

The Comprehensive Chemical Characterization of the Active Pharmaceutical Ingredients in Bituminosulfonates: A Case Study for the Elucidation of Complex Drugs

Cumulative Dissertation

submitted to

the Faculty of Mathematics and Natural Sciences of the University of Rostock

in fulfilment of the requirements for the academic degree

doctor rerum naturalium (Dr. rer. nat.)

by

Lukas Felix Schwalb
born 24.08.1994 in Bad Godesberg now Bonn

Rostock, 2024

Reviewer:

Prof. Dr. Ralf Zimmermann, University of Rostock, Institute of Chemistry

Prof. Dr. Erwin Rosenberg, Vienna University of Technology, Institute of Chemical Technologies and Analytics

Year of submission: 2024

Year of defense: 2024

First Authorships

The following manuscripts were contributed by Lukas Schwalb as the first author and published in peer-reviewed journals. Lukas Schwalb's contribution is declared below.

Title: Analysis of complex drugs by comprehensive two-dimensional gas chromatography and high-resolution mass spectrometry: detailed chemical description of the active pharmaceutical ingredient sodium bituminosulfonate and its process intermediates

Authors: Schwalb, L.; Tiemann, O.; Käfer, U.; Gröger, T.; Rüger, C. P.; Gayko, G.; Zimmermann, R.

Journal: Analytical and Bioanalytical Chemistry

Year: 2023

DOI: <https://doi.org/10.1007/s00216-022-04393-w>

Lukas Schwalb conceived and designed the majority of the study, performed the experiments, processed and evaluated most of the data, and wrote the majority of the manuscript.
--

Title: Applying a risk assessment guided evaluation for verifying comprehensive two-dimensional gas chromatography to analyse complex pharmaceuticals

Authors: Schwalb, L.; Tiemann, O.; Käfer, U.; Gröger, T.; Rüger, C. P.; Zimmermann, R.

Journal: Analytical and Bioanalytical Chemistry

Year: 2024

DOI: <https://doi.org/10.1007/s00216-023-05093-9>

Lukas Schwalb conceived and designed the majority of the study, performed and developed the gas phase analyses, processed and evaluated their data, and wrote most of the manuscript.

Co-authorships:

Submitted:

Title: Rock-to-Pharma: Characterization of Shale Oil-Based Non-Biological Complex Drugs Along the Production Process by High-Resolution Mass Spectrometry

Authors: Tiemann, O.; Rüger, C. P.; Schwalb, L.; Chacón-Patiño, M. L.; Gröger, T.; Zimmermann, R.

Journal: submitted to ACS Analytical Chemistry

Year: submitted on 08.03.2024

DOI: TBD

Lukas Schwalb performed and evaluated the GC×GC experiments, revised and wrote minor parts of the manuscript.

Acknowledgements

I want to express my heartfelt gratitude to the Joint Mass Spectrometry Center (JMSC) team for their invaluable help and guidance during my PhD research. Without their support, I would not have been able to complete my work successfully.

*First, I want to thank **Prof. Dr. Ralf Zimmermann** for his support, supervision, and scientific input and, most importantly, for giving me the opportunity to pursue my doctoral research at JMSC.*

*Furthermore, I am especially grateful for the continuing support of my supervisor, **Thomas Gröger**. For the three and a half years at CMA, I'm grateful; he gave me the possibility to grow, find and follow my research, and travel across the country and abroad to numerous unique cooperation partners and fantastic conferences. I could always rely on his input and help. He is a great mentor; I thank him for his patience and everything he taught me.*

*I also want to thank my supervisor, **Dr. Uwe Käfer**. He introduced me to the device I could count on for the last three and half years, showed me the tricks with the software, and awakened my passion for GC×GC, our community, and data evaluation. I hoped we would have worked together a little bit longer. He greatly impacted me and my thesis.*

*I'm very grateful for the opportunity for this project with outstanding project partners, particularly **Ole Tiemann**. I could not have wished for a better colleague. Together, we managed not only the complex matrices but also handled the challenges of this project. I would also like to thank **Dr. Christopher Rüger**. I always enjoyed the feedback from Rostock and his perspective. I want to thank both of them for the good teamwork!*

*The next project partners I would like to thank are the Ichthyol-Gesellschaft, Cordes, Hermann & Co. (GmbH & Co.). They financially supported the project, provided the sample material, and gave scientific input. In particular, I want to thank **Dr. Guido Gayko** for his cooperation, input, and feedback. It was a pleasure to work with these scientifically fascinating matrices.*

*A special thanks to my colleagues at CMA. You help me with the challenging PhD time. First, I want to thank the entire topic V and the people I had the pleasure to work with: **Elena, Nadine, Ramona, Barbara, and Sandra**. However, so many more people positively influenced me and my work. I want to thank **Elias, Svenja, Jana, Seongho, Anja, and Anusmita** for their support and input over the last three years. Also, besides the work, I want to thank my **friends** for their constant encouragement and friendship.*

*I want to thank my **family**, who always believed in me. To my **parents, brother, and grandparents**, I am very grateful for the opportunities, love, and support they gave me. Last but not least, I would like to thank **Prisca**. She is always by my side, cheering me up and supporting me all the time. For that I am very grateful!*

Table of Contents

Abbreviations and Nomenclature	X
Zusammenfassung.....	XIII
Abstract.....	XV
1 Motivation.....	1
2 Introduction.....	2
2.1 Analysis and Authorization of Drugs in the Regulated Pharmaceutical Industry	2
2.2 The Special Case of Non-Biological Complex Drugs.....	4
2.3 The Complex Active Pharmaceutical Ingredients Bituminosulfonates.....	5
2.4 Analytical Approaches to Address Non-Biological Complex Drugs.....	7
3 Scope.....	9
4 Methods & Instrumentations	10
4.1 High-Resolution Mass Spectrometry (HR-MS)	10
4.1.1 Electron Ionization Time-of-Flight Mass Spectrometry (EI ToF-MS)	11
4.1.2 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR MS).....	13
4.2 Applied Hyphenated Separation Techniques	14
4.2.1 Comprehensive Two-Dimensional Gas Chromatography (GC×GC).....	14
4.2.2 Thermogravimetric Analysis (TGA)	17
4.3 Sample Preparation.....	18
4.3.1 Extraction	19
4.3.2 Derivatization	19
5 Results & Discussion	21
5.1 Risk Assessment-Based Analytical Workflow.....	21
5.1.1 Identification of Critical Parameters via Risk Assessment	21
5.1.2 Overall Study Design	23
5.2 Assessing the Chemical Effects of the Defining Manufacturing Process	24
5.2.1 Vacuum Distillation of the Shale Oil	24
5.2.2 Refining of the Light Distillate.....	27
5.3 Analysis of Sodium Bituminosulfonates	28
5.3.1 Thermal Analysis	28
5.3.2 Sample Preparation.....	29
5.4 Transfer of the Online Derivatization Method to Ammonium Bituminosulfonates.....	34
5.4.1 Chemical Description of Ammonium Bituminosulfonate (ABS)	34
5.4.2 Chemical Description of Ammonium Bituminosulfonate (ABS (mix)).....	36
5.5 Application of an Orthogonal Technique to Cross-Verify Complex Pharmaceuticals	38
6 Conclusion & Outlook.....	40
7 References	41
List of Figures/Tables.....	49
Appendix	51
Curriculum Vitae	51
Conference Participations.....	52
Publications	53

Abbreviations and Nomenclature

Abbreviations and Nomenclature

ABS	Ammonium bituminosulfonate (based on the middle distillate)
ABS (mix)	Ammonium bituminosulfonate (based on the combined distillate phases)
API	Active pharmaceutical ingredient
APPI	Atmosphere pressure photoionization
BDiTSME	Benzodithiophenesulfonate methyl ester
BiPSME	Biphenylsulfonate methyl ester
BiTSME	Bithiophenesulfonate methyl ester
BSME	Benzenesulfonate methyl ester
BTSME	Benzothiophenesulfonate methyl ester
DBE	Double bound equivalent
DiBTSME	Dibenzothiophenesulfonate methyl ester
EI	Electron ionization
ESI	Electrospray ionization
FT-ICR MS	Fourier transform ion cyclotron resonance mass spectrometry
GC	Gas chromatography
GC×GC	Comprehensive two-dimensional gas chromatography
GMP	Good manufacturing practice
HLB	Hydrophilic-lipophilic balance
HR-MS	High-resolution mass spectrometry
HR-ToF-MS	High-resolution time-of-flight mass spectrometry
ICH	International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use
ISME	Indanesulfonate methyl ester
L/L	Liquid/liquid (extraction)
LC	Liquid chromatography
Light distillate	Light distillate fraction of the vacuum distillate shale oil
LOD	Limit of detection
<i>m/z</i>	Mass-to-charge (ratio)
Middle distillate	Middle distillate fraction of the vacuum distillate shale oil
MS	Mass spectrometry
NBCD	Non-biological complex drugs
NSME	Naphthalenesulfonate methyl ester
Ph. Eur.	European Pharmacopoeia
ppm	Parts per million

Abbreviations and Nomenclature

Precursor distillate	Refined light distillate; Process intermediate of SBS
PTSME	Phenylthiophenesulfonate methyl ester
QC	Quality control
SAX	Strong anion exchanger
SBS	Sodium bituminosulfonate
SME	Sulfonate methyl ester
SPE	Solid phase extraction
TGA	Thermogravimetric analyzing
THBTSME	Tetrahydrobenzothiophenesulfonate methyl ester
TMAH	Tetramethylammonium hydroxide
TMSDAM	Trimethylsilyl diazomethane
TSME	Thiophenesulfonate methyl ester
VOC	Volatile organic compounds
WAX	Weak anion exchanger

Zusammenfassung

Seit Jahrhunderten werden komplexe Arzneimittel in der Medizin verwendet. Über diese Zeitspanne belegen zahlreiche Patientenberichte und klinische Studien ihre Wirkung, trotzdem ist ihre chemische Zusammensetzung immer noch weitestgehend unbekannt. Das ist auch der Fall bei der 2014 eingeführten Klasse der nicht biologischen komplexen Arzneimitteln (NBCDs; engl.: Non-biological complex drugs). Diese werden durch die komplexe chemische Zusammensetzung und analytisch unzugänglichen Wirkstoffe (APIs; engl.: Active pharmaceutical ingredients) definiert. Deswegen werden sie durch ihren Herstellungsprozess und simple physikochemischen Parameter beschrieben und kontrolliert, so auch bei Bituminosulfonaten. Arzneimittel mit diesen Wirkstoffen werden heute vorläufig den NBCDs zugeordnet und wie bei anderen Wirkstoffen dieser pharmazeutischen Klasse, ist ihre chemische Zusammensetzung nur teilweise untersucht.

Diese Arbeit befasst sich sowohl mit den analytischen Herausforderungen von Natrium- (SBS; engl.: Sodium bituminosulfonate) und Ammoniumbituminosulfonaten (ABS; engl.: Ammonium bituminosulfonates) als auch den strengen Anforderungen des regulatorischen Umfeldes an die Verlässlichkeit der Ergebnisse. Als Schlüsseltechnik für die Beschreibung der Matrices wurde die umfassende zweidimensionale Gaschromatographie mit hochauflösender Flugzeit-Massenspektrometrie (GC×GC HR-ToF-MS; engl.: Comprehensive two-dimensional gas chromatography high-resolution time-of-flight mass spectrometry) eingesetzt. Zu Beginn der Arbeit wurde eine Risikobewertung durchgeführt, auf deren Basis verschiedene Probenvorbereitungen, thermogravimetrischen Analysen (TGA; engl.: Thermogravimetric analysis) und eine Cross-Verifizierung mittels Fourier-Transform Ionenzyklotronresonanz Massenspektrometrie (FT-ICR MS; engl.: Fourier transform ion cyclotron resonance mass spectrometry) in die Methodenentwicklung mit einbezogen wurden.

Die umfassende Methodenentwicklung mit komplementierenden Methoden konnte die vollständige Sulfonierung von SBS zeigen. Außerdem wies die entwickelte Methode eine hohen Derivatisierungseffizienz auf und führte weder zu thermischen noch zu derivatisierungsbedingten Nebenreaktionen. Darüber hinaus konnte der größte Teil der chemischen Zusammensetzung von SBS mit acht sulfonierte Stoffklassen, inklusive ihrer Grundstruktur und Isomerenverteilung, beschrieben werden. Während SBS nahezu vollständig mit diesen sulfonierten Substanzklassen beschrieben werden konnte, enthalten die beiden ABS-Wirkstoffe einen hohen Anteil an nicht-sulfonierten Verbindungen. Zusätzlich werden durch die umfassende Beschreibung sämtlicher isolierbaren Matrices die chemischen Profile der Wirkstoffe mit ihrer definierenden Herstellung verbunden.

Die Cross-Verifizierung mittels FT-ICR MS und verschiedenen Ionisierungstechniken konnten diese Ergebnisse noch weiter bestätigen. Zusätzlich zeigt der Vergleich sowohl die Vorteile der hohen chromatographischen und massenspektrometrischen Auflösung und der universellen Ionisation, aber auch Limitierung bei der Analyse schwer- oder nicht flüchtiger Verbindungen in ABS.

Diese Arbeit beschreibt nicht nur Bituminosulfonate bis auf die molekulare Ebene, sondern beweist auch die Zuverlässigkeit ihrer Analyseergebnisse.

Abstract

Many complex drugs have been used for centuries. Although their pharmaceutical effect is proven by patient reports and clinical studies, their chemical composition is only partially discovered. The pharmaceutical class of non-biological complex drugs (NBCDs), introduced in 2014, is defined by the inaccessibility of their active pharmaceutical ingredients (APIs) for commonly used methods described by the European Pharmacopoeia (Ph. Eur.). They are defined by their production process and controlled by simple physicochemical parameters. This is also the case for bituminosulfonates. Drugs with these APIs are tentatively allocable to NBCDs, and while their pharmaceutical effects are well described, the chemical composition is only partially known.

This study deals with the analytical challenges of sodium bituminosulfonate (SBS) and ammonium bituminosulfonates (ABS) in consideration of the strict regulatory requirements for the reliability of the results. In addition, it focuses on the analysis via comprehensive two-dimensional gas chromatography hyphenated with high-resolution time-of-flight mass spectrometry (GC×GC HR-ToF-MS). Here, a risk assessment was performed at the start of the method development. Based on this assessment, the analysis of the starting materials and process intermediates, thermogravimetric analysis (TGA), multiple sample preparation methods, and Fourier transform ion cyclotron resonance mass spectrometry (FT-ICR MS) were implemented for the method development.

Within the comprehensive method development, the complete sulfonation of SBS was shown. Additionally, the developed method showed a high derivatization efficiency and the absence of thermally or derivatization agent-induced side reactions. Furthermore, more than 90 % of the abundance of SBS could be described by eight sulfonated compound classes, including their core structural motive and isomeric distribution. For both ABS matrices, three additional classes were identified. However, they show high percentages of non-sulfonated compounds. With the comprehensive chemical description of all isolatable matrices, the chemical profiles of the APIs could be linked to the characteristic manufacturing process.

These results and the applicability of GC×GC as a core analytical technique could be confirmed with complementary sample preparation methods, TGA-MS analysis, and cross-verification with physically orthogonal FT-ICR MS methods. They showed the absence of thermal degradation, derivatization-induced side reaction, and significant shifts caused by the sample preparation. In addition, the beneficial universal properties of the GC×GC HR-ToF-MS and the limitations in the ABS analysis were demonstrated.

As a result, this study not only presents the most in-depth description of bituminosulfonates but also shows the reliability of its result.

