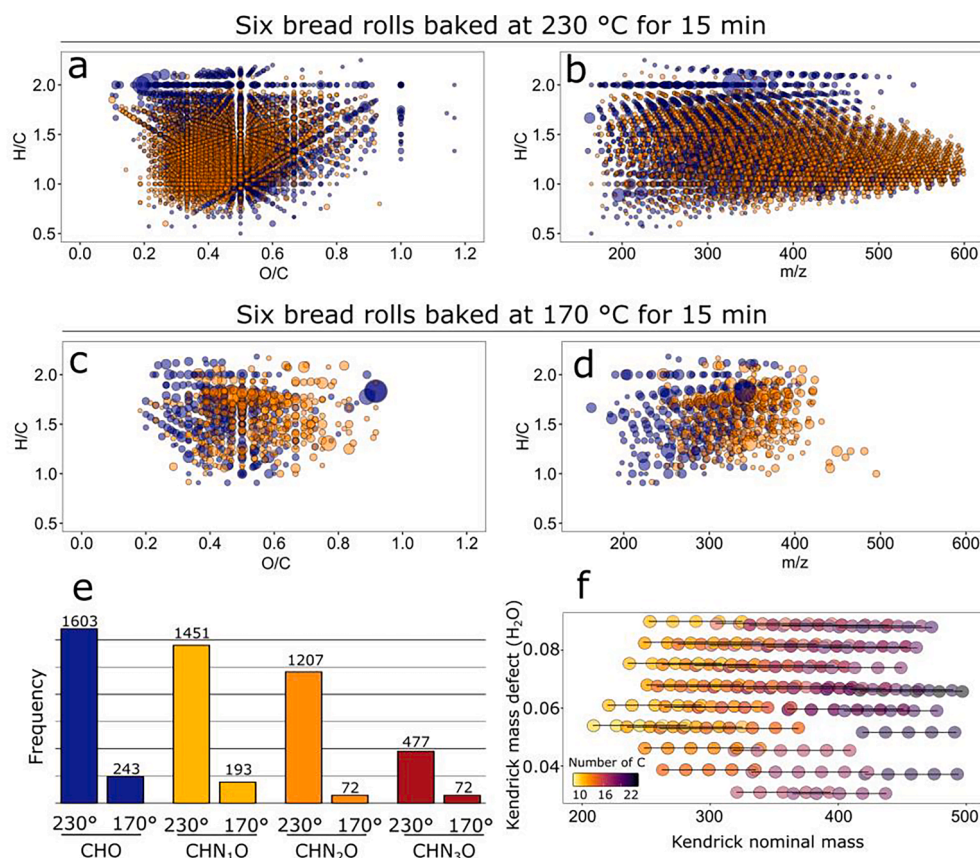
ARTICLE IN PRESS

L. Weidner et al.                                                                                    Food Chemistry xxx (xxxx) xxx



**Fig. 3.** Compositional overviews of oven vapor by Van Krevelen plots and H/C to $m/z$ plots of six bread rolls each baked at 230 °C for 15 min (**a,b**) and at 170 °C for 15 min (**c, d**) (point size proportional to feature intensity). **e** Composition of chemical spaces in both samples. **f** Visualization of dehydration transitions by normalizing IUPAC mass to a water Kendrick Mass scale ($m/z$ * 18/ 18.010565). Series with the same Kendrick Mass Defect and same number of Carbon atoms are connected by straight lines. Only series with at least five consecutive water mass difference increments are shown.

visualization of untargeted ultrahigh-resolution mass spectrometric datasets (Hertkorn et al., 2008; Kim, Kramer, & Hatcher, 2003). Cross-plotting hydrogen:carbon (H/C) and oxygen:carbon (O/C) ratios of all annotated molecular formulas helps illustrating and characterizing the chemical diversity of the vapor samples by associating detected ions with compound classes. Further, straight lines connecting multiple points on the plot can represent chemical reactions, such as (de)hydrogenation, (de)methylation, oxidation or reduction and others. Likewise, cross-plotting of H/C and corresponding $m/z$ values disclose the molecular weight distribution of semi- and non-volatile compounds we detected in oven vapor. At this point, the chemical diversity of oven vapor becomes apparent: the detected features spread over a wide range of polarity and molecular mass. Further, one can visually perceive the impact of increased baking temperatures which leads to the previously discussed larger number of detected compounds and to compounds being present in different sections of the plots. Substance classes such as sugars, amino acids or fatty acids naturally settle into distinct regions in Van-Krevelen-plots as a matter of their molecular composition. All of the above named regions are well populated in samples from bread roll vapor. As data from FT-ICR-MS studies only permit assignment of molecular formulas to exact mass values, we performed further LC-MS/MS experiments to cover some of the annotations with concrete structures (see Supplementary Methods for details). Representative, non-volatile compounds for all of the substance classes named above, e.g., glucose, lysine and stearic acid were confirmed at level one (Sumner et al., 2007) to be present in our samples (see Supplementary Table 4 for further selected and identified compounds). Confirmation of single substances at level one is tedious and only feasible if standard substances are available. However, on one hand, the identifications provide evidence about the presence of semi- and non-volatile compounds in oven vapor. On the other hand, it allows us to demonstrate how the FT-ICR-MS serves as a powerful tool to quickly characterize the molecular

fingerprinting of vapor samples. Through our holistic approach, the entity of the chemical system contained in the headspace is revealed and global response trends to altered processing conditions become apparent. Besides the naturally occurring molecular families named previously, Maillard reaction products (MRP) compose an important and well-studied compound class in cereal products which is induced by thermal treatment (Pozo-Bayón, Guichard, & Cayot, 2006). Consistently Rega et al. reported large quantities of several volatile MRPs to be present in oven vapor released from sponge cake (Rega et al., 2009). As recently shown, discrete compositional areas typical for MRP and caramelization products (CP) were identified based on Maillard model systems (Hemmler et al., 2018). We thrived to connect these findings towards the analysis of food samples. As the Maillard reaction (MR) initiates by the condensation of an amino acid with a carbohydrate, several starting points for the MR cascade on the Van-Krevelen canvas are given as a function of the molecular composition of the amino-moiety. In the course of the MR, intermediate species migrate towards lower H/C and O/C (respectively higher unsaturation) levels. Late stage MRPs such as heterocycles or melanoidins, finally settle towards the bottom left corner of the Van-Krevelen plot (Hemmler et al., 2018). We observed a magnitude of mass signals crowding distinct Maillard-related areas of the Van-Krevelen plot. Further we recognized a clear trend that application of higher baking temperatures pronounces MRP and CP areas at a higher density. Thus, we can confirm the findings of Rega et al. who described a major influence of the MR in oven vapor. As the MR is a reaction sequence enrolling from clearly non-volatile amino acids and carbohydrates, this finding exemplarily demonstrates how FT-ICR-MS analysis of oven vapor is crucial to gain insights about non-volatile compounds contained in this matrix. Another useful tool for the interpretation of ultra-high resolution mass spectra are Kendrick mass defect (KMD) plots. For KMD analysis, molecular building blocks, such as traditionally $CH_2$ in petroleomics, are converted to an integer value to

identify series of homologous species in complex organic mixtures (Kendrick, 1963; Kim et al., 2003). In an adapted KMD analysis we normalized the IUPAC mass of $H_2O$ to a Kendrick mass scale (IUPAC mass $\times$ 18/18.010565). By plotting the Kendrick mass against the Kendrick mass defect of water, the dehydration series are projected onto horizontal lines (see Fig. 3f). In a vapor sample from bread rolls baked at 230 °C for 15 min we detected 37 of such series which show at least five incremental water transitions. These patterns reveal further traces of reaction cascades from (oligo)saccharides towards MRPs and CPs and can correspondingly be observed in studies on melanoidin formation and elucidations of Maillard model systems (Bruhns, Kanzler, Degenhardt, Koch, & Kroh, 2019; Hemmler et al., 2017). We detected dehydration cascades reaching up to a molecular mass of nearly 500 Da in our data set. Certainly FT-ICR-MS gives insight into heavier molecular spaces of oven vapor than accessible before through SPME–GC–MS and PTR-MS studies. For better readability, Van-Krevelen-plots with corresponding H/C to *m/z*-plots grouped into several compositional space can be found in Supplementary Fig. 2. As our baking processes took place in a humid atmosphere, we assume a complex system of transport reactions to take place between dough and its headspace, which are not only limited to the evaporation of hydrophobic volatiles. The presence of steam in the oven's headspace plays an important role and favors the transport of semi- and non-volatile molecules. When Rega et al. studied vapor from sponge cake baking by SPME–GC–MS, they explicitly mentioned the presence of the rather hydrophilic flavor compound 5-hydroxymethylfurfural (HMF) (Rega et al., 2009). Regarding the octanol/water partition coefficient, HMF has a log *P* value of –0.6 (all reported log *P* values according to XLogP3 3.0, PubChem release 2021.05.07). Our FT-ICR-MS analysis of oven vapor now unveils compounds with even higher polarity such as arginine (log *P* = –4.2) or maltose (log *P* = –4.7; both confirmed by LC–MS/MS, see Supplementary Material) to be present in oven vapor. Substances as these need to be transported towards the condensation apparatus by steam as they are not volatile themselves. Further, we confirmed the presence of palmitic and stearic acid (log *P* = 6.4 and 7.2 respectively) to be conveyed by vapor which demonstrates how FT-ICR-MS characterizes oven vapor over a wide range of analyte polarity. For the future, baking experiments with different steam settings are scheduled to investigate the relation between steam content of the oven atmosphere and the polarity of compounds contained in the headspace of a baking chamber. These findings depict a substantial extension of analyte polarity and molecular weight compared to species reported by SPME and PTR studies of oven vapor. In conclusion, we demonstrated how FT-ICR-MS validly characterizes the chemodiversity of oven vapor. It allows the following of global chemical trends triggered by food processing without losing focus in potentially misleading details.

### 3.3. Response surface analysis of molecules in oven vapor

Apart from characterizing the global composition of oven vapor, the scope of this study was to investigate the dynamic release of compounds during food processing to the oven headspace. Therefore, we sampled vapor from six bread rolls each baked at 21 different permutations of temperature and time (see Supplementary Table 2 for exact values). As we followed a cumulative vapor collection strategy, altered baking times in combination with modified baking temperatures yielded different quantities of trapped compounds during the experiments. By fitting second order polynomials to the variable feature intensities (see method section), we describe the features' signal intensity in oven vapor samples as a function of the processing parameters temperature and time. To perform reliable RSM calculations, we only considered features which were present in at least two out of three replicates of a vapor sample and in at least 15 out of 21 experimental temperature time combinations. We were able to calculate models for 991 features with an adjusted $R^2$ value > 0.6. Reflections upon response surface models of selected features in oven vapor permit us to follow the release of molecules during the

cooking of foodstuff similar to Pico et al. who classified VOC based on the shape of their liberation curves during baking and toasting (Pico et al., 2020). Two easily interpretable examples are shown in Fig. 4. We first assessed the conduct of the feature annotated as $C_6H_{12}O_6$ (LC–MS/MS analysis identified glucose to be present in the vapor sample; see Supplementary Table 4). With rising baking temperature, we observed that the oven vapor contains a maximum of glucose, especially after longer processing times. A basic reaction in carbohydrate chemistry are dehydrations reactions. The initially formed hexose dehydration products are key intermediates in early phase Maillard and caramelization chemistry. We monitored this first dehydration product of glucose in oven vapor to demonstrate the traceability of an initial downstream reaction cascade (Fig. 4a). The dehydration products reach their maximum concentration during extensive and hot baking runs. Dividing the recorded intensity of the dehydration products by the one of glucose, the molar proportions become apparent which are lowest at 170 °C and come to a maximum at 224 °C. In the same way as we did for a monosaccharide, we monitored the conduct of the disaccharide maltose under the same conditions (Fig. 4b): at 170 °C a larger quantity of disaccharide was detected in oven vapor, the amount however diminishes with increasing temperature. The effect seems to be independent from the applied processing time. As maltose's relative concentration decreased, its dehydration products were complementarily formed, favorably under hot baking temperatures. A maximum ratio is reached at 242 °C. Initially formed carbohydrate dehydration products such as deoxyglucosones are known to be highly reactive compounds as their dicarbonyl structure is prone to nucleophilic attacks (Jost, Henning, Heymann, & Glomb, 2020). As our samples were cold trapped and frozen right after the baking process and our FT-ICR-MS system is capable of detecting ultra-traces of molecules through the accumulation of multiple scans per sample, we assume that we are able to detect a certain share of unreacted deoxyglucosones in the vapor. By direct comparison of both systems, we observed glucose to accumulate in vapor with higher baking temperature, whereas maltose simultaneously degrades. Recent results on quantitation of carbonyl compounds during industrial bread making processes indeed show that the concentration of glucose rises during baking as oligo- and polysaccharides are degraded to monosaccharides, first by enzymes and during later baking phases by thermal breakdowns. In contrast, after initial enzymatic reproduction of maltose, the major disaccharide in wheat bread, comes to an end, its concentration decreases from the pre-baking phase on with rising temperature, especially during late baking phases (Jost et al., 2020). Both processes are in agreement with our findings for the relative concentrations in the headspace of baked bread rolls. This demonstrates how FT-ICR-MS analysis of oven vapor can serve as a steam-assisted interface to the matrix. Non-volatile compounds released to the headspace of an oven reflect chemical reactions enrolling in the processed matrix. However, it should be noted that different isomers of an annotated molecular formula do not necessarily need to follow the same release pattern as they might be formed in different reaction pathways or even from different precursors. Keeping this bottle-neck in mind, we propose FT-ICR-MS as a sensitive tool for the at-line monitoring of chemical reactions in the underlying matrix.

Evaluation of our response surface models allows us to further understand the background of single molecular components, as the carbohydrates glucose and maltose demonstrate, within vapor as a part of the entire reactome of baked bread: we recognized at which conditions a compound reaches its maximum concentration in the headspace. Again, experiments for clarifying the release dynamics of non-volatiles to the oven headspace are to be performed to understand the underlying transport processes and the role of humidity in the system. The molecular modelling may help to optimize food production. On the one hand, enhancing the formation of compounds which mediate valued sensorial properties helps to increase the consumers' acceptance of a product. On the other hand, the formation of substances with a negative physiological background can be avoided. As oven vapor offers the possibility to
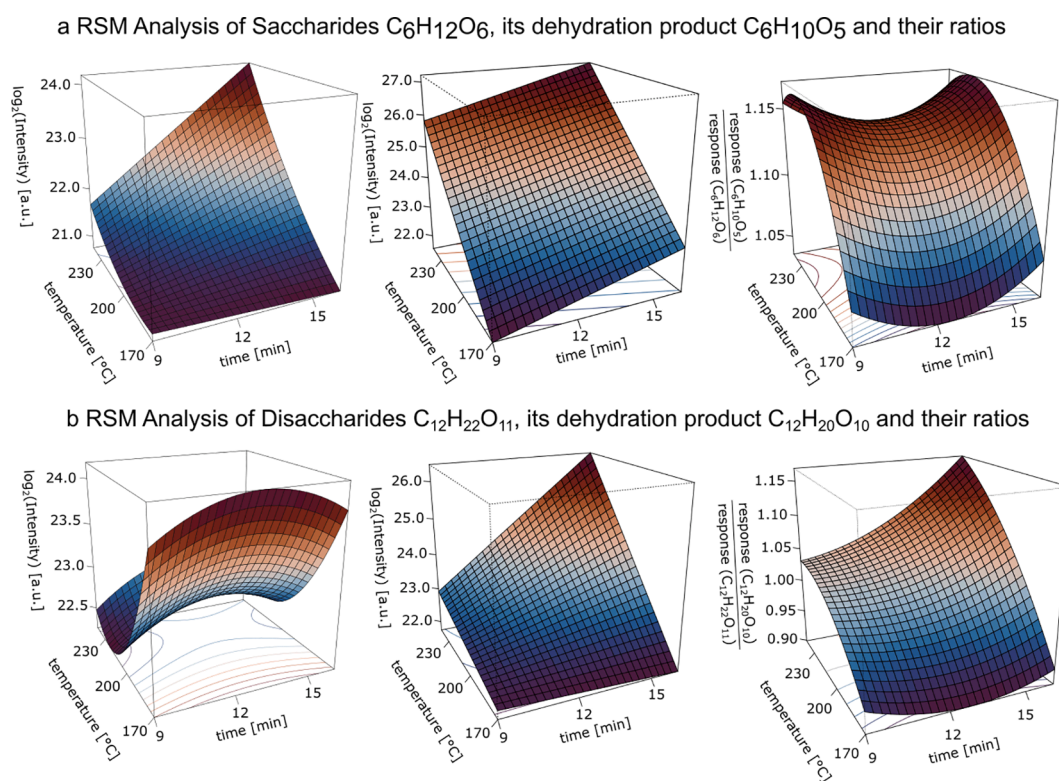
*L. Weidner et al.*

a RSM Analysis of Saccharides $C_6H_{12}O_6$, its dehydration product $C_6H_{10}O_5$ and their ratios

b RSM Analysis of Disaccharides $C_{12}H_{22}O_{11}$, its dehydration product $C_{12}H_{20}O_{10}$ and their ratios

**Fig. 4.** 3-Dimensional plots showing the conduct of the mono- and disaccharide features, their dehydration products and their respective ratios in oven vapor resolved for different baking temperatures and times.

monitor compounds of interest directly in the headspace during the cooking process, extraction protocols for these compounds from the matrix after the baking process become obsolete and we see promising advances for online process supervision by the analysis of oven vapor.

### 3.4. Tracking chemical reactions in oven vapor:

As previously mentioned, we observed a substantial number of features related to the MR processes in our FT-ICR-MS screening of oven vapor. As early as in the 1950's, Hodge et al. categorized the MR into three time-resolved phases of which the last phase describes the formation of compounds with high molecular weight, high degrees of aromaticity and unsaturation along with brown color (Hodge, 1953). This class of compounds is widely termed as advanced glycation end-products (AGE) and is currently under intense research by (bio)

medical and food scientists (Goldberg et al., 2004; Hellwig, Kühn, & Henle, 2018; Kislinger et al., 1999; Ravichandran, Lakshmanan, Raju, Elangovan, Nambirajan, Devanesan, & Thilagar, 2019). In the previous section, we presented RSM models of compounds associable with early phase MR processes. However, FT-ICR-MS can follow complex reaction cascades from precursors *via* intermediates towards the formation of AGEs through accessing the information about non-volatile compounds contained in oven vapor. As an application of our oven vapor analysis in combination with response surface modelling, we describe in the following the formation of the AGE maltosine (6-(3-hydroxy-4-oxo-2-methyl-4(1*H*)-pyridin-1-yl)-l-norleucine). Maltosine develops in bread during baking by the addition of lysine to intermediately formed glucosyl isomaltol, a downstream product of maltose (Hellwig, Kiessling, Rother, & Henle, 2016). Fig. 5a visualizes the course of maltosine formation in our bread roll samples (230 °C, 15 min) through the
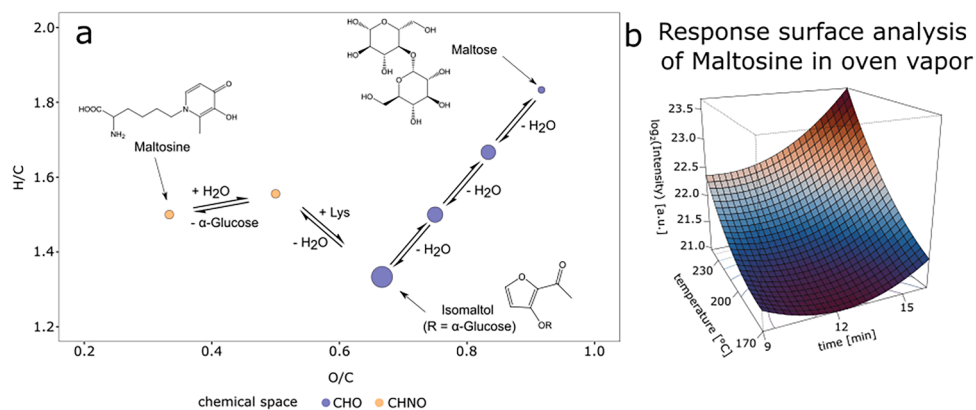
**Fig. 5. a** Traces of the Maillard reaction in oven vapor: the formation of maltosine from maltose and lysine (all identified by MS/MS, see Supplementary Materials and Methods, Supplementary Table 2 and Supplementary Fig. 3) can be followed over all intermediates (point size proportional to feature intensity). **b** Response surface analysis of maltosine in oven vapor from 21 different baking temperatures and durations.

interface of oven vapor in a Van-Krevelen-plot. It was shown that maltosine is an AGE evolving from disaccharides (Pischetsrieder & Severin, 2005), so we observed the initial series of dehydration reactions from maltose, which leads to the formation of glycosyl isomaltol, the most abundant feature in the reaction series. Following, condensation with lysine and cleavage of the glucose moiety leads to the formation of maltosine. We further monitored the evolution of maltosine by response surface modelling (Fig. 5b) as a function of baking temperature and time. During our baking experiments maltosine was formed at highest concentrations when baking at 230 °C for 15 min. Hellwig and co-workers showed that maltosine is predominantly formed in the bread crust up to 19.3 mg/kg (Hellwig et al., 2016). Therefore, it appears reasonable that the non-volatile maltosine is transported from the crust by steam during baking and collected in our sampling device. MRPs play an important role in various research disciplines, including flavor chemistry (Demyttenaere, Tehrani, & De Kimpe, 2002), food safety (Capuano, Ferrigno, Acampa, Ait-Ameur, & Fogliano, 2008) or biomedical research (Poulsen et al., 2013). Maltosine, for example, is known for its metal-chelating properties because of its 3-hydroxy-4-pyridinone structure (Geissler, Hellwig, Markwardt, Henle, & Brandsch, 2011). Further, as a covalent modification of the essential amino acid lysine, its formation alone reduces the biological value of food protein. Studies have shown that maltosine can modify up to 0.4% of lysine (Hellwig et al., 2016). Naturally, most of these modifications are covalently bound to larger peptides and proteins and, as a consequence, we would not expect their transport through steam. Nevertheless, the high sensitivity of the FT-ICR-MS screening not only can detect but also can actively follow the evolution of unbound maltosine through vapor without the need for extraction of the baked dough. In comparison to other oven vapor screenings discussed above, monitoring of entire reaction cascades is hardly feasible as both SPME–GC–MS and PTR-MS fail to detect the highly hydrophilic precursors of the reaction and are constrained to analyze volatile reaction end-products only.

### 3.5. Validation of RSM models:

The calculated models rely on measured feature intensities on a grid of pre-selected processing parameters. To verify that our models yield accurate response predictions, we randomly assigned two further VP within the studied factor ranges (VPI at 223 °C and 10 min, VPII at 180 °C and 14 min). We decided that the deviation between the validation records and respective model predictions must be of the same magnitude as the other model specific prediction errors. We calculated the standard deviation of the 63 datapoints' residuals (21 experimental factor combinations sampled as triplicates) for every model. In a following step we checked if the prediction by our RSM calculations ranged within two standard deviations from the value recorded under both validation conditions. We found 96% of all models to meet this criterion. Only 38 predictions were excluded as invalid by the validation experiments. Initially, we set high quality criteria for features to be selected for modelling based on replicate filtering and later only considered models with adjusted $R^2$ values > 0.6. Further we proved that based on carefully collected data, the prediction of relative concentrations of features in specified oven vapor systems is feasible. Thus, the transport of molecules from wheat bread rolls towards the oven headspace above underlies clearly structured and non-random processes which can be mathematically modelled and predicted.

### 4. Conclusion

By investigating trapped oven vapor arising from wheat bread roll baking with DIA-FT-ICR-MS, we showed that this matrix is a system rich in chemical information. Its chemodiversity reveals insights into the reactome that evolves during food processing with a focus on semi- and non-volatile compounds. Vapor-assisted transport processes give access to several thousands of compounds from the matrix over a broad range

of molecular mass and polarity. Thus, our sampling strategy offers a promising perspective towards at-line sampling and consecutive analysis of cooking processes with a focus on semi- and non-volatile molecules. Besides a holistic description of the molecular entity, we used the contained chemical information for monitoring the evolving and depletion of selected features in the headspace of a professional oven under different processing parameters. Even the complex formation pathway of the AGE maltosine in bread rolls can be followed in oven vapor. Supplementary LC–MS/MS confirmed the identity of selected compounds throughout the entire reactome. We further showed response surface studies to be a useful tool to optimize baking conditions. In a consecutive step, we aim to compare the baked dough itself and its emitted vapor as well as the influence of the humidity content in the oven atmosphere on transport processes. This interlink shall help to understand the relations between new compounds formed in the food matrix and their liberation into the headspace. Altogether we see promising applications in analyzing oven vapor for semi- and non-volatile compounds in (bio-) medical, flavor and food sciences.

### CRediT authorship contribution statement

**Leopold Weidner:** Conceptualization, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Yingfei Yan:** Methodology, Writing – review & editing. **Daniel Hemmler:** Conceptualization, Funding acquisition, Methodology, Project administration, Software, Writing – review & editing. **Michael Rychlik:** Resources, Supervision, Writing – review & editing. **Philippe Schmitt-Kopplin:** Conceptualization, Funding acquisition, Supervision, Project administration, Writing – review & editing.

### Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgement

### Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodchem.2021.131618.

### References

Ait Ameur, L., Rega, B., Giampaoli, P., Trystram, G., & Birlouez-Aragon, I. (2008). The fate of furfurals and other volatile markers during the baking process of a model cookie. *Food Chemistry, 111*(3), 758–763. https://doi.org/10.1016/j.foodchem.2007.12.062

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Bruhns, P., Kanzler, C., Degenhardt, A. G., Koch, T. J., & Kroh, L. W. (2019). Basic structure of melanoidins formed in the maillard reaction of 3-deoxyglucosone and γ-aminobutyric acid. *Journal of Agricultural and Food Chemistry, 67*(18), 5197–5203. https://doi.org/10.1021/acs.jafc.9b00202

Capuano, E., Ferrigno, A., Acampa, I., Ait-Ameur, L., & Fogliano, V. (2008). Characterization of the Maillard reaction in bread crisps. *European Food Research and Technology, 228*(2), 311–319. https://doi.org/10.1007/s00217-008-0936-5

Demyttenaere, J., Tehrani, K. A., & De Kimpe, N. (2002). The chemistry of the most important maillard flavor compounds of bread and cooked rice. *ACS Symposium Series, 826*(Figure 1), 150–165.

Geissler, S., Hellwig, M., Markwardt, F., Henle, T., & Brandsch, M. (2011). Synthesis and intestinal transport of the iron chelator maltosine in free and dipeptide form.

*European Journal of Pharmaceutics and Biopharmaceutics, 78*(1), 75–82. https://doi.org/10.1016/j.ejpb.2010.12.032

Goldberg, T., Cai, W., Peppa, M., Dardaine, V., Baliga, B. S., Uribarri, J., & Vlassara, H. (2004). Advanced glycoxidation end products in commonly consumed foods. *Journal of the American Dietetic Association, 104*(8), 1287–1291. https://doi.org/10.1016/j.jada.2004.05.214

Hellwig, M., Kiessling, M., Rother, S., & Henle, T. (2016). Quantification of the glycation compound (maltosine) in model systems and food samples. *European Food Research and Technology, 242*(4), 547–557. https://doi.org/10.1007/s00217-015-2565-0

Hellwig, M., Kühn, L., & Henle, T. (2018). Journal of Food Composition and Analysis Individual Maillard reaction products as indicators of heat treatment of pasta — A survey of commercial products. *Journal of Food Composition and Analysis, 72*(March), 83–92. https://doi.org/10.1016/j.jfca.2018.06.009

Hemmler, D., Roullier-Gall, C., Marshall, J. W., Rychlik, M., Taylor, A. J., & Schmitt-Kopplin, P. (2017). Evolution of complex Maillard chemical reactions, resolved in time. *Scientific Reports, 7*(1), 3–8. https://doi.org/10.1038/s41598-017-03691-z

Hemmler, D., Roullier-Gall, C., Marshall, J. W., Rychlik, M., Taylor, A. J., & Schmitt-Kopplin, P. (2018). Insights into the chemistry of non-enzymatic browning reactions in different ribose-amino acid model systems. *Scientific Reports, 8*(1), 1–12. https://doi.org/10.1038/s41598-018-34335-5

Hertkorn, N., Frommberger, M., Witt, M., Koch, B. P., Schmitt-Kopplin, P., & Perdue, E. M. (2008). Natural organic matter and the event horizon of mass spectrometry. *Analytical Chemistry, 80*(23), 8908–8919. https://doi.org/10.1021/ac800464g

Hodge, J. E. (1953). Dehydrated foods, chemistry of browning reactions in model systems. *Journal of Agricultural and Food Chemistry, 1*(15), 928–943. https://doi.org/10.1021/jf60015a004

Jost, T., Henning, C., Heymann, T., & Glomb, M. A. (2020). Comprehensive analyses of carbohydrates, 1,2-dicarbonyl compounds and advanced glycation endproducts in industrial bread making. *Journal of Agricultural and Food Chemistry, 69*(12), 3720–3731. https://doi.org/10.1021/acs.jafc.0c07614

Kanawati, B., Bader, T. M., Wanczek, K. P., Li, Y., & Schmitt-Kopplin, P. (2017). Fourier transform (FT)-artifacts and power-function resolution filter in Fourier transform mass spectrometry. *Rapid Communications in Mass Spectrometry, 31*(19), 1607–1615. https://doi.org/10.1002/rcm.7940

Kendrick, E. (1963). A mass scale based on $CH_2 = 14.0000$ for high resolution mass spectrometry of organic compounds. *Analytical Chemistry, 35*(13), 2146–2154. https://doi.org/10.1021/ac60206a048

Kim, S., Kramer, R. W., & Hatcher, P. G. (2003). Graphical method for analysis of ultrahigh-resolution broadband mass spectra of natural organic matter, the Van Krevelen diagram. *Analytical Chemistry, 75*(20), 5336–5344. https://doi.org/10.1021/ac034415p

Kislinger, T., Fu, C., Huber, B., Qu, W., Taguchi, A., Yan, S. D., … Pischetsrieder, M. (1999). N($\varepsilon$)-(carboxymethyl)lysine adducts of proteins are ligands for receptor for advanced glycation end products that activate cell signaling pathways and modulate gene expression. *Journal of Biological Chemistry, 274*(44), 31740–31749. https://doi.org/10.1074/jbc.274.44.31740

Lee, C. C., & Chen, C.-H. (1966). Studies with Radioactive Tracers IX. The Fate of Sucrose-14C during Breadmaking. *In Cereals and Grains Association*.

Marshall, A. G., Hendrickson, C. L., & Jackson, G. S. (1998). Fourier transform ion cyclotron resonance mass spectrometry: a primer. *Mass Spectrometry Reviews, 17*(1), 1–35. https://doi.org/10.1002/(SICI)1098-2787(1998)17:1<1::AID-MAS1>3.0.CO;2-K

Marshall, A. G., & Rodgers, R. P. (2004). Petroleomics: the next grand challenge for chemical analysis. *Accounts of Chemical Research, 37*(1), 53–59. https://doi.org/10.1021/ar020177t

Pence, E.A. (1952). *A Study of baking oven vapors*. Kansas State University of Agriculture and Applied Science.

Pico, J., Khomenko, I., Capozzi, V., Navarini, L., & Biasioli, F. (2020). Real-time monitoring of volatile compounds losses in the oven during baking and toasting of gluten-free bread doughs: A PTR-MS evidence. *Foods, 9*(10), 1498. https://doi.org/10.3390/foods9101498

Pieczonka, S. A., Lucio, M., Rychlik, M., & Schmitt-Kopplin, P. (2020). Decomposing the molecular complexity of brewing. *NPJ Science of Food, 4*(1), 1–10. https://doi.org/10.1038/s41538-020-00070-3

Pischetsrieder, M., & Severin, T. (2005). In *Maillard Reactions in Chemistry, Food and Health* (pp. 37–42). Elsevier. https://doi.org/10.1533/9781845698393.2.37.

Poulsen, M. W., Hedegaard, R. V., Andersen, J. M., de Courten, B., Bügel, S., Nielsen, J., … Dragsted, L. O. (2013). Advanced glycation endproducts in food and their effects on health. *Food and Chemical Toxicology, 60*, 10–37. https://doi.org/10.1016/j.fct.2013.06.052

Pozo-Bayón, M. A., Guichard, E., & Cayot, N. (2006). Flavor control in baked cereal products. *Food Reviews International, 22*(4), 335–379. https://doi.org/10.1080/87559120600864829

Ravichandran, G., Lakshmanan, D. K., Raju, K., Elangovan, A., Nambirajan, G., Devanesan, A. A., & Thilagar, S. (2019). Food advanced glycation end products as potential endocrine disruptors: An emerging threat to contemporary and future generation. *Environment International, 123*(December 2018), 486–500. https://doi.org/10.1016/j.envint.2018.12.032.

Rega, B., Guerard, A., Delarue, J., Maire, M., & Giampaoli, P. (2009). On-line dynamic HS-SPME for monitoring endogenous aroma compounds released during the baking of a model cake. *Food Chemistry, 112*(1), 9–17. https://doi.org/10.1016/j.foodchem.2008.05.028

Rochat, S., & Chaintreau, A. (2005). Carbonyl odorants contributing to the in-oven roast beef top note. *Journal of Agricultural and Food Chemistry, 53*(24), 9578–9585. https://doi.org/10.1021/jf058089l

Roullier-Gall, C., Boutegrabet, L., Gougeon, R. D., & Schmitt-Kopplin, P. (2014). A grape and wine chemodiversity comparison of different appellations in Burgundy: Vintage vs terroir effects. *Food Chemistry, 152*, 100–107. https://doi.org/10.1016/j.foodchem.2013.11.056

Roullier-Gall, C., Signoret, J., Hemmler, D., Witting, M. A., Kanawati, B., Schäfer, B., … Schmitt-Kopplin, P. (2018). Usage of FT-ICR-MS metabolomics for characterizing the chemical signatures of barrel-aged whisky. *Frontiers in Chemistry, 6*(2), 1–11. https://doi.org/10.3389/fchem.2018.00029

Stekhoven, D. J., & Bühlmann, P. (2012). Missforest-Non-parametric missing value imputation for mixed-type data. *Bioinformatics, 28*(1), 112–118. https://doi.org/10.1093/bioinformatics/btr597

Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., … Viant, M. R. (2007). Proposed minimum reporting standards for chemical analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics, 3*(3), 211–221. https://doi.org/10.1007/s11306-007-0082-2

Tziotis, D., Hertkorn, N., & Schmitt-Kopplin, P. (2011). Kendrick-analogous network visualisation of ion cyclotron resonance Fourier transform mass spectra: Improved options for the assignment of elemental compositions and the classification of organic molecular complexity. *European Journal of Mass Spectrometry, 17*(4), 415–421. https://doi.org/10.1255/ejms.1135

Wieland, F., Gloess, A. N., Keller, M., Wetzel, A., Schenker, S., & Yeretzian, C. (2012). Online monitoring of coffee roasting by proton transfer reaction time-of-flight mass spectrometry (PTR-ToF-MS): Towards a real-time process control for a consistent roast profile. *Analytical and Bioanalytical Chemistry, 402*(8), 2531–2543. https://doi.org/10.1007/s00216-011-5401-9

## A.2  Article II: Molecular Characterization of Cooking Processes: A Metabolomics Decoding of Vaporous Emissions for Food Markers and Thermal Reaction Indicators

Article

# Molecular Characterization of Cooking Processes: A Metabolomics Decoding of Vaporous Emissions for Food Markers and Thermal Reaction Indicators

Leopold Weidner,* Jil Vittoria Cannas, Michael Rychlik, and Philippe Schmitt-Kopplin*

Cite This: https://doi.org/10.1021/acs.jafc.3c05383

Read Online

ACCESS | 📊 Metrics & More | 📖 Article Recommendations | 🗒 Supporting Information

**ABSTRACT:** Thermal processing of food plays a fundamental role in everyday life. Whereas most researchers study thermal processes directly in the matrix, molecular information in the form of non- and semivolatile compounds conveyed by vaporous emissions is often neglected. We performed a metabolomics study of processing emissions from 96 different food items to define the interaction between the processed matrix and released metabolites. Untargeted profiling of vapor samples revealed matrix-dependent molecular spaces that were characterized by Fourier-transform ion cyclotron resonance—mass spectrometry and ultra-performance liquid chromatography—mass spectrometry. Thermal degradation products of peptides and amino acids can be used for the differentiation of animal-based food from plant-based food, which generally is characterized by secondary plant metabolites or carbohydrates. Further, heat-sensitive processing indicators were characterized and discussed in the background of the Maillard reaction. These reveal that processing emissions contain a dense layer of information suitable for deep insights into food composition and control of cooking processes based on processing emissions.

**KEYWORDS:** food processing, metabolomics, head space sampling, Maillard reaction, FT-ICR—MS, UPLC—MS/MS

## ◼ INTRODUCTION

Thermal food processing is a fundamental aspect of culinary practices and plays an essential role in our daily lives. The application of heat transforms food from ordinary raw materials into palatable dishes.[1] During the cooking process, a myriad of chemical reactions are triggered, leading to the formation of countless thermal reaction products. Despite the immense importance attributed to these thermal reaction products, the molecular spaces generated during the cooking process are strikingly understudied. It is known that different groups of consumers prefer to cook the same food to a different extent. Especially cultural habits lead to individual preferences, for example, regarding the browning of bakery products[2] or the preferred doneness of steak.[3] As there rarely is a globally accepted standard manufacturing procedure for food, analytical chemistry helps to capture the effects of different food formulations and understand the influence of selected processing temperatures on the chemical reactions in the matrix to provide food according to the consumer's desire.

There are considerably fewer analytical studies about compounds emitted during the cooking process of food than about the cooked food itself, even if head space analysis of cooking emissions offers substantial advantages: compared to extraction-based analysis of food, vaporous emissions offer the possibility of extracting chemical information in real-time during the production process as they are an easily accessible byproduct.[4] In the context of Industry 4.0, where automated production processes play an increasing role in food production,[5] dynamic measurement of processing emissions can be used to save time and improve process control and quality assurance. Further, as extraction solvents become

obsolete, vapor analysis contributes to greener and more sustainable chemistry.[6]

Most analytical works on thermal processing emissions study volatiles lost during cooking processes to the environment. Central points of contemporary headspace-oriented research apply sorbent-trapping of volatiles to analyze aroma-active compounds,[7−9] real-time monitoring protocols by proton transfer reaction time-of-flight mass spectrometry (MS),[10,11] and indoor air quality analysis.[12−14]

Additionally, we recently showed that water evaporating during cooking processes also conveys a layer of non- and semivolatile compounds to the head space, which contain information about the reactome of the processed matrix.[15] This vapor can be used as a direct interface of the matrix to model even complex chemical reactions, such as the generation of advanced glycation end products (AGE).[15] Indeed, many food chemistry studies focus on the investigation of thermal processes by targeting reaction systems such as lipid peroxidation (LPO) and the Maillard reaction (MR) to gain mechanistic insights. In practice, e.g., Hellwig et al. assess the impact on heat treatment of pasta products by targeted measurement of several AGE,[16] and Hu et al. specifically use the MR to develop time—temperature markers in reheated

food like soymilk powder.[17] Therefore, well-studied reactive systems such as the MR and LPO represent a logistic connection to access and define unexplored molecular spaces in matrices, such as processing emissions.

Up to our knowledge, there is no study dedicated yet to concatenating food formulations and different processing parameters with an untargeted screening of cooking head spaces. Therefore, we see that a huge proportion of information contained in cooking vapor currently remains unexploited. By applying (ultra)high-resolution MS foodomics profiling of cooking vapor from 96 different food items, we present an extensive molecular atlas of thermal food processing emissions in this study. By defining matrix-related molecular spaces in the head space, we advance the understanding of processing emissions from a molecular perspective. We focus on the molecular discrimination between plant- and animal-based food sources, food formulations and compositions, and their emitted head space profiles. We further elucidate temperature-indicating features under different processing conditions and in the context of the MR. The distinctive response of diverse processing markers to temperature changes helps decode cooking processes on a molecular level from a head-space-centered perspective.

## ■ EXPERIMENTAL SECTION

**Materials.** L-Amino acids arginine, lysine, glutamate, aspartate, leucine (all >97%), and D-glucose (99.5%), ammonium formate solution (10 M BioUltra), and ammonium acetate solution (5 M BioUltra) were purchased from Sigma-Aldrich (Taufkirchen, Germany). Acetonitrile, 2-propanol, and methanol (all LC−MS grade) were supplied from Merck (Darmstadt, Germany). Ultrapure water was provided by a Milli-Q integral system (Merck, Darmstadt, Germany). Formic acid (for MS) was purchased from Honeywell (Seelze, Germany). An API-TOF Reference Mass Solution Kit, Infinity Lab Deactivator Additive, and ESI-L Low Concentration Tuning Mix were supplied by Agilent Technologies (Waldbronn, Germany).

**Food Samples.** Raw food samples were purchased from local retail stores around Munich, Germany, preferably in deep frozen condition. Food samples were stored at −21 °C until cooking. Fresh food samples were purchased directly before analysis. Selected samples were of plant and animal material. More precisely, the animal-based foods studied were poultry and red meat cuts and fish filets. Plant-based foods studied were bakery and pastry products, vegetables, and plant-based meat alternatives (PBMA). Supporting Information Table S1 gives an overview of the analyzed samples, including their formulation, if available.

**Collection of Vapor Samples.** Vapor from food cooking processes was collected, as recently described:[15] the raw food samples were cooked in a gastronomic convection oven. An overview of the respective processing conditions is given in Supporting Information Table S1. Thermal processing emissions were sampled in a cumulative way by pumping the oven atmosphere from the baking chamber toward a chilled condenser apparatus. Water vapor, gases, and aerosols were collected in a cold trap held at −72 °C. The pump was operated at a constant speed of 1.3 L min$^{-1}$. Per sample, 25 mL of collected head space condensate was lyophilized to complete dryness in a freeze-dryer (Alpha 1-2 LDplus, Martin Christ GmbH, Osterode am Harz, Germany). Dryness was determined by the fact that further lyophilization did not lead to further reductions in sample weight. Dry residues were reconstituted in 250 µL of ultrapure water by vortexing for 2 min and sonication for 10 min. Samples were successively stored at −21 °C until analysis. Samples were prepared in duplicates and in randomized order. Blank samples of the cleaned and empty baking chamber were collected on a daily basis to account for environmental influences.