1 Motivation

The human organism is continually exposed to a large number of chemical compounds, many of whose complexity and effects have not yet been elucidated [1–3]. Their chemical complexity is challenging to address, and biological effects are often not a result of single compounds but synergistic effects of multiple parts in the mixtures [4]. Some complex mixtures have been used as active pharmaceutical ingredients (APIs) in pharmaceutical products for centuries [5, 6]. While their pharmaceutical effects have been observed, investigated, and proven by non-clinical and clinical studies, the chemical composition is often only partially discovered. As a result, these APIs are not limited to single homo molecular ingredients but the entire complex mixture used for the drug formulation. A strict production process defines them, and changes may significantly affect the API's pharmaceutical effects and require a new authorization of the corresponding drug. An exact chemical description of the chemical identity is considered extremely difficult or even impossible and thus avoided [7, 8].

In 2014, the pharmaceutical class of non-biological complex drugs (NBCDs) was introduced [9]. This new class comprises complex drugs not directly obtained from living materials (biological drugs). Due to this definition, drugs allocatable to NBCDs are diverse in their origin, production, composition, and effect. Since the introduction of this class, publications have mainly dealt with the approval of follow-on products by establishing their biological and pharmacological equivalence to the originator [8]. Although this ensures statistically comparable pharmacological effects, this approach requires non-clinical and clinical studies to demonstrate these equivalences but does not aim for the same therapeutic effect for the individual [10].

The in-depth chemical description of the APIs that are allocatable to complex drugs like NBCDs would have multiple advantages [11]. The main benefit is the identification of characteristic compounds. Combined with pharmacological studies, these could elucidate potential pharmaceutical active structures (pharmacophores) and biological mechanisms. As a result, optimized drugs could be designed. In addition, the identity of characteristic compounds allows the development of suitable quality control methods and the chemical distinction from copycat and competitor products. Embedded in the discussion of their authorization, it would enable different approaches without extensive pharmaceutical studies, which allows the interchangeability of originator and follow-on products.

However, the physicochemical complexity of APIs from complex drugs, particularly NBCDs, requires comprehensive and high-resolving analytical methods to address the matrix. In addition, their heterogeneity requires individual approaches to target the unique chemical profile. The possible analytical techniques described by pharmacopoeias or used in the regulated pharmaceutical industry are limited [12] since new state-of-the-art techniques with non-targeted analytical approaches hardly align with current pharmaceutical guidelines and validation procedures.

This study will present a strategy for the chemical characterization of complex APIs, demonstrated with bituminosulfonates, used for multiple drugs today tentatively allocatable to NBCDs. In addition, it will address the particularities of performing such a non-targeted analysis in a pharmaceutical environment.

2 Introduction

2.1 Analysis and Authorization of Drugs in the Regulated Pharmaceutical Industry

Nowadays, there is a wide variety of pharmaceuticals that differ not only in their use but also in their chemical composition, origin, and complexity [13]. Most drugs are classified as “small molecules” and are defined by their homo-molecular active pharmaceutical ingredients (APIs). This pharmaceutical class is the subject of most pharmaceutical guidelines, such as those intended for the chemical analysis of pharmaceuticals.

The chemical composition of a drug is of utmost importance. It determines the physicochemical properties as well as the pharmaceutical effects. Therefore, it is defined and described in the specification of the drug and determined in the early development phase [14]. For the specification, multiple analytical techniques are applied to describe the explicit chemical composition of the drug, exemplary of the combination of chromatographic (retention time) with spectroscopic (*e.g.*, UV-/VIS spectrum) or mass spectrometric information (*m/z* ratio) [14]. These methods should be established with standard materials. The consistency of the drug is then checked using simpler conventional quality control (QC) methods [15]. This guarantees the safety, efficacy, and quality investigated in the non-clinical and clinical trials. Therefore, the known chemical identity of a drug and its pharmaceutical analyses are critical to establish a chemical profile or prove its consistency.

As a result, the quality requirements of analytical procedures are strict and defined in the regulations of the European Pharmacopoeia (Ph. Eur.). Besides these regulations, the guidelines of the International Council for Harmonization of Technical Requirements for Pharmaceuticals for Human Use (ICH) are the most important for the pharmaceutical environment [16]. Although their guidelines are not mandatory, they are or will be implemented in the United States, European Union, and Japan pharmacopoeias. The 2023 updated ICH quality guideline Q2 addresses the validation of analytical methods [17]. Validation must be performed before the pharmaceutical industry can implement an analytical method. Its characteristics (accuracy, precision, and specificity) vary between analytical procedures depending on the intended kind of analysis (*e.g.*, identity, quantitative, or limit testing). In addition to the Q2 required analytical parameters, analytical methods should be easy to implement and allow a high sample throughput. These results in typical chromatographic techniques such as thin layer, gas (GC), and liquid chromatography (LC). They are often hyphenated to a spectroscopic detector (*e.g.*, UV/VIS) or a mass spectrometer [12]. To establish their validity, authentic or even certified reference material or standards are at least highly recommended and almost without alternatives if quantification or limited testing is desired for the QC method. Besides the standard materials, the targets of the method (APIs or critical compounds like residual solvents) must be resolvable by the analytical method (*e.g.*, resolution factor > 1). Both can be realized for the chemically defined APIs of small molecule drugs but are hard to achieve for APIs for which the exact chemical identity cannot be determined.

This is the case for the pharmaceutical classes of biological and non-biological complex drugs (NBCDs). In contrast to the chemically defined APIs of small molecule drugs, the APIs of these two classes are defined as too complex to be fully described by physicochemical means [18]. Biologicals originate from living cells or organisms without or with limited chemical modification (*e.g.*, extractions) [19]. The last group of NBCDs includes many remaining drugs, which are neither classified as small molecules nor biologicals [9]. They do not originate from living cells or organisms and do not comprise single homo-molecular APIs but are often composed of closely related structures. As a result, the pharmaceutical effects of biologicals and NBCDs are not caused by a single molecular structure but

synergistic effects of the complex mixture. Typical APIs for the three pharmaceutical classes are depicted in (Figure 2.1).

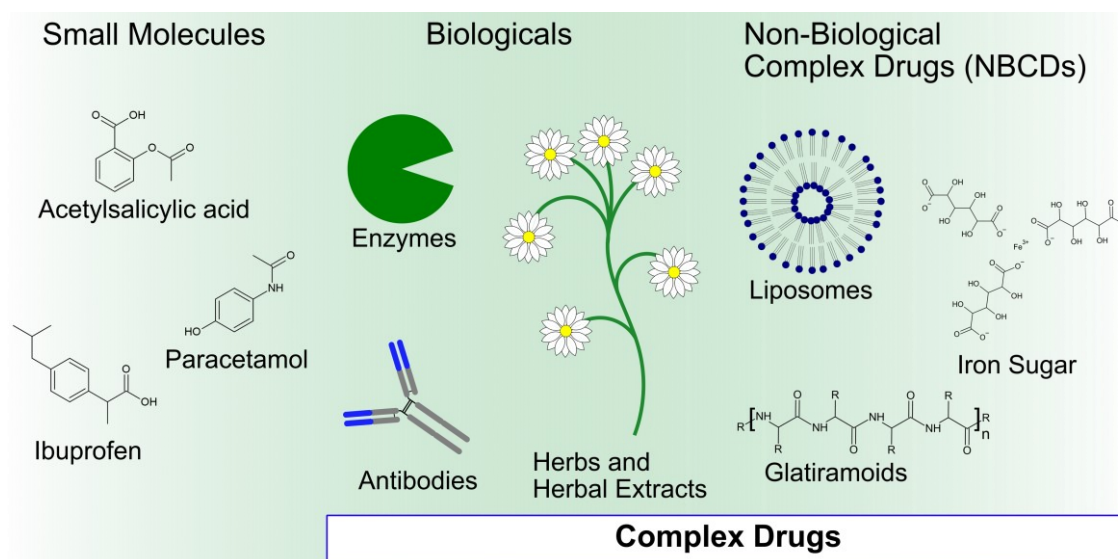


Figure 2.1 Classification of drugs and their active pharmaceutical ingredients (APIs). From left to right with increasing chemical complexity.

The manufacturing process and the starting materials of the APIs of complex drugs are important to ensure their constant chemical composition, quality, and pharmaceutical effects. Based on similarities within all three pharmaceutical classes, typical but not mandatory properties for the pharmaceutical classes can be derived (Table 1).

Table 1 Typical properties of active pharmaceutical ingredients (APIs) allocable to small molecules, biologicals, and non-biological complex drugs (NBCDs) (altered from [20]).

Properties	Small Molecule	Biologicals	NBCDs
Identity	Chemical Identity	Starting materials and manufacturing process	
Molecular weight	Low (<500 Da)	High (range 5-900 kDa)	
Structure	Simple, well-defined	Complex, heterogeneous, defined by the manufacturing process	
Manufacturing	Chemical synthesis	Produced in living cells or organisms	Synthetic technologies (incl. nanotech)
Copy Characteristics (Follow-on or copycat products)	Identical copies can be made	Impossible to ensure an exact copy version	

In particular, the differences in determining the identity of the API (chemical identity vs manufacturing process) and the resulting physicochemical complexities hinder the application of the same pharmaceutical guidelines. NBCDs especially have difficulties aligning with the current guidelines for chemical analysis and authorization [8, 21].

2.2 The Special Case of Non-Biological Complex Drugs

The class NBCD was introduced in 2014 but has not been implemented in the Ph. Eur. until now [22, 23]. However, numerous traditional and modern drugs with different origins, production processes, and physicochemical and pharmaceutical properties are allocable to NBCDs. Typically, glatiramoids, iron carbohydrates, and liposomal formulations are associated with this class [9, 24]. Many drugs, allocable as NBCDs, were released and authorized decades or centuries ago. At these times, the requirements for drug releases were different, and drugs were often only locally (in one country or region) authorized [25, 26]. Since their release, the pharmaceutical effects have often been well observed in patient reports and clinical studies. However, because of their physicochemical complexity, the chemical composition is often only partially investigated, and the entire mixture is defined as API instead of individual molecules. This hinders using the same QC methods, which are applied for small molecule drugs [8, 22, 23], but alternative methods must follow the same guidelines [27, 28].

As a consequence of these analytical and regulatory challenges, the QC methods for complex drugs are often limited to simple physicochemical parameters (*e.g.*, water content, dry substance content) requirements for the pharmaceutical application (*e.g.*, sterility, pyrogen-free). Chemical analyses are limited to a few marker compounds. These markers are usually selected because of their stability and measurability via typical analytical methods like resolved peaks in chromatographic methods [29]. Therefore, defining the complex mixture with a small number of targeted marker compounds that are not necessarily characteristic is not feasible [11]. Alternative validation procedures like cross-validation require preexisting validated methods and profound information on the chemical composition and properties and are often limited to the comparison of two analytical sites [30]. As a result, the standard analytical techniques and the validation guidelines are hard to align with the physicochemical complexity of NBCDs. More general guidelines are applicable for more advanced and powerful analytical techniques and complex matrices. For example, the recently passed ICH guideline Q14 demands the performance of a risk assessment during the method development [31]. Although this approach cannot replace the method validation, it addresses the analytical procedure's potential critical parameters and limits, reducing the gap to a validated method, particularly for complex drugs.

Due to the novelty of NBCDs, there are multiple other aspects for the authorities and manufacturers that require a special or adapted procedure. The most discussed topic deals with the authorization of follow-on products [8, 20, 23]. This debate focuses on biological and pharmacological equivalence, which should show a comparable therapeutical equivalence (biosimilarity) [21]. However, this biosimilarity pathway requires pharmaceutical studies and comparisons to the originator. If they were not linked, the desired follow-on product would be dealt with as a new drug. Due to the analytical challenges, this discussion only briefly mentions the chemical composition and the generic authorization process (same chemical composition). Even a hybrid approach with a complementary chemical and pharmacological comparison often does not decrease the number of studies [32].

Besides the required information for the authorization process, the follow-on products differ in their medical handling [20]. Only generics are interchangeable with each other or the originator and have the same therapeutic effect (same pharmacological and -kinetical effects for individual patients). Although biosimilars prove their biosimilarity, they are not therapeutically equal [10, 23]. However, the lack of detailed chemical descriptions at the molecular level makes it impossible to identify the differences and link them to the drug responses [33]. This shows the gap between biosimilars in medical use and the

need for in-depth chemical descriptions of complex drugs to elucidate the pharmaceutical effects and recreate the same therapeutic effect.

The complexity of the APIs in NBCDs and their diversity within the class make it difficult to follow the same analytical approaches as would be the case for the class of small molecules, and this also applies to the general recommendations and guidelines. Therefore, individual methods must be designed for single or families of APIs to access the complex chemical composition.

2.3 The Complex Active Pharmaceutical Ingredients Bituminosulfonates

The described challenges regarding the complex chemical composition and developing a suitable analytical procedure are present for bituminosulfonates produced by the Ichthyol-Gesellschaft Cordes, Hermann & Co. (GmbH & Co.) KG [34]. Bituminosulfonates are often used as APIs (e.g., Ichthyol® LIGHT and Ichthyol® DARK) in ointments to treat various skin conditions [35, 36]. These ointments and other pharmaceutical products are tentatively allocable to NBCDs. These APIs represent the described analytical challenges of complex drugs. In addition, like other complex APIs, bituminosulfonates are defined by their manufacturing process (Figure 2.2) and controlled by simple physicochemical parameters.

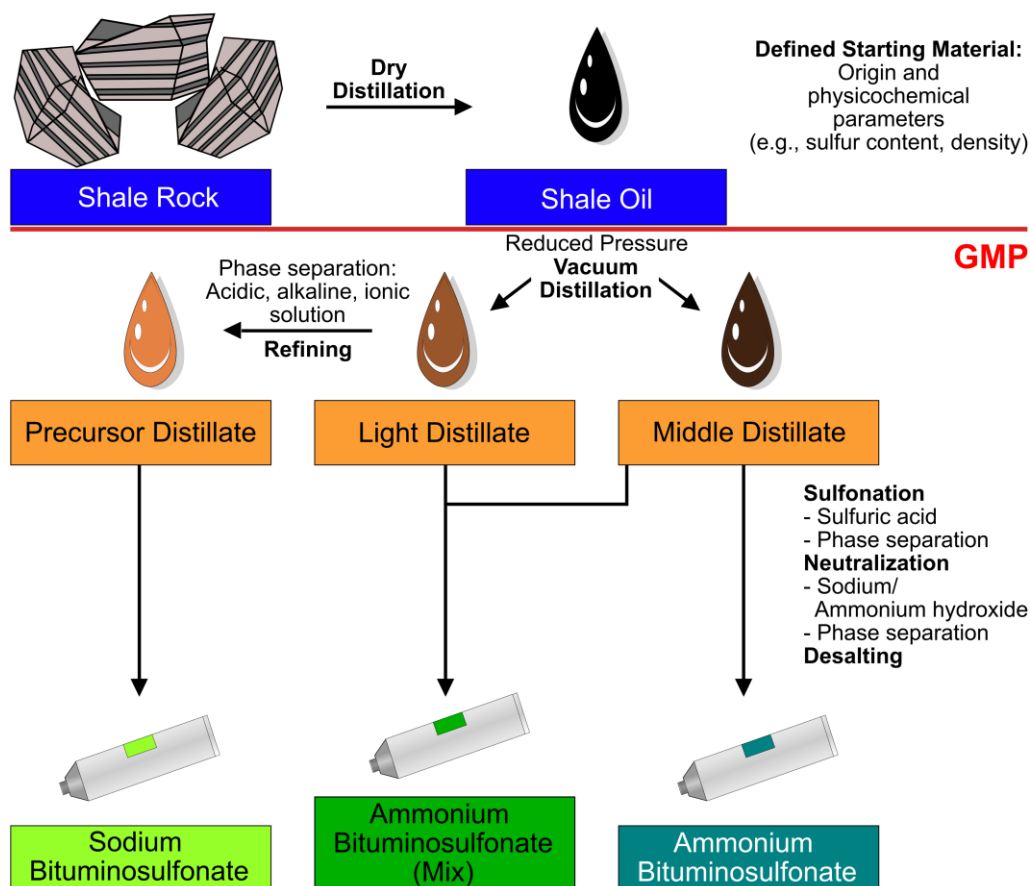


Figure 2.2 Production process of the Bituminosulfonates. From the dry distillation of kerogen-rich rocks to the three active pharmaceutical ingredients (APIs) investigated in this study: sodium bituminosulfonate (SBS), ammonium bituminosulfonate of the combined distillation phases (ABS (mix)), and the middle distillate (ABS). A red line visualizes the good manufacturing practice (GMP) regulated part of the manufacturing process.

Introduction

For the production process, kerogen-containing rocks are mined (*e.g.*, organic matter-rich carbonates such as limestone and dolomite). These rocks are crushed and undergo a dry distillation process. This thermally cracks the sulfur-rich macromolecular kerogen structures into smaller compounds, resulting in shale oil. All the following steps are subject to GMP and define the produced API. In the first regulated step, the shale oil is batch-wise distilled under reduced pressure to obtain two fractions (light and middle distillate) and to separate them from the low- and non-volatile residuals. The distillates produce different APIs. The light distillate is refined with diluted acidic, alkaline, and clay solution for the light product. The refined product (precursor distillate) is sulfonated with concentrated sulfuric acid under controlled temperature conditions, neutralized with sodium hydroxide, separated from the remaining oil phase, and desalted. The formed API is called “sodium bituminosulfonate (SBS)”. For the heavier product, called “ammonium bituminosulfonate (ABS)”, the middle distillate is directly sulfonated without active cooling and neutralized with ammonium hydroxide. The last API is similarly produced with the unified light and middle distillates as starting material. This API is also called “ammonium bituminosulfonate” but to avoid misunderstandings, it is referred to as “ABS (mix)”. For both ABS APIs, the complete phase separation (oil/water), as performed for SBS, is not possible.

Since the introduction of bituminosulfonates (as Ichthammol®) in 1882 [37] and medicinal products such as Ichtholan®, multiple non-clinical and clinical studies have been performed to investigate their therapeutic and biological effects and their related pharmaceutical products [36, 38–48]. Historically, they are used and recommended for treating numerous skin diseases like eczema and psoriasis [36, 38]. Nowadays, the research focuses on the pharmacological and biological mechanistic properties of bituminosulfonates at the cellular level. This research demonstrates *in vitro* and *in vivo* studies of the anti-inflammatory and skin protective effects of bituminosulfonates, which are linked to the recommended therapeutical areas [39, 40]. In addition, multiple studies show the antimicrobial effects of bituminosulfonates against gram-positive bacteria [41, 42]. In particular, the low minimal inhibitory concentration of SBS for staphylococcus aureus and its low potential for resistance development are thoroughly investigated [43]. Furthermore, systematic effects, like rosacea treatment, are investigated [44]. Besides the studies about the various biological effects, the tolerability of the APIs for topical and systemic administrations was also tested [45–48]. Some studies recommend a chemical characterization and the isolation of single compounds to perform more thorough investigations [40].

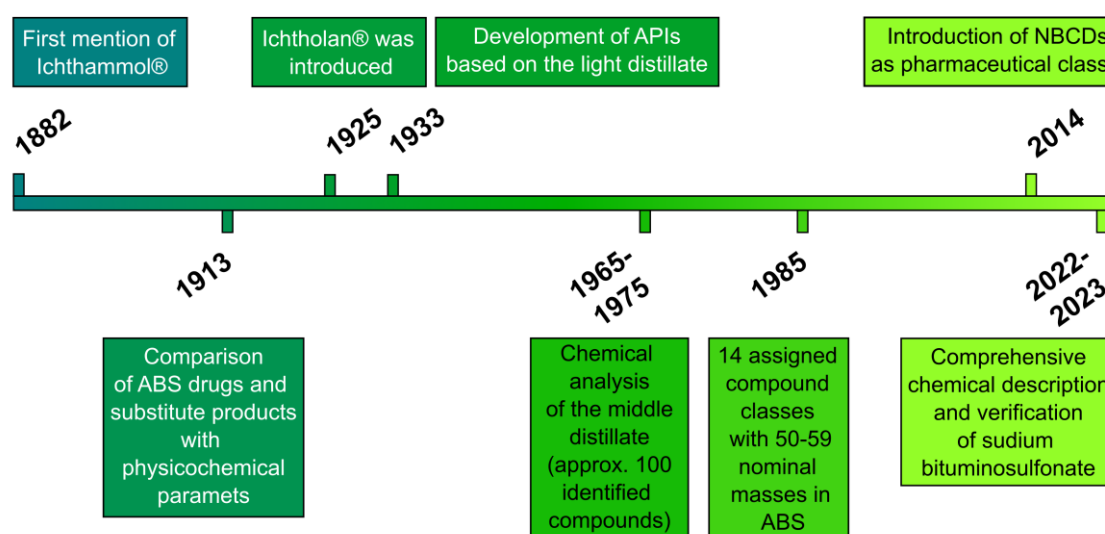


Figure 2.3 Key dates for the development of bituminosulfonates and their pharmaceutical products (on the top) and their most important chemical analysis studies (at the bottom).

Contrary to the numerous studies on efficacy, there are only a few studies on chemical identity. While the first studies from the early 20th century [49, 50] focused on physicochemical properties like dry residue and sulfur content, between the 1970s and 1980s the focus shifted to accessing compounds and compound classes of the distillates and ABS [34, 51–57]. Extensive sample treatment with numerous extraction steps enabled the identification of approximately 100 compounds in the middle distillate [51–57]. The distillate fractions were primarily measured, and compounds were identified via GC-MS. Even with the extensive sample treatment and fractionation, most of the matrix was not identified. The study situation for ABS is similarly limited [34]; the mixture was fractionated with ion exchange and thin-layer chromatography. Afterward, the fractions were derivatized and measured via GC-MS. Here, 14 sulfonated aromatic compound classes with different degrees of alkylation were identified. In total, 50-59 degrees of alkylation (unique molar masses) are reported [34]. However, most of them should be mixtures of different isomers.

These studies were limited by the physicochemical complexity of the matrices and the insufficiently developed analytical techniques. The shale oil origin results in a complex isomeric composition, and the natural sulfur content combined with the sulfonation process forms the isobaric complexity of the APIs. Furthermore, the surfactant properties of aromatic sulfonates and the water content are difficult for most sample preparation procedures [58–60].

Parts of the complex chemical composition is a result of the kerogen. For the kerogen used for the investigated bituminosulfonates, organic sediments from lagoons were enclosed by rocks during the formation of the Alps. In 200 Mio. years, the pressure and temperature combined with absence of oxygen caused the formation of kerogen [61, 62]. In the dry distillation, the large heterogenic kerogen structures break down in diverse and similar small molecules that results in a complex isomeric mixture.

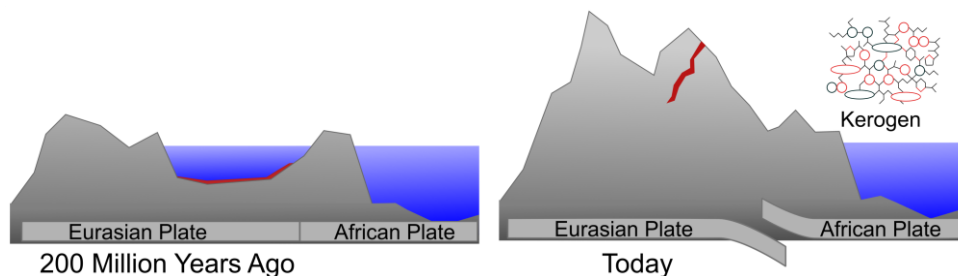


Figure 2. 4 Enclosure of sulfur-rich sediments during the formation of the Alps causing the creation of kerogen.

2.4 Analytical Approaches to Address Non-Biological Complex Drugs

Similar difficulties exist for other complex matrices measured in non-regulated environments. For bituminosulfonates, similarities can be found in petroleum, surfactant, food, environmental, and water analysis [3, 63, 64]. These matrices are often addressed with non-targeted methods. Here, a broad chemical space of the sample is analyzed instead of focusing on a narrow range of compounds with similar physicochemical properties. Non-targeted procedures also aim to cover a broad polarity range to access most of the matrix [65]. This contrasts today's approaches in pharmaceutical analysis, which describe a complex mixture with a small number of targeted compounds or simple physicochemical parameters [12].

Non-targeted studies are primarily qualitative and can be used to establish a characteristic chemical fingerprint [66]. This fingerprint can be utilized to verify the origin of a complex mixture and to

distinguish complex matrices of different origins, perform process control, and compare batches [67]. In addition, non-targeted methods identify characteristic markers in the fingerprint, which can be used as specific and representative targets for targeted QC methods. Quantitative statements that are based directly on a non-targeted analysis usually only reflect relative ratios or do not take into account the specific response of the selected detection system [3]. Therefore, non-targeted studies are often limited to qualitative or semi-quantitative statements [3, 65].

The applicability of the used analytical system strongly determines the reliability of such an analysis. Comprehensive, state-of-the-art analytical techniques are required to address complex matrices with non-targeted methods. These techniques are chosen based on the expected chemical space [68]. However, the unknown chemical composition of complex mixtures and the identity of individual compounds require identifying techniques with universal and selective information. Although spectroscopic methods like infrared and nuclear magnetic resonance spectroscopy meet these requirements, their resolution power for similar compounds in complex mixtures with several hundred or even thousands of compounds is insufficient for their identification [69]. The combination of a mass spectrometer hyphenated to a chromatographic technique (*e.g.*, LC or GC) addresses isomeric and isobaric complexities simultaneously. The mass spectrometric and chromatographic resolution defines the quality and the number of identifying properties (*e.g.*, accurate mass and retention time).

Comprehensive two-dimensional gas chromatography (GC×GC) shows a significantly higher resolution than one-dimensional GC, used in previous ABS studies, for volatile or semi-volatile matrices [70]. GC×GC generates a unique fingerprint that identifies characteristic markers in different matrices and is used to describe complex matrices in numerous areas, like petroleomics, metabolomics, proteomics, and for different environmental samples [71, 72]. Combined with a high-resolving mass spectrometer, the GC×GC can describe complex heteroatom-rich organic matrices, particularly sulfur species that have a low mass split to hydrocarbons ($^{12}\text{C}_3$ vs. $^{32}\text{SH}_4$, 3.4 mDa) [73]. Here, isomers are chromatographically resolved, and the accurate mass information identifies isobaric structures. Based on the accurate masses, characteristic fragmentation, and elution behavior, the compounds in the matrix can be classified instead of individually identified [74]. Besides these advantages, the technique is limited by the volatility (*e.g.*, size and polarity) and thermal stability of the matrix and the contained compounds. As a result, the molecular weight and the polarity are limiting factors for this technique.

Non-targeted measurements with comprehensive techniques like GC×GC generate complex data sets that require extensive data treatments and interpretations [75]. Besides the mostly more complicated hardware, this is one of the main challenges in validating non-targeted methods and barriers to implementation in regulated industries. Recently, there has been a growing focus on the reproducibility and regulation of non-targeted analytical procedures, leading to the publication of guidelines [76, 77]. The procedures include sampling, sample preparation, measurement with a comprehensive analytical technique, and data handling. Also, for GC×GC, initial efforts have been made to develop standardized methods for regulated environments and controlled matrices [78].

In summary, these non-targeted methods with state-of-the-art comprehensive analytical techniques could enable the analysis of complex drugs. Combined with the current focus on regulation and validation in this area, this shows an increase in the analytical capabilities and future applicability in regulated environments like the pharmaceutical industry. As a result, these techniques can narrow the gap in the chemical description of complex drugs and their APIs like bituminosulfonates.

3 Scope

The non-targeted analysis of APIs of the NBCD class presents both analytical and regulatory challenges and is inadequately covered by current Ph. Eur. or ICH guidelines. The APIs bituminosulfonates are examples of very complex chemical matrices that have not been sufficiently chemically characterized. This study aims to demonstrate and apply an analytical approach for the non-targeted analysis of these APIs, which are the basis for NBCDs that consider the specific requirements of a pharmaceutical environment. Therefore, the present study particularly addresses the following points (subjects I-III):

- S I: Development of an untargeted method for an in-depth and comprehensive chemical description of the APIs bituminosulfonates.

The main objective of this study is to uncover the detailed qualitative chemical profile of the different APIs, namely SBS, ABS (mix), and ABS. They present high isobaric and isomeric complexities that surpass the possibilities of common analytical techniques and procedures described by the Ph. Eur. In particular, small mass splits like $^{12}\text{C}_3$ vs. $^{32}\text{SH}_4$ are typical for organic sulfur-rich matrices and require HR-MS. Additionally, the APIs comprise a high number of isomeric compounds with unknown pharmacological modes of action and possible synergistic effects. As a result, targeting prominent or highly abundant individual marker compounds would not be sufficient, and the entire chemical profile must be considered. Based on these requirements and previous studies from the 1970-1980s, the study focuses on GC×GC coupled with a high-resolution time-of-flight mass spectrometer (HR-ToF-MS) as the primary analytical technique to address the bituminosulfonates.

- S II: Evaluation of the method under consideration of the non-targeted approach.

Besides the analytical challenges, a pharmaceutical analysis requires proven and reported reliability. This is usually addressed in the method validation, but today's approaches target individual compounds (*e.g.*, APIs, critical side products, or residuals). This is hardly applicable to non-targeted methods and state-of-the-art comprehensive analytical techniques, where hundreds or even thousands of compounds are analyzed, and neither the pharmaceutical nor the toxicological effects of single constituents are known. Only the effect of the entire mixture is observed in non-clinical and clinical studies. As a result, an alternative pathway must be designed that identifies potential risks, allows evaluation of their impact on the entire matrix, and is based on current alternative industrial guidelines.

- S III: Description of the defining manufacturing process and linking to the chemical identity.

Currently, complex APIs like bituminosulfonates are defined and controlled by their historical manufacturing process. In addition, neither the chemical composition of the starting materials, process intermediates, nor finished APIs are adequately described. However, it is necessary to assess and chemically describe the manufacturing process to expand the definition to a chemical level. In addition, it would provide a better understanding of the APIs. This would narrow the number of possible products and facilitate the identification of constituents. Additionally, the current definition limits the distinction between copycat products and originators to the documentation of the manufacturing process and hinders an evaluation of products from unknown sources. The chemical description of the API allows a retrospective identification of the APIs and their pharmaceutical products.

4 Methods & Instrumentations

4.1 High-Resolution Mass Spectrometry (HR-MS)

Mass spectrometry determines ions based on their mass-to-charge ratio (m/z). Depending on the specific properties of the mass spectrometer used, quantitative and qualitative properties like the abundance, exact mass, elemental composition, and isotopic ratios can be elucidated. For this, the sample is introduced by an inlet into an ion source, where the molecules are ionized and transferred forward in the mass analyzer. Here, the ions are resolved by their m/z value before their ion current is measured, amplified, and visualized in a mass spectrum (Figure 4.1) [79].

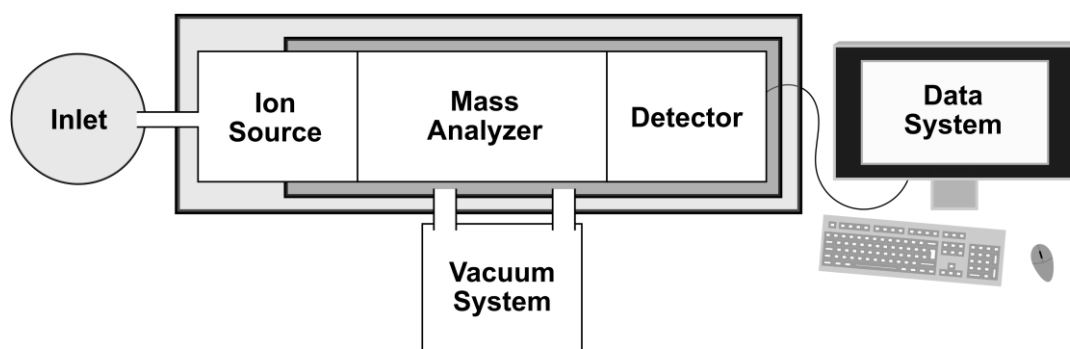


Figure 4.1 Basic composition of a mass spectrometer. Starting from the inlet system, which transfers the sample into the ion source, the molecules in the sample are ionized and directed into the mass analyzer, where they are separated by their mass-to-charge ratio (m/z). After the mass analyzer, the ions reach the detector, and the caused signal is transferred to the data acquisition system. The mass analyzer, detector, and ion source (depending on the ionization principle) operate at reduced pressure (“vacuum”) conditions.