**Food Maillard Reaction Model System Preparation.** A set of five proteinogenic amino acids selected by their relevance based on evaluated data and on their natural abundance in beef[18] (as g/100 g protein) were selected to create a simplified food MR model system[19] to investigate the influence of nonenzymatic glycation in the oven head space of cooked beef. An aqueous solution of amino acids was prepared in respective proportions as amino acids are found in raw beef cuts[18] and reacted with 0.2 M glucose at 100 °C in closed glass vials for 8 h. The final concentrations of amino acids in the reaction mixture were 0.1 M glutamate, 0.061 M asparagine, 0.060 M lysine, 0.055 M leucine, and 0.045 M arginine. Reaction controls of all reagents were prepared and heated in parallel: the single analytes were heated without a reaction partner and treated in the same way as the model systems. Thereby, amino acid and carbohydrate degradation products can be identified. To interrupt the reaction, model systems were cooled on ice and diluted by a factor of 100 (v/v) with 90:10 acetonitrile/water (v/v) and analyzed as condensate samples by hydrophobic liquid interaction chromatography (HILIC) electrospray ionization in the positive ion mode [ESI(+)] UPLC−MS/MS. Samples were prepared in duplicates.

**FT-ICR−MS Data Acquisition.** Vapor reconstitutes were diluted by a factor of 100 with methanol (v/v) prior to the measurement on a solariX Fourier transformation ion cyclotron resonance (FT-ICR−MS) instrument equipped with a 12 T superconducting magnet (Bruker Daltonics, Bremen, Germany). Delivery of the samples was performed in the direct flow injection analysis mode with a PAL RTC system autosampler (CTC Analytics, Zwingen, Switzerland) at a flow rate of 2 µL min$^{-1}$. Per sample, a total of 400 scans with an ion accumulation time of 0.25 s were accumulated from $m/z$ 92 to 1000 with a time domain of 4 megawords. Data was acquired in ESI(−) mode on an APOLLO II source (Bruker Daltonics, Bremen, Germany) with source parameters selected as the following: nitrogen was used as dry gas (4 L min$^{-1}$ at 180 °C) and as a nebulizer gas (2.2 bar at 180 °C). Further, a capillary voltage of 3.600 V and a spray shield voltage of −500 V were applied. External calibration of spectra was performed with clusters of arginine (5 ppm in methanol). Samples were injected in a randomized order.

**FT-ICR−MS Data Processing.** Raw spectra were preprocessed in Compass DataAnalysis 5.0 (Bruker Daltonics, Bremen, Germany). First, internal recalibration of spectra was performed against a reference mass list composed of substances reliably identified in oven head space vapor. Successively, the automatic export of ion signals with $S/N > 6$ to mass lists was performed. All further data processing, analysis, and plotting were performed in the R programming language (version 4.2, R Foundation for Statistical Computing, Vienna, Austria) and Python programming language (version 3.10). Mass lists were consecutively filtered for side bands, heavy isotope peaks ($^{13}$C, $^{15}$N, $^{34}$S, and $^{37}$Cl), and ion adducts.[20] Following, singly charged ions were aligned to a feature matrix with a mass error of ±0.5 ppm. Molecular formulas were assigned to the aligned $m/z$-values via an in-house software tool based on mass difference network analysis.[21] Annotation was restricted to the CHNO elemental space and extended to the CHNOS elemental space for samples where a considerable number of $^{34}$S isotopes were detected. Annotation was performed with a maximum mass error of ±200 ppb.

**UPLC−MS/MS Measurements.** Sample reconstitutes were diluted by a factor of 10 (v/v) with 90:10 acetonitrile/water (v/v) for HILIC analysis and with 10:90 acetonitrile/water (v/v) for RP analysis. Samples were analyzed by a 1290 Infinity II UPLC system coupled to a 6560B ion mobility spectrometry quadrupole time-of-flight mass spectrometer (IMS-QqTOF-MS, both Agilent Technologies, Waldbronn, Germany). For HILIC separation, an Infinity Lab Poroshell 120 HILIC-Z column (2.1 × 150 mm, 1.9 µm, Agilent Technologies, Waldbronn, Germany) was used. Analyte separation was performed in the gradient mode. For HILIC−ESI(+) analysis, eluent A consisted of 10 mM ammonium formate in water with 0.1% formic acid and eluent B of 10 mM ammonium formate in 90:10 acetonitrile/water (v/v) with 0.1% formic acid. For HILIC−ESI(−) analysis, eluent A consisted of 10 mM ammonium acetate in water with 2 µM Infinity Lab Deactivator Additive (adjusted to pH = 9.0)
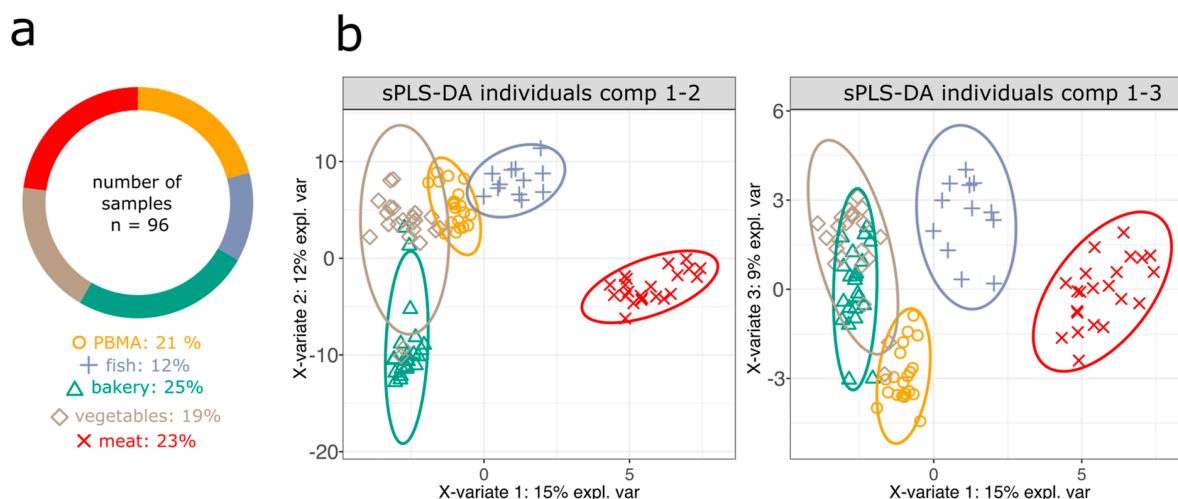
**Figure 1.** (a) Overview of the sample set with sample count by studied food subcategories. (b) sPLS-DA model for discrimination of studied food samples; individuals plotted on dimensions 1−3.

and eluent B of 10 mM ammonium acetate in acetonitrile/water 85:15 (v/v) with 2.5 $\mu$M Infinity Lab Deactivator Additive (adjusted to pH = 9.0). The flow rate was set to 0.25 mL min$^{-1}$. For RP separation, an Infinity Lab Poroshell 120 EC-C8 column (2.1 mm × 150 mm, 2.7 $\mu$m, Agilent Technologies, Waldbronn, Germany) was used. Analyte separation was performed in gradient mode. For RP analysis in the ESI(+) and ESI(−)-mode, eluent A consisted of acetonitrile with 0.1% formic acid and eluent B consisted of water with 0.1% formic acid. The flow rate was set to 0.25 mL min$^{-1}$. For gradient tables, see Supporting Information Table S2. A column equilibration with initial mobile phase conditions during 3 min prior to HILIC analysis and during 2 min prior to RP analysis was performed. Columns were held at 40 °C during analysis. A partial loop injection of 5 $\mu$L of diluted sample was performed automatically. Mass spectra were recorded in the QqTOF-only mode in a range from $m/z$ 50−1700 in data-dependent acquisition at an acquisition rate of 5 Hz. Fragmentation of selected analytes was performed by collision-induced dissociation with a collision energy of 20 eV. Tuning of the ion optics and calibration of the mass analyzer were performed with an Agilent ESI-L Low Concentration Tuning Mix. mass spectra were internally recalibrated by a reference mass solution continuously delivered by a Dual AJS ESI source (Agilent Technologies, Waldbronn, Germany). Source parameters were set as follows: for HILIC-(+) analysis, capillary voltage was set to 3.000 V, nozzle voltage to 0 V, fragmentor to 125 V, and octupole 1 RF voltage to 450 V; dry gas used was nitrogen at 6 L min$^{-1}$ and 225 °C, nebulizer pressure was 40 psi, and the sheath gas used was nitrogen at 10 L min$^{-1}$ and 225 °C. For HILIC-(−) analysis, capillary voltage was set to 3.500 V, nozzle voltage to 0 V, fragmentor to 125 V, octupole 1 RF voltage to 750 V; dry gas was nitrogen at 13 L min$^{-1}$ and 225 °C, nebulizer pressure was 35 psi, and sheath gas used was nitrogen at 12 L min$^{-1}$ and 350 °C. For RP($\pm$) analysis, the capillary voltage was set to 3.500 V, nozzle voltage to 0 V, fragmentor to 400 V, and octupole 1 RF voltage to 750 V. Dry gas used was nitrogen at 13 L min$^{-1}$ and 225 °C, nebulizer pressure was 35 psi, and sheath gas used was nitrogen at 12 L min$^{-1}$ and 350 °C.

**LC−MS/MS Data Processing.** The acquired LC−MS/MS data was converted to open file formats by MSconvert[22] (version 3) and further processed in MZmine[23] (version 3.4.16). In brief, mass detection was performed by using the ADAP chromatogram builder to create mass traces. Feature detection was performed with the local minimum feature resolver. Feature tables were filtered for heavy isotope peaks by the isotopic peak finder, followed by the application of the join aligner module. Missing values were filled with the gap-filling peak finder module. A library search of MS/MS spectra was performed against the MoNA[24] and GNPS[25] libraries. Feature tables were exported to .csv files, and MS/MS spectra were exported to .mgf

files. Metabolites with acquired MS/MS spectra were analyzed by SIRIUS[26] (version 5.6.3) to putatively assign further missing molecular formulas[27] and compound classes.[28,29] Feature-based molecular networks were computed by GNPS[25] (release 30), refined by MolNetEnhancer[30] (release 22), and visualized in Cytoscape[31] (version 3.9.1). Reported metabolites are putatively annotated as level two (unless else mentioned) in the style proposed by Sumner et al.[32] All further data analysis was performed in the R programming language (version 4.2, R Foundation for Statistical Computing, Vienna, Austria) and the Python programming language (version 3.10). The handling of MS/MS spectra was performed by Matchms[33] (version 0.18.0).

**Statistical Analysis.** Statistical analysis was performed in the R programming language (version 4.2, R Foundation for Statistical Computing, Vienna, Austria) and the Python programming language (version 3.10). Multivariate statistical analysis was performed by mixOmics[34] (version 6.1.1) and FactoMineR[35] (version 2.7). Data visualization was performed in ggplot2[36] (version 3.4.0), Complex-Heatmap[37] (version 2.14.0), and ggraph (version 2.1.0).

## RESULTS AND DISCUSSION

**Molecular Classification of Food Processing Emissions.** To expand our knowledge about the chemical composition of oven head space vapor, we convection-cooked 96 raw food samples, partially under multiple processing conditions, and collected their processing emissions. The sample set was designed to depict frequently processed foods consumed in a Western-style diet and to cover a broad chemical range of matrix compositions; see Figure 1a and Supporting Information Table S1 for details on the composition of the sample set. Our untargeted analytical approach unites FT-ICR−MS fingerprinting to profile the chemical complexity of oven head space samples (Supporting Information Figure SI) in combination with UPLC−MS/MS analysis for improved metabolite characterization.

Based on their biological and process-based origin as well as their legal classification, we grouped the samples into the following five food categories: on the one hand, we studied animal-derived foods, subdivided into meat or fish products. On the other hand, plant-derived foods were subgrouped as bakery goods, nonprocessed raw vegetables, and highly preprocessed PBMA. We differentiated the sample set based on mass spectrometric feature intensity detected by HILIC- and RP-($\pm$) UPLC−MS/MS analysis, see Supporting In-
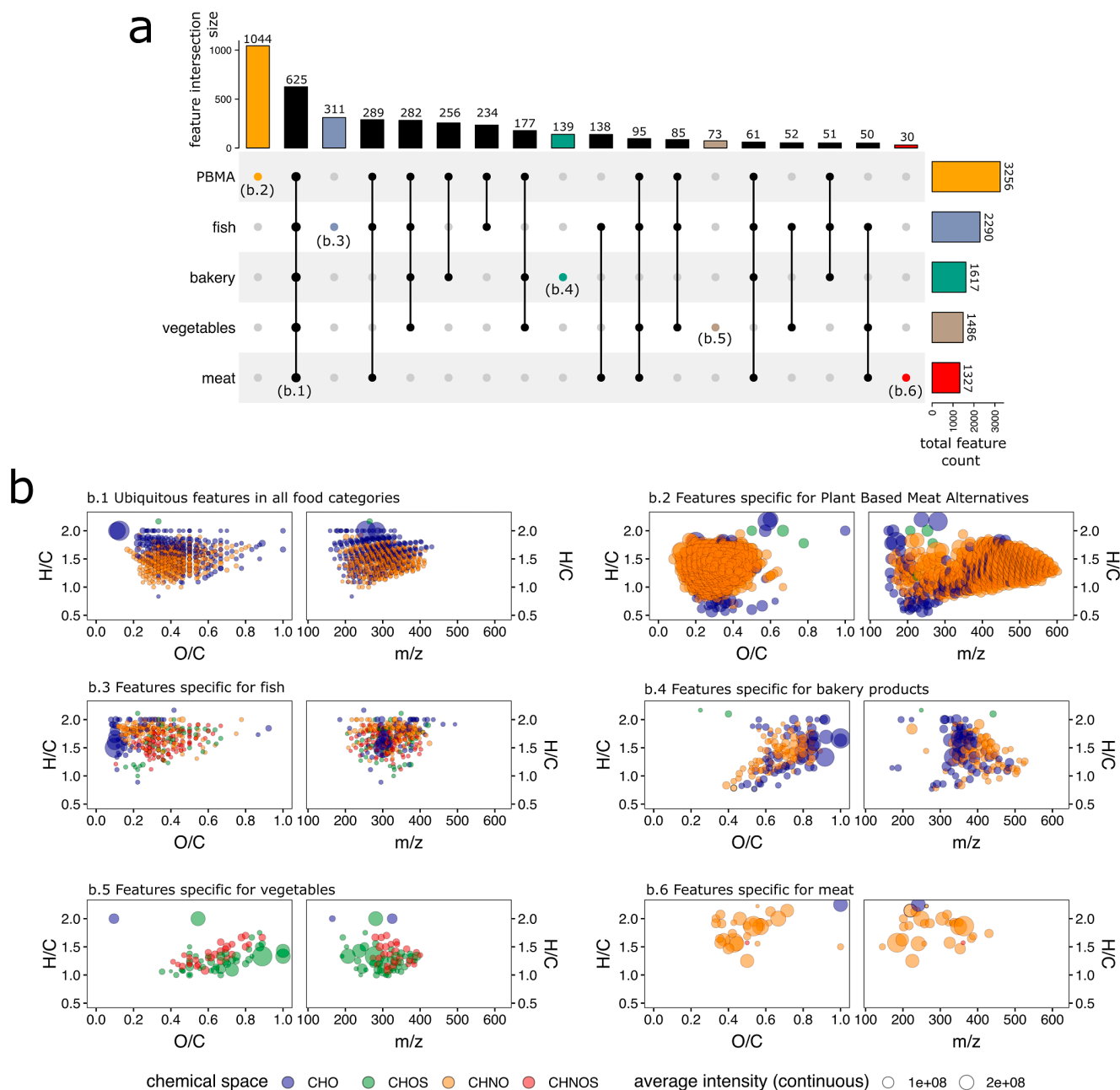
**Figure 2.** (a) UpSet-plot visualizing the number of annotated MS1-features (found in at least 80% of all samples of a food group), group intersections, and unique features by group detected by FT-ICR−MS. (b) Van-Krevelen and respective H/C to $m/z$ plots of features detected by FT-ICR−MS in all studied foods (b1) and detected explicitly in distinct food categories (b2−b6).

formation Figure SII for UPLC−MS feature maps. The principal component analysis of food vapor samples showed solely imprecise differentiation of plant-derived and animal-derived goods on the first three dimensions (Supporting Information Figure SIII). To better probe our classification, we mapped the food categories as the Y matrix of a sparse partial least-squares discriminative analysis (sPLS-DA) for a refined deconvolution of food-specific molecular spaces. The projection of the individual samples can be seen in Figure 1b. The definition of distinct, food-specific clusters can be achieved for both major and further for all five food subgroups in the first three components. Component 1 contributes with a 15% explained variance toward the separation of animal- and plant-based foods. Application of a sparse PLS-DA algorithm allows

us to recognize the most discriminative features between these two groups in the oven headspace, see Supporting Information Figure SIV. Among the most selective features, we found that the presence of leucine, nicotinamide, or carnitine in the oven head space is highly indicative of animal-derived products being processed. These metabolites are abundant in muscle tissue and meet good physicochemical requirements for head space release. By contrast, 5-hydroxymethyl furfural (HMF), maltol, hydroxyphenyl lactic acid, and phenolic compounds such as methoxy salicylic acid and hydroxylated or methoxylated cinnamic acids imply that plant-based foods are being processed. The two earlier-named compounds are linked to carbohydrate-rich food sources, whereas the latter-named compounds are known secondary metabolites in plants

D

and indicate vegetables or PBMA being processed. On a lower level, the clusters of bakery goods, fresh vegetables, and PBMA overlap partially for some samples; however, separation of cluster centers is achieved on components 2 and 3 with 12 and 9% of explained variance, respectively.

We reason that discrimination of plant-based food from the head space without à priori background information about the emitting matrix represents an analytical challenge for three reasons. First, ingredients derived from the same plant can emit the same chemical signature to the head space but are often used as base ingredients for food items of various natures. For example, wheat flour is the base ingredient for wheat bread, heterogeneous pastries, or bread crumb coatings. Therefore, consideration of all food ingredients is important when discussing further differentiation based on head space profiles. Second, therefore, many products which are composed of multiple components might include constituents that strongly coimpact the head space next to the base ingredient. An example from our sample set is natural potatoes versus industrial French fries, where added lipids were found to have a strong impact as free fatty acids were emitted into the head space of French fries. Third, highly processed ingredients often are specialized and remarkably derived to fit product-specific applications. For example, some meat surrogates use wheat protein extruded at high heat and pressure as a base ingredient. Its manufacturing is strongly dependent on the manufacturer's protocols, even if the source for protein might be of the same biological origin, which leads to chemical heterogeneity during preprocessing.

Together, different preprocessing aspects of plant material in combination with complex formulations make the head space classification of plant-based food much more complex than in the case of animal-derived food. Under these aspects, the discriminative model achieves satisfactory separation of all food classes and suggests single features suitable for class distinction.

**Molecular Fingerprinting Defines Chemical Spaces as a Function of the Emitting Matrix.** Next to the characterization on a metabolite-based level, FT-ICR−MS fingerprinting of condensate samples further offers a macroscopic perspective of molecular spaces uniquely found in condensate samples of different food categories. The profiling of all samples revealed 10.600 annotated formulas to be present in oven head space. For further discussion, we considered only features that were present in at least 80% of samples of a food class. We looked for shared features in all classes to construct feature spaces distinctive for single food classes to better understand the influence of matrix effects on the head space. The number of detected features per food group and the intersections between these groups can be seen in Figure 2a. Respective spaces were visualized in Van-Krevelen and H/C to $m/z$ plots, see Figure 2b. Van-Krevelen plots can be used to illustrate untargeted data from ultrahigh-resolution MS experiments: annotated features are cross-plotted by their elemental ratios and mass-to-charge ratio to deconvolute the chemistry of complex matrices in a visually accessible way.[38]

Notably, 625 compounds were omnipresent in the oven head space, independent of the food being processed (Figure 2b1). Within a defined mass range of 200 to 400 Da, these features predominantly consist of relatively saturated compounds with H/C > 1.0. The CHO compound space takes up 43% of the features and mainly consists of saturated fatty acids and oxidized lipids as annotations spread toward higher O/C

rates. Minor numbers of saccharide-like molecular formulas were annotated. The remaining 57% of features fall to the CHNO space and populate spaces in the Van-Krevelen plot being characteristic for amino acids, di- to tripeptides, and their thermal Maillard modification and degradation products. The background of these ubiquitous features in the matrix and their presence in the atmosphere are reasonable. Lipids are essential as building blocks of cell membranes and for energy metabolism and storage. Exposed to thermal energy in an oxygenic atmosphere, they undergo lipid oxidation, leading to intensively studied advanced lipoxidation (end) products (ALE).[39,40] Here, sampling oven head space vapor offers an opportunity for expedient analysis of LPO processes during food processing, which traditionally comes with several challenges.[41] Likewise, amino acids and short peptides can be found in every living organism. We presume processing-related breakdown of larger peptides and associated thermal modification reactions, such as the MR, populate the observed CHNO space. Further, the well-defined chemical frame of these 625 features allows us to draw conclusions about the selectivity of the vapor arising from the processed matrix, which acts as a conveyor for non- and semivolatile compounds toward the oven atmosphere. From the observed molecular composition, we suggest the most favorable release for compounds with a molecular mass from 200 to 400 Da (61% of all annotated mass signals) and a saturation of $1.0 \leq$ H/C $\leq 2.0$ (83% of all annotated mass signals). These intrinsic parameters meet a good compromise between volatility and solubility for vapor-mediated release to the oven head space.

PBMA shows the highest number of unique compounds, as seen in Figure 2b2. Among more than 1.000 features, 89% belong to the CHNO compositional space, which focuses on a restricted range of elemental composition and continuously reaches up to a molecular mass of 600 Da. Modern meat analogs belong to a class of ultraprocessed foods where extrusion techniques restructure plant proteins from various sources under very harsh conditions to mimic fibrous protein structures.[42,43] PBMA, which we investigated in this study, was based on wheat, soy, and pea protein. Considering the various biological sources, the number of shared features seems to be surprisingly high at first sight. However, we conclude that the extreme conditions during the extrusion process, namely, high temperature and pressure, trigger degradation, and thermal modification of plant protein, lead to a uniform population of the spaces seen in the Van-Krevelen plot. The CHO space (11%) ranges from unsaturated fatty acids to a range of phenolic and condensed aromatic substances. Taken together, compounds characteristic for plant-based meat differ remarkably from features that are indicative of animal meat, which stands in accordance with recent comparative metabolic profiling studies of both matrices[44] and according to flavor analysis.[45] Due to huge variations in protein sources and rapid product development in the field of PBMA, future investigations should consider larger sample sets to assess the entire spectrum of this food category. By contrast, we could identify only 30 relatively saturated peptide-like compounds in the oven head space as characteristic features of meat being convection-cooked (see Figure 2b6).

Contrarily, we observed 311 exclusive compounds in vapor arising during the cooking of fish samples. The release of an abundant series of long-chain polyunsaturated fatty acids into the head space is especially noteworthy for this matrix. They originate in the matrix and are a well-studied and nutritionally
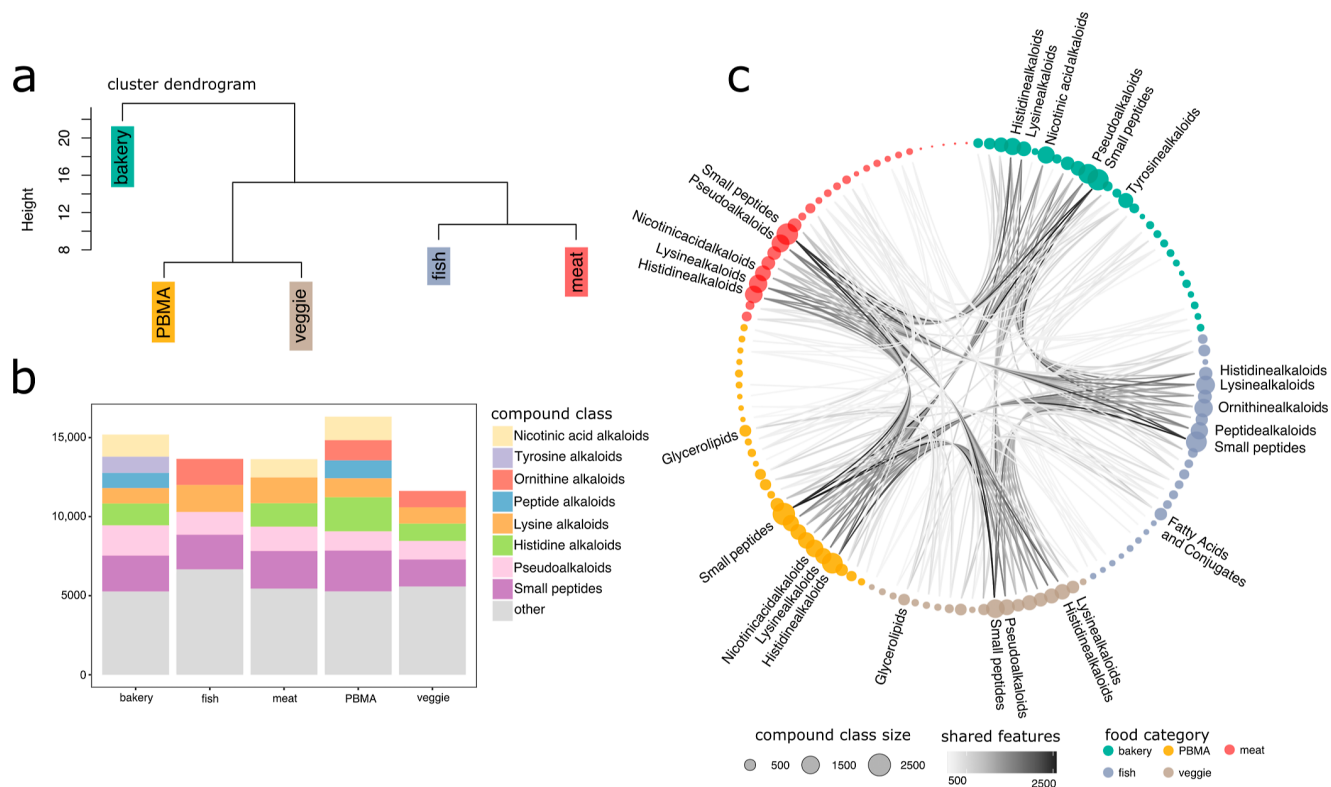
E

**Figure 3.** (a) Hierarchical cluster analysis of NPC compound super classes annotated to tandem MS spectra from cooking emissions. (b) Bar plot of assigned compound class frequencies by the food category. (c) Hierarchical edge graph showing class relations and connectivity (class labels only shown for major classes).

valued characteristic of sea fish.[46] Fish vapor samples further show a prominent signature in sulfurous compound spaces (15% CHOS and 33% CHNOS), which have a common focus at high levels of saturation and a molecular mass of 200 to 400 Da in a region common for peptide-related formulas. Compared to meat, fish vapor was found to show a 10-fold increase in exclusive molecular features and a 1.7-fold increase in total feature numbers. Compared to meat, fish muscle is known for its lower proportion of connective tissue, which makes it more susceptible to disintegration when cooked.[47] Therefore, increased feature counts in the oven atmosphere during the processing of fish are rationalized. This morphological trend continues in the same sense as our observations made on PBMA vapor, where the cellular structure of the matrix is completely ruptured during production, leading to high feature counts in the head space.

The 139 features uniquely occurring in bakery products show a remarkably different chemical background compared to that of the discussed peptide-related spaces. We observe an abundant series of saturated, oxygenated compounds in a carbohydrous space. From there on, reactional trends were described in the background of the caramelization reaction (CR) and MR lead feature series toward the bottom-left corner of the pane.[48,49] Especially dehydration reactions create higher levels of unsaturation, and the introduction of nitrogen by the MR shapes this cluster. These observations directly relate to the matrix, as carbohydrates are a major constituent of bakery goods. Mono- and disaccharides are directly released into the atmosphere, while thermal and enzymatic reactions during the baking process lead to starch degradation to form smaller saccharides. These reducing sugars become available for the reactive pool and for release into the head space. Recently, the impact of the CR and MR in vapor from bakery products was determined.[15]

Vegetable samples generally emit a relatively unabundant chemical profile to the head space due to their elevated water content. Frequently, the dry matter of vegetables only makes up 10−20 weight %.[46] The entire feature space specific for vegetables (97%) belongs to sulfurous compounds; see Figure 2b5. Next to water, carbohydrates are the second most prominent constituent of vegetables,[46] which further restricts the possibility of identifying unique CHO features as these are also abundantly present in other matrices. The CHOS and CHNOS compound spaces offer a reliable option to differentiate between unprocessed vegetables and ultraprocessed PBMA from the perspective of the head space. Sulfurous flavor compounds are of fundamental importance to the aroma of meat. Essential aroma compounds in cooked meat originate from pentose-cysteine MR processes and thiamine degradation.[45,50] We reason that sulfurous compounds are strongly degraded under harsh conditions during the fabrication of restructured plant protein. The thermal decomposition of sulfurous species during processing has been shown and discussed in this context earlier.[51,52] An investigation into gaseous cooking emissions demonstrated that thermal processing releases dimethyl sulfide from vegetables.[53] Up to date, the flavor of PBMA lags behind the flavor of meat.[45,52] Even though plants are known to produce many sulfurous aroma compounds and cysteine, as well as thiamine, which are present in plants,[46,54] our finding contributes to explaining the still-existing differences between meat and PBMA flavor. To better mimic the original taste of
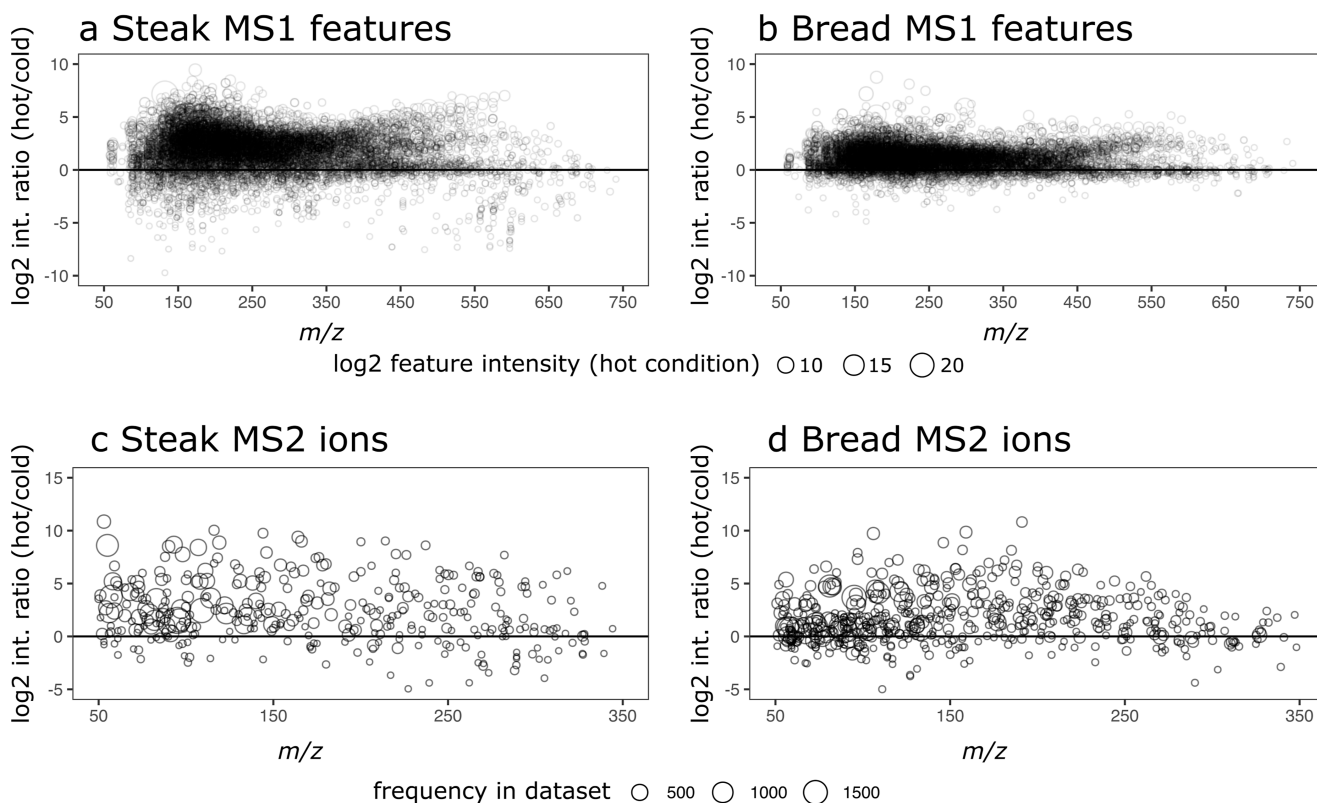
**Figure 4.** UPLC−MS/MS analysis of vapor from rump steaks processed at 210 and 270 °C and wheat bread rolls processed at 200 and 240 °C to visualize temperature dependency of heating markers (a,b) MS1 feature intensity change rates; point size refers to signal abundancy under hot processing conditions. (c,d) MS2 fragment ion intensity change rates, point size refers to the frequency of fragment ion detection.

meat, sulfurous reaction precursors are spiked into PBMA raw material to compensate for this deficiency.[43,45] Analysis of PBMA vapor might help to find better formulations or additives to achieve better consumer acceptance by creating a head space profile closer to meat.

The comparison of molecular patterns emitted by different food types, on the one hand, discloses that distinct chemical spaces with high typicity for a group of foods exist. On the other hand, multiple spaces were noticeably pronounced in the head space of different food types.

To further decode the molecular framework and chemical interplay between food items, we annotated compound classes to our tandem MS fragmentation data by a deep neuronal network classifyer.[28,29] Compared to our FT-ICR−MS-based approach, we here used chromatographically resolved features for detailed insights into the molecular structure. Hierarchical clustering analysis (HCA) of annotated natural product classification (NPC)[29] super classes by food category, see Figure 3a, recognized animal-based and plant-based food to be adjacent based on their NPC spectrum, which confirms the results from sPLS-DA clustering. The bar plot in Figure 3b shows the frequency of assigned NPC super classes. Through all food types, small peptides (less than six units of amino acids) make up the most frequently annotated class. Further, nicotinic acid-, peptide-, and several alpha amino acid-derived alkaloid families frequently occur across all food classes, which is supported by FT-ICR−MS profiling.

In a successive step, hierarchical edge bundling was used to investigate if the same features are responsible for the observed pattern or whether different members of respective compound classes contribute toward this vital expression of nitrogenous

compound classes. Especially small peptides appear to be very homogeneously spread among food classes, with 45% of small peptide features present in all five food groups, see Figure 3c.

The same consistent compound class distribution is perceived for other major compound classes such as peptide, lysine, histidine, and ornithine alkaloids. These evolve from the thermal degradation pathways of the respective peptides and amino acids. As feature numbers per class decrease, specificity for food classes rises, e.g., saccharides are characteristic for bakery products, and fatty acids and conjugates are specific for fish (see Figure 3c in accordance with Figure 2b). This observation suggests that minor compound classes drive the distinction of food categories and matches the results of FT-ICR−MS fingerprinting (Figure 2b), which pinpointed food-specific and miscellaneous feature spaces. These characteristic signatures can originate from two possible sources: first, different transport reactions from the matrix to the head space are known and described in the literature.[55−58] Additionally, secondary chemical reactions between released compounds can further adulterate primary compounds after their initial release.[59,60] Together, both shape a characteristic fingerprint containing molecular information about the chemistry of the matrix and the applied processing conditions. Future extraction-based studies would clarify the relationship between the matrix, corresponding processing emissions, and release processes. Synoptically, we showed that oven vapor contains information suitable for the differentiation of food based on its main component and formulation by suitable computational methods. Even though the studied matrices are of different chemical compositions, oven vapor analysis proves to be a universal and extraction-free method for sampling across

different food matrices. Under the aspect of advanced process control, this knowledge can be used to surveil the quality of production processes, detect fraudulent food, and adapt processes to fluctuating raw-material properties. To our knowledge, we drew an initial picture of the molecular head space profile of PBMA during the cooking process.

**Temperature-Sensitive Features as Thermal Processing Indicators.** To assess the effect of different temperature levels on the processing emissions, we baked wheat bread to a low degree of coloration at 200 °C and to a deep-brown coloration at 240 °C (same processing duration) and convection-cooked rump steak at 210 °C to yield a rare core and at 270 °C to a well-done core (same processing durations).

After vapor collection and UPLC−MS/MS analysis, we found that the feature intensity of compounds released from wheat bread showed a 3.0-fold mean (2.0-fold median) change when increasing the temperature by 40 °C. Increasing the cooking temperature of rump steak by 60 °C led to a 7.5-fold mean (3.7-fold median) increase in feature intensity. As the right-skewed distributions of the intensity ratios suggest, a considerable number of features show a substantially stronger gain in feature intensity as the average factor, see Figure 4a,b. Multiple detected features show a several hundredfold intensity gain during the experiment. We see a promising potential to suggest compounds as head space heating indicators that show high temperature change rates. The most thermoresponsive feature for rump steak, with a 695-fold increase, was annotated as suberic acid. Suberic acid can evolve as a degradation product of oleic acid by LPO[61] and was formerly identified in organic aerosols from charbroiling and meat cooking.[14] Further annotated substances in rump steak vapor with high change rates were pyridoxine (factor 161), leucine (factor 144), methylhistamine (factor 63), prolinamide (factor 47), and valyl proline (factor 34). Remarkably, most of these compounds are small peptides, amino acids, or degradation products of the latter. Their strong abundance supports the hypothesis about the selectivity of head space release processes made earlier and further supports the importance of peptideous compound spaces in oven vapor. Contrarily, features more abundant under cold processing conditions are sparse and can be seen in Figure 4a,b with a log2 change ratio <0. The most distinct feature here is ethanolamine, which was 20 times more abundant under cold processing conditions.

Even after extensive library searches, we could not annotate further meaningful candidates as heating markers for rump steak and none in vapor from wheat bread. The widely known phenomena of uncharacterized spectra, termed "dark matter" in metabolomics,[62] is a critical bottleneck in understanding our untargeted metabolomics data sets. We compared all putatively annotated molecular formulas found by FT-ICR−MS in all food samples to FooDB (version 1.0, accessed June 2023). The comparison revealed only 7.7% of detected metabolites were present in the database. This demonstrates the proportion of unknown data prevailing in metabolomics experiments, along with the need to supply metabolomic data to public repositories. Nevertheless, to gain structural insights into these unannotated features, we also performed a change-ratio analysis of product ions detected during the screening. Therefore, we aligned recorded MS2 scans of the condensate samples to assess their temperature response and frequency of detection. Both matrices show product ions with notable

higher intensity increase rates than detected MS1 features, see Figure 4c,d.

For rump steak, we found that diagnostic fragment ions of the N-heterocyclic aromatic core structure of pyrazines ($m/z$ 53.0134 annotated as $C_2HN_2^+$ and $m/z$ 55.0291 annotated as $C_2H_3N_2^+$) are substantially more abundant and more frequently present under hot cooking conditions. Pyrazines are widely recognized as important flavor compounds of heat-treated meat in the literature.[63] It was shown that they are formed from Strecker degradation of alimentary amino acids, preferably under hot processing conditions.[9] Further, we identified stronger traces of pyrrole core-specific fragments ($m/z$ 80.0495 annotated as $C_5H_6N^+$) under the influence of higher temperatures. Pyrrole cores, known as chromophores, evolve during the MR of amino acids and pentoses by thermal treatment, contributing to a pleasant brown color during cooking.[64]

In rump steak as well as in wheat bread, we found significantly stronger signatures of the imidazole-specific fragment ($m/z$ 81.045 annotated as $C_4H_5N_2^+$) under hot conditions. Imidazole is a diazole core that is under intense study as it is generated in the late phase of the MR and is a commonly found structural motif of AGE. Imidazoles evolve by cross-linking two amino acids, most frequently by condensation of lysine and arginine with a carbonyl moiety.[65,66] For wheat bread, we likewise detected abundant traces of $C_2H_5$-substituted pyrazine cores ($m/z$ 106.0526 annotated as $C_6H_6N_2^+$) and further aromatic, nitrogenous fragment ions ($m/z$ 157.0972 annotated as $C_7H_{12}N_2O_2^+$ and $m/z$ 191.0815 annotated as $C_{10}H_{10}N_2O_2^+$). Earlier studies of vapor from wheat bread showed that application of higher processing temperature triggers abundant nitrogenous signatures in the head space with increasing levels of unsaturation and aromaticity, which is in accordance with the chemical background of the three identified structural motifs.[15] The description of these thermosensitive fragment ions was used as a base for further mechanistic investigations of heating markers in vapor from rump steak.

With this proof of concept, we suggest pursuing the study of specific and highly sensitive thermal processing markers from the head space in favor of extraction-based protocols. Vapor analysis can be performed with time-saving online and at-line setups and contributes toward greener chemistry by saving extraction solvents and material. The evaluation of temperature-sensitive features contained in oven vapor can be used for chemical reaction-guided process control with regard to industry 4.0.