The obtained mass resolution and accuracy define the performance of the mass spectrometer. The mass resolution is calculated with two ions that can be separated (at 10 or 50% full-width half maximum, depending on the definition) [80]. Here, the smaller m/z ratio is divided by their difference ($m_1/(m_2-m_1)$). Therefore, the mass resolution depends on the calculated m/z value, which should be stated beside the resolution. For mass accuracy, the difference between the exact mass and measured accurate mass is calculated, also referred to as mass error, and given in parts per million (ppm). The exact mass for every elemental composition, derived from the determined mass of an ion, is unique. The basis is the so-called mass defect of an isotope. Only ^{12}C is defined with an exact mass equal to its nominal mass (12.00000 Da). Every other isotope has either a positive or a negative deviation from the nominal mass (mass defect), depending on the relative nuclear binding energy compared to ^{12}C (Table 2).

Table 2 List of the isotopes with their atomic mass, mass defect, and natural abundances relevant to this study.

Isotope	Atomic mass [Da]	Mass defect [Da]	Relative abundance [%]
^1H	1.00783	0.00783	99.9885
^{12}C	12.00000	0.00000	98.93
^{14}N	14.00307	0.00307	99.632
^{16}O	15.99491	-0.00509	99.757
^{32}S	31.97207	-0.02793	94.93

These mass defects result in unique exact masses for each molecule of different elemental composition. Therefore, it is possible to calculate the sum formula of the ions with the measured m/z value. However, different elemental compositions ($C_cH_hN_nO_oS_sX_x$) differ only by a small mass. For example, $^{12}C_3$ and $^{32}SH_4$ species have the same nominal mass (isobars) and deviate in their exact mass only by 3.4 mDa. Here, the Fourier-transform MS techniques or Time-of-Flight MS (ToF-MS) are often used as high-resolving and accurate mass spectrometers that can resolve such low mass splits [81]. High-resolution mass spectrometry (HR-MS) resolves ions with small mass splits, combined with high mass accuracy, which enables the calculation of the elemental composition even in complex isobaric mixtures.

The sum formula can be used to group the data and to extract structural information. Here, the double-bond equivalent (DBE) as an integer value for the hydrogen deficiency of an organic compound can be calculated. This value summarizes the number of double bonds and cycles in the compound and indicates its aromaticity [82].

$$DBE = \frac{2*C+2+N-H-X}{2}$$

The information content can be increased by connecting different inlet systems to the mass spectrometer. As an alternative to direct sample introduction, chromatographic separation techniques such as LC or GC can be applied before MS analysis [83]. These techniques provide an additional chromatographic dimension that can address the isotopic composition (see section 4.2). However, the coupling with a chromatographic method limits the applicable ionization techniques. The ionization techniques themselves address different structures within the molecule and can be more general or selective. As a result, the sample introduction and ionization technique highly affect the measured chemical composition [84].

4.1.1 Electron Ionization Time-of-Flight Mass Spectrometry (EI ToF-MS)

The ToF mass spectrometer is a widely used mass analyzer. It is based on the flight time (t) of the ions to reach the detector at the end of the flight tube. This depends on the length of the flight tube (d), the applied electric potential (U), and the m/z value of the ions. Because the electric potential and the flight length remain constant, the flight time directly correlates with the m/z values [81].

$$t = \frac{d}{\sqrt{2U}} \sqrt{\frac{m}{z}}$$

The mass resolution of a ToF is limited by the flight length and affected by deviation of the ions' starting location, time, and kinetic energy before the acceleration into the flight tube. In addition, the ToF analyzer requires a defined starting point and time. Therefore, the ions must be pulsed into the flight tube in ion packages. For continuous ionization techniques like electron ionization (EI), an orthogonal accelerator can be used to fulfill this requirement. Here, a pulsed acceleration voltage directs ion packages in an orthogonal direction from the continuous ion beam into the mass analyzer. The orthogonal accelerator reduces ionization-induced spatial and velocity spread of the ions. As a result, it also increases the mass resolution of the mass spectrometer [85, 86].

Increasing the flight length and minimizing deviations of the kinetic energies in the ion packages are other options to improve the mass resolution. Here, a reflectron can be added to the system [81, 87]. It

consists of multiple regions with increasing electric potentials that form an electric field. The ions penetrate the electrical field at different depths, depending on their kinetic energy, and accelerate in the opposite direction. As a result, multi-reflecting HR-ToF-MS can reach mass resolution above 100,000 (Figure 4.2) [81, 83].

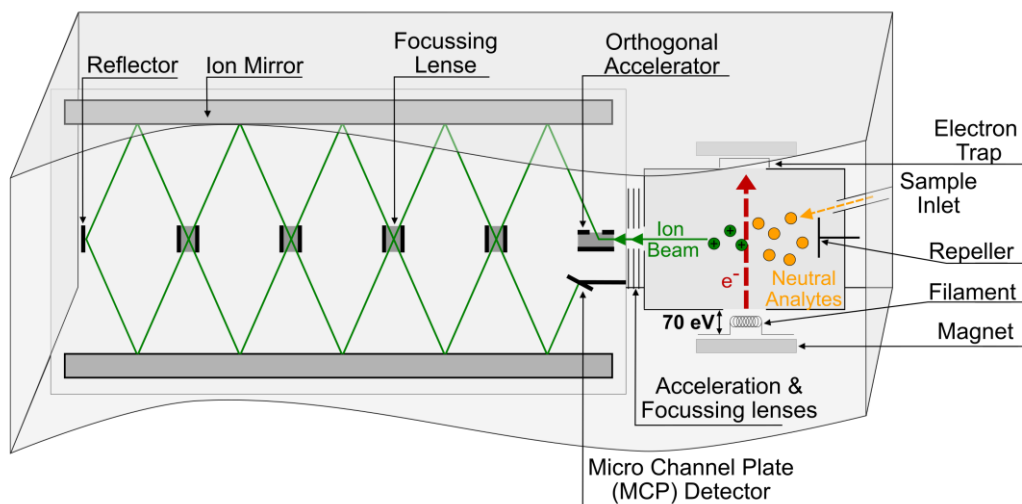


Figure 4.2 Schematic illustration of high-resolution multi-reflecting time-of-flight mass spectrometer with an electron ionization source.

One of the main advantages of the pulsed technique is the high acquisition rate of full mass spectra, which allows the hyphenation with fast eluting techniques with narrow peak widths like GC or GC×GC and covering their broad mass range [88]. In combination with gas phase analyses, the molecules are often ionized by electron ionization (EI) [84]. Depending on the kinetic energy of the applied electrons, EI is a so-called hard ionization technique that induces severe fragmentation of the molecular ion. The electron energy is standardized to 70 eV in most libraries; at this energy, the ionization efficiency reaches a plateau [89]. This is caused by the de Broglie wavelength of the electron, which is similar to the average bond length in an organic molecule and, therefore, in resonance [90]. At this energy, the fragmentation behavior is statistically reproducible and characteristic and allows the identification of core structural motives and functional groups. As a disadvantage of this technique, the high ionization energy may induce a complete fragmentation in unstable ions, which causes a missing molecular ion $[M]^+$. This hinders the calculation of the sum formula for the intact molecule and may hinder the identification of unknown compounds. However, for ions with known chemical composition, the EI fragmentation pattern can allow the distinction of isomers (Figure 4.3).

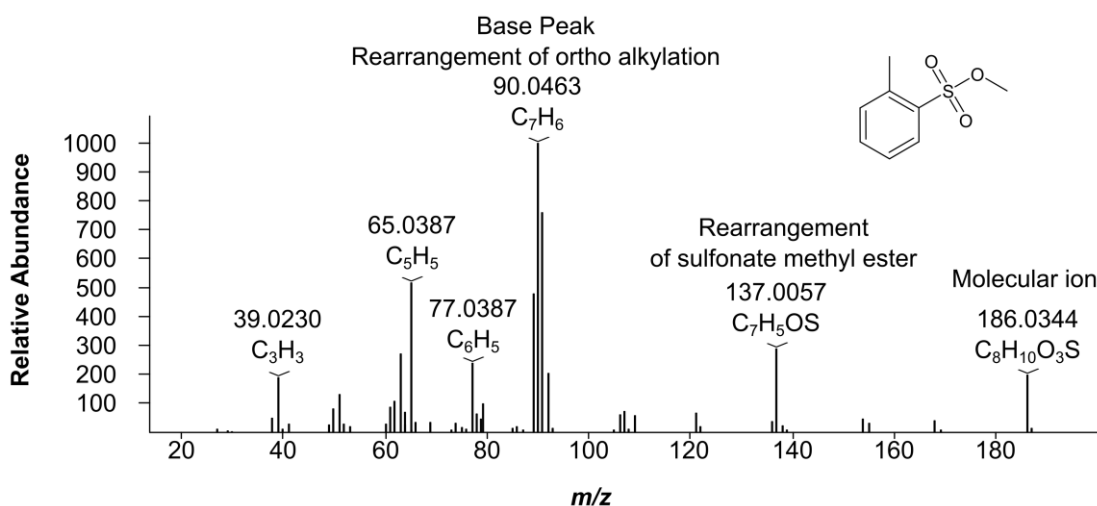


Figure 4.3 Measured electron ionization fragmentation pattern of ortho-toluenesulfonate methyl ester. Labeled with the accurate masses of the peaks and the calculated elemental composition of their neutral species.

4.1.2 Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FT-ICR MS)

Fourier-transform ion cyclotron resonance mass spectrometry (FT-ICR MS) achieves unbeaten resolving power and mass accuracy [81, 91]. Here, the m/z ratio of the ions is determined by their motion in a magnetic field in an ICR cell. This cyclotron motion results from the Lorentz force, which is formed by the movement of an ion in a magnetic field perpendicular to its magnetic field lines.

A commonly used cell is the infinity cell [91]. It consists of three pairs of electrodes: trapping plates that retain the ions inside the cell, excitation plates that form an alternating electrical field, which accelerates the ion packages in a spiral motion resulting in an increased cyclotron radius, and detection plates that record the ICR signal. The cell is placed in a superconducting magnet, creating a magnetic field mostly between 7-15 T.

This setup allows to record the cyclotron motion of the ion packages for a long period of time and thus record many 100,000 or millions of events as transients within typically 100 ms up to a few seconds. As a result, mass resolutions above 10^6 and accuracies in ppb regions are routinely reported for FT-ICR MS. However, the low acquisition rate hinders the hyphenation to fast eluting techniques. Although the FT-ICR MS principle is not limited in the detectable m/z -range, due to challenges with small ions, usually FT-ICR MS measurements record m/z values > 100 .

FT-ICR MS can be coupled with many different ion sources [91]. In recent years, atmospheric pressure ionization techniques like electrospray (ESI) and atmospheric pressure photoionization (APPI) have often been used. In contrast to EI, these two techniques are soft and cause minimal fragmentation, maintaining the molecular ion. However, while EI is relatively universal, ESI is more selective for polar acidic or alkaline compounds (depending on negative or positive ionization), and APPI ionizes electron-rich aromatic structures more strongly.

4.2 Applied Hyphenated Separation Techniques

4.2.1 Comprehensive Two-Dimensional Gas Chromatography (GC×GC)

GC is a commonly used technique for analyzing volatile and semi-volatile compounds. The sample is generally vaporized in the injector, which either operates at a constant temperature or is ramped after introducing the sample to the desired injection temperature. Afterward, some or all evolved gases are transferred to a separation column (split or splitless injection). The vaporized sample is guided by a carrier gas flow (mobile phase) from the injector through the column to the detector. While the mobile phase is an inert gas, usually helium, hydrogen, or nitrogen, the stationary phase is a film that coats the inner surface of the separation column. These mostly capillary columns achieve an excellent separation performance. The compounds in the sample diffuse into the stationary phase and are retained differently depending on the strength of the interaction. This separates the compounds and causes their different retention times. The retention time is also affected by the temperature and carrier gas flow of the measuring program. GC has a relatively high peak capacity and can, depending on the matrix, resolve up to dozens of compounds [92].

For more complex mixtures with several hundreds or even thousands of compounds, GC×GC can be applied [72]. Its two-dimensional separation is achieved by directing eluting compounds from one column (first dimension) through a modulator onto a much shorter second column (second dimension) (Figure 4.4).

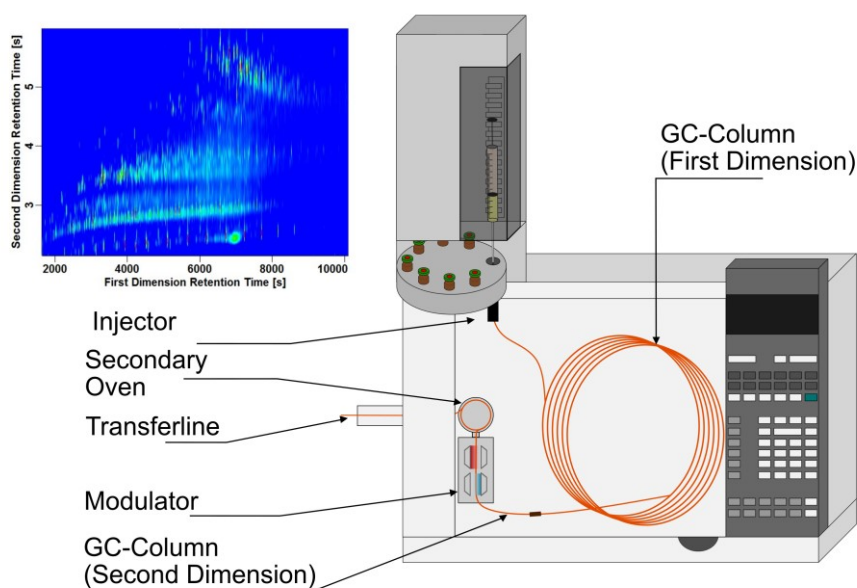


Figure 4.4 Illustration of the instrumental setup of a thermally modulated comprehensive two-dimensional gas chromatograph (GC×GC) with the measured contour plot of ammonium bituminosulfonate (ABS). The technique is often hyphenated with a mass spectrometer.

While the primary column's length and retention time are comparable to one-dimensional GC, the second is much shorter (e.g., 0.5-3 m), resulting in separation times of seconds. The short retention time in the second dimension enables maintaining most of the separation of the first dimension while creating a second separation dimension. This enables the combination of stationary phases, which resolve compounds by different physicochemical properties and interactions [70, 71]. Nowadays, combinations of different stationary phases are published for different matrices and targets [93, 94]. One of the most used combinations consists of a non-polar column in the first and a mid-polar or polar column in the

second dimension. As a result, compounds are resolved according to their volatility in the first and their polarity (e.g., π - π interactions) in the second dimension (Figure 4.5).

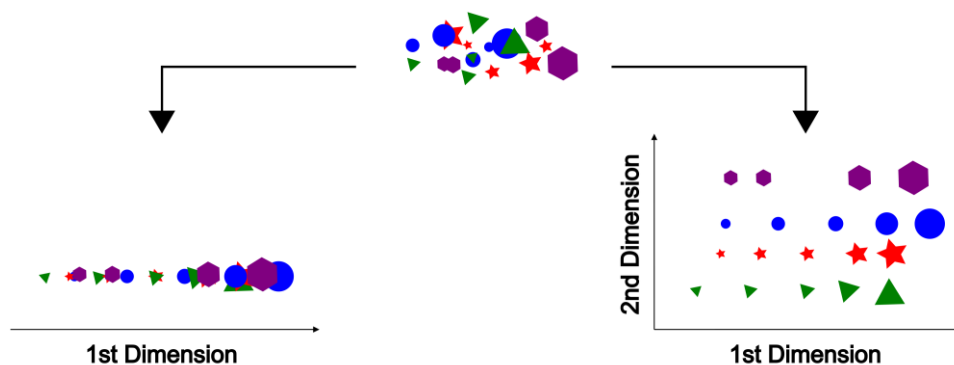


Figure 4.5 Comparison of a gas chromatography (GC) separation (left) and a comprehensive two-dimensional gas chromatograph (GC \times GC) (right). By applying two different mechanisms for separation, compounds that co-elute in the first dimension (depicted by size, usually volatility) can be separated in the second (depicted by color and shape, usually polarity).

A modulator is placed between both dimensions. It collects and directs the eluting compounds of the first into the second dimension. The modulator and the resulting second injection into a second column increase the peak capacity of GC \times GC compared to the GC. There are multiple modulators in the market with different physicochemical approaches. One commonly used modulator is the quad jet cryogenic modulator [95]. Here, the eluting compounds of the first dimension are fixed by a cryogen (cooled by either an electric chiller or liquid nitrogen) and released by a hot jet stream. This cycle repeats with a second pair of jets to build up a gate before the last hot stream “re-injects” the compounds in the second dimension (Figure 4.6).

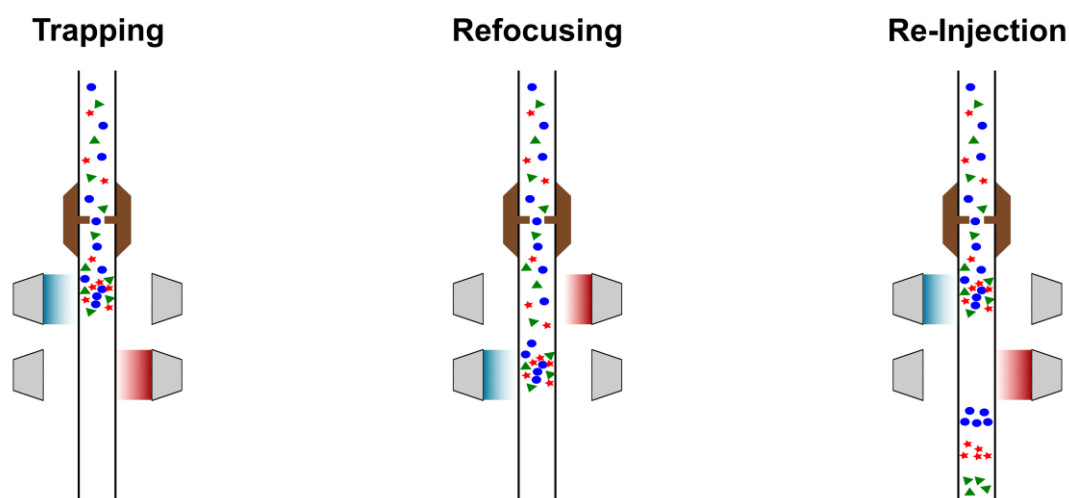


Figure 4.6 Illustration of the operation principle of a quad jet cryogenic modulator. The modulator traps analytes, eluting from the first dimension, by locally and rapidly reducing the temperature with a cold nitrogen stream of a cold jet (Trapping). It re-mobilizes them with a hot stream and traps them for a second time directly afterward to refocus the analytes (Refocusing). In the last step, the modulator releases the re-focused analytes in the second separation dimension with a second hot nitrogen stream (Re-Injection).

The time from reaching the modulator and the “re-injection” into the second dimension (modulation time) should be long enough to enable the elution of the compounds in the second dimension before the next refocused analytes are released, but also as small as possible to maintain the separation of the first

dimension. Practically, the modulation time should be 3-5 times smaller than the peak width of compounds in the first dimension, and the elution in the second dimension should take place in a few seconds. The cryogenic modulation results in a narrow peak width (approx. 50-250 ms), which is a

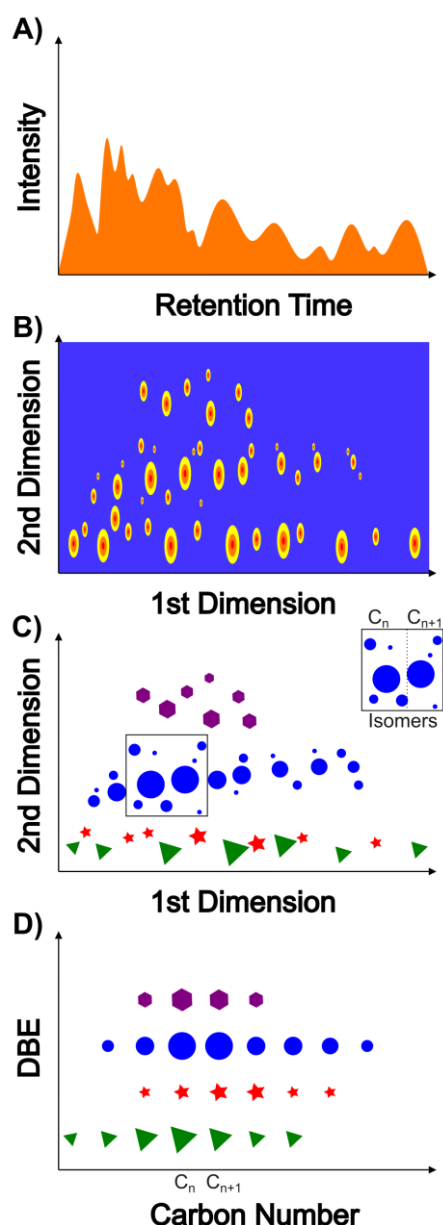


Figure 4.7 Data visualization of a comprehensive two-dimensional gas chromatography high-resolution time-of-flight mass spectrometry (GC×GC-HR-ToF-MS) data set. **A**: one-dimensional chromatogram. **B**: Illustration of a contour plot. **C**: Illustration of the peaks in a color-coded scatter plot with highlighted isomeric separation. **D**: Peaks aligned and isomers summed up in a double bound equivalent (DBE) vs. carbon number plot.

Compared to the resolution on the normal phase in GC×GC, the summed peaks in the DBE vs. carbon number plot are comparably aligned with increasing carbon numbers on the x-axis and aromaticity on

a direct consequence of the focusing and sharp “re-injection” of the thermal modulation. As a side effect, the narrow peak width increases the signal-to noise (S/N) ratio and the limit of detection (LOD) of the GC×GC [95].

GC×GC can be coupled to different mass analyzers and detectors. However, the number of suitable techniques is limited due to the narrow peak widths. A TOF mass spectrometer has a sufficient acquisition rate, and the hyphenation with an EI-HR-ToF-MS significantly increases the information [73, 88]. The detected peaks can be classified and identified with their accurate mass, EI-induced fragmentation pattern, and elution behavior. In addition, combining a chromatographic separation with high mass resolution and accuracy allows the resolution of mass splits that the individually used HR-ToF-MS does not separate. This has special importance for the 3.4 mDa mass split of $^{32}\text{SH}_4$ versus $^{12}\text{C}_3$ in sulfur-rich matrices [73].

Based on this complex information, the data must be rearranged to use and visualize the multidimensional data set. The detector records a signal after the second dimension, visualizing the time-resolved intensity (chromatogram) (Figure 4.7 A). Here, the known modulation time allows the rearranging of the chromatogram into a contour plot (Figure 4.7 B) [94]. With the underlying mass spectrometric information and the characteristic elution profile, the peaks can be classified and either color-coded in the contour plot or transferred in a scatter plot (Figure 4.7 C). These classes can be summarized in bar or pie plots to compare the relative abundance between different samples.

Another way of classifying peak information is primarily based on the mass spectrometric information of the derived sum formula and the calculated DBE values (Figure 4.7 D). The DBE versus carbon number plot is one of the standard visualizations in mass spectrometry [96]. It illustrates the carbon number range versus hydrogen deficiency of the measured compound. Usually, the plots are limited to a fixed number of heteroatom compositions, but color-coded plots with different elemental compositions are also possible.

the y-axis. This plot allows comparing the complex GC×GC HR-ToF-MS data set to other mass spectrometric methods. However, it does not depict the isomeric information of the GC×GC.

While the GC×GC EI HR-ToF-MS technique provides much qualitative information, enabling the classification and identification of unknown compounds, the quantification is challenging [97]. Although the EI is a relatively universal ionization method, compared to other ionization techniques, deviation of the response factors hinders a full quantification. Here, the core structure, alkylation, and functional groups influence the ionization efficiency. Therefore, standards or literature must be available to obtain quantitative information. As a result, the relative abundance of compounds is used to describe complex matrices semi-quantitatively.

In summary, the high peak capacity and broad chemical space make this technique ideal for qualitatively characterizing and describing complex volatile and semi-volatile matrices. In addition to obtaining the chemical fingerprint, the chemical information also enables the classification and identification of unknown compounds. Specifically, the behavior of compounds during the elution process allows for the determination of characteristic elution windows. Compounds with similar elution windows can be classified using mass spectrometric information, such as the fragmentation pattern and accurate masses (Figure 4.8).

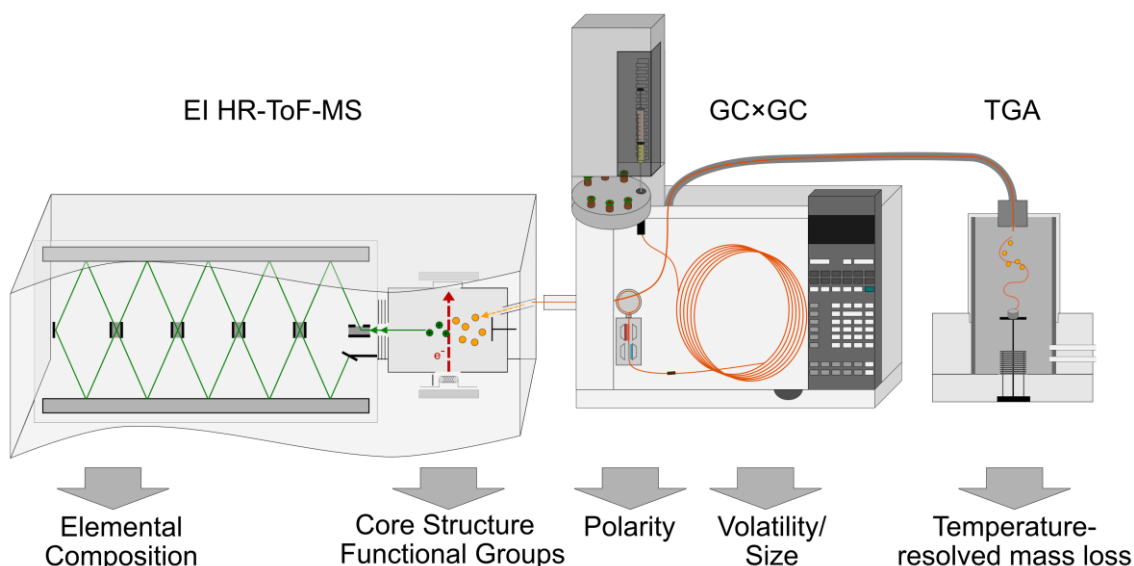


Figure 4.8 Illustration of a comprehensive two-dimensional gas chromatograph (GC×GC) and thermobalance (see 4.2.2) hyphenated with an electron ionization high-resolution time-of-flight mass spectrometer (EI-HR-ToF-MS) with the obtained information of the individual parts.

As a result, GC×GC is a valuable technique to address and characterize complex matrices with a wide variety of isomers and compound classes. Combined with a high-resolving mass spectrometer, it can also address complex isobaric mixtures. The technique is only limited by the vaporability of the matrix and, without suitable standard materials, in the quantitative information.

4.2.2 Thermogravimetric Analysis (TGA)

The thermogravimetric analysis (TGA) measures the mass loss of the sample as a function of the applied temperature and time [98]. Thereby, the sample is placed in a sample holder placed on an analytical

balance. The sample is heated by the surrounding oven under a controlled atmosphere (Figure 4.9). During the measurement the temperature of sample holder and the mass of the sample are monitored to access its thermal behavior.

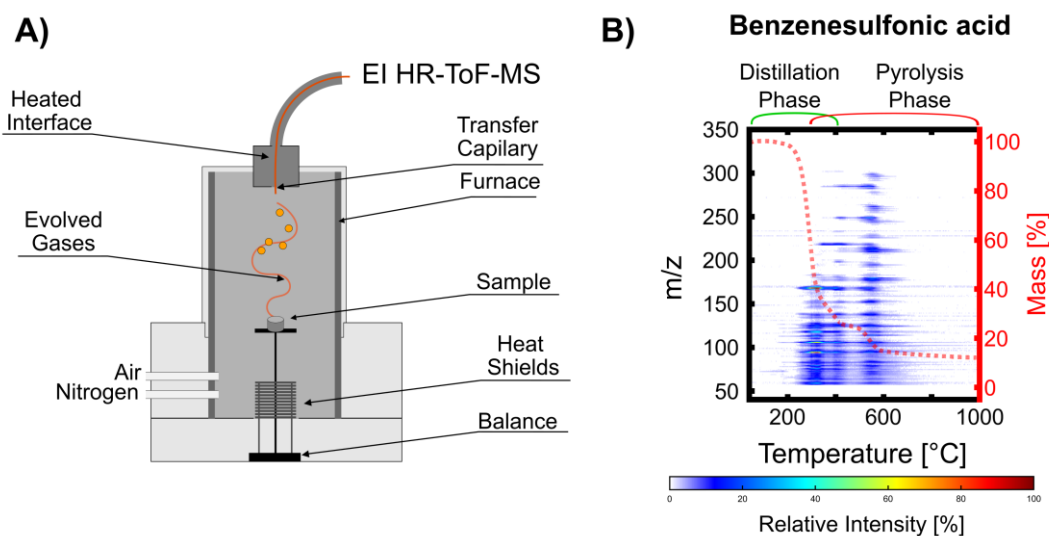


Figure 4.9 Setup and data visualization in thermogravimetric analysis (TGA). **A:** Labeled illustration of a thermobalance with a heated interface for the coupling with a mass spectrometer. **B:** Thermogram of benzenesulfonic acid, visualizing the temperature-resolved mass and the relative intensity of the detected mass-to-charge (m/z) ratios of the ionized evaporated gases analyzed by mass spectrometry.

The thermal behavior of a sample can be used to determine volatility and to evaluate the suitability of analytical methods such as GC. Typically, organic compounds evaporate below 400 °C (volatile and semi-volatile) or above this temperature thermally degrade. Thereby, depending on the structure either volatile compounds or char can be formed, which remains in the sample holder. As a result, thermograms are divided into a distillation and a pyrolysis phase [99]. However, unstable compounds can decompose at lower temperatures. Here, volatile thermal fragments and semi-volatile compounds would evolve simultaneously [100].

Besides the mass loss of the sample, the evolving gases can also be analyzed. Here, the thermobalance can be coupled to a detector, like a mass spectrometer. This set up analyzes the chemical composition of the evolved gases simultaneously with the temperature-resolved mass loss [101]. This allows a quantitative and qualitative investigation of the evolved gases and conclusions on the composition of the sample. The TGA visualizes thermal processes but is limited by the relatively low resolution of an evaporation process, which is comparable with a distillation process based on one theoretical plate [102].

4.3 Sample Preparation

Often, samples cannot be directly measured due to low concentrations, physicochemical properties (e.g., polarity, volatility), or matrix effects. Here, the sample must be modified to enable the applicability of a suitable analytical technique. In particular, aqueous samples are often complex to analyze and must be pre-treated before the analysis [58, 103].

4.3.1 Extraction

Water can disturb analytical procedures like GC and is often removed by extracting target compounds [103]. One method for aqueous samples is solid phase extraction (SPE) [104, 105]. The SPE uses cartridges packed with extraction sorbents to trap and remove the target compounds from the matrix. In the procedure, the sorbents are conditioned and equilibrated before the sample is rinsed through the cartridge to trap the target compounds. Afterward, the sorbent is washed before the trapped compounds are eluted by a suitable solvent.