**Maillard Reaction as a Framework for Understanding Thermal Profiles in the Oven Head Spaces.** Based on the results from our temperature change ratio analysis, we found by interpretation of tandem MS fragmentation data that N-heterocyclic aromatic compounds such as pyrazines, pyrroles, or imidazoles, which evolve during the MR, are significantly more strongly expressed in the cooking head space of steak under hot conditions. In 1994, Lothar Kroh found by analysis of different aqueous Maillard model systems that asparagine has the highest contribution to the formation of pyrazines and that lysine is strongly involved in the generation of pyridines and pyrroles.[67] Later, the role of arginine in cross-linking processes was described.[65] These three amino acids are among the most abundantly present in bovine meat.[18] Further, FT-ICR−MS profiling indicated strong metabolic signatures of the MR as well. Therefore, we decided to investigate the
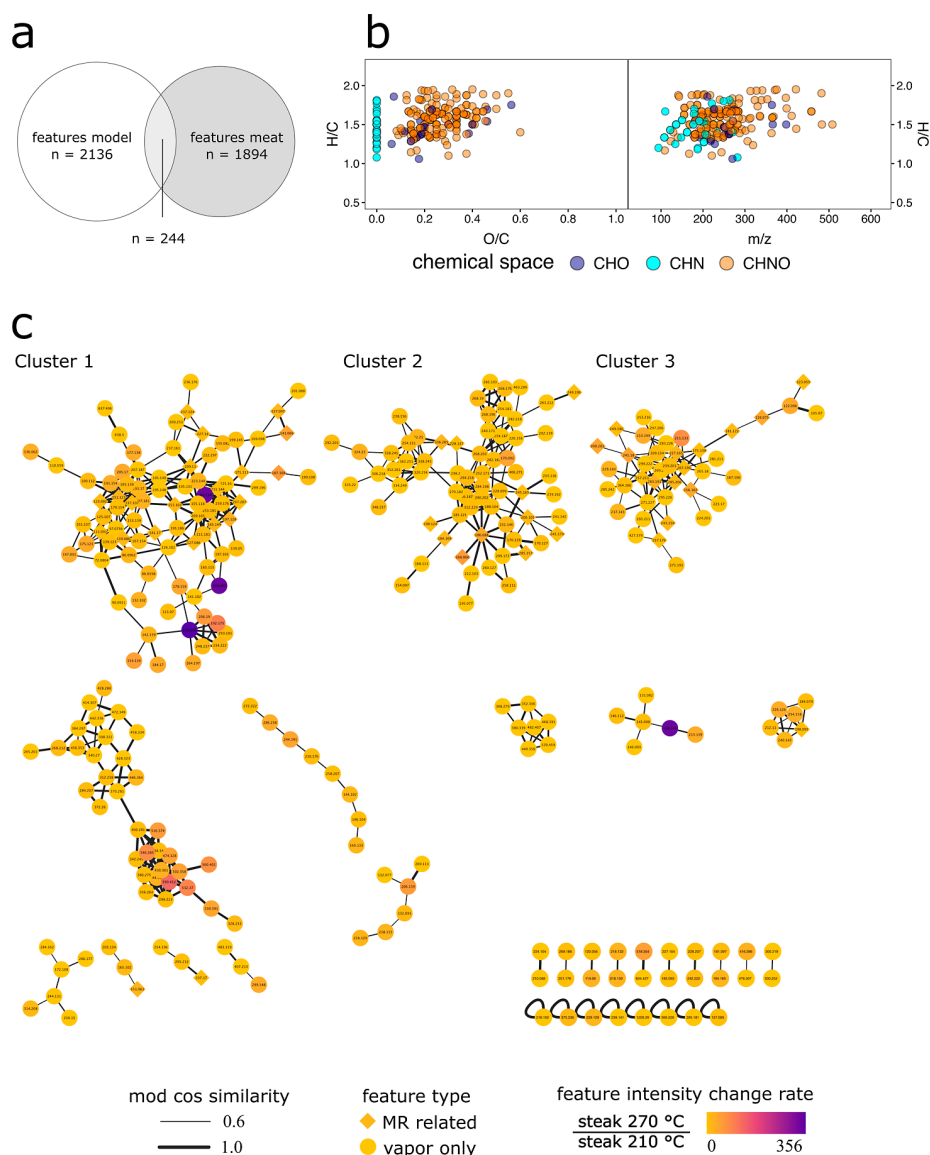
H

**Figure 5.** (a) Number of MS2 features found in vapor from rump steak and a MR model system matched by mass, retention time, and spectral similarity. (b) Van Krevelen and the respective H/C to *m/z* plot of the 244 intersecting features. (c) Enhanced molecular network analysis of features found in vapor from a rump steak. Spectra are clustered by spectral similarity and molecular family. Temperature sensitivity and MR background are indicated.

processing emissions of rump steaks in the background of the MR.

We used an aqueous Maillard model solution (see Food Maillard Reaction Model System Preparation section for the composition) to mimic nonenzymatic glycation processes during the cooking process and to further gain structural knowledge about metabolites in the head space that remained unannotated. We selected arginine, asparagine, and lysine due to their postulated significance and complemented them with leucine and glutamate, which results in a set of the five most predominant amino acids found in bovine meat reacting with glucose. Reactional products were analyzed by HILIC-(+) UPLC−MS/MS under the same conditions as those for beef vapor samples. By incorporating data from this simplified MR model into molecular networking analysis of vapor from rump steak, we introduced seeds of structural knowledge to the study of uncharacterized thermoindicative features in oven vapor.

We analyzed the mixed-model system and the reaction controls under the same conditions. Related to an approach proposed by Yaylayan,[68] we classified amino acid- and carbohydrate fragmentation pools in the control samples. The MR model generated 2.136 features with MS/MS spectra that did not evolve in any reaction control. These were classified as MRP and further refined by feature-based molecular network analysis and fragment-ion search. From 1.894 features with fragment spectra found in the vapor of rump steak cooked at 270 °C, we could augment 244 features (matched by mass, retention time, and spectral similarity) with background information that these were in vitro regenerated by our simplified meat MR model system, see Figure 5a. Respective 244 features were plotted in a Van-Krevelen with H/C- to *m/z* plot to visualize the overlapping chemical spaces of rump steak vapor and of the reaction model, see Figure 5b. By MS2 fragment ion analysis, we identified abundant signatures of pyrazine core structures under hot processing

conditions. The MR model now confirms the generation of 34 CHN pyrazine structures in the vapor sample. A major part of the intersecting features belongs to the CHNO space. These 169 features are initial Maillard reaction products (MRP) and further downstream AGE from the condensation of amino acids and glucose. Additionally, 18 compounds belong to the CHO space. Their absence in the reaction controls is especially noteworthy, as this means that an amine moiety is necessary as a reaction partner for their generation. Amine groups are known to act as catalysts via the formation of Amadori products with carbohydrates in MR pools, which favors $\beta$-eliminations and leads to enolization or isomerization by mitigation of carbonyl groups.[68]

With this background knowledge, we mapped a layer of MR-related information to an enhanced molecular network graph, depicted in Figure 5c. Compounds in rump steak vapor are grouped by their spectral similarity and compound class. The three major clusters are discussed below. Cluster 1, the largest cluster, unites 85 features of substituted benzenes and heterocyclic aromatic compounds. We observed an aggregation of pyrazine and imidazole cores in the cluster. Their formation in the peptide-rich matrix of meat is well studied; various degradation pathways of carbohydrate lead to the formation of reactive $\alpha$-dicarbonyl intermediates. These can initiate the formation of imidazole cores via the cross-linking of amino acids and small peptides.[69] Especially glyoxal and methylglyoxal have been shown to be generated in large quantities in meat, triggering the formation of AGE while cooking,[70] which explains the prominent occurrence of imidazole core structures in the vapor from heated steak. Alkylpyrazines are formed by the intermolecular cyclization of two Strecker aldehydes via intermediate dihydropyrazines formed under thermal influence.[71] Further, cluster 1 contains the most thermoresponsive features of the network.

Cluster 2 contains 66 predominantly aliphatic nitrogenous structures derived from amino acids and small peptides. Among them is pyroglutamic acid (pGlu). pGlu is readily formed under the thermal influence of intramolecular cyclization of N-terminal or free glutamate and glutamine (Glx).[72] Formation of pGlu has an important nutritional aspect as it goes along with the potential loss of the conditionally essential amino acid glutamine. Its abundant signature in oven vapor is reasonable due to the high content of Glx in the muscle tissue.[18] Cluster 3 consists of 45 organoheterocyclic compounds. In contrast to cluster 1, common structural elements are purine and pyridine cores.

The importance of the network-based approach lies in its ability to approximate formerly uncharacterized features found in the cooking vapor by embedding them into known reactional contexts. The spatial proximity of MRP suggests that the surrounding compounds have similar chemical or structural properties. The MR plays an essential role in forming flavor compounds[73] and compounds with implications for human health and nutrition.[74,75] Spectral similarity clustering enables the characterization of analogue molecules or upstream precursor compounds of unknowns.

Combined with the evaluation of the molecular graph and of heating markers via the proposed change-ratio methodology, both methods offer possibilities for a better understanding of which cooking methods and conditions lead to the minimization of undesired substances or the boosting of pleasant flavors.

However, we need to stress that the presented computational approach is able to resolve only some of the complexity of oven vapor samples. The share of uncharacterized features can be reduced by network embedding, but missing links between subnetworks set boundaries. Even though the MR offers a clarifying perspective on 11% of detected features in steak cooking emissions, food is a far too complex system to be understood by such a simplified reactional model. Further, the comparison of molecules detected in cooking fumes with the FooDB emphasized that a huge proportion of the metabolome of food remains undiscovered.

Concludingly, we used a controlled mixed MR model to introduce structural anchors to a continuum of uncharacterized metabolites to further explore the molecular information supplied by oven vapor. It was possible to link the chemical background of 11% of the features found in vapor from rump steak with only a set of six precursor compounds to the MR. In particular, N-heterocyclic aromatic compounds like pyrazine, pyrroles, and imidazoles offer applications for flavor control or optimization of processing settings to preserve nutritional food constituents from a head space prospect. The chemical adulterations triggered in food due to processing still remain fragmentarily understood. In this study, we approximate its complexity by untargeted profiling. Approaches such as the mixed Maillard model help to connect patches of the molecular landscape by contextualizing different molecular families. We propose to classify primary compounds found in the oven head space into two main groups: first, endogenous metabolites from the processed matrix that can also be found in raw, unprocessed food material. Their presence in the oven atmosphere is dependent on the composition of the food and the release efficiency of the food into the head space. Among them are, for example, amino acids or small peptides. Second, there are in situ-generated compounds whose formation is triggered by thermal energy. These compounds, absent in the raw product, contain chemical information about the applied thermal processing. As an example, amino acid degradation products, or AGE, shall be named here.

## ■ ASSOCIATED CONTENT

### Data Availability Statement

Data from UPLC−MS/MS measurements is publicly available in the open-repository MetaboLights under the ID MTBLS8077.

### Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.jafc.3c05383.

> Formulation of foods and processing conditions; chromatographic gradient tables for UPLC separation; exemplary FT-ICR−MS spectrum; UPLC−MS feature maps; PCA of food samples; and sPLS-DA correlation circle plot and contribution component 1 (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Authors

**Leopold Weidner** − *Comprehensive Foodomics Platform, Chair of Analytical Food Chemistry, TUM School of Life Sciences, Technical University of Munich, 85354 Freising, Germany; Analytical BioGeoChemistry, Helmholtz Zentrum München, 85764 Neuherberg, Germany;* ⦿ orcid.org/0000-0002-6801-3647; Phone: +498931873232; Email: leopold.weidner@tum.de

**Philippe Schmitt-Kopplin** − *Comprehensive Foodomics Platform, Chair of Analytical Food Chemistry, TUM School of Life Sciences, Technical University of Munich, 85354 Freising, Germany; Analytical BioGeoChemistry, Helmholtz Zentrum München, 85764 Neuherberg, Germany;* Phone: +498931873246; Email: schmitt.kopplin@tum.de

**Authors**

   **Jil Vittoria Cannas** − *Comprehensive Foodomics Platform, Chair of Analytical Food Chemistry, TUM School of Life Sciences, Technical University of Munich, 85354 Freising, Germany*

   **Michael Rychlik** − *Comprehensive Foodomics Platform, Chair of Analytical Food Chemistry, TUM School of Life Sciences, Technical University of Munich, 85354 Freising, Germany*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.jafc.3c05383

**Author Contributions**

Leopold Weidner: conceptualization, methodology, formal analysis, investigation, software, validation, visualization, project administration, and writing—original draft. Jil Vittoria Cannas: formal analysis and investigation. Michael Rychlik: resources and supervision. Philippe Schmitt-Kopplin: conceptualization, methodology, formal analysis, funding acquisition, and supervision. All authors: writing—review and editing. All authors have given approval to the final version of the manuscript.

**Notes**

The authors declare no competing financial interest.

## ■ REFERENCES

(1) Wrangham, R.; Conklin-Brittain, N. "Cooking as a biological trait". *Comp. Biochem. Physiol.* **2003**, *136* (1), 35−46.

(2) Krutulyte, R.; Costa, A. I.; Grunert, K. G. A Cross-Cultural Study of Cereal Foods Quality Perception. *98th EAAE Seminar "Marketing Dynamics within the Global Trading System: New Perspectives"*; European Association of Agricultural Economists: Chania, Crete, 2006.

(3) Cox, R. J.; Cunial, C. M.; Winter, S. Eliminating Cross Cultural Variations in Consumer Perceptions of Cooked Beef Steaks. *J. Foodserv. Bus. Res.* **2005**, *8* (3), 81−94.

(4) Pratt, K. A.; Prather, K. A. Mass Spectrometry of Atmospheric Aerosols—Recent Developments and Applications. Part II: On-Line Mass Spectrometry Techniques. *Mass Spectrom. Rev.* **2012**, *31* (1), 17−48.

(5) Hassoun, A.; Jagtap, S.; Garcia-Garcia, G.; Trollman, H.; Pateiro, M.; Lorenzo, J. M.; Trif, M.; Rusu, A. V.; Aadil, R. M.; Šimat, V.; Cropotova, J.; Câmara, J. S. Food Quality 4.0: From Traditional Approaches to Digitalized Automated Analysis. *J. Food Eng.* **2023**, *337*, 111216.

(6) Zimmerman, J. B.; Anastas, P. T.; Erythropel, H. C.; Leitner, W. Designing for a Green Chemistry Future. *Science* **2020**, *367* (6476), 397−400.

(7) Fehaili, S.; Courel, M.; Rega, B.; Giampaoli, P. An Instrumented Oven for the Monitoring of Thermal Reactions during the Baking of Sponge Cake. *J. Food Eng.* **2010**, *101* (3), 253−263.

(8) Rega, B.; Guerard, A.; Delarue, J.; Maire, M.; Giampaoli, P. On-Line Dynamic HS-SPME for Monitoring Endogenous Aroma Compounds Released during the Baking of a Model Cake. *Food Chem.* **2009**, *112* (1), 9−17.

(9) Mottram, D. S. The Effect of Cooking Conditions on the Formation of Volatile Heterocyclic Compounds in Pork. *J. Sci. Food Agric.* **1985**, *36* (5), 377−382.

(10) Pico, J.; Khomenko, I.; Capozzi, V.; Navarini, L.; Biasioli, F. Real-Time Monitoring of Volatile Compounds Losses in the Oven during Baking and Toasting of Gluten-Free Bread Doughs: A PTR-MS Evidence. *Foods* **2020**, *9* (10), 1498.

(11) Biasioli, F.; Yeretzian, C.; Märk, T. D.; Dewulf, J.; Van Langenhove, H. Direct-Injection Mass Spectrometry Adds the Time Dimension to (B)VOC Analysis. *TrAC, Trends Anal. Chem.* **2011**, *30* (7), 1003−1017.

(12) Barro, R.; Regueiro, J.; Llompart, M.; Garcia-Jares, C. Analysis of Industrial Contaminants in Indoor Air: Part 1. Volatile Organic Compounds, Carbonyl Compounds, Polycyclic Aromatic Hydrocarbons and Polychlorinated Biphenyls. *J. Chromatogr. A* **2009**, *1216* (3), 540−566.

(13) Lee, S. C.; Li, W. M.; Yin Chan, L. Indoor Air Quality at Restaurants with Different Styles of Cooking in Metropolitan Hong Kong. *Sci. Total Environ.* **2001**, *279* (1−3), 181−193.

(14) Rogge, W. F.; Hildemann, L. M.; Mazurek, M. A.; Cass, G. R.; Simoneit, B. R. T. Sources of Fine Organic Aerosol. 1. Charbroilers and Meat Cooking Operations. *Environ. Sci. Technol.* **1991**, *25* (6), 1112−1125.

(15) Weidner, L.; Yan, Y.; Hemmler, D.; Rychlik, M.; Schmitt-Kopplin, P. Elucidation of the Non-Volatile Fingerprint in Oven Headspace Vapor from Bread Roll Baking by Ultra-High Resolution Mass Spectrometry. *Food Chem.* **2022**, *374*, 131618.

(16) Hellwig, M.; Kühn, L.; Henle, T. Individual Maillard Reaction Products as Indicators of Heat Treatment of Pasta—A Survey of Commercial Products. *J. Food Compos. Anal.* **2018**, *72*, 83−92.

(17) Hu, B.; Li, L.; Hu, Y.; Zhao, D.; Li, Y.; Yang, M.; Jia, A.; Chen, S.; Li, B.; Zhang, X. Development of a Novel Maillard Reaction-Based Time Temperature Indicator for Monitoring the Fluorescent AGE Content in Reheated Foods. *RSC Adv.* **2020**, *10* (18), 10402−10410.

(18) Wu, G.; Cross, H. R.; Gehring, K. B.; Savell, J. W.; Arnold, A. N.; McNeill, S. H. Composition of Free and Peptide-Bound Amino Acids in Beef Chuck, Loin, and Round Cuts. *J. Anim. Sci.* **2016**, *94* (6), 2603−2613.

(19) Pieczonka, S. A.; Hemmler, D.; Moritz, F.; Lucio, M.; Zarnkow, M.; Jacob, F.; Rychlik, M.; Schmitt-Kopplin, P. Hidden in Its Color: A Molecular-Level Analysis of the Beer's Maillard Reaction Network. *Food Chem.* **2021**, *361*, 130112.

(20) Kanawati, B.; Bader, T. M.; Wanczek, K. P.; Li, Y.; Schmitt-Kopplin, P. Fourier Transform (FT)-Artifacts and Power-Function Resolution Filter in Fourier Transform Mass Spectrometry. *Rapid Commun. Mass Spectrom.* **2017**, *31* (19), 1607−1615.

(21) Tziotis, D.; Hertkorn, N.; Schmitt-Kopplin, P. Kendrick-Analogous Network Visualisation of Ion Cyclotron Resonance Fourier Transform Mass Spectra: Improved Options for the Assignment of Elemental Compositions and the Classification of Organic Molecular Complexity. *Eur. J. Mass Spectrom.* **2011**, *17* (4), 415−421.

(22) Chambers, M. C.; MacLean, B.; Burke, R.; Amodei, D.; Ruderman, D. L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; Hoff, K.; Kessner, D.; Tasman, N.; Shulman, N.; Frewen, B.; Baker, T. A.; Brusniak, M. Y.; Paulse, C.; Creasy, D.; Flashner, L.; Kani, K.; Moulding, C.; Seymour, S. L.; Nuwaysir, L. M.; Lefebvre, B.; Kuhlmann, F.; Roark, J.; Rainer, P.; Detlev, S.; Hemenway, T.; Huhmer, A.; Langridge, J.; Connolly, B.; Chadick, T.; Holly, K.; Eckels, J.; Deutsch, E. W.; Moritz, R. L.; Katz, J. E.; Agus, D. B.; MacCoss, M.; Tabb, D. L.; Mallick, P. A Cross-Platform Toolkit for Mass Spectrometry and Proteomics. *Nat. Biotechnol.* **2012**, *30*, 918−920.

(23) Schmid, R.; Heuckeroth, S.; Korf, A.; Smirnov, A.; Myers, O.; Dyrlund, T. S.; Bushuiev, R.; Murray, K. J.; Hoffmann, N.; Lu, M.; Sarvepalli, A.; Zhang, Z.; Fleischauer, M.; Dührkop, K.; Wesner, M.;

Hoogstra, S. J.; Rudt, E.; Mokshyna, O.; Brungs, C.; Ponomarov, K.; Mutabdžija, L.; Damiani, T.; Pudney, C. J.; Earll, M.; Helmer, P. O.; Fallon, T. R.; Schulze, T.; Rivas-Ubach, A.; Bilbao, A.; Richter, H.; Nothias, L. F.; Wang, M.; Orešič, M.; Weng, J. K.; Böcker, S.; Jeibmann, A.; Hayen, H.; Karst, U.; Dorrestein, P. C.; Petras, D.; Du, X.; Pluskal, T. Integrative Analysis of Multimodal Mass Spectrometry Data in MZmine 3. *Nat. Biotechnol.* **2023**, *41*, 447−449.

(24) Horai, H.; Arita, M.; Kanaya, S.; Nihei, Y.; Ikeda, T.; Suwa, K.; Ojima, Y.; Tanaka, K.; Tanaka, S.; Aoshima, K.; Oda, Y.; Kakazu, Y.; Kusano, M.; Tohge, T.; Matsuda, F.; Sawada, Y.; Hirai, M. Y.; Nakanishi, H.; Ikeda, K.; Akimoto, N.; Maoka, T.; Takahashi, H.; Ara, T.; Sakurai, N.; Suzuki, H.; Shibata, D.; Neumann, S.; Iida, T.; Tanaka, K.; Funatsu, K.; Matsuura, F.; Soga, T.; Taguchi, R.; Saito, K.; Nishioka, T. MassBank: A Public Repository for Sharing Mass Spectral Data for Life Sciences. *J. Mass Spectrom.* **2010**, *45* (7), 703−714.

(25) Wang, M.; Carver, J. J.; Phelan, V. V.; Sanchez, L. M.; Garg, N.; Peng, Y.; Nguyen, D. D.; Watrous, J.; Kapono, C. A.; Luzzatto-Knaan, T.; Porto, C.; Bouslimani, A.; Melnik, A. V.; Meehan, M. J.; Liu, W. T.; Crüsemann, M.; Boudreau, P. D.; Esquenazi, E.; Sandoval-Calderón, M.; Kersten, R. D.; Pace, L. A.; Quinn, R. A.; Duncan, K. R.; Hsu, C. C.; Floros, D. J.; Gavilan, R. G.; Kleigrewe, K.; Northen, T.; Dutton, R. J.; Parrot, D.; Carlson, E. E.; Aigle, B.; Michelsen, C. F.; Jelsbak, L.; Sohlenkamp, C.; Pevzner, P.; Edlund, A.; McLean, J.; Piel, J.; Murphy, B. T.; Gerwick, L.; Liaw, C. C.; Yang, Y. L.; Humpf, H. U.; Maansson, M.; Keyzers, R. A.; Sims, A. C.; Johnson, A. R.; Sidebottom, A. M.; Sedio, B. E.; Klitgaard, A.; Larson, C. B.; Boya, C. A. P.; Torres-Mendoza, D.; Gonzalez, D. J.; Silva, D. B.; Marques, L. M.; Demarque, D. P.; Pociute, E.; O'Neill, E. C.; Briand, E.; Helfrich, E. J. N.; Granatosky, E. A.; Glukhov, E.; Ryffel, F.; Houson, H.; Mohimani, H.; Kharbush, J. J.; Zeng, Y.; Vorholt, J. A.; Kurita, K. L.; Charusanti, P.; McPhail, K. L.; Nielsen, K. F.; Vuong, L.; Elfeki, M.; Traxler, M. F.; Engene, N.; Koyama, N.; Vining, O. B.; Baric, R.; Silva, R. R.; Mascuch, S. J.; Tomasi, S.; Jenkins, S.; Macherla, V.; Hoffman, T.; Agarwal, V.; Williams, P. G.; Dai, J.; Neupane, R.; Gurr, J.; Rodríguez, A. M. C.; Lamsa, A.; Zhang, C.; Dorrestein, K.; Duggan, B. M.; Almaliti, J.; Allard, P. M.; Phapale, P.; Nothias, L. F.; Alexandrov, T.; Litaudon, M.; Wolfender, J. L.; Kyle, J. E.; Metz, T. O.; Peryea, T.; Nguyen, D. T.; VanLeer, D.; Shinn, P.; Jadhav, A.; Müller, R.; Waters, K. M.; Shi, W.; Liu, X.; Zhang, L.; Knight, R.; Jensen, P. R.; Palsson, B.; Pogliano, K.; Linington, R. G.; Gutiérrez, M.; Lopes, N. P.; Gerwick, W. H.; Moore, B. S.; Dorrestein, P. C.; Bandeira, N. Sharing and Community Curation of Mass Spectrometry Data with Global Natural Products Social Molecular Networking. *Nat. Biotechnol.* **2016**, *34*, 828−837.

(26) Dührkop, K.; Fleischauer, M.; Ludwig, M.; Aksenov, A. A.; Melnik, A. V.; Meusel, M.; Dorrestein, P. C.; Rousu, J.; Böcker, S. SIRIUS 4: A Rapid Tool for Turning Tandem Mass Spectra into Metabolite Structure Information. *Nat. Methods* **2019**, *16* (4), 299−302.

(27) Dührkop, K.; Shen, H.; Meusel, M.; Rousu, J.; Böcker, S. Searching Molecular Structure Databases with Tandem Mass Spectra Using CSI:FingerID. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112* (41), 12580−12585.

(28) Dührkop, K.; Nothias, L. F.; Fleischauer, M.; Reher, R.; Ludwig, M.; Hoffmann, M. A.; Petras, D.; Gerwick, W. H.; Rousu, J.; Dorrestein, P. C.; Böcker, S. Systematic Classification of Unknown Metabolites Using High-Resolution Fragmentation Mass Spectra. *Nat. Biotechnol.* **2021**, *39* (4), 462−471.

(29) Kim, H. W.; Wang, M.; Leber, C. A.; Nothias, L. F.; Reher, R.; Kang, K. B.; Van Der Hooft, J. J. J.; Dorrestein, P. C.; Gerwick, W. H.; Cottrell, G. W. NPClassifier: A Deep Neural Network-Based Structural Classification Tool for Natural Products. *J. Nat. Prod.* **2021**, *84* (11), 2795−2807.

(30) Ernst, M.; Kang, K. B.; Caraballo-Rodríguez, A. M.; Nothias, L. F.; Wandy, J.; Chen, C.; Wang, M.; Rogers, S.; Medema, M. H.; Dorrestein, P. C.; van der Hooft, J. J. J. Molnetenhancer: Enhanced Molecular Networks by Integrating Metabolome Mining and Annotation Tools. *Metabolites* **2019**, *9* (7), 144.

(31) Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **2003**, *13* (11), 2498−2504.

(32) Sumner, L. W.; Amberg, A.; Barrett, D.; Beale, M. H.; Beger, R.; Daykin, C. A.; Fan, T. W. M.; Fiehn, O.; Goodacre, R.; Griffin, J. L.; Hankemeier, T.; Hardy, N.; Harnly, J.; Higashi, R.; Kopka, J.; Lane, A. N.; Lindon, J. C.; Marriott, P.; Nicholls, A. W.; Reily, M. D.; Thaden, J. J.; Viant, M. R. Proposed Minimum Reporting Standards for Chemical Analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* **2007**, *3* (3), 211−221.

(33) Huber, F.; Verhoeven, S.; Meijer, C.; Spreeuw, H.; Castilla, E.; Geng, C.; van der Hooft, J.; Rogers, S.; Belloum, A.; Diblen, F.; Spaaks, J. Matchms - Processing and Similarity Evaluation of Mass Spectrometry Data. *J. Open Source Softw.* **2020**, *5* (52), 2411.

(34) Rohart, F.; Gautier, B.; Singh, A.; Lê Cao, K. A. MixOmics: An R Package for 'omics Feature Selection and Multiple Data Integration. *PLoS Comput. Biol.* **2017**, *13* (11), No. e1005752.

(35) Lê, S.; Josse, J.; Husson, F. FactoMineR: An R Package for Multivariate Analysis. *J. Stat. Softw.* **2008**, *25* (1), 1−18.

(36) Wickham, H. *Ggplot2 Elegant Graphics for Data Analysis*; Springer-Verlag: New York, 2016.

(37) Gu, Z.; Eils, R.; Schlesner, M. Complex Heatmaps Reveal Patterns and Correlations in Multidimensional Genomic Data. *Bioinformatics* **2016**, *32* (18), 2847−2849.

(38) Kim, S.; Kramer, R. W.; Hatcher, P. G. Graphical Method for Analysis of Ultrahigh-Resolution Broadband Mass Spectra of Natural Organic Matter, the Van Krevelen Diagram. *Anal. Chem.* **2003**, *75* (20), 5336−5344.

(39) Kanner, J. Dietary Advanced Lipid Oxidation Endproducts Are Risk Factors to Human Health. *Mol. Nutr. Food Res.* **2007**, *51*, 1094−1101.

(40) Guéraud, F.; Atalay, M.; Bresgen, N.; Cipak, A.; Eckl, P. M.; Huc, L.; Jouanin, I.; Siems, W.; Uchida, K. Chemistry and Biochemistry of Lipid Peroxidation Products. *Free Radical Res.* **2010**, *44* (10), 1098−1124.

(41) Barriuso, B.; Astiasarán, I.; Ansorena, D. A Review of Analytical Methods Measuring Lipid Oxidation Status in Foods: A Challenging Task. *Eur. Food Res. Technol.* **2013**, *236*, 1−15.

(42) Bohrer, B. M. An Investigation of the Formulation and Nutritional Composition of Modern Meat Analogue Products. *Food Sci. Hum. Wellness* **2019**, *8*, 320−329.

(43) He, J.; Evans, N. M.; Liu, H.; Shao, S. A Review of Research on Plant-Based Meat Alternatives: Driving Forces, History, Manufacturing, and Consumer Attitudes. *Compr. Rev. Food Sci. Food Saf.* **2020**, *19* (5), 2639−2656.

(44) van Vliet, S.; Bain, J. R.; Muehlbauer, M. J.; Provenza, F. D.; Kronberg, S. L.; Pieper, C. F.; Huffman, K. M. A Metabolomics Comparison of Plant-Based Meat and Grass-Fed Meat Indicates Large Nutritional Differences despite Comparable Nutrition Facts Panels. *Sci. Rep.* **2021**, *11* (1), 13828.

(45) Hernandez, M. S.; Woerner, D. R.; Brooks, J. C.; Legako, J. F. Descriptive Sensory Attributes and Volatile Flavor Compounds of Plant-Based Meat Alternatives and Ground Beef. *Molecules* **2023**, *28* (7), 3151.

(46) Belitz, H.-D.; Grosch, W.; Schieberle, P. *Lehrbuch Der Lebensmittelchemie*; Springer, 2008.

(47) Dunajski', E. Texture of Fish Muscle. *J. Texture Stud.* **1980**, *10*, 301−318.

(48) Hemmler, D.; Roullier-Gall, C.; Marshall, J. W.; Rychlik, M.; Taylor, A. J.; Schmitt-Kopplin, P. Insights into the Chemistry of Non-Enzymatic Browning Reactions in Different Ribose-Amino Acid Model Systems. *Sci. Rep.* **2018**, *8* (1), 16879.

(49) Hemmler, D.; Roullier-Gall, C.; Marshall, J. W.; Rychlik, M.; Taylor, A. J.; Schmitt-Kopplin, P. Evolution of Complex Maillard Chemical Reactions, Resolved in Time. *Sci. Rep.* **2017**, *7* (1), 3227.

(50) Pavlidis, D. E.; Mallouchos, A.; Ercolini, D.; Panagou, E. Z.; Nychas, G. J. E. A Volatilomics Approach for Off-Line Discrimination

of Minced Beef and Pork Meat and Their Admixture Using HS-SPME GC/MS in Tandem with Multivariate Data Analysis. *Meat Sci.* **2019**, *151* (2018), 43−53.

(51) Abe, K.; Hori, Y.; Myoda, T. Characterization of Key Aroma Compounds in Aged Garlic Extract. *Food Chem.* **2020**, *312*, 126081.

(52) He, J.; Liu, H.; Balamurugan, S.; Shao, S. Fatty Acids and Volatile Flavor Compounds in Commercial Plant-Based Burgers. *J. Food Sci.* **2021**, *86* (2), 293−305.

(53) Klein, F.; Platt, S. M.; Farren, N. J.; Detournay, A.; Bruns, E. A.; Bozzetti, C.; Daellenbach, K. R.; Kilic, D.; Kumar, N. K.; Pieber, S. M.; Slowik, J. G.; Temime-Roussel, B.; Marchand, N.; Hamilton, J. F.; Baltensperger, U.; Prévôt, A. S. H.; El Haddad, I. Characterization of Gas-Phase Organics Using Proton Transfer Reaction Time-of-Flight Mass Spectrometry: Cooking Emissions. *Environ. Sci. Technol.* **2016**, *50* (3), 1243−1250.

(54) Marcinkowska, M. A.; Jeleń, H. H. Role of Sulfur Compounds in Vegetable and Mushroom Aroma. *Molecules* **2022**, *27* (18), 6116.

(55) Abdullahi, K. L.; Delgado-Saborit, J. M.; Harrison, R. M. Emissions and Indoor Concentrations of Particulate Matter and Its Specific Chemical Components from Cooking: A Review. *Atmos. Environ.* **2013**, *71*, 260−294.

(56) Sioutas, C.; Delfino, R. J.; Singh, M. Exposure Assessment for Atmospheric Ultrafine Particles (UFPs) and Implications in Epidemiologic Research. *Environ. Health Perspect.* **2005**, *113* (8), 947−955.

(57) Kulmala, M.; Vehkamäki, H.; Petäjä, T.; Dal Maso, M.; Lauri, A.; Kerminen, V. M.; Birmili, W.; McMurry, P. H. Formation and Growth Rates of Ultrafine Atmospheric Particles: A Review of Observations. *J. Aerosol Sci.* **2004**, *35* (2), 143−176.

(58) Liu, P.; Yuan, Z.; Zhang, S.; Xu, Z.; Li, X. Experimental Study of the Steam Distillation Mechanism during the Steam Injection Process for Heavy Oil Recovery. *J. Pet. Sci. Eng.* **2018**, *166*, 561−567.

(59) Zhang, Z.; Zhu, W.; Hu, M.; Wang, H.; Chen, Z.; Shen, R.; Yu, Y.; Tan, R.; Guo, S. Secondary Organic Aerosol from Typical Chinese Domestic Cooking Emissions. *Environ. Sci. Technol. Lett.* **2021**, *8* (1), 24−31.

(60) Xiang, W.; Wang, W.; Du, L.; Zhao, B.; Liu, X.; Zhang, X.; Yao, L.; Ge, M. Toxicological Effects of Secondary Air Pollutants. *Chem. Res. Chin. Univ.* **2023**, *39* (3), 326−341.

(61) Jin, S.; Tserng, K. Y. Biogenesis of Dicarboxylic Acids in Rat Liver Homogenate Studied by 13C Labeling. *Am. J. Physiol.* **1991**, *261* (6), E719−E724.

(62) Da Silva, R. R.; Dorrestein, P. C.; Quinn, R. A. Illuminating the Dark Matter in Metabolomics. *Proc. Natl. Acad. Sci. U.S.A.* **2015**, *112* (41), 12549−12550.

(63) Zhao, C.; Cao, H.; Xiao, J. Pyrazines in Food. *Handbook of Dietary Phytochemicals*; Springer: Singapore, 2020; pp 1−25.

(64) Hofmann, T. Identification of Novel Colored Compounds Containing Pyrrole and Pyrrolinone Structures Formed by Maillard Reactions of Pentoses and Primary Amino Acids. *J. Agric. Food Chem.* **1998**, *46* (10), 3902−3911.

(65) Lederer, M. O.; Klaiber, R. G. Cross-Linking of Proteins by Maillard Processes: Characterization and Detection of Lysine-Arginine Cross-Links Derived from Glyoxal and Methylglyoxal. *Bioorg. Med. Chem.* **1999**, *7* (11), 2499−2507.

(66) Henle, T. Maillard Reaction of Proteins and Advanced Glycation End Products (AGEs) in Food. In *Process-Induced Food Toxicants: Occurrence, Formation, Mitigation, and Health Risks*; Stadler, R. H., Linebaker, D. R., Eds.; John Wiley & Sons, Inc, 2009; pp 215−242.

(67) Kroh, L. W. Caramelisation in Food and Beverages. *Food Chem.* **1994**, *51* (4), 373−379.

(68) Yaylayan, V. A. Classification of the Maillard Reaction: A Conceptual Approach. *Trends Food Sci. Technol.* **1997**, *8* (1), 13−18.

(69) Chellan, P.; Nagaraj, R. H. Protein Crosslinking by the Maillard Reaction: Dicarbonyl-Derived Imidazolium Crosslinks in Aging and Diabetes. *Arch. Biochem. Biophys.* **1999**, *368* (1), 98−104.

(70) Eggen, M. D.; Glomb, M. A. Analysis of Glyoxal- And Methylglyoxal-Derived Advanced Glycation End Products during

Grilling of Porcine Meat. *J. Agric. Food Chem.* **2021**, *69* (50), 15374−15383.

(71) Amrani-Hemaimi, M.; Cerny, C.; Fay, L. B. Mechanisms of Formation of Alkylpyrazines in the Maillard Reaction. *J. Agric. Food Chem.* **1995**, *43*, 2818−2822.

(72) Bachhawat, A.; Kumar, A.; Bachhawat, A. K. Pyroglutamic Acid: Throwing Light on a Lightly Studied Metabolite. *Curr. Sci.* **2012**, *102* (2), 288−297.

(73) Danehy, J. P. Maillard Reactions: Nonenzymatic Brownings in Food Systems with Special Reference to the Development of Flavor. *Adv. Food Res.* **1986**, *30*, 77−138.

(74) Labuza, T. P.; Monnier, V.; Baynes, J.; O'Brien, J. *Maillard Reactions in Chemistry, Food and Health*, 1st ed.; Elsevier, 1998.

(75) Hellwig, M.; Henle, T. Baking, Ageing, Diabetes: A Short History of the Maillard Reaction. *Angew. Chem., Int. Ed. Engl.* **2014**, *53* (39), 10316−10329.

M

## A.3 Article III: Real-Time Monitoring of Miniaturized Thermal Food Processing by Advanced Mass Spectrometric Techniques.

Article

# Real-Time Monitoring of Miniaturized Thermal Food Processing by Advanced Mass Spectrometric Techniques

Leopold Weidner, Daniel Hemmler, Michael Rychlik, and Philippe Schmitt-Kopplin*

Cite This: https://doi.org/10.1021/acs.analchem.2c04874
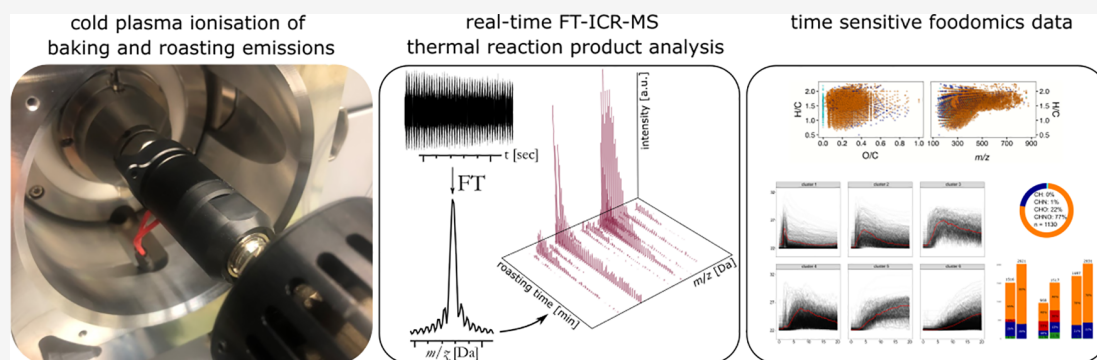
Read Online

ACCESS | 📊 Metrics & More | 📰 Article Recommendations | 🗒 Supporting Information

**ABSTRACT:** Mass spectrometry is a popular and powerful analytical tool to study the effects of food processing. Industrial sampling, real-life sampling, or challenging academic research on process-related volatile and aerosol research often demand flexible, time-sensitive data acquisition by state-of-the-art mass analyzers. Here, we show a laboratory-scaled, miniaturized, and highly controllable setup for the online monitoring of aerosols and volatiles from thermal food processing based on dielectric barrier discharge ionization (DBDI) mass spectrometry (MS). We demonstrate the opportunities offered by the setup from a foodomics perspective to study emissions from the thermal processing of wheat bread rolls at 210 °C by Fourier transformation ion cyclotron resonance MS. As DBDI is an emerging technology, we compared its ionization selectivity to established atmospheric pressure ionization tools: we found DBDI preferably ionizes saturated, nitrogenous compounds. We likewise identified a sustainable overlap in the selectivity of detected analytes with APCI and electrospray ionization (ESI). Further, we dynamically recorded chemical fingerprints throughout the thermal process. Unsupervised classification of temporal response patterns was used to describe the dynamic nature of the reaction system. Compared to established tools for real-time MS, our setup permits one to monitor chemical changes during thermal food processing at ultrahigh resolution, establishing an advanced perspective for real-time mass spectrometric analysis of food processing.

Thermal baking and roasting processes are key fabrication steps during the production of high-value foodstuff such as coffee,[1] cocoa,[2] or various seeds and nuts.[3,4] Likewise, baking often is the final stage during preparation of cereal products.[5,6] Hence, the understanding of baking processes at a molecular level is of immediate interest to several alimental industries with regard to flavor generation and product quality.

The emissions released during and after thermal processing, composed of volatile molecules and aerosols, are *in situ* thermally generated from the dough by complex chemical reaction cascades such as lipid oxidation, caramelization, or Maillard reaction (MR).[7] This transformation of nearly flavorless raw material into pleasant and analytically challenging, processed matrices is subject to intense research. Primarily, mass spectrometric methods are used to investigate and understand the processes used during thermal transformation of foodstuff. Although direct coupling of a mass spectrometer (MS) to a real-life baking plant is accompanied by various

challenges, such an online MS setup permits one to follow baking processes in real time during the heating process. Current commercially available mass spectrometric techniques used for the online monitoring of food baking processes often utilize instrumentation such as MS-electronical noses, proton transfer reaction (PTR) MS, selective ion flow tube (SIFT) MS, or atmospheric pressure chemical ionization (APCI) MS.[8,9] All of these MS-based techniques are direct injection (DI) methods. Their ionization mechanisms depend on the transfer of a proton to the analyte, forming diagnostic pseudo-
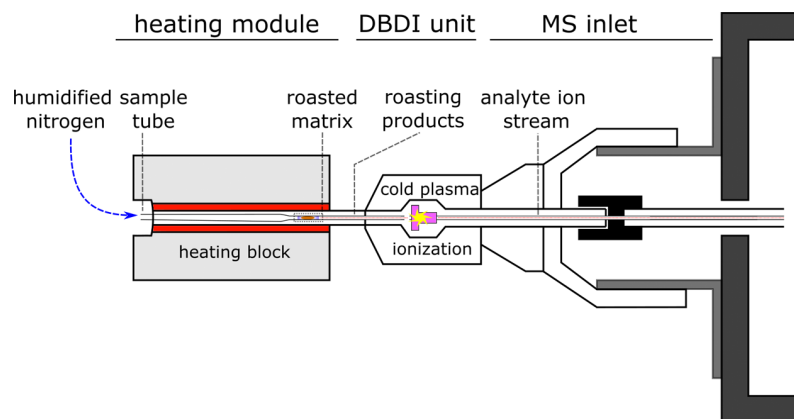
ACS Publications

A

**Figure 1.** Schematic drawing (not proportional) of the ionization setup for real-time monitoring of miniaturized roasting processes attached to a mass spectrometer (DBDI = dielectric barrier discharge ionization).

molecular ions with little fragmentation and come without prior chromatographical separation of analytes. Thus, they all permit the live acquisition of data during food production processes. PTR instruments are commonly coupled to time-of-flight (TOF) mass analyzers, giving the advantage of recording high-resolution data. Combination of PTR and advanced mass analyzers is possible and would further allow the acquisition of tandem mass spectra; however, this option has not been established on the commercial market yet. Whereas SIFT instruments can only achieve unit mass resolution, their strength lies in the opportunity to record quantitative data. An elaborate principle of multiple reagent ion with different reaction constants allows calculation of the analytes under certain circumstances.[10] As recently elaborated by Biasioli et al., another crucial parameter for online measurements is acquisition speed.[9] In this context, especially electronic noses[9] suffer from slow data acquisition and detector poisoning, often hindering application in industrial production. Likewise, the SIFT technology relies on quadrupole technology resulting in extended cycle times when operated in the full scan mode, especially when multiple reagent ion types are selected. All of these technological platforms are successfully applied in environmental and food analyses. Driven by the need to access state-of-the-art mass analyzers for real-time data acquisition, growing attention is given to technical advances in this field.[11]

In this work, we developed an analytical laboratory-scaled thermal processing setup, which is directly coupled to dielectric barrier discharge ionization (DBDI) MS, enabling the real-time analysis of volatiles and aerosols from baking and roasting processes by advanced mass spectrometric techniques. DBDI is a relatively young technology compared to other atmospheric pressure ionization (API) techniques such as APCI. APCI is an established tool for real-time MS analysis popular in breath-, indoor air-, and flavor release analyses.[8,12,13] While DBDI rapidly gained popularity in recent years,[14] its ionization selectivity for thermal baking vapor analysis has not been described before. Easy transferability between different commercial MS instruments as well as the capability of dealing with the smallest amounts of samples was a focus of the developed method. The miniaturized setup further offers the operator the opportunities of a highly controlled environment for thermal food processing studies. Exemplarily, we monitored the baking and roasting of wheat bread rolls on an ultrahigh-resolution Fourier transformation ion cyclotron resonance MS

by an untargeted foodomics approach to demonstrate the potential of the setup.

## ◼ EXPERIMENTAL SECTION

**Materials.** Arginine (reagent grade 98%) was purchased from Sigma Aldrich (Taufkirchen, Germany). Methanol (LC−MS grade) was obtained from Fisher Scientific (Schwerte, Germany).

**Food Samples.** Deep-frozen wheat bread roll samples were purchased from a local bakery. The nutritional values and formulation of the dough can be found in the Supporting Information, Table S1. Prior to the experiments, the raw dough was allowed to rise for 180 min at a temperature of 34 °C and 80% humidity. Following this, the dough was homogenized and portioned into aliquots.

**Real-Time Analysis of Thermal Reaction Products by Dielectric Barrier Discharge Ionization.** Approximately 2 mg of the bread roll sample was placed in a borosilicate glass tube with a piece of glass wool behind the dough. For baking experiments, the tube was placed inside a heated, programmable GC-SPME module (Plasmion GmbH, Augsburg, Germany) used as a laboratory-scale micro-oven during this study. The heater was mounted in such a way that the front opening of the sample tube was placed in direct proximity to the ionization unit. The sample tube was continuously flushed with a stream of 1.0 L min$^{-1}$ humidified nitrogen. Aerosols and volatile baking products evolving from the heated matrix were subject to active and continuous conduct by gas flow and by the negative pressure system of the MS toward a SICRIT SC-30 active capillary low-temperature plasma ionization unit (Plasmion GmbH, Augsburg, Germany). See Figure 1 for a schematic drawing of the setup. Atmospheric pressure DBDI was carried out by applying a high-voltage alternating current of 1.5 kV at a frequency of 15 kHz. Prior to analysis, blank runs of an empty tube, only packed with glass wool, were performed under the same analytical and experimental conditions as during the baking experiment to sample the chemical background of the setup and laboratory air. Baking experiments were carried out at a temperature of 210 °C for 20 min. Online measurement of the experiment by MS was performed in duplicates and as soon as the sample tube was inserted into the preheated setup.

**Ultrahigh-Resolution Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR-MS) Data Acquisition.** The experimental setup described above was
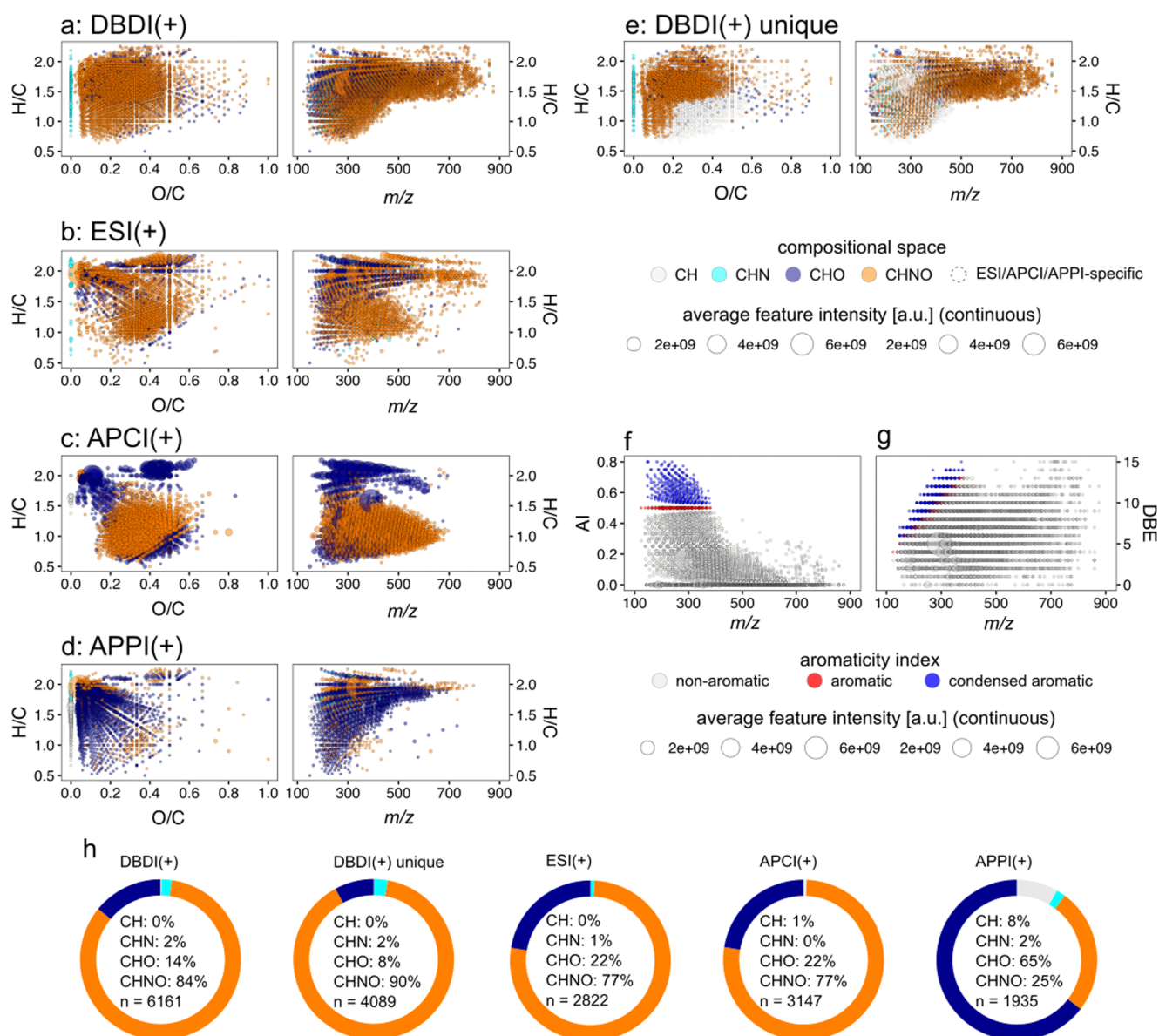
**Figure 2.** (a) Van Krevelen and respective H/C- to $m/z$-plot of wheat bread rolls baked at 210 °C for 20 min monitored by DBDI. (b−d) Baking emissions from the same sample monitored by (b) ESI(+), (c) APCI(+), and (d) APPI(+) coupled to FT-ICR-MS. (e) Features uniquely detected by FT-ICR-MS. (f, g) Aromaticity index (AI) and double bond equivalent (DBE) plots for features presented in (a). (h) Distribution of compositional spaces for studied ionization setups.

installed on a solariX Fourier transformation ion cyclotron resonance (FT-ICR) MS equipped with a 12 T super-conducting magnet (Bruker Daltonics, Bremen, Germany) to record ultrahigh-resolution spectra during the baking process in real time. The instrument was set to the chromatographic mode in the positive ion mode to follow the evolution of baking products. A total number of 10 scans with an ion accumulation time of 0.1 s were accumulated per datapoint from $m/z$ 92 to 1000 with a time domain of 4 megawords. Source parameters were set as follows: dry gas (nitrogen) flow rate 2 L min$^{-1}$ at 180 °C, capillary voltage 400 V, and spray shield voltage −400 V. External calibration of the instrument's mass analyzer was performed by an APOLLO II electrospray ion source (Bruker Daltonics, Bremen, Germany) with clusters of arginine (5 ppm in methanol) in the positive ion mode prior to installation of the SICRIT setup.

**Processing of FT-ICR-MS Data.** Raw data were preprocessed in Compass DataAnalysis 5.0 (Bruker Daltonics, Bremen, Germany). Internal calibration of the acquired spectra was performed with a reference mass list. The reference list was composed of reliably identified masses generated by the SICRIT setup. Following this, ion signals with a signal-to-noise ratio S/N > 6 were manually exported into mass lists for each datapoint during the experiment. Only singly charged ions were considered in this study. Further, data treatment, processing, and analysis were performed in R programming language (Version 4.2, R Foundation for Statistical Computing, Vienna, Austria) and Python programming language (Version 3.10). Initially, the data was cleaned for side band artifacts[15] and heavy isotope peaks. The mass lists were aligned into matrices containing features of exact masses. Molecular formulas were assigned to the features based on an exact

C

pubs.acs.org/ac

mass difference network approach[16] within the CHNO elemental space. Only unambiguously annotated features were further considered. DBDI is a soft ionization technique based on secondary chemical reaction processes and was shown to generate predominantly $[M + H]^+$ pseudo-molecular ions.[17,18] However, DBDI is known to produce a small proportion of other ion species, especially for hydrophobic analytes.[19] For curation of spectra from non-$[M + H]^+$ ions along with in-source fragments, a custom algorithm based on pointwise Pearson correlation and exact mass differences was used to putatively flag multiple ion signals originating from the same compound.[20] Temporal clustering of compounds based on their observed intensity profiles during an experimental run was performed by a self-organizing map (SOM) implementation.[21]

## ■ RESULTS AND DISCUSSION

**Characterization of Aerosols and Thermal Reaction Products Emitted during Roasting Processes by FT-ICR-MS.** We processed 2 mg of wheat bread roll samples at 210 °C for 20 min in our analytical setup. Throughout the entire baking experiment, we detected 6161 annotated mass signals. Ion signals had an average resolution of 578,000 at $m/z$ 200 and of 224,000 at $m/z$ 500. This exceeds the horizons of conventional real-time monitoring setups such as SIFT- or PTR-TOF-MS (resolution below 10,000[5]) and unveils a complex fingerprint of thermal products easily being over-looked but altogether containing valuable chemical information about the studied process. Whereas our foodomics approach aims at a holistic characterization of aerosols and volatiles emitted during the baking process, these mass analyzers focus on a limited number of compounds due to a poorer mass resolution or a slow scan speed. Operated on our FT-ICR-MS platform, our method had an effective scan rate of 0.1 Hz, which is slower than PTR-TOF-MS systems but faster than SIFT-MS in the full scan mode.

As stated before, plasma-based ionization combines numerous secondary ionization mechanisms that are highly dependent on various details like the exact source setup or surrounding atmospheric compositions.[14,22] This gives the precious opportunity to gain novel analytical perspectives on process analysis. To better integrate our DBDI results into existing knowledge on food processing emission research, we first aimed at comparing DBDI with the established API techniques such as electrospray ionization (ESI), atmospheric pressure photoionization (APPI), and APCI. Therefore, we collected baking vapor from the same bread roll samples in a recently described at-line monitoring approach.[23] These condensate samples likewise contain aerosols and thermal reaction products emitted during the baking process and were analyzed by ESI, APCI, and APPI coupled with FT-ICR-MS to contextualize our observations made by DBDI (see Supplementary Methods). The diversity of thermal reaction products detected by DBDI, ESI, APCI, and APPI is summarized by van Krevelen plots and corresponding hydrogen:carbon (H/C)-to-$m/z$ plots in Figure 2a–d. Figure 2e shows DBDI-specific features, which were not detected by any of the other ion sources. Van Krevelen plots are helpful graphical tools to visualize untargeted omics data sets.[24,25] Annotated molecular formulae are cross-plotted by their elemental hydrogen:carbon (H/C) and oxygen:carbon (O/C) or respective $m/z$ ratios on a two-dimensional plane.[25] Based on their coordinates, features settle in regions distinctive for specific compound classes.

Compositional and reactional trends in the analyzed sample become visible on the plot. Figure 2a reveals DBDI preferably ionizes saturated, nitrogenous compounds. We were able to detect compounds with a molecular mass of up to 822 Da, which is a significantly wider mass range than studied before in a time-sensitive approach during roasting. Especially SIFT-MS setups focus on lower molecular masses and are good for detection of gases and small volatile organic compounds. PTR-TOF routinely detects masses of up to 400 Da. This draws attention to heavy condensed compounds conveyed by aerosols and opens a new perspective for online monitoring of baking products as solely offered by volatile molecules.

ESI, APCI, and APPI together lay a stronger focus on CHO-composed compounds than DBDI. Especially photoionized CHO compounds cover a broad range of polarities and saturation. APPI additionally proved to be highly efficient for ionizing hydrocarbons. APCI as the major established real-time monitoring tool revealed predominantly very saturated (H/C ≥ 2.0) and oxygen-rich, unsaturated CHO-species. As opposed to APCI, the main strength of DBDI is the ionization of CHNO compounds. Up to 84% of all annotated features were CHNO compounds (see Figure 2h). Whereas a certain amount of these CHNO features was likewise detected in ESI and APCI measurements, Figure 2e identifies two regions of CHNO compounds uniquely accessed by DBDI. On the one hand, our setup allows the monitoring of saturated nitrogen-containing compounds (H/C > 1.3, O/C < 0.5) with high molecular masses. On the other hand, DBDI selectively revealed a region located at H/C < 1, O/C < 0.2, where highly unsaturated and aromatic heterocyclic compounds are located. Hemmler et al. recently found that late-phase Maillard reaction products (MRPs) settle toward the bottom left corner of the van Krevelen plot.[26] Low-molecular, nitrogen-containing, (condensed) aromatic structures are a diverse compound class[27] and are long known to have an important impact on the flavor of bread.[7] Therefore, we see DBDI as a promising tool for the real-time monitoring of late-phase aromatic N-heterocyclic MRP and their role in flavor chemistry. Further, DBDI showed extended series of aromatic CHN features, of which only parts were ionized by ESI and APPI. Classified by their aromaticity index[28] (AI), the share of aromatic (AI ≥ 0.5) and condensed aromatic (AI ≥ 0.67) compounds locates below 350 Da, as seen in Figure 2f,g, indicating that highly unsaturated compounds only were detected at relatively low molecular masses.

As for all DI methods, possible drawbacks of our setup include ion suppression effects by possible matrix loads. Further, methodological efforts must be made to enable compound identification during thermal processing in real time as routinely applied by chromatographic separation techniques or tandem mass spectrometry. SIFT, PTR as well as APCI are based on $H^+$-transfer from hydroxonium ions to neutral, proton-affine analytes.[8] DBDI connects to these techniques: 1/3 of features ionized by DBDI are confirmed by our studied API techniques. At the same time, its protonating cold plasma mechanisms grant access to further fields of predominantly nitrogenous compounds from process emissions, opening a new perspective on the real-time analysis of thermal food processing.

**Temporal Perspective on the Formation of Baking Products.** One of the main advantages of the developed method is the capability to continuously deliver mass spectrometric data through the baking experiment. Figure 3a
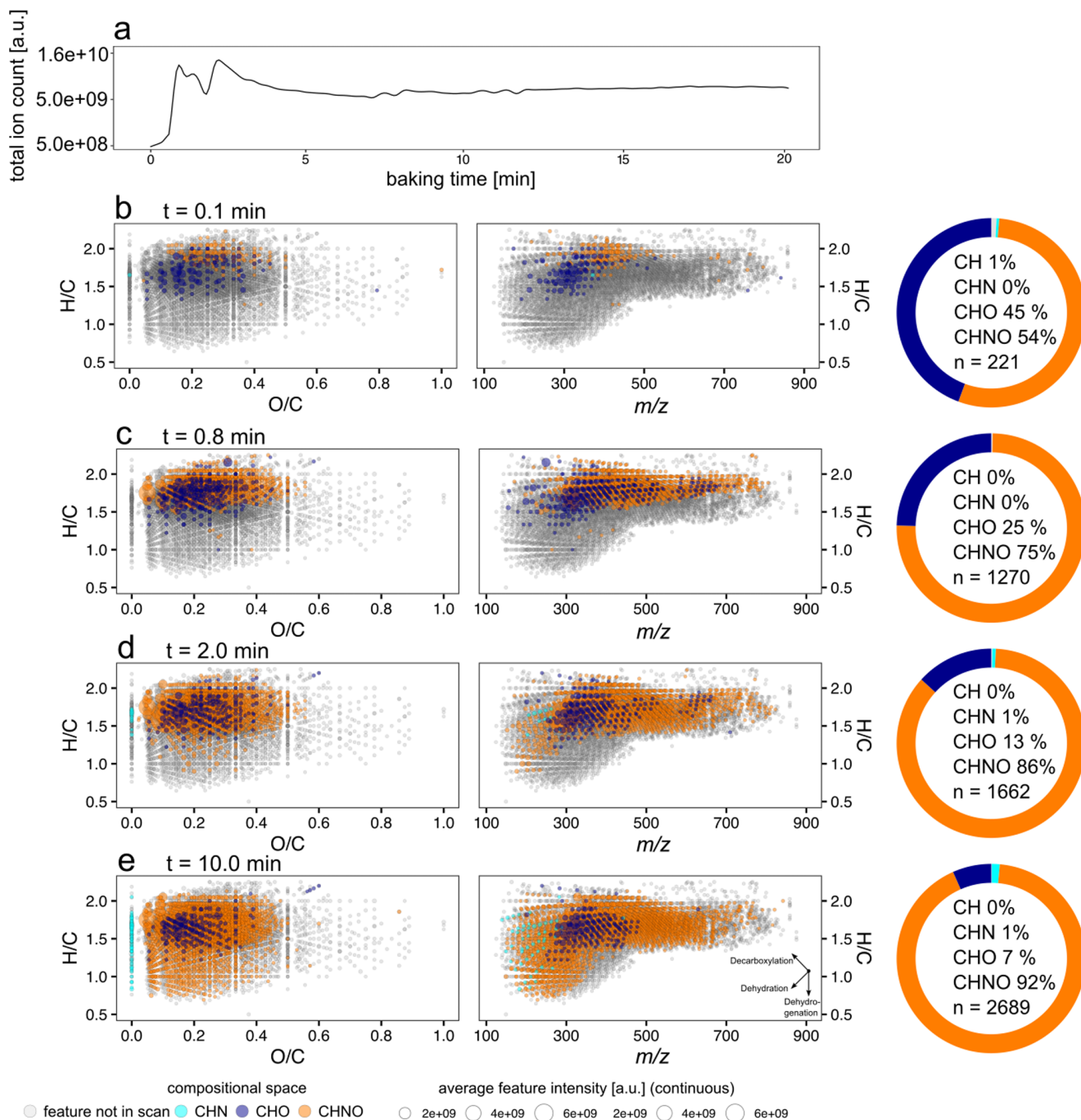
D

**Figure 3.** (a) Total ion count of volatiles and thermal reaction products emerged from a wheat bread roll sample baked at 210 °C monitored by DBDI-FT-ICR-MS. (b–e) Van Krevelen and respective H/C-to-*m/z* plots of four time points extracted over the experiment showing the chemical fingerprint of detected baking and roasting products with compositional information.

shows the total ion count (TIC) of aerosols and thermal reaction products detected by DBDI-FT-ICR-MS from wheat bread roll baked at 210 °C. Immediately after introducing the sample to the setup, we observed an increasing TIC forming two local intensity maxima shortly afterward at 0.8 and 2.0 min baking times. Successively, the TIC decreases to a plateau, which is held until the end of the measurement. As FT-ICR-MS is a very sensitive mass detector, we can study the smallest amounts of samples granting quick heating times to the targeted temperature and avoiding long temperature gradients in the matrix or inhomogeneous temperature profiles through

the sample. This makes our setup an ideal tool for the monitoring of temperature effects on the formation of thermal reaction products or for the application of stable isotope labeling studies. Over the course of the experiment, we can extract a wealth of information about the chemical fingerprint of the emitted aerosols and volatiles. In the Supporting Information, we have attached a time-resolved animation of the studied baking. For the print version, we extracted four time points in Figure 3b,e portraying key time points during the experiment. Immediately after introduction of the sample, 221 reaction products became detectable. The number of
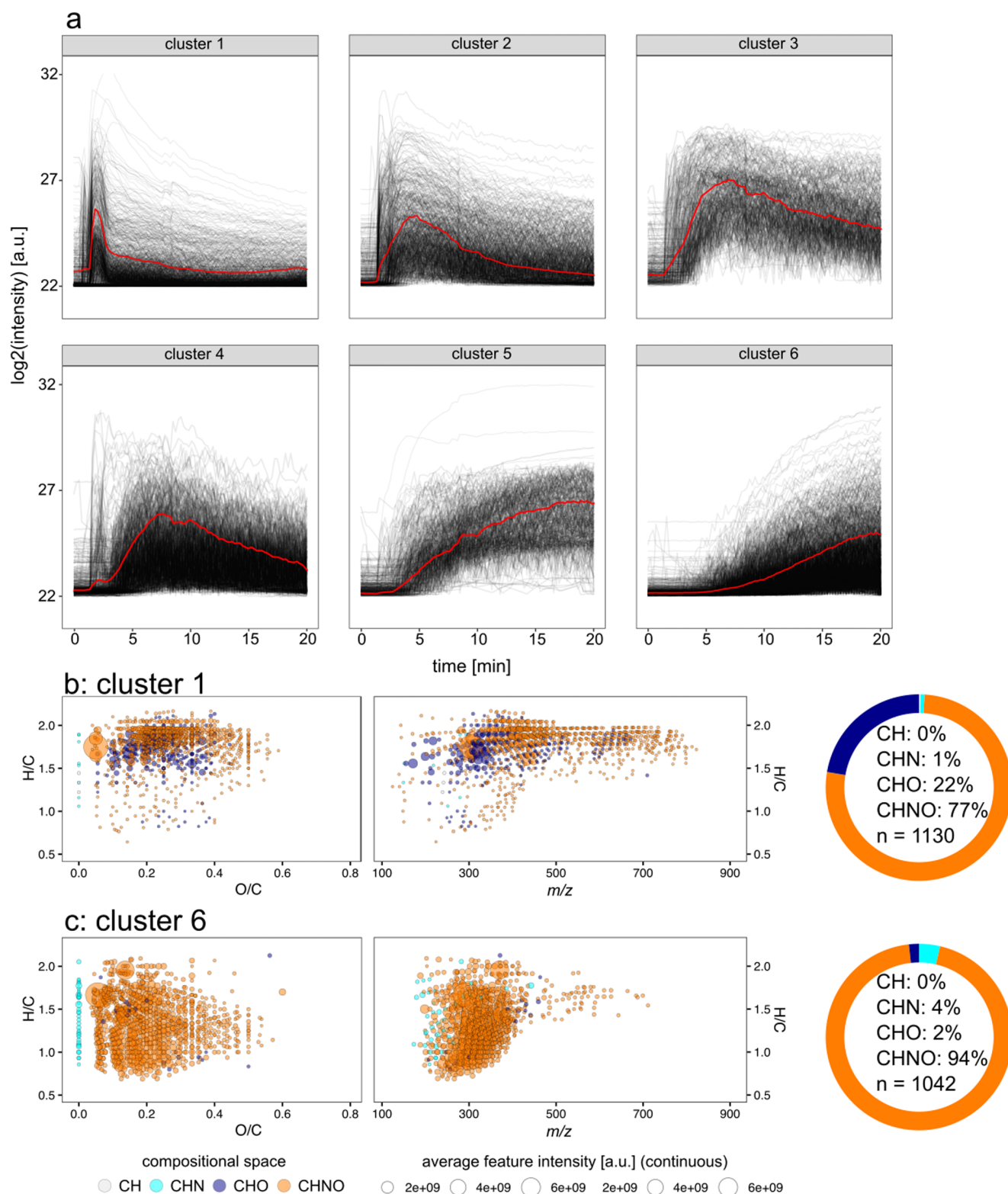
**Figure 4.** (a) Self-organizing map algorithm grouped 6161 annotated extracted ion traces into six clusters based on their temporal profile. Cluster mean is shown in red. (b, c) Van Krevelen plot and respective H/C-to-$m/z$ plots of clusters 1 and 6 with compositional information.

detected compounds continuously increased by a factor of 12 until the end of the experiment to a maximum of 2834 annotated molecular formulae in the last scan. Notably, during the first few seconds of the experiment, the ratio CHO/CHNO is at 1:1.2, which is the highest ratio observed during the entire

experiment. We conclude a globally observed thermally induced introduction of nitrogen into the CHO space. This is supported by increasing nitrogen:carbon ratios along with decreasing oxygen:carbon ratios, indicating the introduction of nitrogen and dehydration reactions (see the Supporting

Information, Figure S1). We observed two major trends triggered by thermal energy, as seen in Figure 3b via c to d. The formation of higher molecular masses, predominantly in the CHNO space with high saturation levels (H/C 1.5–2.0), reached their intensity maximum after 1 min of baking time. We suggest condensation reactions take place between the lower molecular spaces, forming larger molecules being responsible for these compounds. From minute two on, we observed an opposing trend leading to the pronunciation of lower molecular mass spaces below 300 Da: decarboxylation, dehydration, and dehydrogenation reactions lead to populations of the above-described late-phase MRP space. Formerly, these trends toward higher unsaturation induced by thermal energy were discussed as thermal maturation pathways,[29] and in the context of carbonization processes,[30] which correspond to oxidizing processes, we observed them toward late baking phases. This carbonization continues until the end of our observation, leading to low dynamics in the system after 10 min of heating. In the end phase of our experiment, only few new compounds are generated; these together are located either at very high saturation and high molecular masses ($m/z$ > 700 Da) or at high levels of aromaticity, representing the end-points of both observed molecular progressions (see Figure 3e).

**Temporal Classification of Thermal Reaction Product Profiles.** To further understand the temporal dynamics of the system, we extracted the intensity profiles of detected features in the matrix. The ultrahigh resolving power of FT-ICR-MS often reveals more than a dozen different ions per nominal mass, unveiling a complex system of chemical reactions enrolling during baking.

The dynamic entity of this fingerprint-pattern aggregates up to the time-resolved TIC containing valuable information about the thermal processes. To further investigate these, we constructed extracted ion chromatograms (XICs) from our FT-ICR-MS data for all annotated features. Missing values were imputed, and the XICs were scaled to an individual range from 0 to 1. Following this, we used a SOM implementation as an unsupervised, artificial neuronal network to classify all 6161 XICs into groups based on their temporal pattern. Manual inspection of SOM parameters revealed that there are six major groups of comparably behaving emission profiles. Figure 4a shows the grouped XIC traces with cluster means together with compositional information of the two most populated clusters (clusters 1 and 6, Figure 4b,c; corresponding plots for XIC-groups 2–5 are shown in the Supporting Information, Figure S2). Clusters of XICs can be grouped into two major classes.

Features in clusters 1–4 (see Figure 4a) show an early and initial intensity maximum, after which their abundance continuously decreases during further heating. The lower the cluster number, the earlier the associated feature reaches its most abundant point, whereas all cluster means come to a maximum before 10 min of baking. Clusters 1–4 differ in the pace and magnitude of intensity loss: whereas features of cluster 1 (Figure 4b) show a rapid burst in increasing intensity followed by a decrease shortly afterward, those of clusters 2–4 forfeit with a flatter slope. Further, features in clusters 3 and 4 remain detectable with a higher relative intensity level until the end of the experiment. Whether these features still are continuously reprocured after 10 min of baking or signals come from retained traces, which are still detectable due to the high sensitivity of our mass analyzer, remains unclear.

Certainly, their initial generation strongly impacts the first few minutes of baking.

In contrast to clusters 1–4, XICs in clusters 5 and 6 arise with a temporal delay in the experiment and continuously become more abundant over time. Clusters 5 and 6 represent 34% of all mass traces. SOM discriminates a temporal offset between these two groups: features of group 6 (Figure 4c) arise in a later phase than those of group 5. When we extracted the features of group 1 and visualized them on a van Krevelen plane (Figure 4b), it became apparent that these correspond to the saturated molecular families, which were already described above in the context early-stage baking (Figure 3b,c). This validates the unsupervised clustering algorithm. We further recognize that features grouped in cluster 6 (Figure 4c) are nitrogenous, highly unsaturated, and aromatic. In clusters 5 and 6, condensed or heterocyclic aromatic structure elements were found to have an important impact. These features again can be associated with late-phase Maillard reaction processes according to their location on the van Krevelen face.

Based on these groupings, we can divide the microdough experiment into an early "baking phase" (clusters 1–4) with early-arising features. Especially cluster 1 (Figure 4b) shows an immediate intensity increase upon application of heat, dominating the chemical fingerprint during the first few seconds of the experiment. This is followed by a latter "roasting phase" (clusters 5, 6), which is characterized by a set of completely different compound classes.

In a recent study about volatiles lost to the oven head-space during baking, which were measured in real time by PTR-TOF-MS, the authors similarly investigated first the baking, followed by the toasting of gluten-free bread.[5] Similar to our results, they classified different release curves of volatiles and grouped them into nine characteristic patterns. More precisely, fermentation or early MRP was detected earlier during their study and showed analogous release patterns to our cluster 1. By contrast, they observed late MRP to require more time in the oven to be formed when they are consecutively released later during the experiment. This corresponds to our hypothesis about the chemical background of compounds located in our cluster 6. Further, Pico et al. observed other release patterns, which resemble the profiles of XICs in groups 1–4. They exemplarily associated (asymmetric) bell-shaped curves with compounds having a background in fermentation processes.

Recently, it was alike shown by PTR-TOF-MS that during frying of potatoes, some compounds, for example, aldehydes from lipid oxidations processes, also show an early-stage augmentation right after application of heat.[31] Together with our findings, these two studies emphasize the importance of considering thermal food processing as a system of particular, segmented subprocesses and not as an undivided entity.

## ■ CONCLUDING REMARKS

In this work, we present a laboratory-scale, analytical, DBDI-based setup for the time-sensitive analysis of aerosols and thermal reaction products from baking and roasting processes of foodstuff. Key features of the setup are a cold plasma-based ionization together with operability with the smallest amounts of samples in a highly controlled environment. FT-ICR-MS delivered time-resolved ultrahigh-resolution fingerprints during the study. In comparison to established online analytical platforms, we showed that DBDI unveils various molecular spaces populated by 6161 features during thermal processing

of wheat bread rolls at 210 °C. Graphical tools allowed us to identify different molecular and temporal reaction patterns within the system known from the Maillard reaction and caramelization processes. Unsupervised machine learning further helps us decipher the complex temporal system into an early "baking phase" and a later "roasting phase", which shows the potential real-time monitoring of thermal food processing has to offer.

## ■ ASSOCIATED CONTENT

### ⓈI Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.2c04874.

> Nutritional information and formulation of wheat bread rolls; supplementary methods for sampling of oven baking emissions; bulk parameters of molecules detected through a roasting experiment at 210 °C; van Krevelen plots of SOM clusters 2−5; GIF file visualizing the baking process of wheat bread roll dough in real time acquired by DBDI-FT-ICR-MS; photos of the heated sample throughout the experiment; annotated molecular formulas of compared API sources; and raw FT-ICR-MS spectra of all compared source types (PDF)

## ■ AUTHOR INFORMATION

### Corresponding Author

**Philippe Schmitt-Kopplin** − *Comprehensive Foodomics Platform, Chair of Analytical Food Chemistry, TUM School of Life Sciences, Technical University of Munich, 85354 Freising, Germany; Helmholtz Zentrum München, Analytical BioGeoChemistry, 85764 Neuherberg, Germany;* Phone: +498931873246; Email: schmitt.kopplin@tum.de

### Authors

**Leopold Weidner** − *Comprehensive Foodomics Platform, Chair of Analytical Food Chemistry, TUM School of Life Sciences, Technical University of Munich, 85354 Freising, Germany; Helmholtz Zentrum München, Analytical BioGeoChemistry, 85764 Neuherberg, Germany;* ⓞ orcid.org/0000-0002-6801-3647

**Daniel Hemmler** − *Comprehensive Foodomics Platform, Chair of Analytical Food Chemistry, TUM School of Life Sciences, Technical University of Munich, 85354 Freising, Germany; Helmholtz Zentrum München, Analytical BioGeoChemistry, 85764 Neuherberg, Germany*

**Michael Rychlik** − *Helmholtz Zentrum München, Analytical BioGeoChemistry, 85764 Neuherberg, Germany*

Complete contact information is available at:
https://pubs.acs.org/10.1021/acs.analchem.2c04874

### Author Contributions

L.W.: Conceptualization, methodology, formal analysis, investigation, software, validation, visualization, project administration, writing—original draft. D.H.: Conceptualization, methodology, funding acquisition, project administration, supervision. M.R.: Resources, supervision. P.S.-K.: Conceptualization, methodology, formal analysis, funding acquisition, supervision. All authors: Writing—review & editing. All authors have given approval to the final version of the manuscript

### Notes

The authors declare no competing financial interest.

## ■ ACKNOWLEDGMENTS

## ■ REFERENCES

(1) Wieland, F.; Gloess, A. N.; Keller, M.; Wetzel, A.; Schenker, S.; Yeretzian, C. *Anal. Bioanal. Chem.* **2012**, *402*, 2531−2543.
(2) Huang, Y.; Barringer, S. A. *J. Food Sci.* **2011**, *76*, C279−C286.
(3) Fischer, M.; Wohlfahrt, S.; Varga, J.; Matuschek, G.; Saraji-Bozorgzad, M. R.; Walte, A.; Denner, T.; Zimmermann, R. *Food Anal. Methods* **2017**, *10*, 49−62.
(4) Adelina, N. M.; Wang, H.; Zhang, L.; Zhao, Y. *Food Res. Int.* **2021**, *140*, No. 110026.
(5) Pico, J.; Khomenko, I.; Capozzi, V.; Navarini, L.; Biasioli, F. *Foods* **2020**, *9*, No. 1498.
(6) Al-Rashdan, A.; Helaleh, M. I. H.; Nisar, A.; Ibtisam, A.; Al-Ballam, Z. *Int. J. Anal. Chem.* **2010**, *2010*, 1−8.
(7) Cho, I. H.; Peterson, D. G. *Food Sci. Biotechnol.* **2010**, *19*, 575−582.
(8) Taylor, A. J.; Beauchamp, J. D.; Langford, V. S. *Dynamic Flavor: Capturing Aroma Using Real-Time Mass Spectrometry*, ACS Symposium Series; American Chemical Society, 2021; Vol. *1402*, pp 1−16.
(9) Biasioli, F.; Yeretzian, C.; Märk, T. D.; Dewulf, J.; van Langenhove, H. *TrAC, Trends Anal. Chem.* **2011**, *30*, 1003−1017.
(10) Smith, D.; Španěl, P. *Mass Spectrom. Rev.* **2005**, 661−700.
(11) Pratt, K. A.; Prather, K. A. *Mass Spectrom. Rev.* **2012**, *31*, 17−48.
(12) Barro, R.; Regueiro, J.; Llompart, M.; Garcia-Jares, C. *J. Chromatogr. A* **2009**, *1216*, 540−566.
(13) Zehentbauer, G.; Krick, T.; Reineccius, G. A. *J. Agric. Food Chem.* **2000**, *48*, 5389−5395.
(14) Ayala-Cabrera, J. F.; Montero, L.; Meckelmann, S. W.; Uteschil, F.; Schmitz, O. *J. Anal. Chim. Acta* **2022**, *1238*, No. 340353.
(15) Kanawati, B.; Bader, T. M.; Wanczek, K. P.; Li, Y.; Schmitt-Kopplin, P. *Rapid Commun. Mass Spectrom.* **2017**, *31*, 1607−1615.
(16) Moritz, F.; Kaling, M.; Schnitzler, J. P.; Schmitt-Kopplin, P. *Plant Cell Environ.* **2017**, *40*, 1057−1073.
(17) Meyer, C.; Müller, S.; Gurevich, E. L.; Franzke, J. *Analyst* **2011**, *136*, 2427−2440.
(18) Wolf, J. C.; Gyr, L.; Mirabelli, M. F.; Schaer, M.; Siegenthaler, P.; Zenobi, R. *J. Am. Soc. Mass Spectrom.* **2016**, *27*, 1468−1475.
(19) Weber, M.; Wolf, J. C.; Haisch, C. *J. Am. Soc. Mass Spectrom.* **2021**, *32*, 1707−1715.
(20) GitHub. https://github.com/leopold-weidner/DBDIpy (accessed October 27, 2022).
(21) GitHub. https://github.com/JustGlowing/minisom (accessed August 17, 2022).
(22) Gyr, L.; Klute, F. D.; Franzke, J.; Zenobi, R. *Anal. Chem.* **2019**, *91*, 6865−6871.
(23) Weidner, L.; Yan, Y.; Hemmler, D.; Rychlik, M.; Schmitt-Kopplin, P. *Food Chem.* **2022**, *374*, No. 131618.
(24) Hertkorn, N.; Frommberger, M.; Witt, M.; Koch, B. P.; Schmitt-Kopplin, P.; Perdue, E. M. *Anal. Chem.* **2008**, *80*, 8908−8919.
(25) Kim, S.; Kramer, R. W.; Hatcher, P. G. *Anal. Chem.* **2003**, *75*, 5336−5344.
(26) Hemmler, D.; Roullier-Gall, C.; Marshall, J. W.; Rychlik, M.; Taylor, A. J.; Schmitt-Kopplin, P. *Sci. Rep.* **2018**, *8*, No. 16879.
(27) Zvively, M. *Perfumer & Flavorist*, 2006. www.PerfumerFlavorist.com/articles
(28) Koch, B. P.; Dittmar, T. *Rapid Commun. Mass Spectrom.* **2006**, *20*, 926−932.
(29) Peters, K. E.; Xia, X.; Pomerantz, A. E.; Mullins, O. C. Geochemistry Applied to Evaluation of Unconventional Resources. In *Unconventional Oil and Gas Resources Handbook*; Gulf Professional Publishing, 2016.

(30) Parthasarathy, P.; Narayanan, S. K. *Environ. Prog. Sustainable Energy* **2014**, *33*, 676−680.

(31) Majchrzak, T.; Marć, M.; Wasik, A. *Food Res. Int.* **2022**, *160*, No. 111716.

## A.4  Article IV: DBDIpy: a Python library for processing of untargeted datasets from real-time plasma ionization mass spectrometry.

OXFORD

## Systems biology

# DBDIpy: a Python library for processing of untargeted datasets from real-time plasma ionization mass spectrometry

Leopold Weidner [1,2,*], Daniel Hemmler[1,2], Michael Rychlik [1]
and Philippe Schmitt-Kopplin[1,2,*]

[1]Comprehensive Foodomics Platform, TUM School of Life Sciences, Technical University of Munich, Freising 85354, Germany and
[2]Analytical BioGeoChemistry, Helmholtz Zentrum Muenchen, Neuherberg 85764, Germany

*To whom correspondence should be addressed.

Associate Editor: Janet Kelso

## Abstract

**Motivation:** Plasma ionization is rapidly gaining popularity for mass spectrometry (MS)-based studies of volatiles and aerosols. However, data from plasma ionization are delicate to interpret as competing ionization pathways in the plasma create numerous ion species. There is no tool for detection of adducts and in-source fragments from plasma ionization data yet, which makes data evaluation ambiguous.

**Summary:** We developed DBDIpy, a Python library for processing and formal analysis of untargeted, time-sensitive plasma ionization MS datasets. Its core functionality lies in the identification of in-source fragments and identification of rivaling ionization pathways of the same analytes in time-sensitive datasets. It further contains elementary functions for processing of untargeted metabolomics data and interfaces to an established ecosystem for analysis of MS data in Python.

**Availability and implementation:** DBDIpy is implemented in Python (Version $\geq$ 3.7) and can be downloaded from PyPI the Python package repository (https://pypi.org/project/DBDIpy) or from GitHub (https://github.com/leopold-weidner/DBDIpy).