The extraction sorbent and the elution solvent depend on the target compounds and must be chosen accordingly. The sorbents can address a broad range of compounds (*e.g.*, hydrophilic-lipophilic balance (HLB)) or target ionized compounds (*e.g.*, ion exchanger). For example, SPEs are often used to extract acids and their salts with a strong (SAX) or weak anion exchanger (WAX). Depending on the affinity of the targets for complementary charged ions, the sorbents can be loaded permanently (strong ion exchangers) or pH-dependently (weak ion exchangers). The affinity to the sorbent material affects the trapping and elution efficiency. In particular, a strong ion exchanger could hinder the elution of the trapped compounds.

An alternative sample preparation method is the liquid/liquid (L/L) extraction [105]. Here, two immiscible solvents are used, like water and an organic solvent (*e.g.*, dichloromethane, *n*-cyclo-hexane, or toluene), to separate the compounds in the matrix. The solvents form two phases, in which the compounds are distributed according to their affinity. To evaluate the distribution in the chemical equivalence of the compounds between a non-polar and a polar solvent, the octanol/water coefficient can be used. However, the equilibrium can be influenced by additives (*e.g.*, salts) or physicochemical properties (*e.g.*, pH value). In addition, the sample can show a significant influence and even form a mixed phase, particularly salts or detergents, like sulfonates, show this effect.

Both sample preparation methods, SPE and L/L extraction, are primarily designed to concentrate and extract target compounds but can also be used to cover a broad chemical range. However, they can shift the sample's chemical composition depending on the trapping and elution efficiency or the chemical equilibrium of single compounds to the used organic solvent.

4.3.2 Derivatization

Some analytical methods require specific physicochemical properties to be applicable, *e.g.*, the volatility of the matrix for GC analyses. However, some non-volatile polar or thermal sensitive compounds that cannot be measured directly via gas phase analysis, can be chemically modified to lower their boiling point below the decomposition temperature and enable the measurement via GC techniques [106].

Different derivatization reactions and agents can be used to reduce the polarity and increase the volatility. Methylation is a frequently used derivatization reaction. Here, a methyl group exchanges a hydrogen atom of a Brønsted acid (*e.g.*, alcohol, carboxylic acids) or reacts with a Lewis base (*e.g.*, amine) in a nucleophilic substitution. The required polarity and the groups that can be methylated

depend on the strength of the derivatization agent. Typical methylation agents are diazomethane and its derivatives like trimethylsilyl diazomethane (TMSDAM) [107].

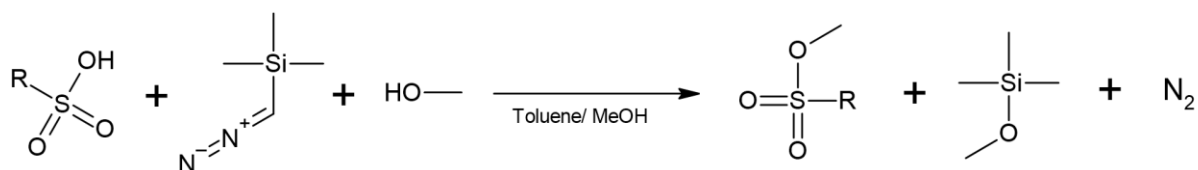


Figure 4.10 Chemical equation of the reaction of sulfonic acids with trimethylsilyl diazomethane (TMSDAM). Derived from the reaction with carboxylic acids [107].

However, derivatization agents, like diazomethanes, demand the absence of water and need time to react with the sample. Alternatively, tetramethylammonium hydroxide (TMAH) can methylate compounds in aqueous matrices during a hot injection. In the injector, the TMAH transfers one methyl group to the functional group, causing the evaporation and injection of the derivatized compounds and trimethylamine on the GC column [108].

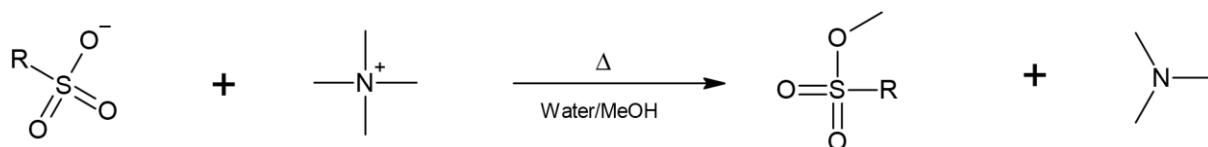


Figure 4.11 Chemical equation of the reaction of sulfonates with tetramethylammonium hydroxide (TMAH). Derived from the reaction with carboxylic acids [108].

The substantial reactivity causing the methylation of polar functional groups, the solubility in various solvents, the tolerance of an aqueous sample, and the absence of a necessary previous extraction make the derivatization agent TMAH applicable to many different matrices and for various compounds. Besides methylation, TMAH can also introduce side reactions. The most prominent is thermochemolysis, which cleaves esters or other functional groups that can be hydrolyzed [108]. In addition, the fast reaction time limits the derivatization efficiencies of steric-hindered or low-reactive compounds. Therefore, the derivatized matrix may not directly reflect the original chemical composition [106].

Although derivatization agents change the analytes into a measurable form, they can induce side reactions and shifts in the chemical composition. Therefore, the derivatization agent must be chosen based on the measured matrix. This minimizes unwanted reactions during the derivatization and facilitates the interpretation of the measured derivatized sample.

5 Results & Discussion

The order of the objectives listed in the scope (S. I-III) reflects the priorities within this study. However, they are intertwined and hard to address individually. Therefore, a holistic approach considering all objectives must be used. For new procedures or applications, one of the most used tools in regulated industries is the risk assessment [109]. It identifies possible limitations and interferences and is the basis for evaluating the reliability of the analytical method [31]. Also, in the pharmaceutical industry, risk assessment is a standard method to identify critical steps or parameters in a process [110]. The recently revised Q2 and newly implemented Q14 guidelines of the ICH also recommend a risk assessment for analytical procedures [17, 31]. Based on these guidelines, the risk assessment is applied before the method development with information on the investigated matrix, the target compounds, and the applied analytical technique. Missing data for critical parameters should be addressed via complementary experiments, and the method should be adjusted accordingly. For this study, with a complex matrix with an unknown chemical composition, a comprehensive state-of-the-art analytical technique, a non-targeted analysis approach, and an unclear validation procedure, a risk assessment-based method development addresses at least the objectives S I and II simultaneously, considers broad industrial guidelines, and represents the starting point.

5.1 Risk Assessment-Based Analytical Workflow

5.1.1 Identification of Critical Parameters via Risk Assessment

To perform a risk assessment for the analysis of bituminosulfonates, the analytical technique and the targets must be chosen beforehand. Based on the promising but very limited results of the previous GC-MS studies, this work focuses on the non-targeted analysis via GC×GC HR-ToF-MS. As a consequence of the non-targeted approach, the entire chemical space of the bituminosulfonates must be considered and cannot be limited to sulfonated constituents or even to the reported approx. 59 sulfonated nominal masses from Koch et al. [34]. Based on these conditions and inspired by the ICH guidelines, the study begins by visualizing the parameters that are directly or indirectly applied to the bituminosulfonates for analysis by GC×GC HR-ToF-MS (Figure 5.1).

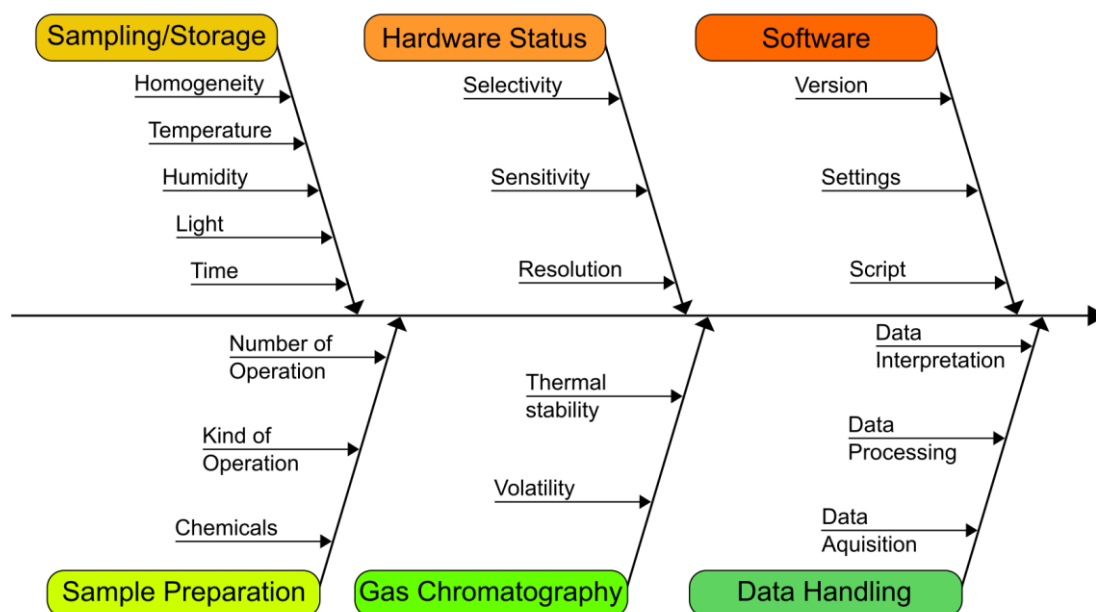


Figure 5.1 Ishikawa diagram of influencing parameters for a qualitative analysis of bituminosulfonates via comprehensive two-dimensional gas chromatograph (GC×GC-HR-ToF-MS) [111]. They are divided into general (sampling/storage, hardware status, and software above the arrow) and operational parameters (sample preparation, gas chromatographic method, and data handling below the arrow).

The visualization shows the numerous parameters in an analytical procedure that could affect its outcome. Many of these characteristics are applicable in various contexts. However, the relevance of these critical parameters may vary depending on the matrix, targets, and analytical technique. As a consequence, the available information on the matrix and analytical system is used to evaluate their potential impact and likelihood in the next step. However, such an evaluation requires extensive knowledge of the properties of the matrix, which is not the case for NBCDs and, in particular, bituminosulfonates. For example, none of the measured values can be compared with a target value (e.g., authentic standard material). Therefore, assessing the criticality of the listed parameters becomes difficult, and additional experiments must be included. Furthermore, because of the missing authentic standard materials, the study and the risk assessment focus on the qualitative information and relative compositions of the bituminosulfonates.

Examples of critical parameters for GC analysis of a low or non-volatile matrix, like bituminosulfonates, are the sample preparation, as well as volatility and thermal stability at the injection temperature. These parameters can be addressed individually with suitable complementary reference methods. Additionally, the combined effects of the entire analysis can be evaluated with an orthogonal analytical technique in a holistic cross-verification approach, which also assesses the covered chemical space. Based on the advantages of these both approaches and the aim of this study to demonstrate the reliability of the results, the method development includes the investigation of the three exemplarily but typical parameters of sample preparation, vaporability, and thermal stability, and it is concluded by a cross-verification with an orthogonal technique (Figure 5.2).

5.1.2 Overall Study Design

The number of critical parameters combined with the lack of information on the physicochemical properties of bituminosulfonates or comparable matrices shows that complementary measurements and experiments are necessary to develop and verify an appropriate analytical method (Figure 5.2).

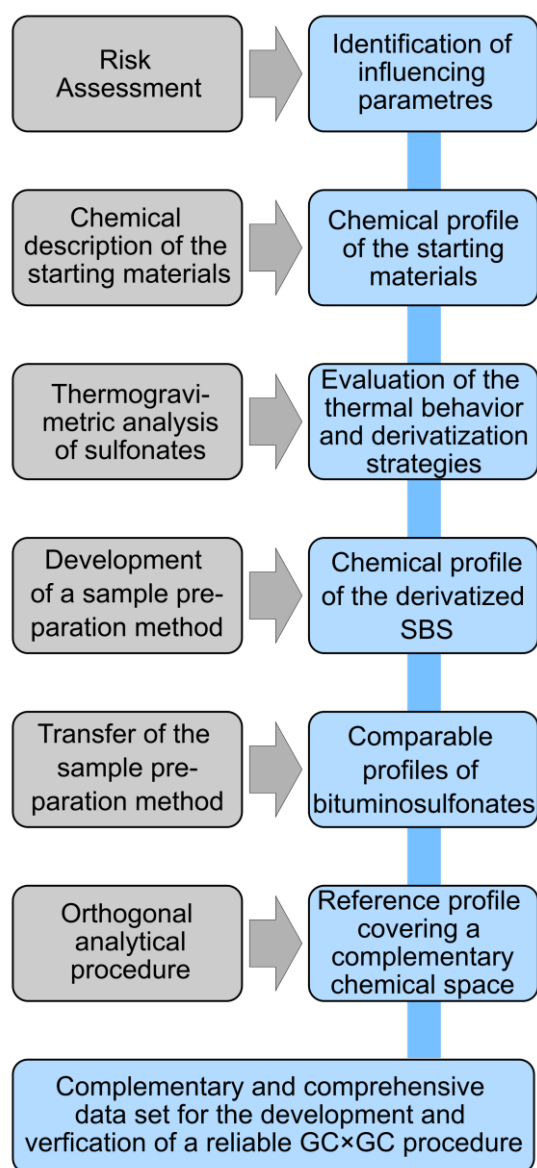


Figure 5.2 Analytical workflow for the development and verification of a comprehensive two-dimensional gas chromatograph (GC×GC) analysis. Including the performed step (left/gray) and the obtained information (right/blue).

While the information on matrices similar to the API is limited, there are many comparable matrices with the shale oil or its distillates that are also reliably measured via GC×GC [93, 97, 112, 113]. Therefore, the description of the starting materials and process intermediates of bituminosulfonates by GC×GC is an established procedure. In addition, their chemical characterization combined with the knowledge of the manufacturing process gives a starting point for the analysis of the APIs. It limits the mass spectrometric challenges by setting the isobaric complexity to the measured and added elements and gives core structural motives and isomeric information of the precursor compounds. Furthermore, this directly links the APIs to their defining manufacturing process (S. III) in order to identify compounds and related side reactions like polymerization and oxidation [114, 115]. Starting with the shale oil, the GMP-regulated manufacturing process is followed to describe the chemical composition of the distillate fractions and the differences between the light distillate and the refined precursor.

Because of the missing literature on the thermal behavior of bituminosulfonates and the limited literature on sulfonates [100], the critical parameters of vaporability and thermal stability of the API can only be evaluated with a complementary method. Here, the TGA-MS can evaluate the quantitative amenability of GC×GC for sulfonates. However, due to the completely unknown thermal behavior and the expected complexity of the APIs combined with the low resolution of the TGA, commercially available

single compounds are the basis for discussing the thermal behavior of the complex matrix. The previous analysis of bituminosulfonate and other GC-MS analyses of sulfonates often applied methylation of the sulfonates [34, 59, 60]. The reference compounds include sodium sulfonate, sulfonic acid, and sulfonate methyl ester (SME) to evaluate their approach.

If the TGA-MS confirms the amenability of SMEs for GC×GC, multiple methylation-based sample preparations should be performed for SBS, compared to each other, and linked with the precursor

distillate. This addresses the third critical parameter, sample preparation, and checks for its effects on the obtained chemical profile. To improve the efficiency of the step, it is performed for SBS, and afterward, the best method is transferred to other APIs.

In the last step, FT-ICR MS with different orthogonal ionization methods is included to prove the method's reliability and to address the suitability of its covered chemical space. The critical comparison of the data sets from both techniques is used to verify the results and to decide if the developed GC×GC analytical procedure is suitable for the detailed chemical characterization of the three bituminosulfonates.