**Contact:** leopold.weidner@tum.de or schmitt-kopplin@tum.de

**Supplementary information:** Supplementary data are available at *Bioinformatics* online.

## 1 Introduction

Application of plasma-based ion sources for real-time mass spectrometry (MS) shows exponential growth over the last years. This is because plasma ionization comes with relatively low instrumental cost, is simple to operate, very sensitive to a wide range of volatile analytes and is used in various fields like food processing-, environmental- or clinical breath research (Ayala-Cabrera *et al.*, 2022). Naturally, vendors of ion sources thrive to construct robust setups which reduce artifacts and maximize the generation of $[M+H]^+$ pseudo-molecular ions in comparison to competing, undesirable adduct ions. However, the chemical composition of plasma is non-trivial (Adamovich *et al.*, 2017), highly dependent of surrounding atmospheric gases and the design of the ion source. These factors account for miscellaneous ionization and fragmentation pathways in plasma ion sources (Gyr *et al.*, 2019; Wolf *et al.*, 2016). Exemplarily, saturated analytes are prone to in-source oxygenation reactions or hydride abstractions. Some alternating ionization pathways still are even not understood up to date

(Ayala-Cabrera *et al.*, 2022). Consequently, the user is tempted to mistake multiple ion species of the same analytes as unrelated compounds, leading to ambiguous conclusions during the evaluation of untargeted datasets. In related analytical disciplines, the community developed tools to simplify complex data structures. Exemplarily, the CAMERA algorithm annotates adduct peaks in LC-MS data to form compound spectra for a better annotation of features (Kuhl *et al.*, 2012). Currently, there is no computational tool to process data from direct infusion plasma-based ionization available; even though the growing community is in need for harmonized data processing pipelines to handle the challenges of miscellaneous plasma ionization pathways. Therefore, we developed the open-source Python package DBDIpy. Inspired by data from dielectric barrier discharge ionization (DBDI), we provide a novel computational algorithm, to automate the interpretation and to reduce the size of datasets generated by plasma ionization. DBDIpy groups multiple ion species of the same analyte, removes spectral artifacts and facilitates the evaluation of convoluted data to the user.

## 2 Implementation

DBDIpy is designed to handle and process data obtained from untargeted high-resolution real-time plasma ionization MS. It has interfaces to the matchms-ecosystem; a popular Python package for processing of MS data (Huber *et al.*, 2020). Data loaded and preprocessed by matchms can be imported by DBDIpy for further analysis. The central algorithm of DBDIpy performs the grouping of systematically occurring ion species from the same.

### 2.1 Core functionality

The grouping of non-$[M + H]^+$-ions with the pseudo-molecular ion is performed by a computational two-step open-search approach using the DBDIpy.identify_adducts() function: first, extracted ion chromatogram (XIC) shape similarities are calculated by computing pointwise Pearson correlation coefficient across all XIC pairs in the dataset. Second, highly correlated XICs are refined by mass difference analysis. Adducts and in-source fragments are identified from a set of pre-defined rules, which the user can flexibly customize. Exemplarily, the presence of two highly correlated XICs with a mass difference of $18.010565 \pm$ the error of the mass spectrometer implicates an in-source water loss. The output of DBDIpy.identify_adducts() is a dictionary holding one data frame for each defined adduct type. The data frames contain information on the corresponding XICs matches such as correlation coefficient or mass difference. A comprehensive description of DBDIpy's functions, the source code and an exemplary data-analysis workflow can be found on the GitHub repository and in the Supplementary Information. In brief, data are loaded and aligned by DBDIpy.import_spectra() via the matchms interface from matchms processing pipelines or from open file formats like .mgf. Following, missing values in the dataset are imputed by DBDIpy.impute_intensities() in preparation for adduct detection. This step consists of interpolating missing intensities within the signal region of the feature and of adding a noisy baseline to form uniform-length XICs. After this pre-treatment, adduct detection is performed as described above. The results of the adduct search can be visually inspected by calling DBDIpy.plot_adducts(): the temporal course of selected XICs, their correlation coefficients, mass differences and optionally supplied metadata will be shown to the user. This serves for validation of results and to investigate the grouped adduct systems. Finally, the DBDIpy.export_to_spectra() function permits data to be submitted to successive matchms data-handling or to be exported to open file formats.

### 2.2 Application

To showcase the utility of DBDIpy, we performed a demonstrational data analysis. The demo data are from a foodomics study where wheat bread was roasted and thermal reaction products were monitored by DBDI-MS (Weidner *et al.*, 2023). It consists of 4196 features. After importing and preparing the data, DBDIpy.identify_adducts() was used to search for in-source water losses and for one to four oxygen adducts (correlation of $r > 0.95$). Figure 1a gives an overview of the quantity of identified adducts. In total, 710 potential adducts were identified, which corresponds to 17% of all features. This finding emphasizes the importance to perform adduct detection on untargeted plasma ionization datasets. Exemplarily, the temporal profile of an in-source oxidation series of one single compound annotated as $[C_{15}H_{17}O_2N+H]^+$ is shown in Figure 1b. An independent, network-based annotation workflow (Moritz *et al.*, 2017) was used to validate the finding. It confirmed the four mass signals to form a cluster of systematically oxidized and structurally related features.

## 3 Conclusion

We introduced DBDIpy as the first software tool for the identification and curation of in-source fragment ions and non-pseudomolecular
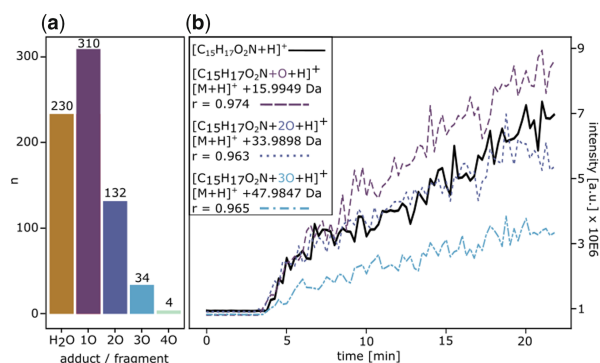


**Fig. 1.** (**a**) Total number of in-source fragments and oxygen adducts detected by DBDIpy.identify_adducts() in an untargeted foodomics dataset of 4196 features ($r > 0.95$). (**b**) Temporal course of oxygen adducts generated during ionization of feature $[C_{15}H_{17}O_2N+H]^+$

ions from time-sensitive plasma ionization MS data. By uniting tied features to groups of adducts, DBDIpy reduces the data size of extensive untargeted datasets and facilitates their interpretation. Analytical chemists from various disciplines can integrate DBDIpy to their workflows as a key processing step to ameliorate their knowledge about convoluted spectra.

## Funding

## Data availability

DBDIpy is an open-source package and publicly available on GitHub https://github.com/leopold-weidner/DBDIpy. A demo dataset can be downloaded on Zendo under doi.org/10.5281/zenodo.7221088.

## References

Adamovich,I. *et al.* (2017) The 2017 plasma roadmap: low temperature plasma science and technology. *J. Phys. D: Appl. Phys.*, **50**, 323001.

Ayala-Cabrera,J.F. *et al.* (2022) Review on atmospheric pressure ionization sources for gas chromatography-mass spectrometry. Part I: Current ion source developments and improvements in ionization strategies. *Anal. Chim. Acta*, **1238**, 340353.

Gyr,L. *et al.* (2019) Characterization of a nitrogen-based dielectric barrier discharge ionization source for mass spectrometry reveals factors important for soft ionization. *Anal. Chem.*, **91**, 6865–6871.

Huber,F. *et al.* (2020) matchms - processing and similarity evaluation of mass spectrometry data. *J. Open Source Softw.*, **5**, 2411.

Kuhl,C. *et al.* (2012) CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. *Anal. Chem.*, **84**, 283–289.

Moritz,F. *et al.* (2017) Characterization of poplar metabotypes via mass difference enrichment analysis. *Plant Cell Environ.*, **40**, 1057–1073.

Weidner,L. *et al.* (2023) Real-time monitoring of miniaturized thermal food processing by advanced mass spectrometric techniques. *Anal. Chem.*, **95**, 1694–1702.

Wolf,J.C. *et al.* (2016) A radical-mediated pathway for the formation of [M + H]+ in dielectric barrier discharge ionization. *J. Am. Soc. Mass Spectrom.*, **27**, 1468–1475.

# B Appendix: Supplementary materials

In the following chapter, I present the official Supplementary Material files accompanying the research articles printed in Appendix A.

Page numbers, layout and citations are conserved from the official publications and do not follow the style of this thesis.

## B.1 Elucidation of the non-volatile fingerprint in oven headspace vapor from bread roll baking by ultra-high resolution mass spectrometry - Supplementary Material

# Elucidation of the non-volatile fingerprint in oven headspace vapor from bread roll baking by ultra-high resolution mass spectrometry.

Weidner Leopold, Yingfei Yan, Hemmler Daniel, Rychlik Michael, Schmitt-Kopplin Philippe

## Supporting Information

Table of content:

**Supplementary Materials and Method for identification of non- and semi volatile compounds by LC-MS/MS:** To demonstrate the capability of the developed FT-ICR-MS screening method, the presence non- and semi- volatile molecules was confirmed by LC-MS/MS. Chemicals used were: Ultra-pure water from a Milli-Q integral system (Merck, Darmstadt, Germany). Arginine, lysine, histidine, proline, glutamine, glutamic acid, glycerol, niacin, glucose, ribose, maltose and succinic acid (all 98% purity or higher) and ammonium formate solution 10 M in water (BioUltra grade) were purchased from Sigma Aldrich (Taufkirchen, Germany). Formic acid (98% for mass spectrometry) was ordered from Honeywell (Seelze, Germany). Acetonitrile and 2-propanol (both hypergrade for LC-MS) were purchased from Merck (Darmstadt, Germany).

ESI-L Low Concentration Tuning Mix and an API-TOF Reference Mass Solution Kit were supplied by Agilent Technologies (Waldbronn, Germany). The analysis was performed on a 1290 Infinity II UPLC system coupled to a 6560B Ion Mobility Q-TOF mass spectrometer, both from Agilent technologies (Waldbronn, Germany). For analysis in ESI-(+) mode, chromatographic separation of vapor condensate was performed on a ZIC-cHILIC analytical column (100 x 2.1 mm, 3 μm, Merck, Darmstadt, Germany) and the operational conditions were as following: eluent A was 5% acetonitrile in ultra-pure water with 5 mM ammonium formate and 0.1% formic acid and eluent B was 95% acetonitrile in ultra-pure water with 5 mM ammonium formate and 0.1% formic acid. The column was held at 40 °C during analysis and a constant flow rate of 0.5 mL min$^{-1}$ was applied. For analysis in ESI-(-) mode, a Hypercarb analytical column was used (100 x 2.1 mm, 3 μm, Thermo Fisher Scientific, Dreiech, Germany). The mobile phase constituted of ultra-pure water with 0.2% formic acid on channel A and of acetonitrile with 0.2% formic acid on channel B. A flow rate of 0.2 mL min$^{-1}$ was selected, and the column oven was set to 40 °C. The gradients used for chromatographic separation can be seen in Supplementary Table 1. A sample volume of 10 μl was injected. The mass spectrometer was calibrated with an ESI-L Low Concentration Tuning Mix prior to analysis and auto recalibration of MS spectra during acquisition was enabled by spraying API-TOF reference masses via the reference mass sprayer of the Dual AJS ESI source. Nitrogen was used as nebulizer gas at a flow rate of 12 L min$^{-1}$ and temperature of 250 °C as well as sheath gas at 11 L min$^{-1}$ and 250 °C. The VCap was set to 4.000 V and the Nozzle Voltage to 10 V. Data dependent acquisition was performed in Q-TOF only mode from 50 Da to 1700 Da at a scan rate of 6 Hz. A maximum number of three precursors per cycle was fragmented by collision induced dissociation with an energy of 20 eV. Standard solutions of non- and semi volatile food related compounds were prepared in 95% acetonitrile and injected at a concentration of 5 ppm under the same analytical condition as the condensate was analyzed. MS2 spectra were manually searched and compared with acquired standard reference spectra, the human metabolome database (HMDB) and the MassBank of North America (MoNA).

Supplementary Table 1 Nutritional composition of bread roll samples (per 100 g)

| Nutritional fact | |
|---|---|
| Calories | 1067 kJ / 252 kcal |
| Fat | 2.0 g |
| - Saturated | 0.4 g |
| Carbohydrates | 49.0 g |
| - Sugar | 1.9 g |
| Fiber | 2.8 g |
| Protein | 8.1 g |
| Salt | 1.4 g |

Dough ingredients were as the following:
Wheat flour, water, yeast, salt, rapeseed oil, dry rye sour dough, dextrose, mono- and diacetyl esters of tartaric acid from mono- and diglycerides of fatty acids, mono- and diglycerides of fatty acids, wheat malt flour, maltodextrin, powdered sweet whey, wheat gluten, starch

Supplementary Table 2 Information on baking temperatures and times

| coded level | - α | -1 | - 0.5 | 0 | 0.33 | 0.5 | 0.66 | 1 | α |
|---|---|---|---|---|---|---|---|---|---|
| time [min] | 7.8 | 9.0 | 10.5 | 12.0 | * | 13.5 | * | 15.0 | 16.2 |
| temp. [°C] | 158 | 170 | * | 200 | 210 | * | 220 | 230 | 242 |

*not processed

Validation experiments: VPI: 223 °C, 10 min, VPII : 180 °C, 14 min

Supplementary Table 3 Constitution of mobile phase for UPLC separation of vapor condensate

| cHILIC gradient for ESI(+)-analysis | | | Hypercarb gradient for ESI(-)-analysis | | |
|---|---|---|---|---|---|
| time [min] | %A | %B | time [min] | %A | %B |
| 0.0 | 2 | 98 | 0.0 | 90 | 10 |
| 3.0 | 2 | 98 | 1.0 | 90 | 10 |
| 11.0 | 30 | 70 | 6.0 | 0 | 100 |
| 12.0 | 40 | 60 | 9.5 | 0 | 100 |
| 16.0 | 95 | 5 | 10.0 | 90 | 10 |
| 18.0 | 95 | 5 | | | |
| 19.0 | 2 | 98 | | | |
| 20.0 | 2 | 98 | | | |

Supplementary Table 4 Food related compounds in oven vapor from bread rolls baked at 230 °C for 15 min identified by LC-MS/MS

| substance name | molecular formula | monoisotopic mass [Da] | log P[1] | confidence of identification[2] |
|---|---|---|---|---|
| Glucose | $C_6H_{12}O_6$ | 180.063388 | -2.6 | identified |
| Ribose | $C_5H_{10}O_5$ | 150.052823 | -2.5 | identified |
| Maltose | $C_{12}H_{22}O_{11}$ | 342.116212 | -4.7 | identified |
| 5-Hydroxymethyl furfural | $C_6H_6O_3$ | 126.031694 | -0.6 | identified |
| Lysine | $C_6H_{14}N_2O_2$ | 146.105528 | -3.0 | identified |
| Trimethyl lysine | $C_9H_{21}N_2O_2^+$ | 189.159754 | -2.1 | putatively annotated |
| Arginine | $C_6H_{14}N_4O_2$ | 174.111676 | -4.2 | identified |
| N,N-Dimethyl arginine | $C_8H_{18}N_4O_2$ | 202.142976 | -3.6 | putatively annotated |
| Ornithine | $C_5H_{12}N_2O_2$ | 132.089878 | -4.4 | putatively annotated |
| Histidine | $C_6H_9N_3O_2$ | 139.074562 | -3.2 | identified |
| Histamine | $C_5H_9N_3$ | 111.079647 | -0.7 | putatively annotated |
| Proline | $C_5H_9NO_2$ | 115.063329 | -2.5 | identified |
| Glutamic acid | $C_5H_9NO_4$ | 147.053158 | -3.7 | identified |
| Glutamine | $C_5H_{10}N_2O_3$ | 146.069142 | -3.1 | identified |
| Succinic acid | $C_4H_6O_4$ | 118.026609 | -0.6 | identified |
| Palmitic acid | $C_{16}H_{32}O_2$ | 256.240230 | 6.4 | identified |
| Stearic acid | $C_{18}H_{36}O_2$ | 284.271530 | 7.2 | identified |
| Oleic acid | $C_{18}H_{34}O_2$ | 282.255880 | 6.5 | putatively annotated |
| Glycerol | $C_3H_8O_3$ | 92.047344 | -1.8 | identified |
| Niacin | $C_6H_5NO_2$ | 123.032028 | 0.4 | identified |
| Betaine | $C_5H_{11}NO_2$ | 117.078979 | 0.5 | putatively annotated |

1 XLogP3 3.0 (PubChem release 2021.05.07), 2 according to reporting standards for metabolomic experiments (Sumner et al., 2007). **Identified compounds**: analysis of pure chemical reference standard with same analytical LC-MS/MS method: Comparison of retention time, exact mass and MS2 fragmentation pattern. **Putatively annotated compound**: exact mass and comparison of MS2 spectral similarity with public repositories (e.g. Human Metabolome Data Base or MassBank of North America)
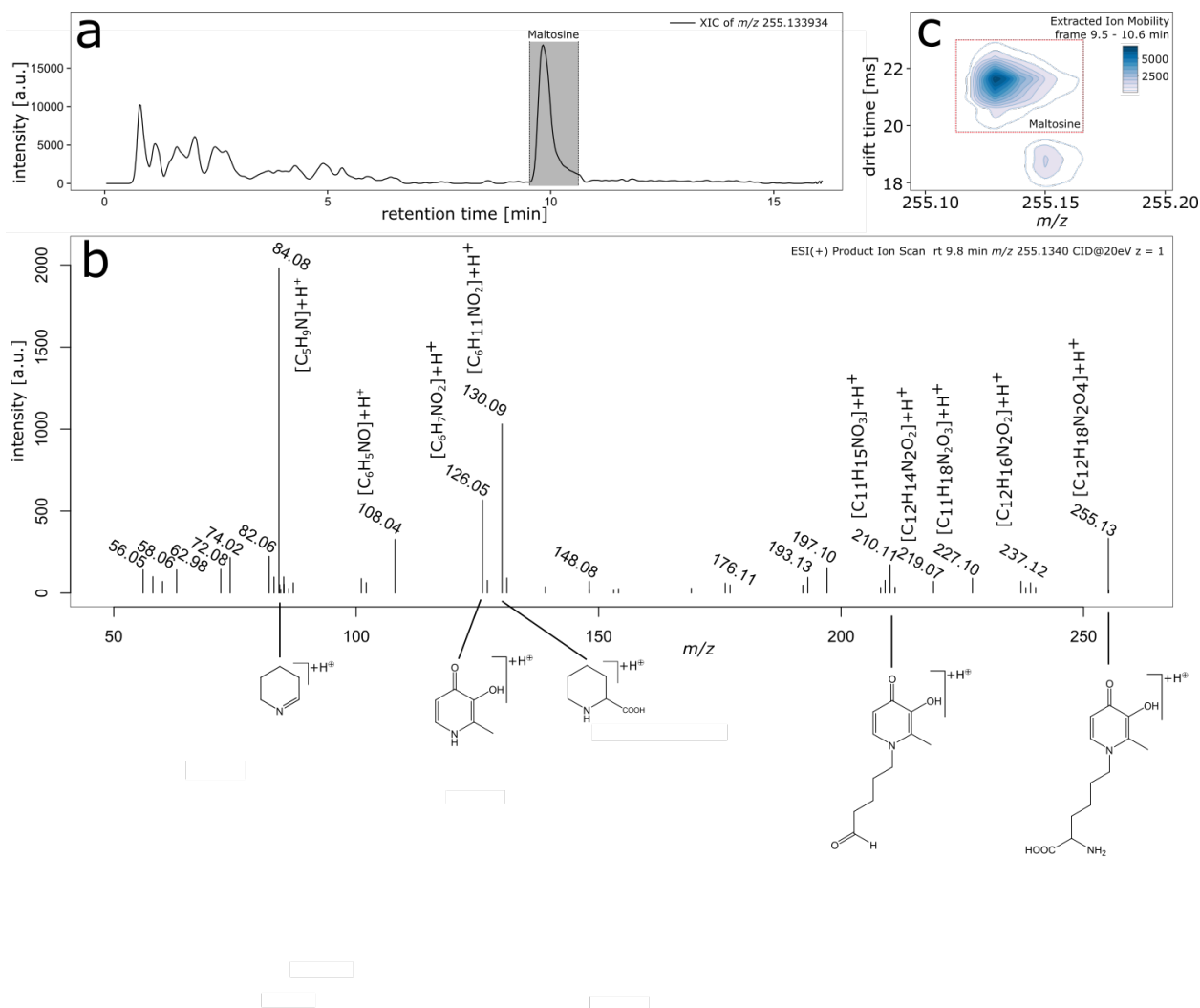
# Bulk data of all experiments
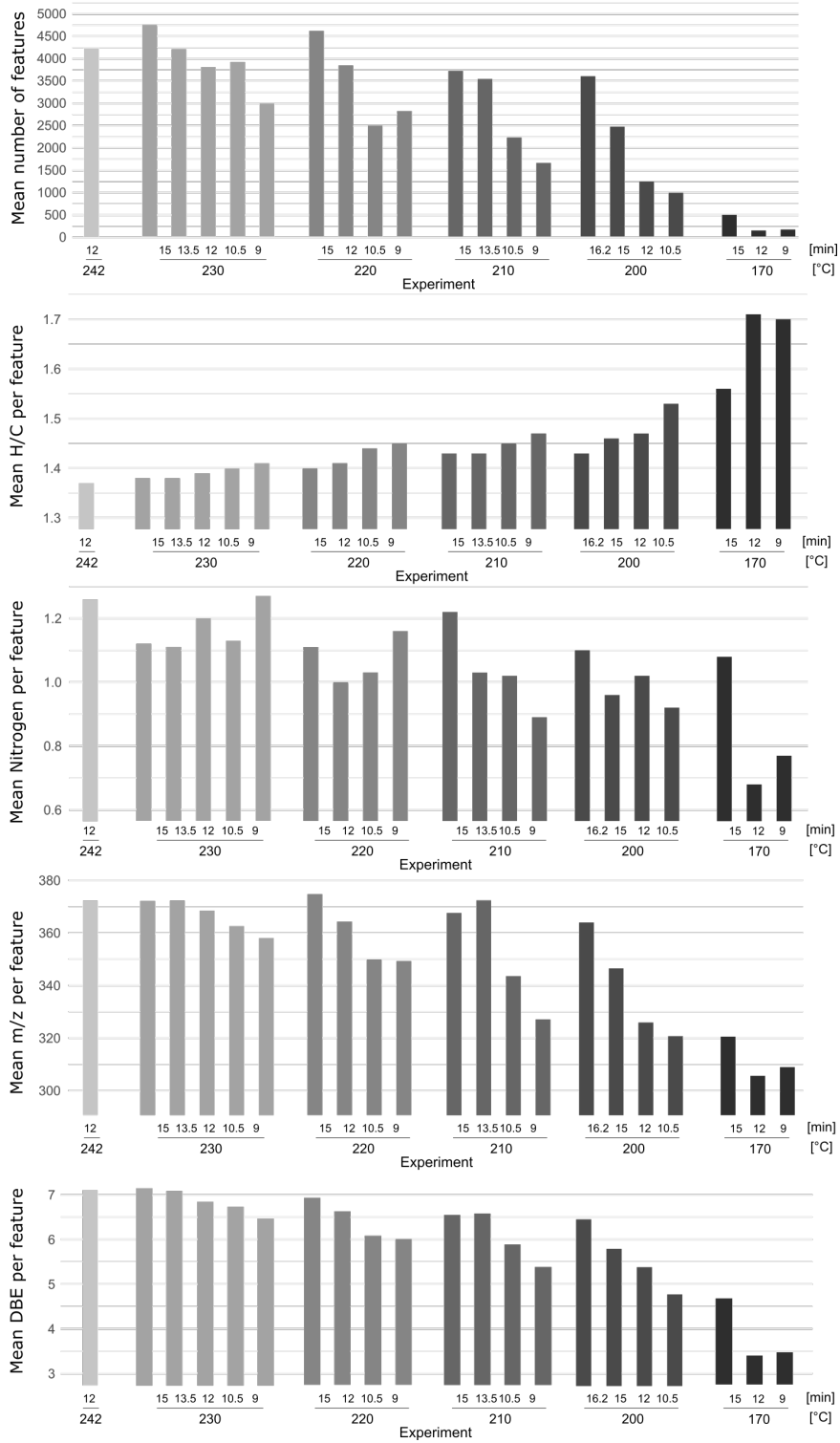


Supplementary Figure 1 Bulk data calculated for all detected features calculated throughout different baking conditions.
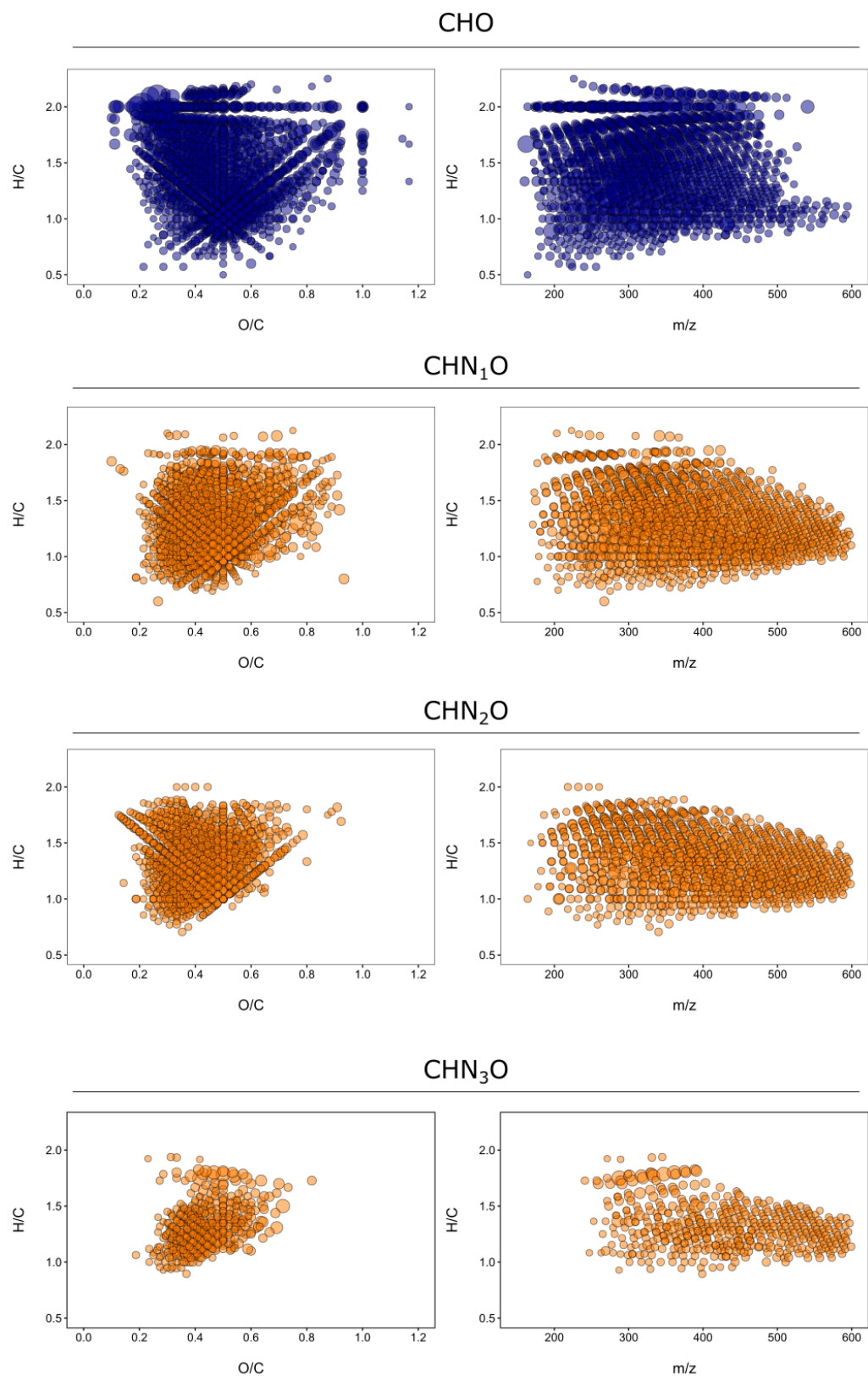
Supplementary Figure 2 Van-Krevelen- with corresponding H/C to $m/z$-plots of six bread rolls baked at 230 °C for 15 min divided by chemical spaces (point size proportional to signal intensity).

Supplementary Figure 3 Validation of maltosine: a condensate sample of wheat bread rolls baked at 230 °C for 15 min measured by LC-MS/MS (see supplementary method for details). **a** Extracted ion chromatogram of *m/z* 255.133934 (± 10 ppm) with the maltosine peak on grey background. **b** An annotated MS2 spectrum (CID, 20 eV) of maltosine. The fragmentation pattern stands in accordance with literature and spectra obtained from model systems. **c** Ion mobility drift plot of the extracted frame from 9.5 to 10.6 min reveal no coeluting isomeric compounds in the condensate sample.

## B.2 Molecular characterization of cooking processes: a metabolomics decoding of vaporous emissions for food markers and thermal reaction indicators.

# Molecular characterization of cooking processes: a metabolomics decoding of vaporous emissions for thermal reaction indicators and food markers

Leopold Weidner, Jil Vittoria Cannas, Michael Rychlik, Philippe Schmitt-Kopplin

## Supporting Information

Supplementary Table I: Formulation of foods and processing conditions

| Food Category | Sample | Temperature [°C] | Time [min] | Ingredients / Species |
|---|---|---|---|---|
| Bakery products | Wheat bread, Froneri, cold | 200 | 12 | Wheat flour, water, yeast, salt, rapeseed oil, dry rye sour dough, dextrose, mono- and diacetyl esters of tartaric acid from mono- and diglycerides of fatty acids, mono- and diglycerides of fatty acids, wheat malt flour, maltodextrin, powdered sweet whey, wheat gluten, starch |
| | Wheat bread, Froneri, hot | 240 | 12 | |
| | Wheat bread, Aryzta | 200 | 12 | Wheat flour, water, wheat sourdough (wheat flour, water, rye flour), yeast, rye flour, salt, canola oil, barley malt extract, dextrose, wheat malt flour, wheat grain, maltodextrin, sweet whey powder (milk), emulsifier soy lecithins |
| | Wheat bread, Coppenrat & Wiese | 190 | 12-20 | wheat flour, water, barley malt flour, iodized table salt (table salt, potassium iodate), yeast, sugar, vegetable oil (canola) |
| | Rye bread, Coppenrat & Wiese | | | water, rye flour, wheat flour, wheat gluten, salt, yeast, barley malt flour, barley malt extract, rye dough, sugar, emulsifier (soya lecithin) |
| | Full grain bread, Coppenrat & Wiese | | | Wheat flour (44%), water, sunflower seeds (8%), rye flour and whole grain meal (4%), sesame (3%), soybean meal, blue poppy seeds (2%), linseed (2%), iodized table salt (table salt, potassium iodate), yeast, canola oil, wheat malt flour, spice |
| | Kartoffelrösti | | | Potatoes, sun flower oil, potato starch, table salt, roasted onions (onions, palm oil, wheat flour, salt), glucose, spice |
| | Pain au Chocolat | | | Wheat flour, butter, water, chocolate (sugar, cocoa mass, cocoa butter, emulsifier (soya lecithin)), sugar, yeast, wheat gluten, egg, table salt, flour treatment agent (ascorbic acid, enzymes) |
| | Croissant | | | E300, wheat flour, butter: 23%, amylase, hemicellulase, water, invert sugar syrup, yeast, eggs, salt, gluten, flour treatment agent, emulsifier E322: rapeseed lecithin. |
| | Brezn | | | Wheat flour, water, yeast, table salt, baking agent (wheat flour, wheat malt flour, dextrose, emulsifier: lecithins (contains soy), mono- and diglycerides of fatty acids; stabilizer: guar gum; flour treatment agents: amylases, lipases, xylanases, proteases), pretzel liquor (water, acidity regulator: sodium hydroxide) |
| | Apfelstrudel | | | Apples, wheat flour, plant margarine (palm fat, sun flower oil, water, mono- and diglycerides of edible oils, citric acid), water, bread crumbs (wheat flour, sugar, cinnamon, salt, yeast, raisins, sugar, salt, cinnamon, ascorbic acid, natural flavor, citric acid |
| | Topfenstrudel | | | Wheat flour, water, vegetable margarine (palm fat, sunflower oil, water, emulsifier: mono- and diglycerides of fatty acids, salt, acidity regulator: citric acid), curd (milk) 14%, sugar, raisins, modified starch, corn starch, salt, thickener: E464, flavoring, acidifier: citric acid. |

| Category | Item | | | |
|---|---|---|---|---|
| Meat | Krustenbraten without skin | 220 | 16 | *Sus scrofa domesticus* |
| | Krustenbraten with skin | | | *Sus scrofa domesticus* |
| | Barbarie without skin | | | *Cairina moschata* |
| | Barbarie with skin | | | *Cairina moschata* |
| | Rump steak cold | 210 | 8 | *Bos taurus* |
| | Rump steak hot | 270 | 8 | *Bos taurus* |
| | Chicken breasts without skin | 220 | 16 | *Gallus gallus domesticus* |
| | Chicken breasts with skin | | | *Gallus gallus domesticus* |
| | Chicken breasts hot | 250 | 16 | *Gallus gallus domesticus* |
| | Kalbsleberkas | 220 | 8 | Pork, veal, bacon, salt, spice, dextrose, spice extract, E450, E21 |
| | Geflügelleberkas | | | Pork, poultry, bacon, salt, spice, dextrose, spice extract, E450, E21 |
| Fish | Salmon filet | 220 | 10 | *Salmo salar* |
| | Cod fish filet | | | *Gadus morhua* |
| | Pangasius filet | | | *Pangasianodon hypophthalmus* |
| | Tuna steak | | | *Thunnus albacares* |
| | Fish sticks, bread crumb coated | | | Pacific polar cod (*Theragra chalcogramma*) 65%, breadcrumbs (wheat flour, water, spices (paprika, turmeric), yeast, salt), canola oil, wheat flour, starch, water, salt. |
| | Alaska pollock, bread crumb coated | | | Alaska pollock (52%), wheat flour, water, sunflower oil, potato starch, wheat starch, cooking salt, raising agents (E450, E500), dextrose, spices (paprika, garlic powder, turmeric) |
| Vegetables | Carrots | 210 | 10 | *Daucus carota subsp. sativus* |
| | Cabbage | | | *Brassica oleracea convar. capitata var. alba* |
| | Brussel sprouts | | | *Brassica oleracea var. gemmifera* |
| | Onions | | | *Allium cepa* |
| | Zucchini | | | *Cucurbita pepo subsp. pepo convar. giromontiina* |
| | Potatoes (natural) | | 13 | *Solanum tuberosum* |
| | French fries, Gut & Günstig | | 20 | Potatoes, palm oil, dextrose |
| | French fries, Agrafrost | | | Potatoes, 10% sunflower oil, wheat flour, salt, modified starch, spices |

| | | | | |
|---|---|---|---|---|
| | Sweet potatoes fries | | 15 | 86% sweet potato, 8% batter (rice flour, modified starch, potato dextrin, corn flour, table salt, raising agents: diphosphates, sodium carbonates; stabilizer: xanthan gum; spice extract), 6% vegetable oil (sunflower, canola). |
| Plant-based meat alternatives | Jackfruit patties | 220 | 11 | Jackfruit 61%, oat flakes, onions, rapeseed oil, pea flour, tomato paste, sugar, soy sauce (water, soy bones, iodized table salt), herbs, spices, iodized table salt, smoked table garlic, citric acid |
| | Mühlenschnitzel, Rügenwalder | | | Drinking water, wheat flour, 11% soy protein, canola oil, wheat gluten, oat fiber, table salt, thickener: methyl cellulose, corn flour, natural flavor, wheat starch, distilled vinegar, spices, sugar, psyllium husk, yeast. |
| | Mühlenfrickadelle, Rügenwalder | | | Drinking water, 15% wheat gluten, onions, canola oil, 9% wheat flour, distilled vinegar, table salt, thickener: methyl cellulose, citrus fiber, natural flavor, spice extracts, dextrose, spices, maltodextrin, yeast. |
| | No Mince, Vegetarian Butcher | | | 91% soy base (drinking water, textured soy protein, barley malt extract), spice extracts, spirit vinegar / spirit vinegar , dextrose, table salt, yeast extract, natural flavor, spices, caramelized/caramelized sugar, garlic powder, onion powder, ferric diphosphate (ferric pyrophosphate), vitamin B12. |
| | No Chicken Nuggets, Vegetarian Butcher | | | 58% soy base (drinking water, soy protein), cornflakes (corn, sugar, table salt), vegetable oils (canola, sunflower), flour (corn, rice), oat fiber, modified starch, natural flavors, thickener (methyl cellulose), vinegar powder (spirit vinegar / spirit vinegar, acidifier (sodium hydroxide)), table salt, starch, soy protein, acidifier (citric acid), ferric diphosphate (ferric pyrophosphate), vitamin B12. |
| | No Beef Burger, Vegetarian Butcher | | | 71% soy base (drinking water, textured soy protein, wheat protein, wheat starch), sunflower oil, onions, egg white, starch, spice extract, spices, caramelized / caramelized sugar, natural flavors, thickener (processed euchema seaweed), dextrose, palm fat, flavors, table salt, ferric diphosphate (ferric pyrophosphate), vitamin B12. |
| | No Hot Dog, Vegetarian Butcher | | | 52% soy base (drinking water, soy protein ), vegetable oils (rapeseed, palm), field beans, starch, wheat gluten, potato protein, thickener (methyl cellulose), natural flavors, stabilizers (sodium alginate, calcium chloride, processed euchema seaweed), table salt, spices, smoke flavoring, psyllium husk, yeast extract, citrus fiber, acidifier (citric acid), ferric diphosphate (ferric pyrophosphate), colorant (iron oxide, iron hydroxide), vitamin B12. |
| | Beyond Burgers | | | Water, pea protein (16%), canola oil, coconut oil, rice protein, flavor, stabilizer (methyl cellulose), potato starch, apple extract, coloring (beet), maltodextrin, pomegranate extract, salt, potassium salt, lemon juice concentrate, corn vinegar, carrot powder, emulsifier (sunflower lecithin) |
| | Vegetarisches Knusperfilet, FischVomFeld | | | Salsify (37%), WHEAT flour, jackfruit (8%), water, sunflower oil, cauliflower (4%), partially oiled hemp seed powder (4%), borlotti beans (3%), table salt, linseed oil |

| | | |
|---|---|---|
| | | (1%), thickener methyl cellulose, lemon juice from lemon juice concentrate, yeast, spices (paprika, turmeric). |
| Vegetarischer Backfisch, FischVomFeld | | Black salsify (37%), wheat flour, jackfruit (8%), water, sunflower oil, cauliflower (4%), partially oiled hemp seed powder (4%), borlotti beans (3%), table salt, linseed oil (1%), thickener methyl cellulose, lemon juice from lemon juice concentrate, yeast, spices (paprika, turmeric). |

| HILIC gradient for ESI(+)-analysis | | | HILIC gradient for ESI(-)-analysis | | | C8 gradient for ESI($\pm$)-analysis | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| time [min] | %A | %B | time [min] | %A | %B | time [min] | %A | %B |
| 0.0 | 2 | 98 | 0.0 | 4 | 96 | 0.0 | 95 | 5 |
| 3.0 | 2 | 98 | 2.0 | 4 | 96 | 1.9 | 95 | 5 |
| 11.0 | 30 | 70 | 5.5 | 12 | 88 | 8.9 | 5 | 95 |
| 12.0 | 40 | 60 | 8.5 | 12 | 88 | 13.7 | 2 | 98 |
| 16.0 | 95 | 5 | 9.0 | 14 | 86 | 14.7 | 95 | 5 |
| 18.0 | 95 | 5 | 14.0 | 14 | 86 | | | |
| 19.0 | 2 | 98 | 17.0 | 18 | 82 | | | |
| 20.0 | 2 | 98 | 23.0 | 35 | 65 | | | |
| | | | 24.0 | 35 | 65 | | | |
| | | | 24.5 | 4 | 96 | | | |
| | | | 26.0 | 4 | 96 | | | |
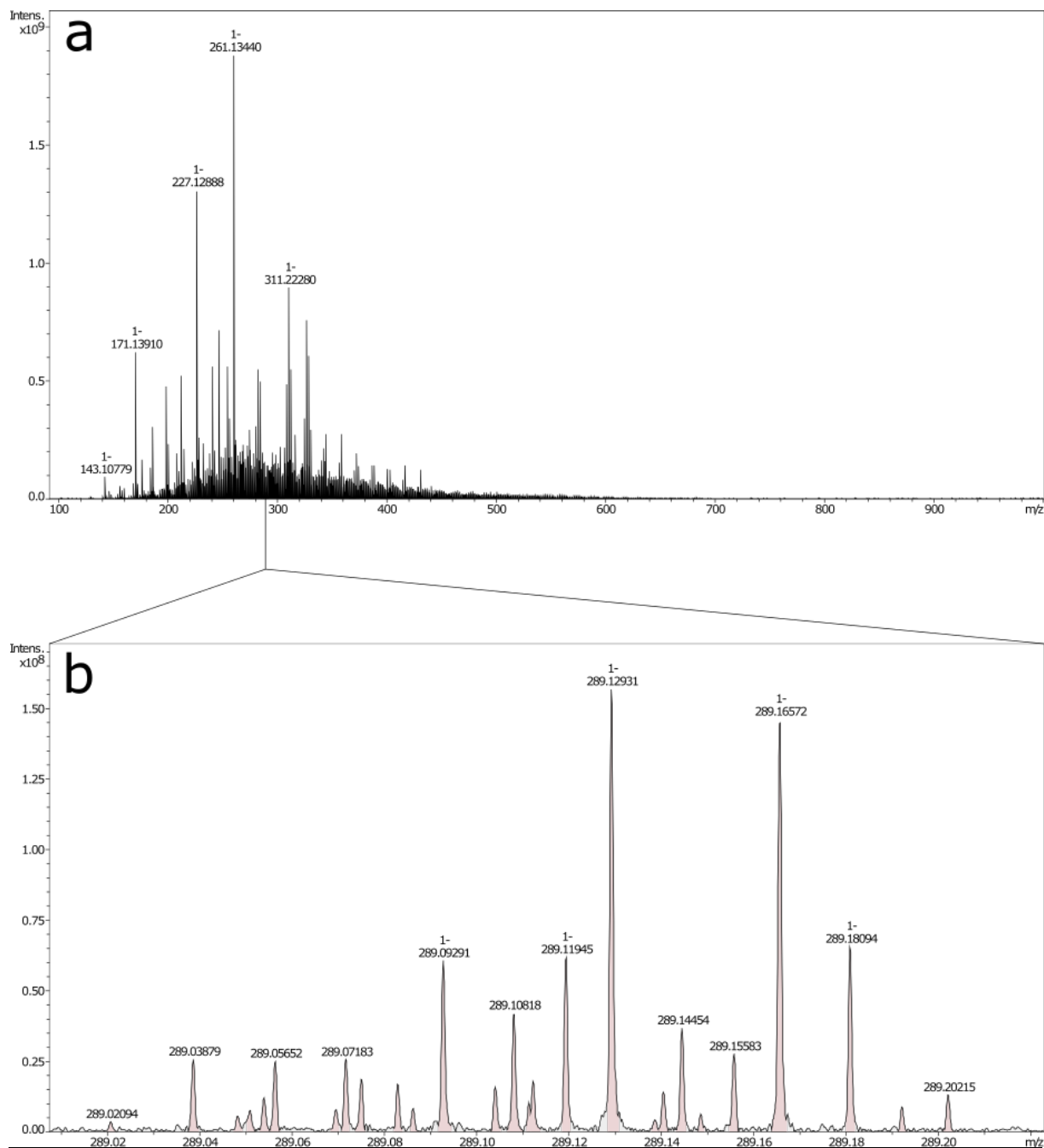
Supplementary Figure I **a** FT-ICR-MS spectrum of oven vapor collected from a cooking process of Beyond Burger patties. In one sample, 3256 monoisotopic mass signals with a resolution of 506.000 at $m/z$ 200 were detected. **b** Crop of the spectrum at $m/z$ 289 showing the molecular fingerprint contained in oven vapor with 25 resolved mass signals in one nominal mass.
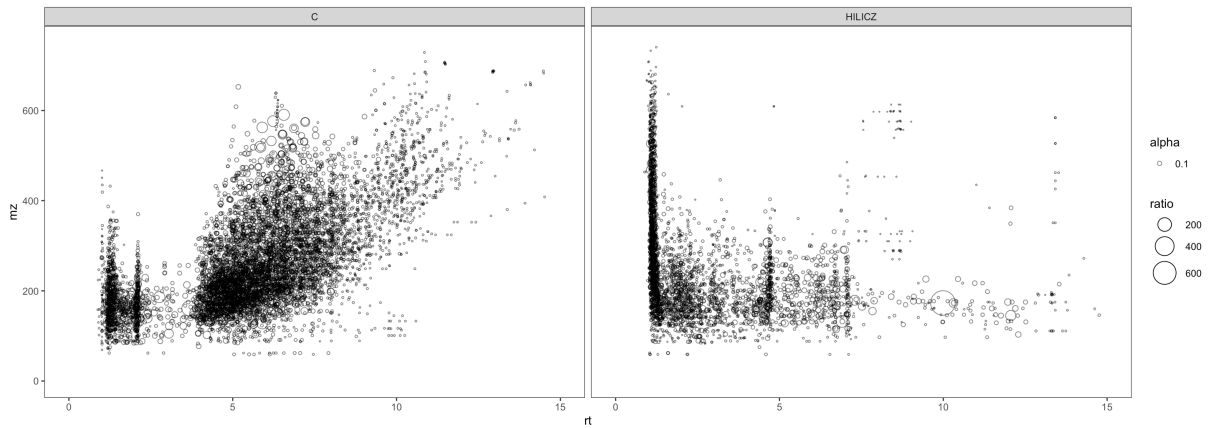
Supplementary Figure II Feature maps showing retention time vs. $m/z$ plots of C8 and HILIC separation of vapor samples.

Supplementary Figure III
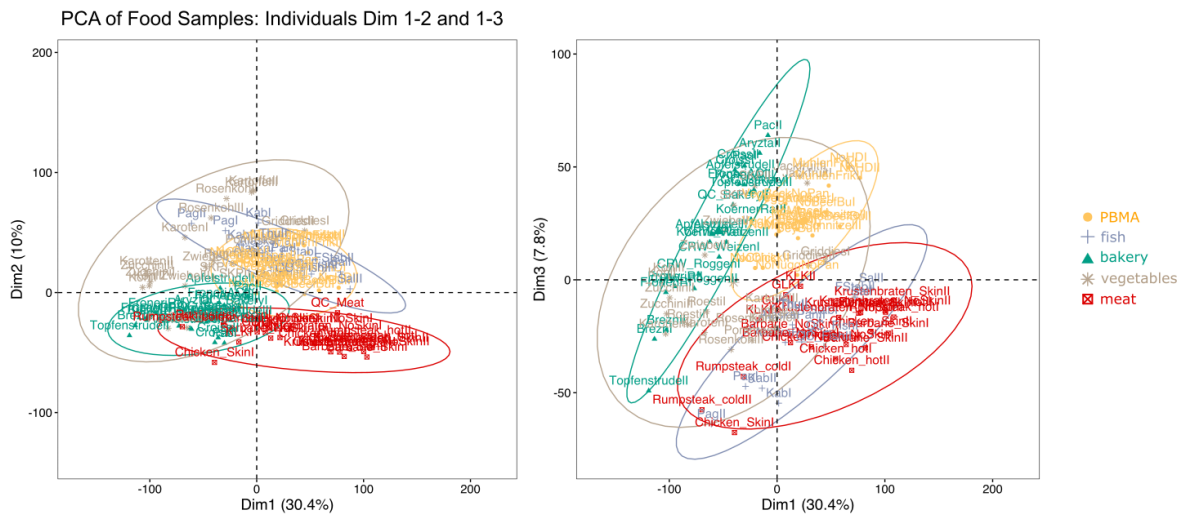


Supplementary Figure III Principal component analysis of food vapor samples on Dimensions 1 to 3

Supplementary Figure IV Correlation circle plot and contribution of variables to sPLS-DA separation of food samples.

## B.3 Real-Time Monitoring of Miniaturized Thermal Food Processing by Advanced Mass Spectrometric Techniques - Supplementary Material

Supporting Information

# Real-time monitoring of miniaturized thermal food processing by advanced mass spectrometric techniques

Leopold Weidner [a, b], Daniel Hemmler [a, b], Michael Rychlik [b], Philippe Schmitt-Kopplin [a, b, *]

a Comprehensive Foodomics Platform, Chair of Analytical Food Chemistry, TUM School of Life Sciences, Technical University of Munich, Maximus-von-Imhof-Forum 2, 85354 Freising, Germany
b Helmholtz Zentrum München, Analytical BioGeoChemistry, Ingolstädter Landstr. 1, 85764 Neuherberg, Germany
* Philippe Schmitt-Kopplin, email: schmitt.kopplin@tum.de, Telephone: +498931873246

Table of content

Supplementary Table I Nutritional information (per 100 g) and formulation of wheat bread rolls used for roasting experiments

| Nutritional fact | Value [per 100 g] |
|---|---|
| Calories | 1204 kJ/288 kcal |
| Fat | 2.0 g |
| -saturated | 0.2 g |
| Carbohydrates | 55.6 g |
| -sugar | 0.6 g |
| Protein | 9.4 g |
| Salt | 1.7 g |

Ingredients: wheat flour, water, yeast, rapeseed oil, roasted barley malt, wheat malt, sea salt

Supplementary Methods:

Vapor samples consisting of aerosols and thermal reaction products from bread roll baking were collected from a gastronomic oven and treated exactly as describes in https://doi.org/10.1016/j.foodchem.2021.131618. The oven was operated at 210 °C corresponding to the same baking conditions as applied for DBDI monitoring. Deviations from the reference were the following:

APCI and APPI measurements were realized by mounting respective APOLLO II ion sources on a 12 T solariX FT-ICR-MS instrument (all Bruker Daltonics, Bremen, Germany). Condensate samples were diluted by a factor of 1000 with methanol (v/v). Analysis was performed in direct flow injection analysis and positive ion mode by continuously delivering an analyte flow of 20 µl min$^{-1}$ with a mechanical syringe pump to the source. Source settings were the following: nebulizer gas flow 2.0 bar of Nitrogen, dry gas 3.0 l min$^{-1}$ of Nitrogen. The dry-temperature was set to 200 °C A voltage of 1500 V was applied to the capillary and an end plate offset of -500 V was selected. For APCI operation the heater was set to 350 °C. Spectra were recorded from 92 to 1000 Da with a time domain of 4 MegaWord. A total number of 500 scans with an ion accumulation time of 0.25 sec were accumulated per sample. Spectra were processed as the ones acquisitioned with electro-spray ion source which can be followed in above-named reference.
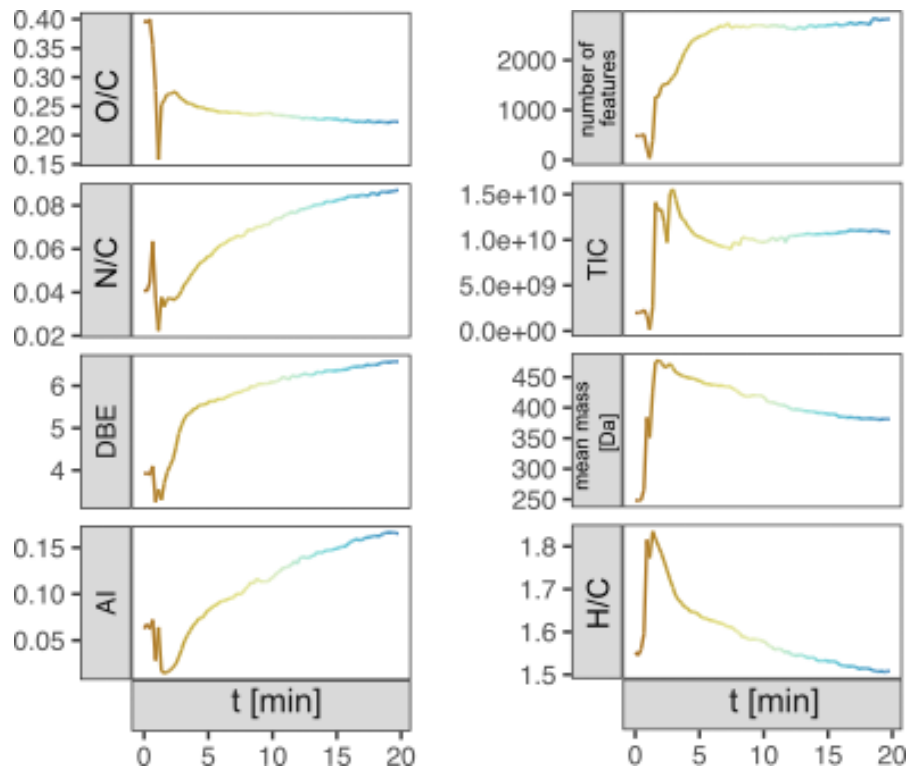
Supplementary Figure I Molecular bulk parameters of a bread roll sample baked at 210 °C for 20 min monitored by DBDI-FT-ICR-MS
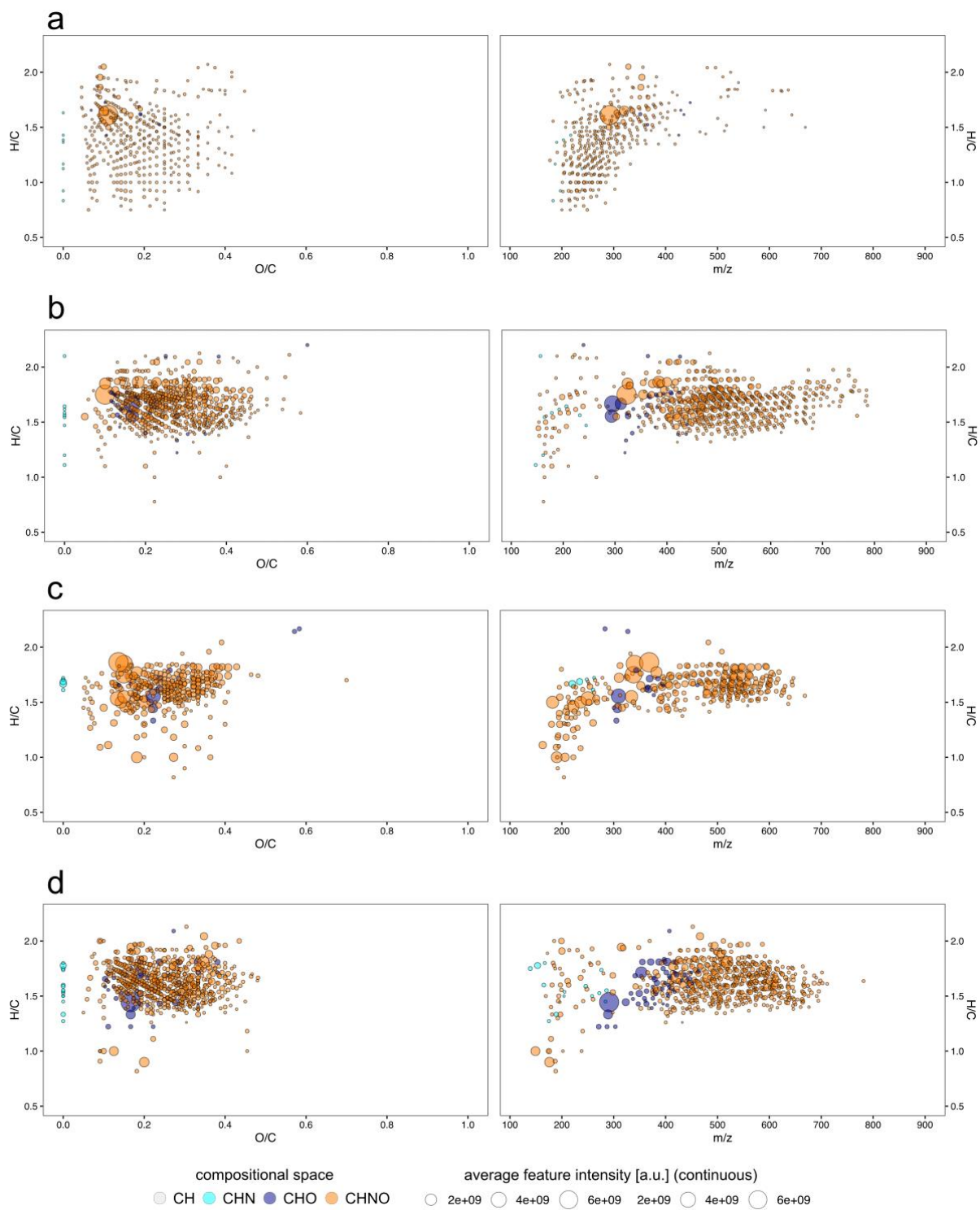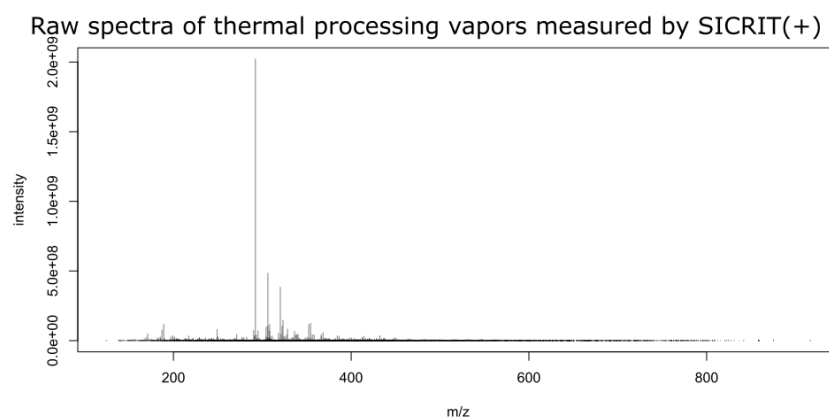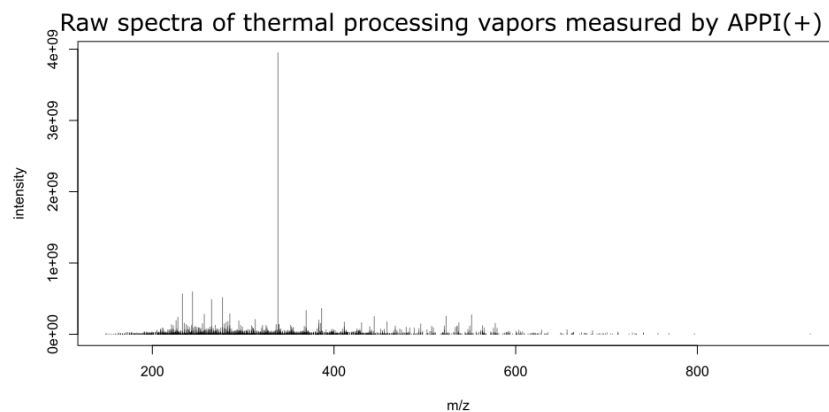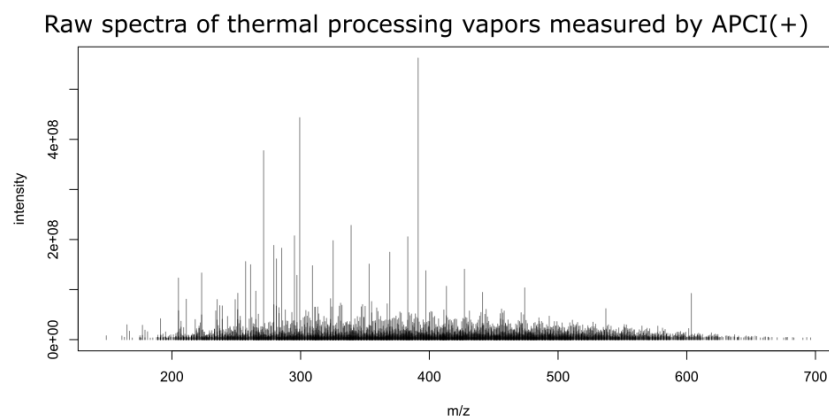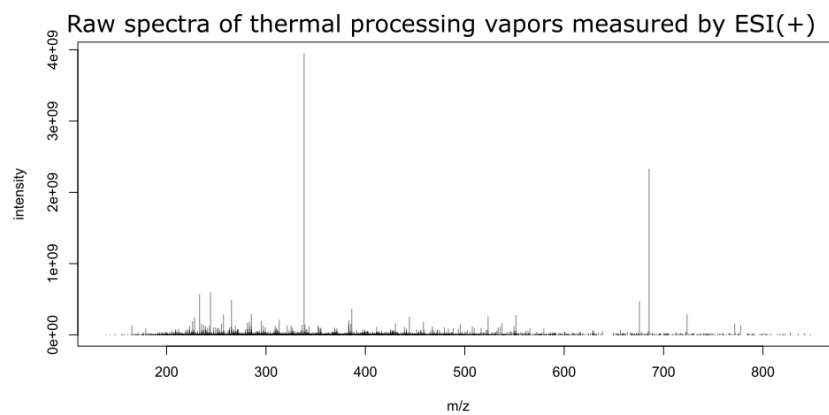
Supplementary Figure II Van-Krevelen- and respective H/C- to *m/z*-plots of Self-organizing-maps clusters 2 to 5
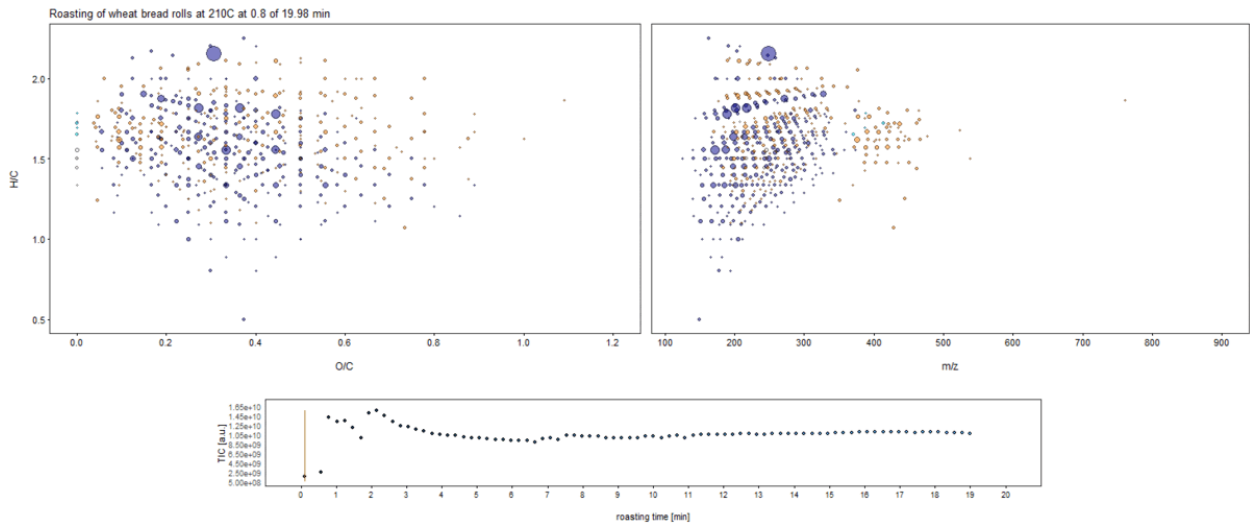
Supplementary Figure III An animated real-time view of a baking process at 210 °C monitored by DBDI-FT-ICR-MS visualized by Van-Krevelen- and respective H/C- to *m/z*-plots.
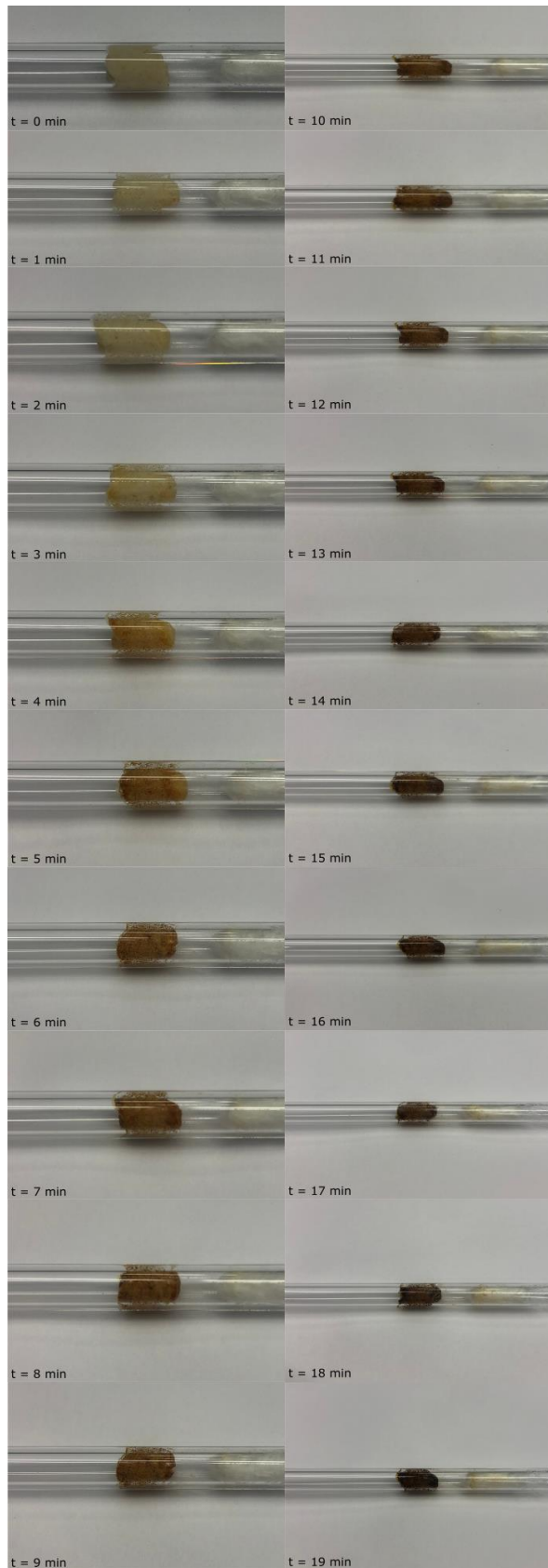
Supplementary Figure IV Photos of dough samples processed at 210 °C every minute of baking

Supplementary Figure V Comparison of FT-ICR-MS spectra of thermal processing emissions of ESI, APCI, APPI and SICRIT. Annotations can be found under https://doi.org/10.5281/zenodo.7426466

## B.4 DBDIpy: a Python library for processing of untargeted datasets from real-time plasma ionization mass spectrometry - Supplementary Material

# DBDIpy: a Python library for processing of untargeted datasets from real-time dielectric barrier discharge ionization mass spectrometry

Weidner Leopold, Hemmler Daniel, Rychlik Michael, Schmitt-Kopplin Philippe

# Supporting Information

Table of content:

## 1 Acquisition of DBDI-MS data:

Approximately 2 mg of wheat bread dough (formulation: wheat flour, water, yeast, rapeseed oil, roasted barley malt, wheat malt, sea salt) from a local bakery were fermented at 34 °C, 80% humidity for 180 min. The samples were baked at 210 °C for 20 min in a laboratory-scale, analytical setup. The sample was fitted inside a borosilicate glass tube and introduced into a controllable GC-SPME-module directly connected to a SICRIT SC-30 ion source (both Plasmion GmbH, Augsburg, Germany). A flow of 1.0 L min$^{-1}$ humidified Nitrogen and the negative pressure system of a 12 T solariX Fourier transformation ion cyclotron resonance mass spectrometer (FT-ICR-MS) (Bruker Daltonics, Bremen, Germany) actively transported thermal reaction products from the dough into the dielectric barrier discharge ion source. Ionization was performed at 1.5 kV and 15 kHz. For real-time data acquisition, the FT-ICR-MS was set to chromatographic mode. Data was acquisitioned between $m/z$ 92 and 1000 at 4 megawords and with an ion accumulation time of 0.1 sec. In total, ten scans were accumulated per datapoint. Nitrogen at 2 L min$^{-1}$ and 180 °C, was used as dry gas. A spray shield voltage of -400 V and a capillary voltage of 400 V were selected.

## 2 Preprocessing of demonstrational data:

Compass DataAnalysis 5.0 (Bruker Daltonics, Bremen, Germany) was used for internal calibration of raw data by reference mass lists. Singly charged ion signals with  S/N > 6 were manually exported to mass lists and further pre-processed by an in-house workflow in R programming

language (Version 4.2, R Foundation for Statistical Computing, Vienna, Austria) consisting of removal of side band artifacts, heavy isotope peaks and application of mass defect filtering. After alignment of mass lists to matrices holding mass spectrometric features, an in-house software tool was used to assign molecular formulae to the averaged $m/z$-values. Finally, 500 features were randomly selected from the matrix to form the demonstrational dataset for DBDIpy.

Demonstrational DBDIpy workflow:

In the following, we provide a comprehensive workflow to demonstrate the functions of DBDIpy (in markdown). It showcases all components from installation, the matchms interfaces and the core functionality.

## Tutorial

The following tutorial showcases an ordinary data analysis workflow by going through all functions of DBDIpy from loading data until visualization of correlation results. Therefore, we supplied a demo dataset which is publicly available [here](https://doi.org/10.5281/zenodo.7221089).

The demo data is from an experiments where wheat bread was roasted for 20 min and monitored by DBDI coupled to FT-ICR-MS. It consists of 500 randomly selected features.

![bitmap](https://user-images.githubusercontent.com/81673643/198022057-8b5da4b9-f6bd-43b7-9b6c-32fd119f93a7.png)
<p align = "center">
Fig.1 - Schematic DBDIpy workflow for in-source adduct and fragment detection: imported MS1 data are aligned, imputed and parsed to combined correlation and mass difference analysis.
</p>

### 1. Importing MS data
DBDIpy core functions utilize 2D tabular data. Raw mass spectra containing *m/z*-intensity-pairs first will need to be aligned to a DataFrame of features. We build features by using the ``align_spectra()`` function. ``align_spectra()`` is the interface to load data from open file formats such as .mgf, .mzML or .mzXML files via ``matchms.importing``.

If your data already is formatted accordingly, you can skip this step.

```python
##loading libraries for the tutorial
import os
import feather
import numpy as np
import pandas as pd
```

```python
import DBDIpy as dbdi
from matchms.importing import load_from_mgf
from matchms.exporting import save_as_mgf


##importing the downloaded .mgf files from demo data by matchms
demo_path = ""                              #enter path to demo dataset
demo_mgf = os.path.join(demo_path, "example_dataset.mgf")
spectrums = list(load_from_mgf(demo_mgf))


##align the listed Spectra
specs_aligned = dbdi.align_spectra(spec = spectrums, ppm_window = 2)
```

We first imported the demo MS1 data into a list of ``matchms.Spectra`` objects. At this place you can run your personal ``matchms`` preprocessing pipelines or manually apply filters like noise reduction.

By aplication of ``align_spectra()``, we transformed the list of spectra objects to a two-dimensional ``pandas.DataFrame``. Now you have a column for each mass spectrometric scan and features are aligned to rows. The first column shows the mean *m/z* of a feature.

If a signal was not detected in a scan, the according field will be set to an instance of ``np.nan``.

Remember to set the ``ppm_window`` parameter according to the resolution of you mass spectrometric system.

We now can inspect the aligned data, e.g. by running:

```python
specs_aligned.describe()
specs_aligned.info()
```

Several metabolomics data processing steps can be applied here if not already performed in ``matchms``. These might include application of noise-cutoffs, feature selection based on missing values, normalization or many others.

``specs_aligned.isnull().values.any()`` will give us an idea if there are missing values in the data. These cannot be handled by successive DBDIpy functions and most machine learning algorithms, so we need to impute them.

### 2. Imputation of missing values

``impute_intensities()`` will assure that after imputation we will have a set of uniform length extracted ion chromatograms (XIC) in our DataFrame. This is an important prerequisite for pointwise correlation calculation and for many tools handling time series data.

Missing values in our feature table will be imputed by a two-stage imputation algorithm.
- First, missing values within the detected signal region are interpolated in between.
- Second, a noisy baseline is generated for all XIC to be of uniform length which the length of the longest XIC in the dataset.

The function lets the user decide which imputation method to use. Default mode is ``linear``, however several others are available.

```python
feature_mz = specs_aligned["mean"]
specs_aligned = specs_aligned.drop("mean", axis = 1)

##impute the dataset
specs_imputed = dbdi.impute_intensities(df = specs_aligned, method = "linear")
```

Now ``specs_imputed`` does not contain any missing values anymore and is ready for adduct and in-source fragment detection.

```python
##check if NaN are present in DataFrame
specs_imputed.isnull().values.any()
Out[]: False
```

### 3. Detection of adducts and in-source fragments

Based on the ``specs_imputed``, we compute pointwise correlation of XIC traces to identify in-source adducts or in-source fragments generated during the DBD ionization process. The identification is performed in a two-step procedure:
- First, calculation of pointwise intensity correlation identifies feature groups with matching temporal intensity profiles through the experiment.
- Second, (exact) mass differences are used to refine the nature of potential candidates.

By default, ``identify_adducts()`` searches for $[M-H_2O+H]^+$, $[M+O_1+H]^+$ and $[M+O_2+H]^+$. For demonstrational purposes we also want to search for $[M+O_3+H]^+$ in this example.
Note that ``identify_adducts()`` has a variety of other parameters which allow high user customization. See the help file of the functions for details.

```python
##prepare a DataFrame to search for O3-adducts
adduct_rule = pd.DataFrame({'deltamz': [47.984744],'motive': ["O3"]})

##identify in-source fragments and adducts
search_res = dbdi.identify_adducts(df = specs_imputed, masses = feature_mz, custom_adducts = adduct_rule,
                   method = "pearson", threshold = 0.9, mass_error = 2)
```

The function will return a dictionary holding one DataFrame for each adduct type that was defined. A typical output looks like the following:

```python
##output search results
search_res
Out[24]:
{'O':  base_mz   base_index match_mz match_index   mzdiff     corr
```

```
19     215.11789        24 231.11280       ID40 15.99491 0.963228
310    224.10699        33 240.10191       ID51 15.99492 0.939139
605    231.11280        39 215.11789       ID25 15.99491 0.963228
1413   240.10191        50 224.10699       ID34 15.99492 0.939139
1668   244.13321        55 260.12812       ID67 15.99491 0.976541,

                   ...

'O2': base_mz   base_index match_mz match_index  mzdiff    corr
1437  240.10191        50 272.09174       ID77 31.98983 0.988866
1677  244.13321        55 276.12304       ID84 31.98983 0.972251
2362  260.12812        66 292.11795      ID100 31.98983 0.964096
3024  272.09174        76 240.10191       ID51 31.98983 0.988866
3354  276.12304        83 244.13321       ID56 31.98983 0.972251,

                   ...

'H2O': base_mz    base_index match_mz match_index   mzdiff    corr
621    231.11280        39 249.12337       ID60 18.01057 0.933640
1883   249.12337        59 231.11280       ID40 18.01057 0.933640
3263   275.13902        82 293.14958      ID102 18.01056 0.948774
4775   293.14958       101 275.13902       ID83 18.01056 0.948774
5573   300.08665       112 318.09722      ID140 18.01057 0.905907

                   ...

'O3': base_mz    base_index match_mz match_index   mzdiff    corr
320    224.10699        33 272.09174       ID77 47.98475 0.924362
1688   244.13321        55 292.11795      ID100 47.98474 0.964896
3013   272.09174        76 224.10699       ID34 47.98475 0.924362
4631   292.11795        99 244.13321       ID56 47.98474 0.964896
13597  438.28502       308 486.26976      ID356 47.98474 0.935359

                   ...
````
```

The ``base_mz`` and ``base_index`` column give us the index of the features which correlates with a correlation partner specified in ``match_mz`` and ``match_index``.
The mass difference between both is given for validation purpose and the correlation coefficient between both features is listed.


Now we can for example search series of Oxygen adducts of a single analyte:

```python
##search for oxygenation series
two_adducts                       =                       np.intersect1d(search_res["O"]["base_index"],
np.intersect1d(search_res["O"]["base_index"],search_res["O2"]["base_index"]))
three_adducts = np.intersect1d(two_adducts , search_res["O3"]["base_index"])

three_adducts
Out[33]: array([55, 99], dtype=int64)
```

This tells us that features 55 and 99 both putatively have $[M+O_{1-3}+H]^+$ adduct ions with correlations of $R^2$ > 0.9 in our dataset.

Let's visualize this finding!

### 4. Visualization of correlation results

Now that we putatively identified some related ions of a single analyte, we want to check their temporal response during the baking experiment.

Therefore, we can use the ``plot_adducts()`` function to conveniently draw XICs.

The demo dataset even comes along with some annotated metadata for our features, so we can decorate the plot and check our previous results!

```python
##load annotation metadta
demo_path = ""                                    #enter path to demo dataset
demo_meta = os.path.join(demo_path, "example_metadata.feather")
annotation_metadata = feather.read_dataframe(demo_meta)

##plot the XIC
dbdi.plot_adducts(IDs = [55,66,83,99], df = specs_imputed, metadata = annotation_metadata,
transform = True)
```

```
<p align="center">
  <img              width="600"              height="288"              src="https://user-
images.githubusercontent.com/81673643/198047792-9a9019ab-5c00-4365-a25c-
2cbcd0d3d20f.png">
</p>
<p align = "center">
```
Fig.2 - XIC plots for features 55, 66, 83 and 99 which have highly correlated intensity profile through the baking experiment.
```
</p>
```

We see that the XIC traces show a similar intensity profile through the experiment. The plot further tells us the correlation coefficients of the identified adducts.

From the metadata we can see that the detected mass signals were previously annotated as $C_{15}H_{17}O_{2-5}N$ which tells us that we most probably found an Oxgen-adduct series.

If MS2 data was recorded during the experiment we now can go on further and compare fragment spectra to reassure the identifications. You might find [ms2deepscore](https://github.com/matchms/ms2deepscore) to be a usefull library to do so in an automated way.

### 5. Exporting tabular MS data to match.Spectra objects

If you want to export your (imputed) tabular data to ``matchms.Spectra`` objects, you can do so by calling the ``export_to_spectra()`` function. We just need to re-add a column containing *m/z* values of the features.
This gives you access to the matchms suite and enables you to safe your mass spectrometric data to open file formats.
Hint: you can manually add some metadata after construction of the list of spectra.

```python
##export tabular MS data back to list of spectrums.
specs_imputed["mean"] = feature_mz

speclist = dbdi.export_to_spectra(df = specs_imputed, mzcol = 88)
```

```
##write processed data to .mgf file
save_as_mgf(speclist, "DBDIpy_processed_spectra.mgf")
```

We hope you liked this quick introduction into DBDIpy and will find its functions helpful and inspiring on your way to work through data from direct infusion mass spectrometry. Of course, the functions are applicable to all sort of ionisation mechanisms and you can modify the set of adducts to search in accordance to your source.

If you have open questions left about functions, their parameter or the algorithms we invite you to read through the built-in help files. If this does not clarify the issues, please do not hesitate to get in touch with us!

4 Documentation and help files for DBDIpy's functions:

Here we provide an extensive description of DBDIpy's functions including their parameters, default values and output objects.

align_spectra(spec, ppm_window = 2):
    """Feature building from a list of spectra.
    Aligns detected peaks in a list of matchms.Spectra objects into two-dimensional tabular data.
    Utilizes data loaded by matchms.importing module and connects to matchMS (pre-)processing
workflows.
    Parameters
    ----------
    spec : list of matchms.Spectra
        Spectra imported by matchms from mass spectrometric experiments
        (MS1) of instance matchms.Spectrum.

    ppm_window : float, optional
          Window for mass alignment in ppm. Default is 2 ppm.
    Returns
    -------
    A two-dimensional pd.DataFrames containing aligned mass spectrometric features.
    The first column contains mean m/z values of peaks across all scans followed by
    column-wise arranged signal intensities.
    Absence of a signal in a scan results in filling the table with nan instead.
    See Also
    --------
    matchms.importing : For information about reading mass spectrometric data into Python.
"""

```
export_to_spectra(df, mzcol = 0):
    """Builds a list of spectra from tabular data.
    Splits a two-dimensional pd.DataFrames containing aligned mass spectrometric features
    to a list of matchms.Spectrum objects for further processing or exporting to .mgf-files.
    Parameters
    ----------
    df : pd.DataFrame
        A DataFrame containing tabular mass spectrometric data.
        The first column (by default) contains mass to charge ratios,
        successive columns contain corresponting signal intensities
        of each mass spectrometric scan.
    mzcol : int, optional
        Position of the column containing m/z information of the features.
        Default is 0.
    Returns
    -------
    A a list containing a matchms.Spectrum object for each column of the input
    DataFrame except for the m/z column.
    See Also
    --------
    matchms.exporting : For information about writing mass spectrometric data
    to open file formats.
    """
```

identify_adducts(df, masses, custom_adducts = None, method = "pearson", threshold = 0.90, mass_error = 2):

"""Finds different ion species of a single analyte molecule.

Computes pointwise correlation of XIC traces to identify in-source adducts or in-source fragments

generated during direct infusion mass spectrometric data aquisition.

Putative identification of adducts and fragments is based on a correlation threshold and mass differences.

Parameters

----------

df : pd.DataFrame

    A DataFrame of equal-length ion traces formated as rows.

    Input DataFrame can be provided by align_spectra() and impute_intensities().

masses : pd.series

    A series of m/z values for specification of adduct or fragment types.

    Use of theoretic masses is strongly recommended for higher precission.

    If theoretic masses are not available, adapt mass_error

    in accordance to the type of your mass analyzer.

adduct_rules : pd.DataFrame, optional

    The function searches for 1-Oxygen and 2-Oxygen adducts

    and in-source water-losses by default.

    Custom adduct rules need to be specified in a DataFrame as following:

    custom_adducts = pd.DataFrame({'deltamz': [mz1, mz2, mz3],

             'motive': ["motive1", "motive2", "motive3"]}).

method : str, {'pearson', 'spearman', 'kendall'}

    Correlation method for pointwise compariation of XIC traces.

threshold : float

    Correlation treshold to associate two ions. Default is 0.9.

mass_error : int or float

    Tolerance of the mass spectrometer in ppm. Default is 2 ppm.

Returns

-------

A dictionary of DataFrames containing pairwise information about correlation of XIC traces and putatively identified in-source adducts or in-source fragments.

Raises

------

ValueError

    If the input DataFrame contains missing values.

See Also

--------

align_spectra(), impute_intensities() : For preparation of the input data.
"""

```
impute_intensities(df, method = "linear"):
    """Fills nan values in data tables.
    Imputes NaN values contained in a DataFrame consisting of aligned mass spectra.
    Input DataFrame can be provided by align_spectra().
    Extracted Ion Chromatograms often are not of the same length. To generate a set of
    uniform-length ion intensity series, a multi-step imputation approach is used:
        I)  Missing values within the detected signal region are interpolated.
        II) A noisy baseline is added for all XIC to be of uniform length which
            the length of the longest XIC in the dataset.
    The returned DataFrame is suitable for time series analysis or other multivariate statistics.