This workflow aims to develop an analytical procedure for the comprehensive chemical analysis of bituminosulfonates targeting their qualitative chemical composition. In addition, it also serves as a basis for describing the manufacturing process based on the composition of the isolatable matrices. Particular consideration is given to the parameters set out in the risk assessment, which must be evaluated. The workflow results in a measuring plan aligned with the manufacturing process (Figure 5.3).

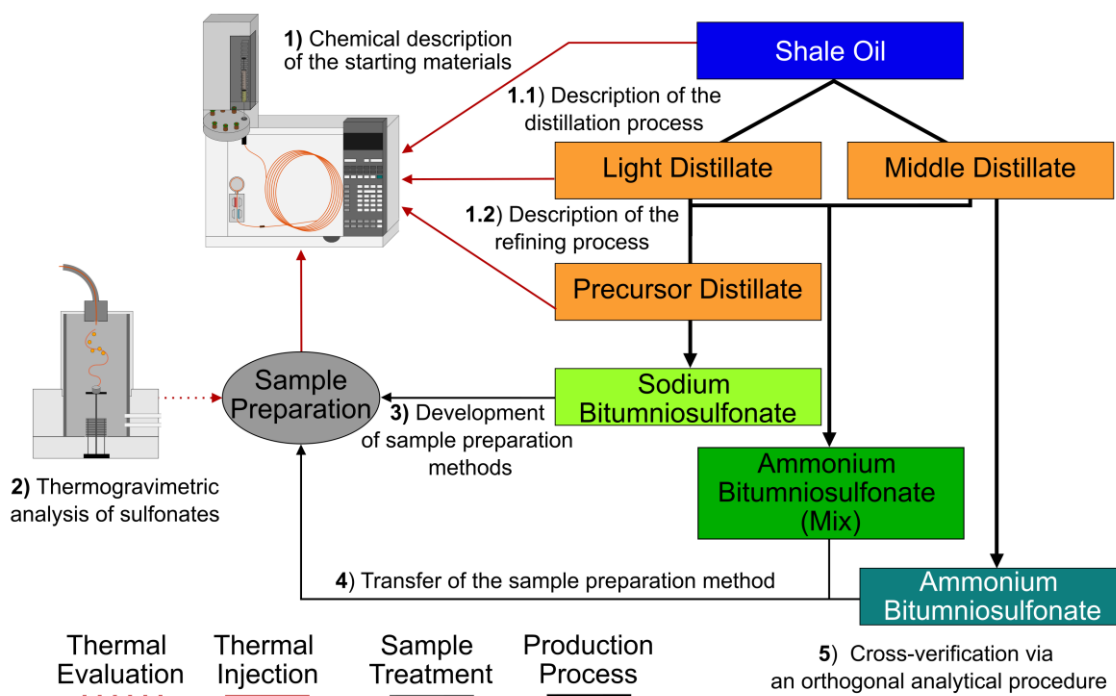


Figure 5.3 Measuring plan derived from the analytical workflow. The numbered steps are placed within the manufacturing process, and the measurements are linked to the used analytical devices. The steps are numbered according to their location in the manufacturing process and in the order of the discussion.

The plan illustrates the comprehensive nature of the study, addressing the entire GMP-regulated manufacturing process and evaluating the limitation of GC×GC via TGA-MS and ESI(-) FT-ICR MS. The numbered steps in Figure 5.3 also visualize the order in the discussion of the following sections.

5.2 Assessing the Chemical Effects of the Defining Manufacturing Process

5.2.1 Vacuum Distillation of the Shale Oil

Like other complex APIs, bituminosulfonates are defined by their manufacturing process. Therefore, it is necessary to elucidate the chemical composition of the starting materials and process intermediates

to show the chemical effects of the individual steps in the specified and regulated process. The manufacturing process of bituminosulfonates starts with the vacuum distillation of shale oil. The defined distillation range (in reduced pressure) allows the evaporation of the matrix and the applicability of the GC×GC technique. With the information on the elution behavior, accurate masses, and fragmentation pattern, the GC×GC accesses the chemical composition of the shale oil and its two distillate fractions. The classified peaks are transferred in a scatter plot to combine the chromatographic information with the elemental composition (Figure 5.4).

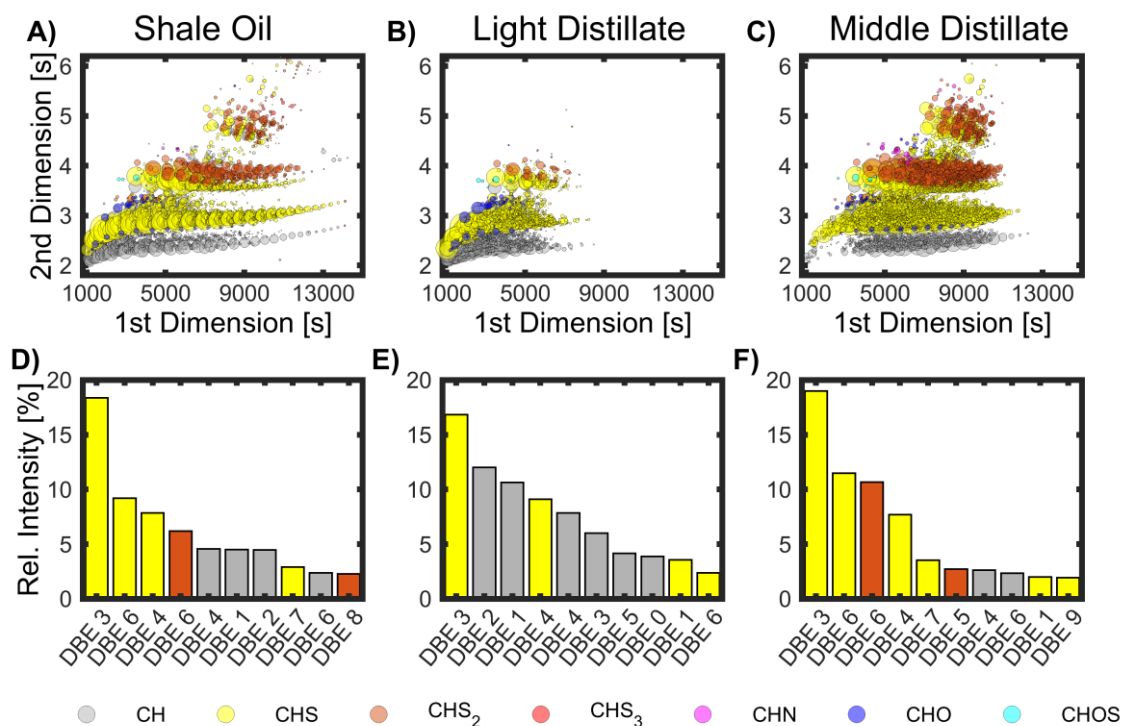


Figure 5.4 Chemical composition of the shale oil and its distillation products [116]. From left to right are the shale oil, light, and middle distillate depicted. **A-C**: Scatter plots of the classified peaks from the comprehensive two-dimensional gas chromatography (GC×GC) measurements. The circles are color-coded according to the elemental composition, and the area illustrates the peak's relative abundance (relative to the most abundant peak). **D-F**: Bar plots of the ten most abundant compound classes. The bars are color-coded according to the elemental composition and represent the relative abundance in the GC×GC measurements (excluded column bleed region).

The scatter plots in Figure 5.4 illustrate the broad chemical profile of the shale oil (A), both boiling point cuts (B and C), and their characteristic elution behavior in the GC×GC. Based on the chromatographic and mass spectrometric information, the abundances of the classified peaks, as values for the detector response, can be summarized in their compound classes (Figure 5.4 D, E, and F). Here, the compound classes are indicated by their elemental composition and DBE value but also the core structural motive of most of them could be identified [74]. This enables a chemical comparison of the starting materials in addition to the characteristic fingerprint. Although all three samples show CHS DBE 3 (Thiophenes) with 15-20 % as the most abundant class, the aromaticity and number of sulfur atoms differ between the three samples. Here, the low boiling point cut off of the light distillate is represented by the relatively high abundance of non-aromatic compound classes. The distillation phase of the middle distillates is also more focused in its elution pattern and concentrates particularly on the two-ring aromatic species (CH: DBE 7; CHS_{1,2}: DBE 5-6). In addition to the differences between the

compound classes, they differ in the carbon number range within these classes. Both can be depicted via DBE vs carbon number plots (Figure 5.5).

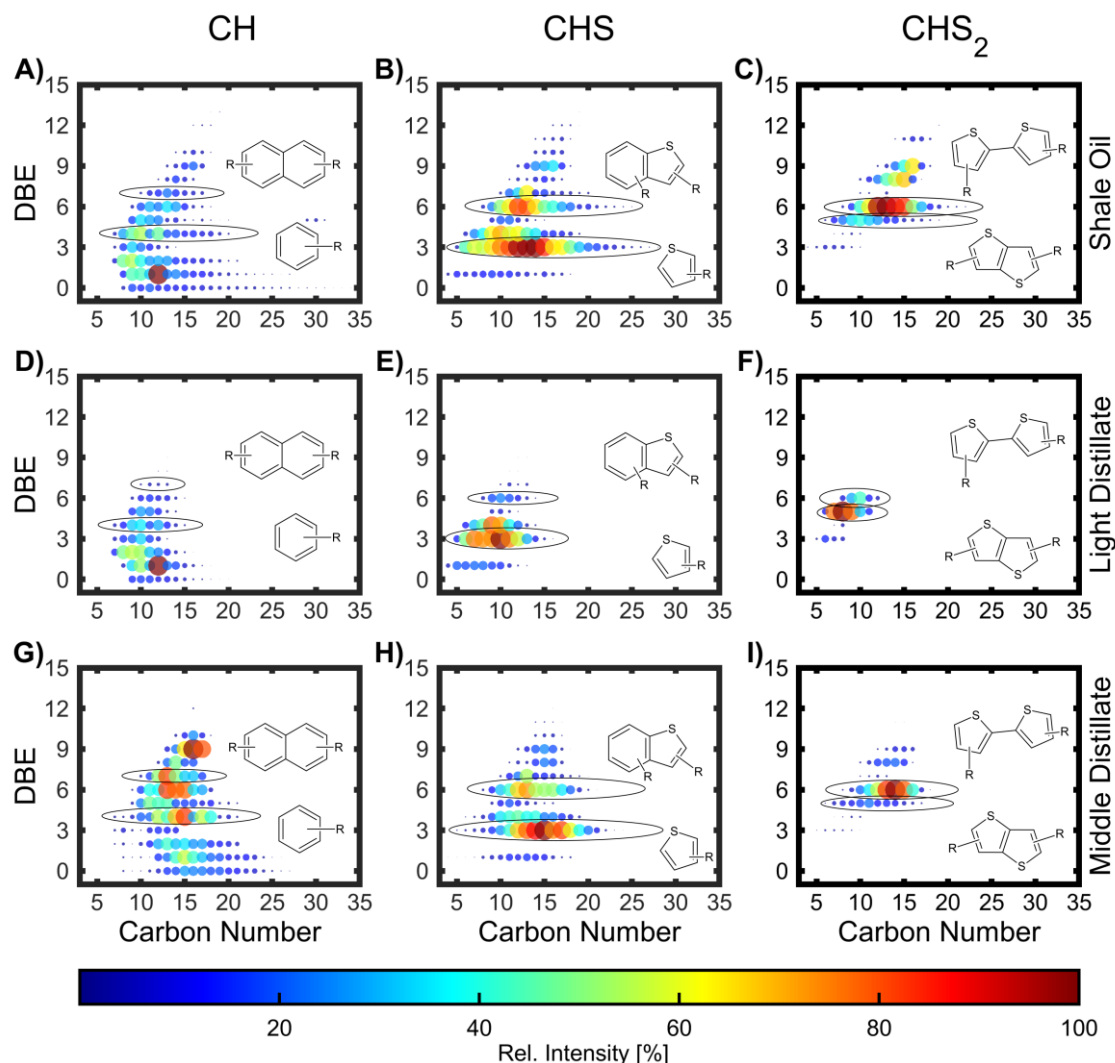


Figure 5.5 Double bond equivalent (DBE) vs. carbon number plots of the shale oil and its distillation products. From left to right are the elemental compositions depicted CH, CHS, and CHS₂, and from the top to the bottom, the matrices: shale oil, light, and middle distillate. The circles' color and size illustrate the relative abundance (to the most abundant sum formula in the individual figures).

While the compound class distribution showed the relative abundance of the ten most abundant compound classes (Figure 5.4 E and F), the DBE vs carbon number plots depict the cut-off points in the aromaticity (CH: DBE 10) and chain length (CH: C₂₇) for the middle distillate and light distillate (CH: DBE 7; C₁₈) (Figure 5.5 A, D, and G). As a result of the distillation, the highest eluting compounds are removed and not present in either of the distillates. In addition, while the light distillate shows a focus on one-ring aromatic compounds (DBE 3-5) with about ten carbon atoms, the middle distillate is shifted to two-ring aromatics (DBE ≥ 6) and carbon numbers around 15 (for thiophenes). The profile of the shale oil is more diverse and covers a broad range of carbon atoms and DBE values.

As a result of the chemical description of the distillates, the cut-off points of the distillation process and the chemical differences of the starting materials could be visualized. This allows the establishment of core structural motives and carbon number ranges of compounds that can react with sulfuric acid.

5.2.2 Refining of the Light Distillate

While distillation processes are well-described and performed in different industries for comparable matrices, refining processes are diverse and difficult to compare [117, 118]. In the manufacturing process of SBS, the light distillate is treated with a diluted acidic, alkaline, and clay solution to obtain the precursor distillate. This historic process visually clarifies the sample from the oily dark brown light distillate to the clear yellow refined precursor. To address the chemical effect of this step, both matrices were measured five times with an optimized oven program and modulation time (Figure 5.6).

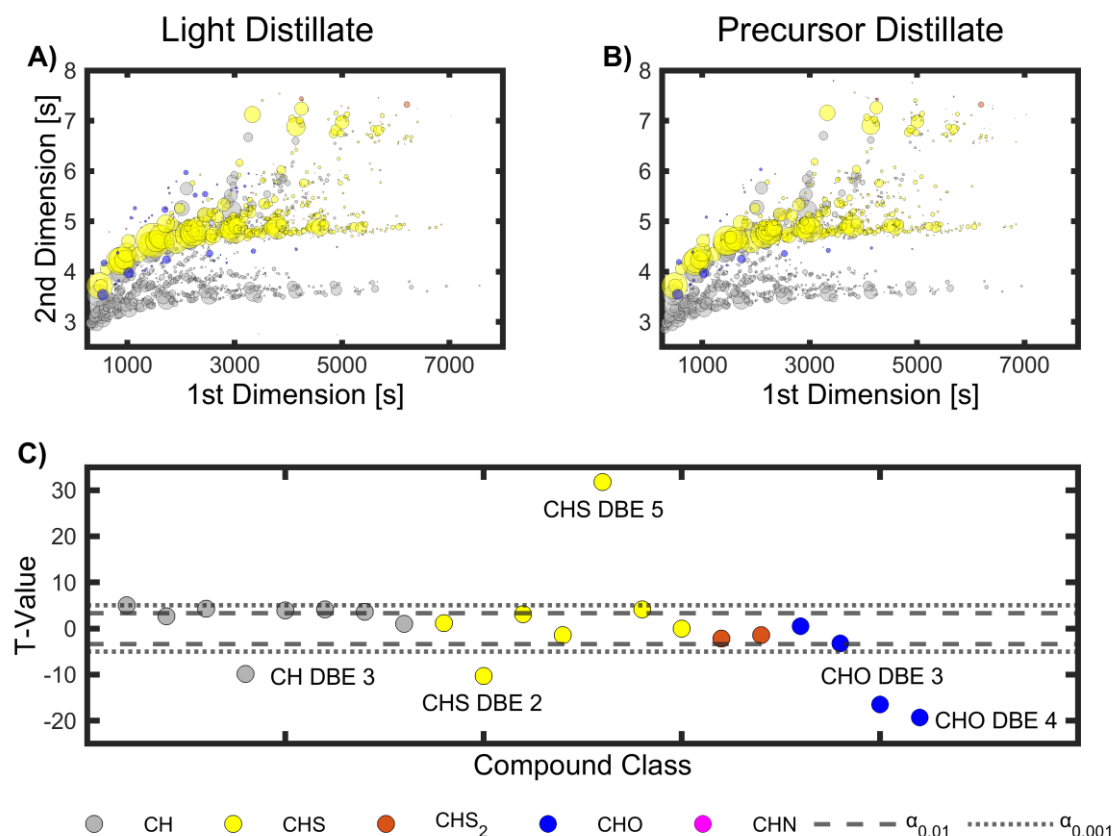


Figure 5.6 Statistical comparison of the light distillate and precursor distillate [74]. **A** and **B**: Scatter plot of the classified peaks from the comprehensive two-dimensional gas chromatography (GC×GC) measurement of the light distillate and refined precursor. **C**: Illustration of the non-absolute T-values of a two-sided T-test comparing refined precursor to the light distillate. The colored circles represent the substance classes with increasing double bond equivalent (DBE) values (X-axis from left to right: CH DBE 0-7, CHS DBE 1-7, CHS₂ DBE 5-6, and CHO DBE 1-4, CHN DBE 3-4 are not present in the refined precursor), and the horizontal lines visualize the significance threshold values. A negative T-value shows higher abundance in the light distillate and a positive in the refined precursor.