    Parameters
    ----------
    df : pd.DataFrame
        A DataFrame containing missing (NaN) values to be imputed.
        Input DataFrame can be provided by align_spectra().
    method : str
            Intepolation method to be used; default is "linear".
            Supported imputation methods are all  methods from
            pandas.DataFrame.interpolate() and
            scipy.interpolate.interp1d().
    Returns
    -------
    A DataFrame of equal length ion intensity series without NaN.
    See Also
    --------
    align_spectra() : For preparation of the input data.
    """
```

```python
plot_adducts(IDs, df, metadata = None, transform = False):
    """Visualizes identified adducts or in-source fragments.
    A graphical tool for the visualization of correlated XIC traces identified by
    identify_adducts().
    The temporal evolution of selected features is plotted to inspect correlation results
    and adduct information.
    Parameters
    ----------
    IDs : list
        A list of IDs (indices) to select correlated features from df.
    df: pd.DataFrame
        A two-dimensional DataFrame containing aligned mass spectrometric features
        e.g. generated by align_spectra().


    metadata : pd.DataFrame, optional
            A DataFrame containing annotated metadata for the intensity Data.Frame.
            Should contain a column called "mol_formula" for annotation of the plot.
    transform : bool, optional
             Whether plotted intensities should be scaled by a log2 function.
    Returns
    -------
    Shows a 2D-scatterplot of XIC and returns it as a matplotlib.figure object for
    individual modification.
    See Also
    --------
    align_spectra() : For preparation of the input data.
    identify_aducts() : For identification of multiple ion species from one compound.
    matplotlib.plt : For further customization of the returned plot object
    """
```

# C Appendix: Further publications with co-authorship contributions

Beyond the immediate scope of this thesis, I contributed to two further research projects which will be briefly named in the following chapter. I contributed as a co-author to one article which is currently under peer review and to one published full research article.

## C.1 Co-authorship I: Glycerinyl-modified arginine and lysine are putative biomarkers for PARK7-related early-onset Parkinson's disease.

Prudente De Mello N., Berger M. T., Lehmann K., Yan Y., <u>Weidner L.</u>, Tokarz J., Möller G., Cheng Y., Keipert S., Ciciliot S., Artati A., Wettmarshausen J., Nilsson R., Brandt D, Kutschke M., Leimpek A., Vogt Weisenhorn D., Wurst W., Adamski J., Jain M., Jastroch M., Mandemakers W., Bonifati V., Schmitt-Kopplin P., Perocchi F. & Dyar K. A., *Glycerinyl-modified arginine and lysine are putative biomarkers for PARK7-related early-onset Parkinson's disease.*
This manuscript was submitted to Molecular Neurodegeneration under manuscript number *MOND-D-24-00051* and is under peer review at the time of submitting this thesis.

*Candidate's contributions:* L.W. identified and characterized the novel metabolites, developed the extraction protocols for fibroblast and muscle tissue, prepared and measured fibroblast and muscle samples, performed data analysis, wrote and revised sections of the manuscript.

## C.2 Co-authorship II: C2-addition patterns emerging from acetylene and nickel sulfide in simulated prebiotic hydrothermal conditions.

This work has been published as Diederich P., Ruf A., Geisberger T., <u>Weidner, L.</u>, Seitz C., Eisenreich W., Huber C. & Schmitt-Kopplin, P. (2023). *C2-addition patterns emerging from acetylene and nickel sulfide in simulated prebiotic hydrothermal conditions* communications chemistry, 2023, Volume 6, Article 220.
doi.org/10.1038/s42004-023-01021-1

*Candidate's contributions:* L.W. implemented the self orientating maps algorithm, performed formal analysis, prepared the corresponding figure and revised the manuscript.

# Bibliography

(1)    Carmody, R. N., and Wrangham, R. W. (2009). The energetic significance of cooking. *Journal of Human Evolution 57*, 379–391.

(2)    James, S. R., Dennell, R. W., Gilbert, A. S., Lewis, H. T., Gowlett, J. A. J., Lynch, T. F., Mcgrew, W. C., Peters, C. R., Pope, G. G., and Stahl, A. B. (1989). Hominid Use of Fire in the Lower and Middle Pleistocene: A Review of the Evidence. *Current Anthropology 30*, 1–26.

(3)    Wrangham, R. In *Early Hominin Diets: The Known, The Unknown, and The Unknowable*, Ungar, P., Ed.; Oxford University Press: New York, NY, 2006, pp 308–323.

(4)    Parthasarathy, P., and Narayanan, S. K. (2014). Effect of Hydrothermal Carbonization Reaction Parameters on. *Environmental Progress & Sustainable Energy 33*, 676–680.

(5)    Wobber, V., Hare, B., and Wrangham, R. (2008). Great apes prefer cooked food. *Journal of Human Evolution 55*, 340–348.

(6)    Belitz, H.-D., Grosch, W., and Schieberle, P., *Lehrbuch der Lebensmittelchemie*; 6, 2008.

(7)    Megías-Rangil I, García-Lorda P, Torres-Moreno M, Bulló M, and Salas-Salvadó J (2004). Nutrient content and health effects of nuts. *Arch Latinoam Nutrition 54*.

(8)    Wu, G., Cross, H., Gehring, K., Savell, J., Arnold, A., and McNeill, S. (2016). Composition of free and peptide-bound amino acids in beef chuck, loin, and round cuts. *Journal of Animal Science 94*, 2603–2613.

(9)    Wandsnider, L. (1997). The Roasted and the Boiled: Food Composition and Heat Treatment with Special Emphasis on Pit-Hearth Cooking. *Journal of Anthropological Archaeology 16*, 1–48.

(10)    Jenkins, A. D., Stepto, R. F., Kratochvíl, P., and Suter, U. W. (1996). Glossary of basic terms in polymer science (IUPAC Recommendations 1996). *Pure and Applied Chemistry 68*, 2287–2311.

(11)    Hendry, G. A. F., and Wallace, R. K. (1993). The origin, distribution, and evolutionary significance of fructans. *Science and Technology of Fructans*, 119–139.

(12)    Domene-López, D., García-Quesada, J. C., Martin-Gullon, I., and Montalbán, M. G. (2019). Influence of starch composition and molecular weight on physicochemical properties of biodegradable films. *Polymers 11*.

(*13*) Stick, R. V., and Williams, S. J. In *Carbohydrates: The Essential Molecules of Life*, 2nd ed.; Elsevier: 2009, pp 321–341.

(*14*) El Khadem, H. S. In *Encyclopedia of Physical Science and Technology*, 3rd ed.; Academic Press: 2003, pp 369–416.

(*15*) Hamm, R. (1977). Postmortem breakdown of ATP and glycogen in ground muscle: A review. *Meat Science 1*, 15–39.

(*16*) Carmody, R. N. et al. (2019). Cooking shapes the structure and function of the gut microbiome. *Nature Microbiology 4*, 2052–2063.

(*17*) Hyötyläinen, T., Bondia-Pons, I., and Orešič, M. (2013). Lipidomics in nutrition and food research. *Molecular Nutrition and Food Research 57*, 1306–1318.

(*18*) Chen, H., Wei, F., Dong, X. y., Xiang, J. q., Quek, S. y., and Wang, X. (2017). Lipidomics in food science. *Current Opinion in Food Science 16*, 80–87.

(*19*) Lottenberg, A. M., Afonso, M. d. S., Lavrador, M. S. F., Machado, R. M., and Nakandakare, E. R. (2012). The role of dietary fatty acids in the pathology of metabolic syndrome. *The Journal of Nutritional Biochemistry 23*, 1027–1040.

(*20*) McWilliams, M., *Foods: experimental perspectives.* MacMillian Publishing Company: New York, NY, 1989.

(*21*) Rehner, G., and Daniel, H. (2010). Der Gastrointestinaltrakt — Vermittler zwischen Außen- und Innenwelt des Organismus. *Biochemie der Ernährung*, 307–361.

(*22*) Fichtner, M., Voigt, K., and Schuster, S. (2017). The tip and hidden part of the iceberg: Proteinogenic and non-proteinogenic aliphatic amino acids. *Biochimica et Biophysica Acta (BBA) - General Subjects 1861*, 3258–3269.

(*23*) Toldrá, F., Reig, M., Aristoy, M. C., and Mora, L. (2018). Generation of bioactive peptides during food processing. *Food Chemistry 267*, 395–404.

(*24*) Bailey, A. J. (1989). The Chemistry of Collagen Cross-Links and Their Role in Meat Texture. *Reciprocal Meat Conference Proceedings 42*, 127–135.

(*25*) Chen, L., Ma, L., Zhou, M., Liu, Y., and Zhang, Y. (2014). Effects of pressure on gelatinization of collagen and properties of extracted gelatins. *Food Hydrocolloids 36*, 316–322.

(*26*) Weston, A. R., Rogers, R. W., and Althen, T. G. (2002). Review: The Role of Collagen in Meat Tenderness. *The Professional Animal Scientist 18*, 107–111.

(*27*) Haus, J. M., Carrithers, J. A., Trappe, S. W., and Trappe, T. A. (2007). Collagen, crosslinking, and advanced glycation end products in aging human skeletal muscle. *Journal of Applied Physiology 103*, 2068–2076.

(*28*) Pieczonka, S. A., Lucio, M., Rychlik, M., and Schmitt-Kopplin, P. (2020). Decomposing the molecular complexity of brewing. *npj Science of Food 4*, 1–10.

(*29*) Roullier-Gall, C., Signoret, J., Hemmler, D., Witting, M. A., Kanawati, B., Schäfer, B., Gougeon, R. D., and Schmitt-Kopplin, P. (2018). Usage of FT-ICR-MS Metabolomics for Characterizing the Chemical Signatures of Barrel-Aged Whisky. *Frontiers in Chemistry 6*, 1–11.

(*30*) Roullier-Gall, C., Witting, M., Gougeon, R. D., and Schmitt-Kopplin, P. (2014). High precision mass measurements for wine metabolomics. *Frontiers in Chemistry 2*, 1–9.

(*31*) Álvarez-Rivera, G., Valdés, A., León, C., and Cifuentes, A. In *Food Chemistry, Function and Analysis No. 26*, Barros-Velazquez, J., Ed.; The Royal Society of Chemistry: 2021, pp 1–53.

(*32*) Shahidian, A., Ghassemi, M., Mohammadi, J., and Hashemi, M. In *Bio-Engineering Approaches to Cancer Diagnosis and Treatment*; Elsevier: 2020, pp 12–16.

(*33*) Rehner, G., and Daniel, H. (2010). Die Gewinnung biologischer Energie aus Nährstoffen. *Biochemie der Ernährung*, 275–294.

(*34*) Rutenberg, M. W., and Solarek, D. (1984). Starch derivatives: production and uses. *Starch: Chemistry and Technology*, 311–388.

(*35*) Kennedy, H., and Fischer, A. (1984). Starch and dextrins in prepared adhesives. *Starch: Chemistry and Technology*, 593–610.

(*36*) Guarás, M. P., Ludueña, L. N., and Alvarez, V. A. (2017). Development of Biodegradable Products from Modified Starches. *Starch-Based Materials in Food Packaging: Processing, Characterization and Applications*, 77–124.

(*37*) Jost, T., Henning, C., Heymann, T., and Glomb, M. A. (2021). Comprehensive Analyses of Carbohydrates, 1,2-Dicarbonyl Compounds and Advanced Glycation Endproducts in Industrial Bread Making. *Journal of Agricultural and Food Chemistry 69*, 3720–3731.

(*38*) Rehner, G., and Daniel, H. (2010). Wahrnehmung des Geschmacks und des Geruchs der Nahrung. *Biochemie der Ernährung*, 199–216.

(*39*) Bautista, D. A. In *Encyclopedia of Food Microbiology: Second Edition*; Academic Press: 2014, pp 465–470.

(*40*) Tapal, A., and Tiku, P. K. (2019). Nutritional and Nutraceutical Improvement by Enzymatic Modification of Food Proteins. *Enzymes in Food Biotechnology: Production, Applications, and Future Prospects*, 471–481.

(*41*) Toldrá, F., Flores, M., and Aristoy, M.-C. In *Encyclopedia of Meat Sciences*; Elsevier: 2014; Vol. 2, pp 206–211.

(*42*) Mazorra-Manzano, M. A., Ramírez-Suárez, J. C., Moreno-Hernández, J. M., and Pacheco-Aguilar, R. (2018). Seafood proteins. *Proteins in Food Processing, Second Edition*, 445–475.

*(43)*   Sulser, H., DePizzol, J., and Buchi, W. (1967). A Probable Flavoring Principle in Vegetable-Protein Hydrolysates. *Journal of food science 32*.

*(44)*   Clemente, A. (2000). Enzymatic protein hydrolysates in human nutrition. *Trends in Food Science & Technology 11*, 254–262.

*(45)*   Yang, J., Bai, W., Zeng, X., and Cui, C. (2019). Gamma glutamyl peptides: The food source, enzymatic synthesis, kokumi-active and the potential functional properties – A review. *Trends in Food Science & Technology 91*, 339–346.

*(46)*   Fritz, H., Trautschold, I., and Werle, E. (1974). Protease Inhibitors. *Methods of Enzymatic Analysis*, 1064–1080.

*(47)*   Danehy, J. P. (1986). Maillard reactions: nonenzymatic brownings in food systems with special reference to the development of flavor. *Advances in food research 30*.

*(48)*   Pozo-Bayón, M. A., Guichard, E., and Cayot, N., *Flavor control in baked cereal products*; 4, 2006; Vol. 22, pp 335–379.

*(49)*   Demyttenaere, J., Tehrani, K. A., and De Kimpe, N. (2002). The chemistry of the most important maillard flavor compounds of bread and cooked rice. *ACS Symposium Series 826*, 150–165.

*(50)*   Starowicz, M., and Zieliński, H. (2019). How Maillard Reaction Influences Sensorial Properties (Color, Flavor and Texture) of Food Products? *Food reviews international 35*, 707–725.

*(51)*   Rizzi, G. P. (1997). Chemical structure of colored maillard reaction products. *Food Reviews International 13*, 1–28.

*(52)*   Lin, L., Sun, Y., and Zhang, P. (2021). Color development kinetics of maillard reactions. *Industrial and Engineering Chemistry Research 60*, 3495–3501.

*(53)*   Reyes, F. G., Poocharoen, B., and Wrolstad, R. E. (1982). Maillard Browning Reaction of Sugar-Glycine Model Systems: Changes in Sugar Concentration, Color and Appearance. *Journal of Food Science 47*, 1376–1377.

*(54)*   Labuza, T. P., Monnier, V., Baynes J., and O'Brien, J., *Maillard reactions in chemistry, food and health*, 1st; Elsevier: 1998.

*(55)*   Poulsen, M. W., Hedegaard, R. V., Andersen, J. M., de Courten, B., Bügel, S., Nielsen, J., Skibsted, L. H., and Dragsted, L. O. (2013). Advanced glycation endproducts in food and their effects on health. *Food and Chemical Toxicology 60*, 10–37.

*(56)*   Erbersdobler, H. F., and Somoza, V. (2007). Forty years of furosine - Forty years of using Maillard reaction products as indicators of the nutritional quality of foods. *Molecular Nutrition and Food Research 51*, 423–430.

(57) Maillard, L. (1912). Action des acides aminés sur les sucres; formation des mélanoïdines par voie methodique. *Comptes Rendus hebdomadaires de l'Académie des Sciences 154*, 66–68.

(58) Fayle, S. E., and Gerrard, J. A., *The Maillard reaction*; Royal Society of Chemistry: 2002, p 120.

(59) Ames, J. M. (1992). The Maillard Reaction. *Biochemistry of Food Proteins*, 99–153.

(60) Ledl, F., Schleicher, E., and Nursten, H. E. (1990). New Aspects of the Maillard Reaction in Foods and in the Human Body. *Angewandte Chemie International Edition in English 29*, 565–594.

(61) Henle, T. In *Process-Induced Food Toxicants: Occurrence, Formation, Mitigation, and Health Risks*, Stadler, R. H., and Linebaker, D. R., Eds.; 2; John Wiley & Sons, Inc: 2009, pp 215–242.

(62) Thornalley, P. J. (2005). Dicarbonyl intermediates in the Maillard reaction. *Annals of the New York Academy of Sciences 1043*, 111–117.

(63) Pischetsrieder, M., and Severin, T. In *Maillard Reactions in Chemistry, Food and Health*; Woodhead Publishing Ltd: 2005, pp 37–42.

(64) Prestes Fallavena, L., Poerner Rodrigues, N., Damasceno Ferreira Marczak, L., and Domeneghini Mercali, G. (2022). Formation of advanced glycation end products by novel food processing technologies: A review. *Food chemistry 393*.

(65) Hartmann, S., and Schieberle, P. (2016). On the role of Amadori rearrangement products as precursors of aroma-active strecker aldehydes in cocoa. *ACS Symposium Series 1237*, 1–13.

(66) Hellwig, M., and Henle, T. (2020). Maillard Reaction Products in Different Types of Brewing Malt. *Journal of Agricultural and Food Chemistry 68*, 14274–14285.

(67) Glomb, M. A., Gobert, J., and Voigt, M. (2010). Dicarbonyls from maillard degradation of glucose and maltose. *ACS Symposium Series 1042*, 35–44.

(68) Hodge, J. E. (1953). Dehydrated foods, Chemistry of Browning Reactions in Model Systems. *Journal of Agricultural and Food Chemistry 1*, 928–943.

(69) Wrodnigg, T., and Eder, B., *Glycoscience: Epimerisation, Isomerisation and Rearrangement Reactions of Carbohydrates*; Stütz, A., Ed.; Springer: 2001, pp 115–152.

(70) Yaylayan A, V. (1997). Classification of the Maillard reaction: A conceptual approach. *Trends in Food Science & Technology 8*, 13–18.

(71) Hemmler, D., Roullier-Gall, C., Marshall, J. W., Rychlik, M., Taylor, A. J., and Schmitt-Kopplin, P. (2018). Insights into the Chemistry of Non-Enzymatic Browning Reactions in Different Ribose-Amino Acid Model Systems. *Scientific Reports 8*, 1–10.

(72)   Yan, Y., Hemmler, D., and Schmitt-Kopplin, P. (2022). HILIC-MS for Untargeted Pro-
       filing of the Free Glycation Product Diversity. *Metabolites 12.*

(73)   Ajandouz, E. H., and Puigserver, A. (1999). Nonenzymatic Browning Reaction of Es-
       sential Amino Acids: Effect of pH on Caramelization and Maillard Reaction Kinetics.
       *Journal of Agricultural and Food Chemistry 47,* 1786–1793.

(74)   Rizzi, G. P. (2005). The Maillard Reaction in Foods. *Maillard Reactions in Chemistry,
       Food and Health,* 11–19.

(75)   Yu, X., Zhao, M., Liu, F., Zeng, S., and Hu, J. (2013). Antioxidants in volatile Maillard
       reaction products: Identification and interaction. *LWT - Food Science and Technology
       53,* 22–28.

(76)   Nursten, H. E. (1981). Recent developments in studies of the maillard reaction. *Food
       Chemistry 6,* 263–277.

(77)   Farmer, L. J., and Mottram, D. S. (1990). Interaction of lipid in the maillard reaction
       between cysteine and ribose: The effect of a triglyceride and three phospholipids on
       the volatile products. *Journal of the Science of Food and Agriculture 53,* 505–525.

(78)   Han, L., Lin, Q., Liu, G., Han, D., and Niu, L. (2022). Review of the formation and
       influencing factors of food-derived glycated lipids. *Critical reviews in food science and
       nutrition 62,* 3490–3498.

(79)   Alaiz, M., Hidalgo, F. J., and Zamora, R. (1999). Effect of pH and Temperature on
       Comparative Antioxidant Activity of Nonenzymatically Browned Proteins Produced
       by Reaction with Oxidized Lipids and Carbohydrates. *Journal of Agricultural and Food
       Chemistry 47,* 748–752.

(80)   De Lange, P., and Dekker, L. P. (1954). A Browning Reaction between Thiamine and
       Glucose. *Nature 1954 173:4413 173,* 1040–1041.

(81)   Cerny, C., and Briffod, M. (2007). Effect of pH on the Maillard reaction of 13C5-xylose,
       cystein, and thiamin. *Journal of Agricultural and Food Chemistry 55,* 1552–1556.

(82)   Kitayama, S., Igoshi, A., Shimamura, Y., Noda, K., and Murata, M. (2022). Forma-
       tion scheme and some properties of a thiamine-derived pigment, pyrizepine, formed
       through the Maillard reaction. *Bioscience, Biotechnology, and Biochemistry 86,* 672–
       680.

(83)   Kurata, T., Fujimaki, M., and Sakurai, Y. (1973). Red Pigment Produced by the Reaction
       of Dehydro-l-ascorbic Acid with $\alpha$-Amino Acid. *Agricultural Biological Chemistry 37,*
       1471–1477.

(84)   Rogacheva, S. M., Kuntcheva, M. J., Panchev, I. N., and Obretenov, T. D. (1995). L-
       Ascorbic acid in nonenzymatic reactions - I. Reaction with glycine. *Zeitschrift für
       Lebensmittel-Untersuchung und -Forschung 200,* 52–58.

(*85*) Glomb, M. A., and Henning, C. (2016). Formation of reactive fragmentation products during the Maillard degradation of reducing sugars- A review. *ACS Symposium Series 1237*, 117–131.

(*86*) Yu, A. N., Tan, Z. W., and Wang, F. S. (2013). Mechanistic studies on the formation of pyrazines by Maillard reaction between l-ascorbic acid and l-glutamic acid. *LWT - Food Science and Technology 50*, 64–71.

(*87*) Ames, J. M. (1990). Control of the Maillard reaction in food systems. *Trends in Food Science & Technology 1*, 150–154.

(*88*) Wong, C. W., Wijayanti, H. B., and Bhandari, B. R. (2015). Maillard reaction in limited moisture and low water activity environment. *Food Engineering Series*, 41–63.

(*89*) Warmbier, H. C., Schnickels, R. A., and Labuza, T. P. (1976). Effects of glycerol on non-enzymatic browning in a solid intermediate moisture model food system. *Journal of Food Science 41*, 528–531.

(*90*) Villamiel, M., del Castillo, M. D., and Corzo, N. In *Food Biochemistry and Food Processing*, Hui, H., Ed.; Blackwell Publishing: 2006; Chapter 4.

(*91*) Greenshields, R. (1973). Caramel 2. manufacture, composition and properties. *Process biochemistry*.

(*92*) Myers, D. V., and Howell, J. C. (1992). Characterization and specifications of caramel colours: An overview. *Food and Chemical Toxicology 30*, 359–363.

(*93*) Hardt, R., and Baltes, W. (1987). Zur Analytik von Zuckercouleuren 1. Mitteilung: Differenzierung der Couleurklassen mittels Curiepunkt-Pyrolyse-Capillargaschromatographie-Massenspektrometrie. *Zeitschrift für Lebensmittel Untersuchung und Forschung 185*, 275–280.

(*94*) Quintas, M. A., Fundo, J. F., and Silva, C. L. (2010). Sucrose in the Concentrated Solution or the Supercooled "State": A Review of Caramelisation Reactions and Physical Behaviour. *Food Engineering Reviews 2*, 204–215.

(*95*) Kroh, L. W. (1994). Caramelisation in food and beverages. *Food Chemistry 51*, 373–379.

(*96*) deBruijn, J., Kieboom, A., vanBekkum H., and vander Poel, P. (1986). Reactions of monosaccharides in aqueous alkaline solutions. *Sugar Technology Reviews*.

(*97*) Golon, A., and Kuhnert, N. (2012). Unraveling the chemical composition of caramel. *Journal of Agricultural and Food Chemistry 60*, 3266–3274.

(*98*) Tosi, E., Ciappini, M., Ré, E., and Lucero, H. (2002). Honey thermal treatment effects on hydroxymethylfurfural content. *Food Chemistry 77*, 71–74.

(99)    Pereira, L. L., Debona, D. G., Pinheiro, P. F., de Oliveira, G. F., ten Caten, C. S., Mok-sunova, V., Kopanina, A. V., Vlasova, I. I., Talskikh, A. I., and Yamamoto, H. In *Quality Determinants In Coffee Production*, Pereira, L. L., and Moreira, T. R., Eds.; Springer Cham: 2020, pp 303–372.

(100)   Cho, I. H., and Peterson, D. G. (2010). Chemistry of bread aroma: A review. *Food Science and Biotechnology 19*, 575–582.

(101)   Patton, S. (1955). Browning and Associated Changes in Milk and Its Products: A Review. *Journal of Dairy Science 38*, 457–478.

(102)   Dransfield, E. (2008). The taste of fat. *Meat Science 80*, 37–42.

(103)   Besnard, P., Passilly-Degrace, P., and Khan, N. A. (2015). Taste of fat: A sixth taste modality? *Physiological Reviews 96*, 151–176.

(104)   Uchida, K. (2007). Future of toxicology–lipid peroxidation in the future: from biomarker to etiology. *Chemical research in toxicology 20*, 3–5.

(105)   Guéraud, F., Atalay, M., Bresgen, N., Cipak, A., Eckl, P. M., Huc, L., Jouanin, I, Siems, W., and Uchida, K. (2010). Chemistry and biochemistry of lipid peroxidation products. *Free Radical Research 44*, 1098–1124.

(106)   Niki, E., Yoshida, Y., Saito, Y., and Noguchi, N. (2005). Lipid peroxidation: Mechanisms, inhibition, and biological effects. *Biochemical and Biophysical Research Communications 338*, 668–676.

(107)   Bacellar, I. O., Itri, R., Rodrigues, D. R., and Baptista, M. S. (2022). Photosensitized Lipid Oxidation: Mechanisms and Consequences to Health Sciences. *Lipid Oxidation in Food and Biological Systems: A Physical Chemistry Perspective*, 305–337.

(108)   Kanner, J. (2007). Dietary advanced lipid oxidation endproducts are risk factors to human health. *Mol. Nutr. Food Res. 51*, 1094–1101.

(109)   Frankel, E. N. (2005). Lipid oxidation. *Lipid Oxidation*, 1–470.

(110)   Mottram, D. S., and Whitfield, F. B. (1995). Maillard-Lipid Interactions in Nonaqueous Systems: Volatiles from the Reaction of Cysteine and Ribose with Phosphatidylcholine. *Journal of Agriculture and Food Chemistry 43*, 1302–1306.

(111)   Sohail, A., Al-Dalali, S., Wang, J., Xie, J., Shakoor, A., Asimi, S., Shah, H., and Patil, P. (2022). Aroma compounds identified in cooked meat: A review. *Food research international 157*.

(112)   Ait Ameur, L., Rega, B., Giampaoli, P., Trystram, G., and Birlouez-Aragon, I. (2008). The fate of furfurals and other volatile markers during the baking process of a model cookie. *Food Chemistry 111*, 758–763.

(*113*)  Mesías-García, M., Guerra-Hernández, E., and García-Villanova, B. (2010). Determination of furan precursors and some thermal damage markers in baby foods: Ascorbic acid, dehydroascorbic acid, hydroxymethylfurfural and furfural. *Journal of Agricultural and Food Chemistry 58*, 6027–6032.

(*114*)  Wu, L., Du, B., Vander Heyden, Y., Chen, L., Zhao, L., Wang, M., and Xue, X. (2017). Recent advancements in detecting sugar-based adulterants in honey – A challenge. *TrAC - Trends in Analytical Chemistry 86*, 25–38.

(*115*)  Perez-Locas, C., and Yaylayan, V. A. (2010). The Maillard reaction and food quality deterioration. *Chemical Deterioration and Physical Instability of Food and Beverages*, 70–94.

(*116*)  Hellwig, M., Kühn, L., and Henle, T. (2018). Individual Maillard reaction products as indicators of heat treatment of pasta — A survey of commercial products. *Journal of Food Composition and Analysis 72*, 83–92.

(*117*)  Ramasamy, R., Vannucci, S. J., Yan, S. S. D., Herold, K., Yan, S. F., and Schmidt, A. M. (2005). Advanced glycation end products and RAGE: A common thread in aging, diabetes, neurodegeneration, and inflammation. *Glycobiology 15*, 16–28.

(*118*)  Guilbaud, A., Niquet-Leridon, C., Boulanger, E., and Tessier, F. J. (2016). How Can Diet Affect the Accumulation of Advanced Glycation End-Products in the Human Body? *Foods 5*, 1–14.

(*119*)  Ravichandran, G., Lakshmanan, D. K., Raju, K., Elangovan, A., Nambirajan, G., Devanesan, A. A., and Thilagar, S. (2019). Food advanced glycation end products as potential endocrine disruptors: An emerging threat to contemporary and future generation. *Environment International 123*, 486–500.

(*120*)  Sugimura, T., Nagao, N., Kawachi, T., Honda, M., Yahagi, T., Seino, Y., Stao, S., Matsukura, N., Matsushima, T., Shirai, A., Sawamura, M., and Matsumoto, H. In *Origins of Human Cancer*, Hiatt, H., Watson, J., and Winstein, J., Eds.; Cold Spring Harbor Laboratory: Cold Spring Harbor, NY, 1977, pp 1561–1577.

(*121*)  Turesky, R. J. (2007). Formation and biochemistry of carcinogenic heterocyclic aromatic amines in cooked meats. *Toxicology Letters 168*, 219–227.

(*122*)  Turesky, R. J., and Le Marchand, L. (2011). Metabolism and biomarkers of heterocyclic aromatic amines in molecular epidemiology studies: lessons learned from aromatic amines. *Chemical research in toxicology 24*, 1169–1214.

(*123*)  Alaejos, M. S., and Afonso, A. M. (2011). Factors That Affect the Content of Heterocyclic Aromatic Amines in Foods. *Comprehensive Reviews in Food Science and Food Safety 10*, 52–108.

(*124*)  Barzegar, F., Kamankesh, M., and Mohammadi, A. (2019). Heterocyclic aromatic amines in cooked food: A review on formation, health risk-toxicology and their analytical techniques. *Food Chemistry 280*, 240–254.

(*125*)  Morales, F. J., Somoza, V., and Fogliano, V. (2012). Physiological relevance of dietary melanoidins. *Amino Acids 42*, 1097–1109.

(*126*)  Bruhns, P., Kanzler, C., Degenhardt, A. G., Koch, T. J., and Kroh, L. W. (2019). Basic Structure of Melanoidins Formed in the Maillard Reaction of 3-Deoxyglucosone and γ-Aminobutyric Acid. *Journal of Agricultural and Food Chemistry 67*, 5197–5203.

(*127*)  Wang, H. Y., Qian, H., and Yao, W. R. (2011). Melanoidins produced by the Maillard reaction: Structure and biological activity. *Food Chemistry 128*, 573–584.

(*128*)  Langner, E., and Rzeski, W. (2014). Biological properties of melanoidins: A review. *International Journal of Food Properties 17*, 344–353.

(*129*)  Lee, J., Bousquières, J., Descharles, N., Roux, S., Michon, C., Rega, B., and Bonazzi, C. (2020). Potential of model cakes to study reaction kinetics through the dynamic on-line extraction of volatile markers and TD-GC-MS analysis. *Food Research International 132*, 109087.

(*130*)  Rega, B., Guerard, A., Delarue, J., Maire, M., and Giampaoli, P. (2009). On-line dynamic HS-SPME for monitoring endogenous aroma compounds released during the baking of a model cake. *Food Chemistry 112*, 9–17.

(*131*)  Zviely, M. (2006). Heterocyclic Nitrogen-and Sulfur-Containing Aroma Chemicals. *Perfumer & Flavorist 31*, 20–35.

(*132*)  Shahidi, F., and Hossain, A. (2022). Role of Lipids in Food Flavor Generation. *Molecules 27*.

(*133*)  Cepeda-Vázquez, M., Rega, B., Descharles, N., and Camel, V. (2018). How ingredients influence furan and aroma generation in sponge cake. *Food Chemistry 245*, 1025–1033.

(*134*)  Mondal, A., and Datta, A. K. (2008). Bread baking - A review. *Journal of Food Engineering 86*, 465–474.

(*135*)  Wrangham, R., and Conklin-Brittain, N. L. (2003). Cooking as a biological trait. *Comparative Biochemistry and Physiology 136*, 35–46.

(*136*)  Yu, X., and Zuo, T. (2021). Editorial: Food Additives, Cooking and Processing: Impact on the Microbiome. *Frontiers in Nutrition 8*, 1–2.

(*137*)  Cowan, M. M. (1999). Plant products as antimicrobial agents. *Clinical microbiology reviews 12*, 564–582.

(138)   Shah, S., Brown, P. D., Mayengbam, S., Gänzle, M. G., Wang, W., Mu, C., Lettrari, S., Bertagnolli, C., and Shearer, J. (2020). Metabolic and Gut Microbiota Responses to Sourdough Pasta Consumption in Overweight and Obese Adults. *Frontiers in Nutrition 7*, 615003.

(139)   Bautista, R., Fernández, E., and Falqué, E. (2007). Effect of the contact with fermentation-lees or commercial-lees on the volatile composition of white wines. *European Food Research and Technology 224*, 405–413.

(140)   Chu, F. S. (1991). Mycotoxins: food contamination, mechanism, carcinogenic potential and preventive measures. *Mutation Research/Genetic Toxicology 259*, 291–306.

(141)   Adeyeye, S. A. (2016). Fungal mycotoxins in foods: Fungal mycotoxins in foods: A review. *Cogent Food and Agriculture 2*.

(142)   Tola, M., and Kebede, B. (2016). Occurrence, importance and control of mycotoxins: A review. *Cogent Food and Agriculture 2*.

(143)   Hennekinne, J. A., De Buyser, M. L., and Dragacci, S. (2012). Staphylococcus aureus and its food poisoning toxins: characterization and outbreak investigation. *FEMS Microbiology Reviews 36*, 815–836.

(144)   Granum, P. E., and Lund, T. (1997). Bacillus cereus and its food poisoning toxins. *FEMS Microbiology Letters 157*, 223–228.

(145)   Songer, J. G. (2004). The emergence of Clostridium difficile as a pathogen of food animals. *Animal Health Research Reviews 5*, 321–326.

(146)   Ting, P. T., and Freiman, A. (2004). The story of Clostridium botulinum: from food poisoning to Botox. *Clinical Medicine 4*, 258–261.

(147)   Amit, S. K., Uddin, M. M., Rahman, R., Islam, S. M., and Khan, M. S. (2017). A review on mechanisms and commercial aspects of food preservation and processing. *Agriculture and Food Security 6*.

(148)   Farkas, J., and Mohácsi-Farkas, C. (2011). History and future of food irradiation. *Trends in Food Science & Technology 22*, 121–126.

(149)   Varghese, K. S., Pandey, M. C., Radhakrishna, K., and Bawa, A. S. (2014). Technology, applications and modelling of ohmic heating: a review. *Journal of Food Science and Technology 51*, 2304–2317.

(150)   Tao, Y., Sun, D. W., Hogan, E., and Kelly, A. L. (2014). High-Pressure Processing of Foods: An Overview. *Emerging Technologies for Food Processing*, 3–24.

(151)   Barbosa-Cánovas, G. V., and Altunakar, B. (2006). Pulsed electric fields processing of foods: An overview. *Food Engineering Series*, 3–26.

(*152*)   Guzik, P., Kulawik, P., Zajac, M., and Migdał, W. (2022). Microwave applications in the food industry: an overview of recent developments. *Critical Reviews in Food Science and Nutrition 62*, 7989–8008.

(*153*)   Khaire, R. A., Thorat, B. N., and Gogate, P. R. (2022). Applications of ultrasound for food preservation and disinfection: A critical review. *Journal of Food Processing and Preservation 46*.

(*154*)   Leistner, L. (1994). Further developments in the utilization of hurdle technology for food preservation. *Journal of Food Engineering 22*, 421–432.

(*155*)   Leistner, L. (1992). Food preservation by combined methods. *Food Research International 25*, 151–158.

(*156*)   Leistner, L. (2000). Basic aspects of food preservation by hurdle technology. *International Journal of Food Microbiology 55*, 181–186.

(*157*)   Murphy, R. Y., Johnson, E. R., Duncan, L. K., Clausen, E. C., Davis, M. D., and March, J. A. (2001). Heat Transfer Properties, Moisture Loss, Product Yield, and Soluble Proteins in Chicken Breast Patties During Air Convection Cooking. *Poultry Science 80*, 508–514.

(*158*)   Sablani, S. S., Marcotte, M., Baik, O. D., and Castaigne, F. (1998). Modeling of simultaneous heat and water transport in the baking process. *LWT - Food Science and Technology 31*, 201–209.

(*159*)   Chen, H., Marks, B. P., and Murphy, R. Y. (1999). Modeling coupled heat and mass transfer for convection cooking of chicken patties. *Journal of Food Engineering 42*, 139–146.

(*160*)   Minkowycz, W. J., Sparrow, E. M., Schneider, G. E., and Pletcher, R. H., *Handbook of numerical heat transfer*; Wiley-Interscience: New York, NY, 2000, pp 53–123.

(*161*)   Broyart, B., and Trystram, G. (2002). Modelling heat and mass transfer during the continuous baking of biscuits. *Journal of Food Engineering 51*, 47–57.

(*162*)   Rabeler, F., and Feyissa, A. H. (2018). Kinetic modeling of texture and color changes during thermal treatment of chicken breast meat. *Food and Bioprocess Technology 11*, 1495–1504.

(*163*)   Rabeler, F., and Feyissa, A. H. (2018). Modelling the transport phenomena and texture changes of chicken breast meat during the roasting in a convective oven. *Journal of Food Engineering 237*, 60–68.

(*164*)   Portanguen, S., Ikonic, P., Clerjon, S., and Kondjoyan, A. (2014). Mechanisms of Crust Development at the Surface of Beef Meat Subjected to Hot Air: An Experimental Study. *Food and Bioprocess Technology 7*, 3308–3318.

(*165*)    Zanoni, B., Peri, C., and Pierucci, S. (1993). A study of the bread-baking process. I: A phenomenological model. *Journal of Food Engineering 19*, 389–398.

(*166*)    Dagerskog, M. (1979). Pan frying of meat patties. : 1. A study of heat and mass transfer. *Lebensmittel-Wissenschaft und Technologie 12*, 217–224.

(*167*)    Manley, D. In *Biscuit, Cookie and Cracker Manufacturing Manuals*, Green, G., Ed.; Woodhead PublishingLimited: Cambridge, England, 1998, pp 15–20.

(*168*)    Dominguez-Hernandez, E., Salaseviciene, A., and Ertbjerg, P. (2018). Low-temperature long-time cooking of meat: Eating quality and underlying mechanisms. *Meat Science 143*, 104–113.

(*169*)    Marsh, B. B. (1977). The nature of tenderness. *Proceedings of the 30th annual reception of meat conference 30*, 30–69.

(*170*)    Pico, J., Khomenko, I., Capozzi, V., Navarini, L., and Biasioli, F. (2020). Real-time monitoring of volatile compounds losses in the oven during baking and toasting of gluten-free bread doughs: A PTR-MS evidence. *Foods 9*.

(*171*)    Pico, J., Antolín, B., Román, L., Gómez, M., and Bernal, J. (2018). Analysis of volatile compounds in gluten-free bread crusts with an optimised and validated SPME-GC/QTOF methodology. *Food Research International 106*, 686–695.

(*172*)    Pravisani, C. I., Califano, A. N., and Calvelo, A. (1985). Kinetics of Starch Gelatinization in Potato. *Journal of Food Science 50*, 657–660.

(*173*)    Pérez, S., and Bertoft, E. (2010). The molecular structures of starch components and their contribution to the architecture of starch granules: A comprehensive review. *Starch - Stärke 62*, 389–420.

(*174*)    Kutz, J. R., Rientsma ; Loney, L. M., Meredith ; R, P. G., Axford, D. W. E., Elton, G. A. H., Hollinger, G., Kuniak, L., Marchessnult, R. H., and Billmeyer, F. W. (1982). Tentative Hypothesis to Explain How Electrolytes Affect the Gelatinization Temperature of Starches in Water. *Starch - Stärke 34*, 233–239.

(*175*)    Donovan, J. W. (1979). Phase transitions of the starch–water system. *Biopolymers 18*, 263–275.

(*176*)    Pence, E. A. A Study of baking oven vapors, 1952.

(*177*)    Wang, H., Xiang, Z., Wang, L., Jing, S., Lou, S., Tao, S., Liu, J., Yu, M., Li, L., Lin, L., Chen, Y., Wiedensohler, A., and Chen, C. (2018). Emissions of volatile organic compounds (VOCs) from cooking and their speciation: A case study for Shanghai with implications for China. *Science of the Total Environment 621*, 1300–1309.

(*178*)    Nolte, C. G., Schauer, J. J., Cass, G. R., and Simoneit, B. R. (1999). Highly polar organic compounds present in meat smoke. *Environmental Science and Technology 33*, 3313–3316.

(*179*)  Abdullahi, K. L., Delgado-Saborit, J. M., and Harrison, R. M. (2013). Emissions and indoor concentrations of particulate matter and its specific chemical components from cooking: A review. *Atmospheric Environment 71*, 260–294.

(*180*)  Xiang, W., Wang, W., Du, L., Zhao, B., Liu, X., Zhang, X., Yao, L., and Ge, M. (2023). Toxicological Effects of Secondary Air Pollutants. *Chemical Research in Chinese Universities 39*, 326–341.

(*181*)  Klein, F. et al. (2016). Characterization of Gas-Phase Organics Using Proton Transfer Reaction Time-of-Flight Mass Spectrometry: Cooking Emissions. *Environmental Science and Technology 50*, 1243–1250.

(*182*)  Zeng, J., Yu, Z., Mekic, M., Liu, J., Li, S., Loisel, G., Gao, W., Gandolfo, A., Zhou, Z., Wang, X., Herrmann, H., Gligorovski, S., and Li, X. (2020). Evolution of Indoor Cooking Emissions Captured by Using Secondary Electrospray Ionization High-Resolution Mass Spectrometry. *Environmental Science & Technology Letters 7*, 76–81.

(*183*)  Wang, Q. et al. (2020). Hourly Measurements of Organic Molecular Markers in Urban Shanghai, China: Primary Organic Aerosol Source Identification and Observation of Cooking Aerosol Aging. *ACS Earth and Space Chemistry 4*, 1670–1685.

(*184*)  Zhu, L., and Wang, J. (2003). Sources and patterns of polycyclic aromatic hydrocarbons pollution in kitchen air, China. *Chemosphere 50*, 611–618.

(*185*)  Rogge, W. F., Hildemann, L. M., Mazurek, M. A., Cass, G. R., and Slmoneit, B. R. T. (1991). Sources of Fine Organic Aerosol. 1. Charbroilers and Meat Cooking Operations. *Environ. Sci. Technol 25*, 1112–1125.

(*186*)  Atkinson, R. (2000). Atmospheric chemistry of VOCs and NOx. *Atmospheric Environment 34*, 2063–2101.

(*187*)  Ko, Y. C., Cheng, L. S. C., Lee, C. H., Huang, J. J., Huang, M. S., Kao, E. L., Wang, H. Z., and Lin, H. J. (2000). Chinese Food Cooking and Lung Cancer in Women Nonsmokers. *American Journal of Epidemiology 151*, 140–147.

(*188*)  Lee, S. C., Li, W. M., and Yin Chan, L. (2001). Indoor air quality at restaurants with different styles of cooking in metropolitan Hong Kong. *Science of The Total Environment 279*, 181–193.

(*189*)  Rochat, S., and Chaintreau, A. (2005). Carbonyl odorants contributing to the in-oven roast beef top note. *Journal of Agricultural and Food Chemistry 53*, 9578–9585.

(*190*)  Rosati, J. A., Krebs, K. A., and Liu, X. (2007). Emissions from Cooking Microwave Popcorn. *Emissions from Cooking Microwave Popcorn 47*, 701–709.

(*191*)  Rega, B., Guerard, A., Maire, M., and Giampaoli, P. (2006). Searching the missed flavour: chemical and sensory characterisation of flavour compounds released during baking. *Developments in Food Science 43*, 605–608.

(*192*)  Fehaili, S., Courel, M., Rega, B., and Giampaoli, P. (2010). An instrumented oven for the monitoring of thermal reactions during the baking of sponge cake. *Journal of Food Engineering 101*, 253–263.

(*193*)  Mcgorrin, R. J. (2009). One Hundred Years of Progress in Food Analysis. *Journal of Agricultural and Food Chemistry 57*, 8076–8088.

(*194*)  Cajak, T., Hajaslova, J., and Mastovska K In *Handbook of food analysis instruments*, Ötles, S., Ed.; CRC Press: Boca Raton, 2016, pp 197–228.

(*195*)  Knolhoff, A. M., and Croley, T. R. (2016). Non-targeted screening approaches for contaminants and adulterants in food using liquid chromatography hyphenated to high resolution mass spectrometry. *Journal of Chromatography A 1428*, 86–96.

(*196*)  Sharma, K., and Paradakar, M. (2010). The melamine adulteration scandal. *Food Security 2*, 97–107.

(*197*)  García, N., and Schwarzinger, S. In *Food Fraud: A Global Threat with Public Health and Economic Consequences*, Hellberg, R. S., Everstine, K., and Sklare, S. A., Eds.; Academic Press: 2021, pp 309–334.

(*198*)  Casadei, E., Valli, E., Panni, F., Donarski, J., Farrús Gubern, J., Lucci, P., Conte, L., Lacoste, F., Maquet, A., Brereton, P., Bendini, A., and Gallina Toschi, T. (2021). Emerging trends in olive oil fraud and possible countermeasures. *Food Control 124*.

(*199*)  Sanger, F., Air, G. M., Barrell, B., Brown, N. L., Coulson, A. R., Fiddes, C. A., Hutchison, C. A., Slocombe, P. M., and Smith, M. (1977). Nucleotide sequence of bacteriophage phi X174 DNA. *Nature 256*, 687–695.

(*200*)  Oliver, S. G., Winson, K., Michael, Kell, D. B., and Baganz, F. (1998). Systematic functional analysis of the yeast genome. *Trends in Biotechnology 16*.

(*201*)  Wilkins, M. R., Sanchez, J. C., Gooley, A. A., Appel, R. D., Humphery-Smith, I., Hochstrasser, D. F., and Williams, K. L. (1996). Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnology & genetic engineering reviews 13*, 19–50.

(*202*)  Patterson, S. D., and Aebersold, R. H. (2003). Proteomics: the first decade and beyond. *Nature genetics 33 Suppl*, 311–323.

(*203*)  Craig Venter, J. et al. (2001). The sequence of the human genome. *Science 291*, 1304–1351.

(*204*)  Alseekh, S., and Fernie, A. R. (2018). Metabolomics 20 years on: what have we learned and what hurdles remain? *Plant Journal 94*, 933–942.

(*205*) Schmitt-Kopplin, P., Hemmler, D., Moritz, F., Gougeon, R. D., Lucio, M., Meringer, M., Müller, C., Harir, M., and Hertkorn, N. (2019). Systems chemical analytics: introduction to the challenges of chemical complexity analysis. *Faraday Discussions 218*, 9–28.

(*206*) Uppal, K., Walker, D. I., Liu, K., Li, S., Go, Y. M., and Jones, D. P. (2016). Computational Metabolomics: A Framework for the Million Metabolome. *Chemical Research in Toxicology 29*, 1956–1975.

(*207*) Smedsgaard, J., and Nielsen, J. (2005). Metabolite profiling of fungi and yeast: from phenotype to metabolome by MS and informatics. *Journal of Experimental Botany 56*, 273–286.

(*208*) Fiehn, O. (2002). Metabolomics-the link between genotypes and phenotypes. *Plant Molecular Biology 48*, 155–171.

(*209*) Grainger, D. J., and Nicholson, J. K. In *Encyclopedia of Molecular Cell Biology and Molecular Medicine*, Meyers, R. A., Ed., 2nd ed.; Wiley-VCH Verlag GmbH & Co. KGaA: Weinheim, 2005, pp 189–205.

(*210*) https://metabolomicssociety.org/board-committees/scientific-task-groups/, 2023.

(*211*) Fiehn, O. et al. (2007). The metabolomics standards initiative (MSI). *Metabolomics 3*, 175–178.

(*212*) Sumner, L. W. et al. (2007). Proposed minimum reporting standards for chemical analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics 3*, 211–221.

(*213*) Dangour, A. D., Lock, K., Hayter, A., Aikenhead, A., Allen, E., and Uauy, R. (2010). Nutrition-related health effects of organic foods: A systematic review. *American Journal of Clinical Nutrition 92*, 203–210.

(*214*) Cifuentes, A. (2009). Food analysis and foodomics. *Journal of Chromatography A 1216*, 7109.

(*215*) García-Cañas, V., Simó, C., Herrero, M., Ibáñez, E., and Cifuentes, A. (2012). Present and future challenges in food analysis: Foodomics. *Analytical Chemistry 84*, 10150–10159.

(*216*) Valdés, A., Álvarez-Rivera, G., Socas-Rodríguez, B., Herrero, M., Ibáñez, E., and Cifuentes, A. (2022). Foodomics: Analytical Opportunities and Challenges. *Analytical Chemistry 94*, 366–381.

(*217*) Yang, Y., Li, J., Jia, X., Zhao, Q., Ma, Q., Yu, Y., Tang, C., and Zhang, J. (2022). Characterization of the Flavor Precursors and Flavor Fingerprints in Grazing Lambs by Foodomics. *Foods 11*, 191.

(*218*)  Man, K. Y., Chan, C. O., Tang, H. H., Dong, N. p., Capozzi, F., Wong, K. H., Kwok, K. W. H., Chan, H. M., and Mok, D. K. W. (2021). Mass spectrometry-based untargeted metabolomics approach for differentiation of beef of different geographic origins. *Food Chemistry 338*, 127847.

(*219*)  Josić, D., Peršurić, Rešetar, D., Martinović, T., Saftić, L., and Kraljević Pavelić, S. (2017). Use of Foodomics for Control of Food Processing and Assessing of Food Safety. *Advances in Food and Nutrition Research 81*, 187–229.

(*220*)  Álvarez, G., Montero, L., Llorens, L., Castro-Puyana, M., and Cifuentes, A. (2018). Recent advances in the application of capillary electromigration methods for food analysis and Foodomics. *Electrophoresis 39*, 136–159.

(*221*)  Da Silva, R. R., Dorrestein, P. C., and Quinn, R. A. (2015). Illuminating the dark matter in metabolomics. *Proceedings of the National Academy of Sciences of the United States of America 112*, 12549–12550.

(*222*)  Rappaport, S. M., Barupal, D. K., Wishart, D., Vineis, P., and Scalbert, A. (2014). The Blood Exposome and Its Role in Discovering Causes of Disease. *Environmental Health Perspectives 122*, 769–774.

(*223*)  https://lipidomicstandards.org/reporting_checklist/, 2023.

(*224*)  Haug, K., Cochrane, K., Nainala, V. C., Williams, M., Chang, J., Jayaseelan, K. V., and O'Donovan, C. (2020). MetaboLights: a resource evolving in response to the needs of its scientific community. *Nucleic Acids Research 48*, D440–D444.

(*225*)  Sud, M., Fahy, E., Cotter, D., Azam, K., Vadivelu, I., Burant, C., Edison, A., Fiehn, O., Higashi, R., Nair, K. S., Sumner, S., and Subramaniam, S. (2016). Metabolomics Workbench: An international repository for metabolomics data and metadata, metabolite standards, protocols, tutorials and training, and analysis tools. *Nucleic Acids Research 44*, D463–D470.

(*226*)  Haug, K., Salek, R. M., and Steinbeck, C. (2017). Global open data management in metabolomics. *Current Opinion in Chemical Biology 36*, 58–63.

(*227*)  Kang, M., Ko, E., and Mersha, T. B. (2022). A roadmap for multi-omics data integration using deep learning. *Briefings in Bioinformatics 23*.

(*228*)  Conesa, A., and Beck, S. (2019). Making multi-omics data accessible to researchers. *Scientific Data 2019 6:1 6*, 1–4.

(*229*)  Antonucci, F., Figorilli, S., Costa, C., Pallottino, F., Raso, L., and Menesatti, P. (2019). A review on blockchain applications in the agri-food sector. *Journal of the Science of Food and Agriculture 99*, 6129–6138.

(*230*) Galvez, J. F., Mejuto, J. C., and Simal-Gandara, J. (2018). Future challenges on the use of blockchain for food traceability analysis. *TrAC Trends in Analytical Chemistry 107*, 222–232.

(*231*) Niessen, W. M., and Correa C., R. A., *Interpretation of MS-MS Mass Spectra of Drugs and Pesticides*; Desiderio, D. M., and Loo, J. A., Eds.; John Wiley & Sons, Inc: Hoboken, NJ, 2017; Vol. 1, pp 1–388.

(*232*) Pratt, K. A., and Prather, K. A. (2012). Mass spectrometry of atmospheric aerosols—Recent developments and applications. Part II: on-line mass spectrometry techniques. *MassSpectrometryReviews 31*, 17–48.

(*233*) Adamovich, I. et al. (2017). The 2017 Plasma Roadmap: Low temperature plasma science and technology. *Journal of Physics D: Applied Physics 50*.

(*234*) Chu, P. K., and Lu, X., *Low temperature plasma technology methods and applications*; CRC Press: 2014.

(*235*) Smoluch, M., Mielczarek, P., and Silberring, J. (2016). Plasma-based ambient ionization mass spectrometry in bioanalytical sciences. *Mass Spectrometry Reviews 35*, 22–34.

(*236*) Weber, M., Wolf, J. C., and Haisch, C. (2021). Gas Chromatography-Atmospheric Pressure Inlet-Mass Spectrometer Utilizing Plasma-Based Soft Ionization for the Analysis of Saturated, Aliphatic Hydrocarbons. *Journal of the American Society for Mass Spectrometry 32*, 1707–1715.

(*237*) Harris, G. A., Nyadong, L., and Fernandez, F. M. (2008). Recent developments in ambient ionization techniques for analytical mass spectrometry. *Analyst 133*, 1297–1301.

(*238*) Meyer, C., Müller, S., Gurevich, E. L., and Franzke, J. (2011). Dielectric barrier discharges in analytical chemistry. *Analyst 136*, 2427–2440.

(*239*) Marshall, A. G., and Guant, S. (1996). Advantages of High Magnetic Field for Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Rapid Communications in Mass Spectrometry 10*, 1819–1823.

(*240*) Marshall, A. G., Hendrickson, C. L., and Jackson, G. S. (1998). Fourier transform ion cyclotron resonance mass spectrometry: A primer. *Mass Spectrometry Reviews 17*, 1–35.

(*241*) Kind, T., and Fiehn, O. (2007). Seven Golden Rules for heuristic filtering of molecular formulas obtained by accurate mass spectrometry. *BMC Bioinformatics 8*, 1–20.

(*242*) Tziotis, D., Hertkorn, N., and Schmitt-Kopplin, P. (2011). Kendrick-analogous network visualisation of ion cyclotron resonance Fourier transform mass spectra: Improved options for the assignment of elemental compositions and the classification of organic molecular complexity. *European Journal of Mass Spectrometry 17*, 415–421.

(243) Rama, E. C., Costa-García, A., and Fernández-Abedul, M. T. (2016). Pin-Based Flow Injection Electroanalysis. *Analytical Chemistry 88*, 9958–9963.

(244) Kolev, S. D., and McKelvie, I. D. (2008). Advances in flow injection analysis and related techniques. *Comprehensive analytical chemistry 54*, ed. by Barcelo, O.

(245) Zhu, J., and Cole, R. B. (2001). Ranking of gas-phase acidities and chloride affinities of monosaccharides and linkage specificity in collision-induced decompositions of negative ion electrospray-generated chloride adducts of oligosaccharides. *Journal of the American Society for Mass Spectrometry 12*, 1193–1204.

(246) Boutegrabet, L., Kanawati, B., Gebefügi, I., Peyron, D., Cayot, P., Gougeon, R. D., and Schmitt-Kopplin, P. (2012). Attachment of chloride anion to sugars: mechanistic investigation and discovery of a new dopant for efficient sugar ionization/detection in mass spectrometers. *Chemistry 18*, 13059–13067.

(247) Kuhl, C., Tautenhahn, R., Böttcher, C., Larson, T. R., and Neumann, S. (2012). CAMERA: An integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. *Analytical Chemistry 84*, 283–289.

(248) Pieczonka, S. A., Rychlik, M., and Schmitt-Kopplin, P. (2021). Metabolomics in Brewing Research. *Comprehensive Foodomics*, 116–128.

(249) Aharoni, A., De Vos, C. H., Verhoeven, H. A., Maliepaard, C. A., Kruppa, G., Bino, R., and Goodenowe, D. B. (2002). Nontargeted metabolome analysis by use of Fourier Transform Ion Cyclotron Mass Spectrometry. *Omics : a journal of integrative biology 6*, 217–234.

(250) Pieczonka, S. A., Hemmler, D., Moritz, F., Lucio, M., Zarnkow, M., Jacob, F., Rychlik, M., and Schmitt-Kopplin, P. (2021). Hidden in its color: A molecular-level analysis of the beer's Maillard reaction network. *Food Chemistry 361*, 130112.

(251) Pieczonka, S. A., Paravicini, S., Rychlik, M., and Schmitt-Kopplin, P. (2021). On the Trail of the German Purity Law: Distinguishing the Metabolic Signatures of Wheat, Corn and Rice in Beer. *Frontiers in Chemistry 9*, 715372.

(252) Pieczonka, S. A., Zarnkow, M., Diederich, P., Hutzler, M., Weber, N., Jacob, F., Rychlik, M., and Schmitt-Kopplin, P. (2022). Archeochemistry reveals the first steps into modern industrial brewing. *Scientific reports 12*.

(253) Hemmler, D., Roullier-Gall, C., Marshall, J. W., Rychlik, M., Taylor, A. J., and Schmitt-Kopplin, P. (2017). Evolution of Complex Maillard Chemical Reactions, Resolved in Time. *Scientific Reports 7*, 3–8.

(254) Roullier-Gall, C., Signoret, J., Coelho, C., Hemmler, D., Kajdan, M., Lucio, M., Schäfer, B., Gougeon, R. D., and Schmitt-Kopplin, P. (2020). Influence of regionality and maturation time on the chemical fingerprint of whisky. *Food Chemistry 323*, 126748.

(255) Kuhnert, N., D'souza, R. N., Behrends, B., Ullrich, M. S., and Witt, M. (2020). Investigating time dependent cocoa bean fermentation by ESI-FT-ICR mass spectrometry. *Food Research International 133*, 109209.

(256) Maia, M., Figueiredo, A., Cordeiro, C., and Sousa Silva, M. (2021). FT-ICR-MS-based metabolomics: A deep dive into plant metabolism. *Mass Spectrometry Reviews*, 1–22.

(257) Plumb, R. S., Gethings, L. A., Rainville, P. D., Isaac, G., Trengove, R., King, A. M., and Wilson, I. D. (2023). Advances in high throughput LC/MS based metabolomics: A review. *TrAC Trends in Analytical Chemistry 160*, 116954.

(258) Klinkenberg, A., and Zuiderweg, F. J. (1956). Longitudinal diffusion and resistance to mass transfer as causes of nonideality in chromatography. *Chemical Engineering Science 5*, 271–289.

(259) Asensio-Ramos, M., Hernández-Borges, J., Rocco, A., and Fanali, S. (2009). Food analysis: A continuous challenge for miniaturized separation techniques. *Journal of Separation Science 32*, 3764–3800.

(260) Žuvela, P., Skoczylas, M., Liu, J. J., Baczek, T., Kaliszan, R., Wong, M. W., Buszewski, B., and Héberger, K. (2019). Column Characterization and Selection Systems in Reversed-Phase High-Performance Liquid Chromatography. *Chemical Reviews 119*, 4818.

(261) Roca, M., Alcoriza, M. I., Garcia-Cañaveras, J. C., and Lahoz, A. (2021). Reviewing the metabolome coverage provided by LC-MS: Focus on sample preparation and chromatography-A tutorial. *Analytica Chimica Acta 1147*, 38–55.

(262) McCalley, D. V. (2017). Understanding and manipulating the separation in hydrophilic interaction liquid chromatography. *Journal of Chromatography A 1523*, 49–71.

(263) Dugo, P., Cacciola, F., Kumm, T., Dugo, G., and Mondello, L. (2008). Comprehensive multidimensional liquid chromatography: Theory and applications. *Journal of Chromatography A 1184*, 353–368.

(264) Goryński, K., Bojko, B., Nowaczyk, A., Buciński, A., Pawliszyn, J., and Kaliszan, R. (2013). Quantitative structure–retention relationships models for prediction of high performance liquid chromatography retention time of small molecules: Endogenous metabolites and banned compounds. *Analytica Chimica Acta 797*, 13–19.

(265) Forcisi, S., Moritz, F., Kanawati, B., Tziotis, D., Lehmann, R., and Schmitt-Kopplin, P. (2013). Liquid chromatography–mass spectrometry in metabolomics research: Mass analyzers in ultra high pressure liquid chromatography coupling. *Journal of Chromatography A 1292*, 51–65.

(266) Nagana Gowda, G. A., and Djukovic, D. (2014). Overview of mass spectrometry-based metabolomics: Opportunities and challenges. *Methods in Molecular Biology 1198*, 3–12.

(*267*)  Wang, R., Yin, Y., and Zhu, Z. J. (2019). Advancing untargeted metabolomics using data-independent acquisition mass spectrometry technology. *Analytical and Bioanalytical Chemistry 411*, 4349–4357.

(*268*)  Guo, J., and Huan, T. (2020). Comparison of Full-Scan, Data-Dependent, and Data-Independent Acquisition Modes in Liquid Chromatography-Mass Spectrometry Based Untargeted Metabolomics. *Analytical Chemistry 92*, 8072–8080.

(*269*)  Nye, L. C. et al. (2019). A comparison of collision cross section values obtained via travelling wave ion mobility-mass spectrometry and ultra high performance liquid chromatography-ion mobility-mass spectrometry: Application to the characterisation of metabolites in rat urine. *Journal of Chromatography A 1602*, 386–396.

(*270*)  Levy, A. J., Oranzi, N. R., Ahmadireskety, A., Kemperman, R. H., Wei, M. S., and Yost, R. A. (2019). Recent progress in metabolomics using ion mobility-mass spectrometry. *TrAC Trends in Analytical Chemistry 116*, 274–281.

(*271*)  Gaspari, M., and Cuda, G. (2011). Nano LC–MS/MS: A Robust Setup for Proteomic Analysis. *Nanoproteomics 790*, 115–126.

(*272*)  Li, W., Wu, Z., Xu, Y., Long, H., Deng, Y., Li, S., Xi, Y., Li, W., Cai, H., Zhang, B., and Wang, Y. (2023). Emerging LC-MS/MS-based molecular networking strategy facilitates foodomics to assess the function, safety, and quality of foods: recent trends and future perspectives. *Trends in Food Science & Technology 139*, 104114.

(*273*)  Hernández-Mesa, M., Ropartz, D., García-Campaña, A. M., Rogniaux, H., Dervilly-Pinel, G., and Le Bizec, B. (2019). Ion Mobility Spectrometry in Food Analysis: Principles, Current Applications and Future Trends. *Molecules 2019, Vol. 24, Page 2706 24*, 2706.

(*274*)  Taylor, A. J., Beauchamp, J. D., and Langford, V. S. (2021). Non-destructive and High-Throughput - APCI-MS, PTR-MS and SIFT-MS as Methods of Choice for Exploring Flavor Release. *ACS Symposium Series 1402*, 1–16.

(*275*)  Zoccali, M., Tranchida, P. Q., and Mondello, L. (2019). Fast gas chromatography-mass spectrometry: A review of the last decade. *TrAC Trends in Analytical Chemistry 118*, 444–452.

(*276*)  Kaplitz, A. S., Kresge, G. A., Selover, B., Horvat, L., Franklin, E. G., Godinho, J. M., Grinias, K. M., Foster, S. W., Davis, J. J., and Grinias, J. P. (2019). High-Throughput and Ultrafast Liquid Chromatography. *Analytical Chemistry 34*, 5.

(*277*)  Biasioli, F., Yeretzian, C., Märk, T. D., Dewulf, J., and Van Langenhove, H. (2011). Direct-injection mass spectrometry adds the time dimension to (B)VOC analysis. *TrAC - Trends in Analytical Chemistry 30*, 1003–1017.

(278)   Wieland, F., Gloess, A. N., Keller, M., Wetzel, A., Schenker, S., and Yeretzian, C. (2012). Online monitoring of coffee roasting by proton transfer reaction time-of-flight mass spectrometry (PTR-ToF-MS): Towards a real-time process control for a consistent roast profile. *Analytical and Bioanalytical Chemistry 402*, 2531–2543.

(279)   Gloess, A. N., Vietri, A., Wieland, F., Smrke, S., Schönbächler, B., López, J. A., Petrozzi, S., Bongers, S., Koziorowski, T., and Yeretzian, C. (2014). Evidence of different flavour formation dynamics by roasting coffee from different origins: On-line analysis with PTR-ToF-MS. *International Journal of Mass Spectrometry 365-366*, 324–337.

(280)   Majchrzak, T., Marć, M., and Wasik, A. (2022). Understanding the early-stage release of volatile organic compounds from rapeseed oil during deep-frying of tubers by targeted and omics-inspired approaches using PTR-MS and gas chromatography. *Food Research International 160*.

(281)   Liu, N., Koot, A., Hettinga, K., de Jong, J., and van Ruth, S. M. (2018). Portraying and tracing the impact of different production systems on the volatile organic compound composition of milk by PTR-(Quad)MS and PTR-(ToF)MS. *Food Chemistry 239*, 201–207.

(282)   Smith, D., and Španěl, P. (2005). Selected ion flow tube mass spectrometry (SIFT-MS) for on-line trace gas analysis. *Mass Spectrometry Reviews 24*, 661–700.

(283)   Huang, Y., and Barringer, S. A. (2011). Monitoring of Cocoa Volatiles Produced during Roasting by Selected Ion Flow Tube-Mass Spectrometry (SIFT-MS). *Journal of Food Science 76*.

(284)   Hertkorn, N., Frommberger, M., Witt, M., Koch, B. P., Schmitt-Kopplin, P., and Perdue, E. M. (2008). Natural organic matter and the event horizon of mass spectrometry. *Analytical Chemistry 80*, 8908–8919.

(285)   ElMasry, G., Morsy, N., Al-Rejaie, S., Ayed, C., Linforth, R., and Fisk, I. (2019). Real-time quality authentication of honey using atmospheric pressure chemical ionisation mass spectrometry (APCI-MS). *International Journal of Food Science & Technology 54*, 2983–2997.

(286)   Lancioni, C., Castells, C., Candal, R., and Tascon, M. (2022). Headspace solid-phase microextraction: Fundamentals and recent advances. *Advances in Sample Preparation 3*, 100035.

(287)   Lord, H., and Pawliszyn, J. (2000). Evolution of solid-phase microextraction technology. *Journal of Chromatography A 885*, 153–193.

(288)   Poinot, P., Grua-Priol, J., Arvisenet, G., Rannou, C., Semenou, M., Bail, A. L., and Prost, C. (2007). Optimisation of HS-SPME to study representativeness of partially baked bread odorant extracts. *Food Research International 40*, 1170–1184.

(*289*) Adelina, N. M., Wang, H., Zhang, L., and Zhao, Y. (2021). Comparative analysis of volatile profiles in two grafted pine nuts by headspace-SPME/GC–MS and electronic nose as responses to different roasting conditions. *Food Research International 140*, 110026.

(*290*) Harwood, A. D., Nutile, S. A., Landrum, P. F., and Lydy, M. J. (2015). Tenax extraction as a simple approach to improve environmental risk assessments. *Environmental Toxicology and Chemistry 34*, 1445–1453.

(*291*) Ranau, R., and Steinhart, H. (2005). Identification and evaluation of volatile odor-active pollutants from different odor emission sources in the food industry. *European Food Research and Technology 220*, 226–231.

(*292*) Fullana, A., Carbonell-Barrachina, Á. A., and Sidhu, S. (2004). Volatile aldehyde emissions from heated cooking oils. *Journal of the Science of Food and Agriculture 84*, 2015–2021.

(*293*) Wampler, T. P. In *Techniques for Analyzing Food Aroma*, Marsili, R., Ed., 1st; CRC Press: Boca Raton, 1996, pp 27–58.

(*294*) Schmid, R. et al. (2023). Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nature Biotechnology 41*, 447–449.

(*295*) Dunn, W. B., Wilson, I. D., Nicholls, A. W., and Broadhurst, D. (2012). The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans. *Bioanalysis 4*, 2249–2264.

(*296*) Briscoe, C. J., Stiles, M. R., and Hage, D. S. (2007). System suitability in bioanalytical LC/MS/MS. *Journal of Pharmaceutical and Biomedical Analysis 44*, 484–491.

(*297*) Evans, A. M. et al. (2020). Dissemination and analysis of the quality assurance (QA) and quality control (QC) practices of LC–MS based untargeted metabolomics practitioners. *Metabolomics 16*, 1–16.

(*298*) Beger, R. D. et al. (2019). Towards quality assurance and quality control in untargeted metabolomics studies. *Metabolomics 15*, 1–5.

(*299*) Neumann, S., and Böcker, S. (2010). Computational mass spectrometry for metabolomics: Identification of metabolites and small molecules. *Analytical and Bioanalytical Chemistry 398*, 2779–2788.

(*300*) Wishart, D. S. et al. (2022). HMDB 5.0: the Human Metabolome Database for 2022. *Nucleic Acids Research 50*, D622–D631.

(*301*) Wang, M. et al. (2016). Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnology 34*, 828–837.

(*302*)   Horai, H. et al. (2010). MassBank: A public repository for sharing mass spectral data for life sciences. *Journal of Mass Spectrometry 45*, 703–714.

(*303*)   de Jonge, N. F., Mildau, K., Meijer, D., Louwen, J. J., Bueschl, C., Huber, F., and van der Hooft, J. J. (2022). Good practices and recommendations for using and benchmarking computational metabolomics metabolite annotation tools. *Metabolomics 18.*

(*304*)   Demuth, W., Karlovits, M., and Varmuza, K. (2004). Spectral similarity versus structural similarity: mass spectrometry. *Analytica Chimica Acta 516*, 75–85.

(*305*)   Ernst, M., Kang, K. B., Caraballo-Rodríguez, A. M., Nothias, L. F., Wandy, J., Chen, C., Wang, M., Rogers, S., Medema, M. H., Dorrestein, P. C., and van der Hooft, J. J. (2019). Molnetenhancer: Enhanced molecular networks by integrating metabolome mining and annotation tools. *Metabolites 9.*

(*306*)   Amara, A., Frainay, C., Jourdan, F., Naake, T., Neumann, S., Novoa-del-Toro, E. M., Salek, R. M., Salzer, L., Scharfenberg, S., and Witting, M. (2022). Networks and Graphs Discovery in Metabolomics Data Analysis and Interpretation. *Frontiers in Molecular Biosciences 9*, 841373.

(*307*)   Dührkop, K., Shen, H., Meusel, M., Rousu, J., and Böcker, S. (2015). Searching molecular structure databases with tandem mass spectra using CSI:FingerID. *Proceedings of the National Academy of Sciences of the United States of America 112*, 12580–12585.

(*308*)   Dührkop, K., Nothias, L. F., Fleischauer, M., Reher, R., Ludwig, M., Hoffmann, M. A., Petras, D., Gerwick, W. H., Rousu, J., Dorrestein, P. C., and Böcker, S. (2020). Systematic classification of unknown metabolites using high-resolution fragmentation mass spectra. *Nature Biotechnology 2020 39:4 39*, 462–471.

(*309*)   Ludwig, M., Nothias, L. F., Dührkop, K., Koester, I., Fleischauer, M., Hoffmann, M. A., Petras, D., Vargas, F., Morsy, M., Aluwihare, L., Dorrestein, P. C., and Böcker, S. (2020). Database-independent molecular formula annotation using Gibbs sampling through ZODIAC. *Nature Machine Intelligence 2020 2:10 2*, 629–641.

(*310*)   Djoumbou Feunang, Y., Eisner, R., Knox, C., Chepelev, L., Hastings, J., Owen, G., Fahy, E., Steinbeck, C., Subramanian, S., Bolton, E., Greiner, R., and Wishart, D. S. (2016). ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. *Journal of Cheminformatics 8*, 1–20.

(*311*)   Huber, F., Ridder, L., Verhoeven, S., Spaaks, J. H., Diblen, F., Rogers, S., and Van Der Hooft, J. J. (2021). Spec2Vec: Improved mass spectral similarity scoring through learning of structural relationships. *PLOS Computational Biology 17.*

(*312*)   Neumann, J. (2022). FAIR Data Infrastructure. *Advances in Biochemical Engineering Biotechnology 182*, 195–207.

(313)  Chokkathukalam, A., Kim, D. H., Barrett, M. P., Breitling, R., and Creek, D. J. (2014). Stable isotope-labeling studies in metabolomics: new insights into structure and dynamics of metabolic networks. *Bioanalysis 6*, 511–524.

(314)  Schieberle, P. (2005). The Carbon Module Labeling (CAMOLA) Technique: A Useful Tool for Identifying Transient Intermediates in the Formation of Maillard-Type Target Molecules. *Annals of the New York Academy of Sciences 1043*, 236–248.

(315)  Desiere, F., German, B., Watzke, H., Pfeifer, A., and Saguy, S. (2001). Bioinformatics and data knowledge: the new frontiers for nutrition and foods. *Trends in Food Science & Technology 12*, 215–229.

(316)  Lasi, H., Fettke, P., Kemper, H.-G., Feld, T., and Hoffmann, M. (2014). Industrie 4.0. *Wirtschaftsinformatik 56*, 261–264.

(317)  Hassoun, A., Jagtap, S., Garcia-Garcia, G., Trollman, H., Pateiro, M., Lorenzo, J. M., Trif, M., Rusu, A. V., Aadil, R. M., Šimat, V., Cropotova, J., and Câmara, J. S. (2023). Food quality 4.0: From traditional approaches to digitalized automated analysis. *Journal of Food Engineering 337*, 111216.

(318)  Hassoun, A., Jagtap, S., Trollman, H., Garcia-Garcia, G., Abdullah, N. A., Goksen, G., Bader, F., Ozogul, F., Barba, F. J., Cropotova, J., Munekata, P. E., and Lorenzo, J. M. (2023). Food processing 4.0: Current and future developments spurred by the fourth industrial revolution. *Food Control 145*, 109507.

(319)  Wang, X., Bouzembrak, Y., Lansink, A. G., and van der Fels-Klerx, H. J. (2022). Application of machine learning to the monitoring and prediction of food safety: A review. *Comprehensive Reviews in Food Science and Food Safety 21*, 416–434.

(320)  Jeevanandam, J., Agyei, D., Danquah, M. K., and Udenigwe, C. (2022). Food quality monitoring through bioinformatics and big data. *Future Foods: Global Trends, Opportunities, and Sustainability Challenges*, 733–744.

(321)  Deng, X., Cao, S., and Horn, A. L. (2021). Emerging Applications of Machine Learning in Food Safety. *Annual Review of Food Science and Technology 12*, 513–538.

(322)  Dalabasmaz, S., Dittrich, D., Kellner, I., Drewello, T., and Pischetsrieder, M. (2019). Identification of peptides reflecting the storage of UHT milk by MALDI-TOF-MS peptide profiling. *Journal of Proteomics 207*, 103444.

(323)  Jiménez-Carvelo, A. M., González-Casado, A., Bagur-González, M. G., and Cuadros-Rodríguez, L. (2019). Alternative data mining/machine learning methods for the analytical evaluation of food quality and authenticity – A review. *Food Research International 122*, 25–39.

(324)  Erban, A., Fehrle, I., Martinez-Seidel, F., Brigante, F., Más, A. L., Baroni, V., Wunderlin, D., and Kopka, J. (2019). Discovery of food identity markers by metabolomics and machine learning technology. *Scientific Reports 2019 9:1 9*, 1–19.

(*325*) Davy, H., Sinclair, G., and Bedford, J. R., *Elements of agricultural chemistry, in a course of lectures for the Board of agriculture*; Eastburn, Kirk and Co: New York, 1815, pp 1–452.

(*326*) McDonald, J. D., Zielinska, B., Fujita, E. M., Sagebiel, J. C., Chow, J. C., and Watson, J. G. (2003). Emissions from charbroiling and grilling of chicken and beef. *Journal of the Air and Waste Management Association 53*, 185–194.

(*327*) Commins, B. T. (1969). Formation of polycyclic aromatic hydrocarbons during pyrolysis and combustion of hydrocarbons. *Atmospheric Environment 3*, 565–572.

(*328*) Engel, W., Bahr, W., and Schieberle, P. (1999). Solvent assisted flavour evaporation - A new and versatile technique for the careful and direct isolation of aroma compounds from complex food matrices. *European Food Research and Technology 209*, 237–241.

(*329*) Zhao, Y., Hu, M., Slanina, S., and Zhang, Y. (2006). Chemical Compositions of Fine Particulate Organic Matter Emitted from Chinese Cooking. *Environmental Science and Technology 41*, 99–105.

(*330*) Zhao, W., Hopke, P. K., Gelfand, E. W., and Rabinovitch, N. (2007). Use of an expanded receptor model for personal exposure analysis in schoolchildren with asthma. *Atmospheric Environment 41*, 4084–4096.

(*331*) Zimmerman, J. B., Anastas, P. T., Erythropel, H. C., and Leitner, W. (2020). Designing for a green chemistry future. *Science 367*, 397–400.

(*332*) De Roos, K. B. (2003). Effect of texture and microstructure on flavour retention and release. *International Dairy Journal 13*, 593–605.

(*333*) Kulmala, M., Vehkamäki, H., Petäjä, T., Dal Maso, M., Lauri, A., Kerminen, V. M., Birmili, W., and McMurry, P. H. (2004). Formation and growth rates of ultrafine atmospheric particles: a review of observations. *Journal of Aerosol Science 35*, 143–176.

(*334*) Sioutas, C., Delfino, R. J., and Singh, M. (2005). Exposure Assessment for Atmospheric Ultrafine Particles (UFPs) and Implications in Epidemiologic Research. *Environmental Health Perspectives 113*, 947.

(*335*) Masango, P. (2005). Cleaner production of essential oils by steam distillation. *Journal of Cleaner Production 13*, 833–839.

(*336*) Liu, P., Yuan, Z., Zhang, S., Xu, Z., and Li, X. (2018). Experimental study of the steam distillation mechanism during the steam injection process for heavy oil recovery. *Journal of Petroleum Science and Engineering 166*, 561–567.

(*337*) Norwitz, G., Nataro, N., and Keliher, P. N. (1987). Steam distillation of phenolic compounds in the presence of a large amount of sodium chloride. *Microchemical Journal 35*, 240–243.

(*338*)   Rijks, J., Curvers, J., Noy, T., and Cramers, C. (1983). Possibilities and limitations of steam distillation—extraction as a pre-concentration technique for trace analysis of organics by capillary gas chromatography. *Journal of Chromatography A 279*, 395–407.

(*339*)   Lee, C. C., and Chen, C.-H. (1966). Studies with Radioactive Tracers IX. The Fate of Sucrose-14C during Breadmaking. *Cereals and Grains Association.*

(*340*)   Seydel, C. (2021). Single-cell metabolomics hits its stride. *Nature Methods 2021 18:12 18*, 1452–1456.

(*341*)   Rychlik, M., and Schmitt-Kopplin, P. (2020). Reading From the Crystal Ball: The Laws of Moore and Kurzweil Applied to Mass Spectrometry in Food Analysis. *Frontiers in Nutrition 7*, 516095.

(*342*)   Murkovic, M. (2004). Formation of heterocyclic aromatic amines in model systems. *Journal of Chromatography B 802*, 3–10.

(*343*)   van den Berg, F., Lyndgaard, C. B., Sørensen, K. M., and Engelsen, S. B. (2013). Process Analytical Technology in the food industry. *Trends in Food Science & Technology 31*, 27–35.

(*344*)   Alves, V. M., Auerbach, S. S., Kleinstreuer, N., Rooney, J. P., Muratov, E. N., Rusyn, I., Tropsha, A., and Schmitt, C. (2021). Curated Data In — Trustworthy In Silico Models Out: The Impact of Data Quality on the Reliability of Artificial Intelligence Models as Alternatives to Animal Testing. *Alternatives to Laboratory Animals 49*, 73–82.

(*345*)   Jain, A., Patel, H., Nagalapatti, L., Gupta, N., Mehta, S., Guttula, S., Mujumdar, S., Afzal, S., Sharma Mittal, R., and Munigala, V. (2020). Overview and Importance of Data Quality for Machine Learning Tasks. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 3561–3562.

(*346*)   Lin, C. S., Pan, Y. C., Kuo, Y. X., Chen, C. K., and Tien, C. L. (2021). A Study of Automatic Judgment of Food Color and Cooking Conditions with Artificial Intelligence Technology. *Processes 2021, Vol. 9, Page 1128 9*, 1128.

(*347*)   Wishart, D. https://foodb.ca, 2023.

# List of Figures

# Curriculum Vitae

**Leopold Weidner**
leopold.weidner[@]tum.de

## ACADEMIC AND PROFESSIONAL EXPERIENCE

08/2019 - 10/2023 | **Research Associate** at the Chair of Analytical Food Chemistry, Technical University of Munich and at Helmholtz Munich Research Unit Analytical BioGeoChemistry

10/2018 - 04/2019 | **Final Thesis in Food Chemistry** at „l'Institute des Sciences du Vin et de la Vigne" titled "Characterization of the aromatic notes associated with the qualitative aging of Cognac eaux-de-vie"

| | |
|---|---|
| 10/2016 - 09/2018 | **Zweiter Prüfungsabschnitt der Staatsprüfung für Staatlich geprüfte Lebensmittelchemikerinnen und Lebensmittelchemiker** at the Friedrich-Alexander-University Erlangen-Nürnberg, graduated as "Lebensmittel-chemiker" |
| 10/2016 - 03/2018 | **Teaching Assistant for Stereochemistry** at the Friedrich-Alexander-University Erlangen-Nürnberg |
| 10/2014 - 09/2016 | **Erster Prüfungsabschnitt der Staatsprüfung für Staatlich geprüfte Lebensmittelchemikerinnen und Lebensmittelchemiker** at the Friedrich-Alexander-University Erlangen-Nürnberg |
| 08/2012 - 09/2016 | **Laboratory Assistant** at Gesellschaft für Schadstoffuntersuchung mbH |

## VOLUNTARY COMMITMENT, MEMBERSHIPS AND ADDITIONAL ACTIVITIES DURING ACADEMIC EDUCATION

| | |
|---|---|
| since 2021 | **Metabolomics Society:** Active Member |
| since 2019 | **Member of the TUM Graduate School**: at the School of Life Sciences, Graduate Center HELENA |
| since 2019 | **Student supervision**: practical courses, research internships and Master's theses |
| since 2014 | **AG Junge Lebensmittelchemie:** Devision of the Lebensmittelchemische Gesellschaft, active member and temporarily head of the local committees in Erlangen-Nürnberg and München-Freising |
| since 2010 | **Deutscher Alpenverein:** Sections Nürnberg and München, active member |

## AWARDS

05/2021 | **Best Poster Award** in "Analytical tools"
| at the $16^{th}$ Weurman Flavour Research Symposium

# Eidstaatliche Erklärung

Ich, **Leopold Weidner**, erkläre an Eides statt, dass ich die bei der promotionsführenden Einrichtung TUM School of Life Sciences der Technischen Universität München zur Promotionsprüfung vorgelegte Arbeit mit dem Titel:

**Metabolomic analysis to profile thermal food processing emissions**

am Lehrstuhl für Analytische Lebensmittelchemie unter der Anleitung und Betreuung durch **apl.-Prof. Dr. Philippe Schmitt-Kopplin** ohne sonstige Hilfe erstellt und bei der Abfassung nur die gemäß § 7 Abs. 6 und 7 angegebenen Hilfsmittel benutzt habe.

(x) Ich habe keine Organisation eingeschaltet, die gegen Entgelt Betreuerinnen und Betreuer für die Anfertigung von Dissertationen sucht, oder die mir obliegenden Pflichten hinsichtlich der Prüfungsleistungen für mich ganz oder teilweise erledigt.

(x) Ich habe die Dissertation in dieser oder ähnlicher Form in keinem anderen Prüfungsverfahren als Prüfungsleistung vorgelegt.

(x) Teile der Dissertation wurden in Food Chemistry, Analytical Chemistry, Bioinformatics und Journal of Agricultural and Food Chemistry veröffentlicht.

(x) Ich habe den angestrebten Doktorgrad noch nicht erworben und bin nicht in einem früheren Promotionsverfahren für den angestrebten Doktorgrad endgültig gescheitert.

(x) Ich habe keine Kenntnis über ein strafrechtliches Ermittlungsverfahren in Bezug auf wissenschaftsbezogene Straftaten gegen mich oder eine rechtskräftige strafrechtliche Verurteilung mit Wissenschaftsbezug.

Die öffentlich zugängliche Promotionsordnung sowie die Richtlinien zur Sicherung guter wissenschaftlicher Praxis und für den Umgang mit wissenschaftlichem Fehlverhalten der TUM sind mir bekannt, insbesondere habe ich die Bedeutung von § 27 PromO (Nichtigkeit der Promotion) und § 28 PromO (Entzug des Doktorgrades) zur Kenntnis genommen.

Ich bin mir der Konsequenzen einer falschen Eidesstattlichen Erklärung bewusst.

Mit der Aufnahme meiner personenbezogenen Daten in die Alumni-Datei bei der TUM bin ich einverstanden.

_____

(Leopold Weidner)          München, den 24. Februar 2024