The results of the replicates show high repeatability with relative standard deviations below 10 % for the majority of the detected compound classes [74]. Also, the light distillate and its refined product are comparable. They show a similar elution pattern (Figure 5.6 **A** and **B**) and chemical composition, with approximately 60 % of the relative abundance assigned to aromatic and 50 % to organic sulfur compounds [74]. Multiple statistical tests were performed based on the high repeatability and the similarity of the matrices. The data sets of the compound classes were tested for outliers (Grubb's test), compared by their variances (F-test), and their average values (T-test). Thereby, significant chemical differences could be shown for the polar oxygen and nitrogen species. The refining procedure caused a significant reduction partially below the LOD of compounds assigned to these classes; they include

phenol and pyridine. Eliminating these restricted and potentially harmful compounds indicates a detoxification effect of the refining procedure. Besides this expected effect, a significant decrease in the low intense and non-aromatic CH DBE 3 and CHS DBE 2 classes and an increase in CHS DBE 5 classes are shown. This may be caused by side reactions of these classes during the refining process.

This comparison showed, for the first time, the chemical effect of the refining process in the strictly regulated manufacturing process of SBS. It removes potentially harmful compounds like phenol and pyridine species limited by the Ph. Eur. before the sulfonation process. In addition, the study elucidates the chemical composition of the precursor distillate, which is sulfonated, in the following manufacturing process. Therefore, it limits the possible products and the elemental composition of $C_cH_hO_oS_s$ and gives a chemical profile to refer to in the analysis of SBS.

5.3 Analysis of Sodium Bituminosulfonates

5.3.1 Thermal Analysis

Shale oil and its distillates are typical matrices for GC×GC [93, 112], but sulfonated products like SBS are more challenging for gas phase analysis [60]. The applicability of GC analysis requires the vaporability of the matrix, which is questionable for the sodium sulfonates in the APIs. For the

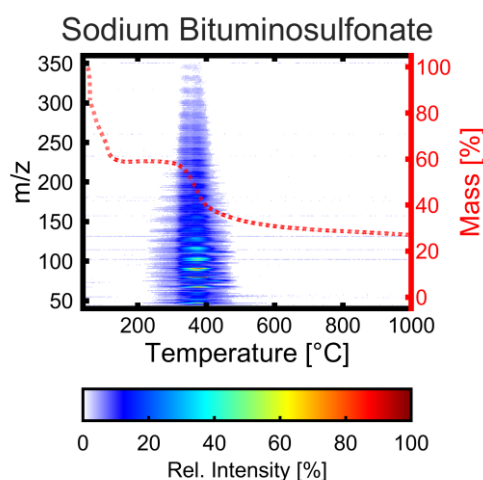


Figure 5.7 Thermogram of the measurement from the sodium bituminosulfonate (SBS). The thermogram is combined with the measured mass loss (red).

development of a suitable analytical procedure, the thermal behavior of the SBS was addressed via TGA-MS (Figure 5.7).

The obtained thermogram does not show volatile organic compounds (VOCs) and only above 350 °C evolved gases were detected ($> 40 m/z$). This indicates at least a strong sulfonation of the refined precursor or removal of non-sulfonated compounds in the phase separation after the neutralization. The mass loss of approximately 40 w-% below 100 °C results from the evaporation of water, which is out of the measured m/z range (40-500). The mass spectrum of the second mass lost event mainly consists of fragments from alkylated aromatic core structures (thiophenes, benzenes, benzothiophenes, etc.). However, accurate masses assignable to sulfonates were not detected. This points towards thermal instability above

350 °C and missing volatility of the sodium sulfonates in the API. As a result of the uncertainties in the thermal behavior of SBS, the thermal analysis must be extended.

Due to the complexity of the APIs, sodium benzenesulfonate was chosen as a commercially available compound to indicate the thermal behavior of SBS. Besides the sodium salt, benzenesulfonic acid and benzenesulfonate methyl ester were used to represent targets or intermediates of sample preparation of previous sulfonate GC- MS studies [34, 59, 60]. The TGA-MS measured these single compounds to monitor the mass loss in parallel with detecting the evolved gases (Figure 5.8).

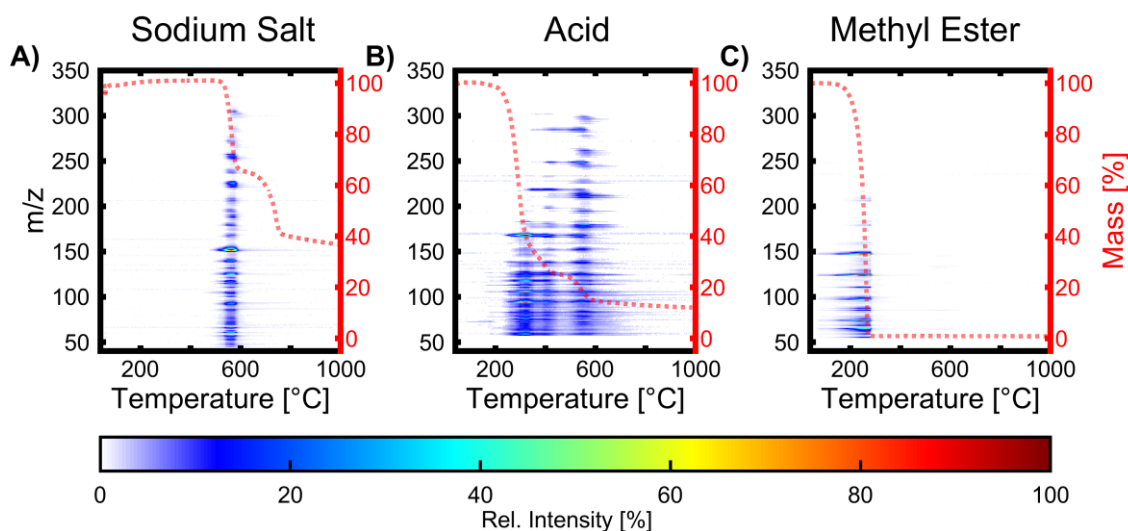


Figure 5.8 Thermogram of the benzenesulfonates [111]. Sodium benzenesulfonate (A), benzenesulfonic acid (B), and benzenesulfonate methyl ester (C). The thermogram is combined with the measured mass loss (red).

The TGA-MS demonstrates the missing volatility of sodium benzenesulfonate (Figure 5.8 A). Here, the first mass loss and evolved gases were detected close to 600 °C with m/z values exceeding the molecular ion, indicating the formation of thermal degradation products like diphenyl sulfide (accurate mass: m/z 186.0504) [100]. The higher temperature of the indicator compared to the API is a result of matrix effects or an increased instability of the sulfur aromatic sulfonates, which should be the major constituents based on the analysis of the precursor distillate. While the sodium benzenesulfonate shows no volatility, the benzenesulfonic acid is partially vaporizable at temperatures below 350 °C. Approximately 60 % of the mass from the acid went into the gas phase below this threshold temperature (Figure 5.8 B). In this temperature range, the MS also detected the molecular ion of benzenesulfonic acid (accurate mass: m/z 158.0037). However, diphenyl sulfone (accurate mass: m/z 218.0399), a thermal degradation product of benzenesulfonic acid, was detected in parallel, indicating simultaneously occurring vaporization and thermal degradation for the benzenesulfonic acid [100]. The benzenesulfonate methyl ester completely evaporates below 300 °C, with a mass spectrum similar to the NIST database and no m/z values above the molecular ion.

This analysis demonstrates the criticality of the injection temperature due to the formation of thermal degradation products in sodium benzenesulfonate and benzenesulfonic acid. Only the benzenesulfonate methyl ester is measurable via GC×GC. This confirms that methylation is a suitable derivatization reaction and should lead to the vaporability of sulfonates in SBS. In addition, the absence of volatile compounds in the thermal analysis of SBS (Figure 5.7) indicates a strong or even complete sulfonation of the precursor distillate.

5.3.2 Sample Preparation

In the sulfonation step, the aromatic compounds in the distillates react with concentrated sulfuric acid and form sulfonates. This was also shown in the 1980s [34]. After the sulfonation and neutralization, the complete separation and removal of the organic phase are reported for SBS, the target of the sample preparation. Also, the absence of volatile compounds in the SBS indicates a completely sulfonated API. Due to its missing volatility, the conversion into a thermally stable and volatile derivative is necessary.

Based on the previous studies and TGA-MS analyses, this can be addressed by the methylation of the sulfonates. Additionally, other characteristics like the water content (close to 50 %), surfactant properties, and required high reproducibility must be considered in the sample preparation [58, 60]. At first, an online derivatization method with tetramethylammonium hydroxide (TMAH) was developed. Here, the sample preparation is limited to adding TMAH to the API and diluting the solution 1:10 with MeOH/H₂O (1:1 V/V) (Figure 5.9).

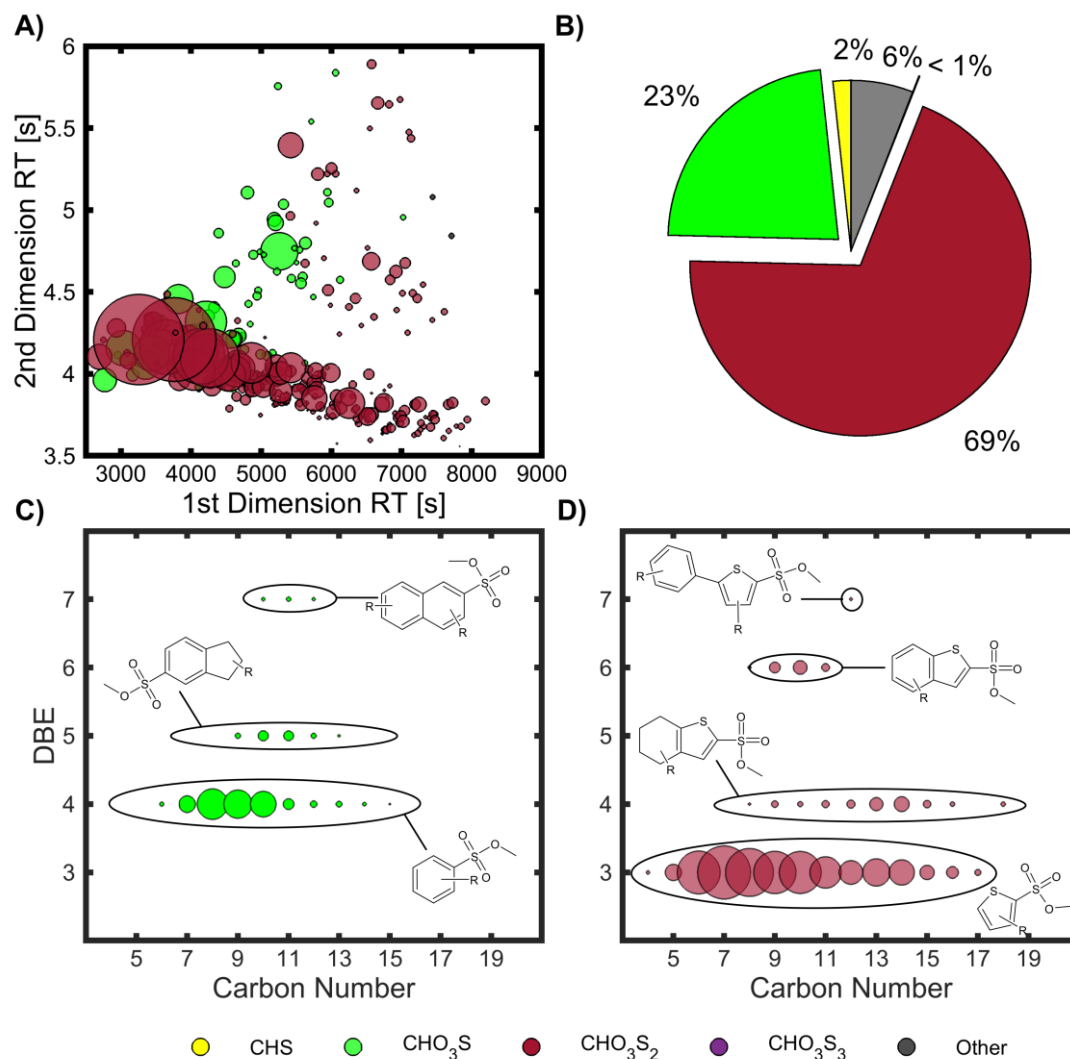


Figure 5.9 Chemical composition of sodium bituminosulfonate (SBS) [74]. **A:** Scatter plot of the classified sulfonated peaks CHO_3S , CHO_3S_2 , and CHO_3S_3 from the comprehensive two-dimensional gas chromatography ($\text{GC}\times\text{GC}$) species. The circle area represents the relative abundance of the peak (relative to the most abundant isomer). **B:** Pie chart of the relative abundance of the classified compound classes with highlighted CHO_3S , CHO_3S_2 , and CHO_3S_3 species. **C and D:** DBE vs. Carbon number plot of the CHO_3S (**C**) and CHO_3S_2 (**D**) species, with the circle area represents the relative abundance (to the most abundant sum formula of this species) of the calculated sum formula (minus the methyl group of the derivatization) and the color the elemental composition.

This online derivatization method shows numerous sulfonated species (> 500) detected as sulfonate methyl esters (SMEs) (Figure 5.9 A), with the most abundant signals assigned to two sulfonated thiophene isomers $\text{C}_7\text{H}_{10}\text{O}_3\text{S}_2$ and $\text{C}_8\text{H}_{12}\text{O}_3\text{S}_2$. Sulfonates represent more than 90 % of the relative abundance and predominantly contain a sulfur atom in their core structural motive (CHO_3S_2) (Figure 5.9 B). They are distributed in eight compound classes: three CHO_3S_1 , four CHO_3S_2 , and one CHO_3S_3 species (Figure 5.9 C). These detected sulfonated compound classes are assigned to the methyl ester of

thiophene- (TSME), benzene- (BSME), tetrahydrobenzothiophenes- (THBTSME), indane- (ISME), benzothiophenes- (BTSME), naphthalene- (NSME), bithiophene- (BiTSME), and phenylthiophene-sulfonate (PTSME) (Figure 5.10). They are listed from the highest to lowest abundant compound classes. Their corresponding precursor compounds also represent the most abundant aromatic compound classes in the light distillate (Figure 5.4 E). However, some measured aromatic compound classes of the precursor distillate are missing, like CH DBE 5 (Indenes), CHS₂ DBE 5 (Thienothiophenes), or the possible side product of the refining process CHS DBE 5. This may be a combination of their low abundance and lower reactivity with sulfuric acid [119].

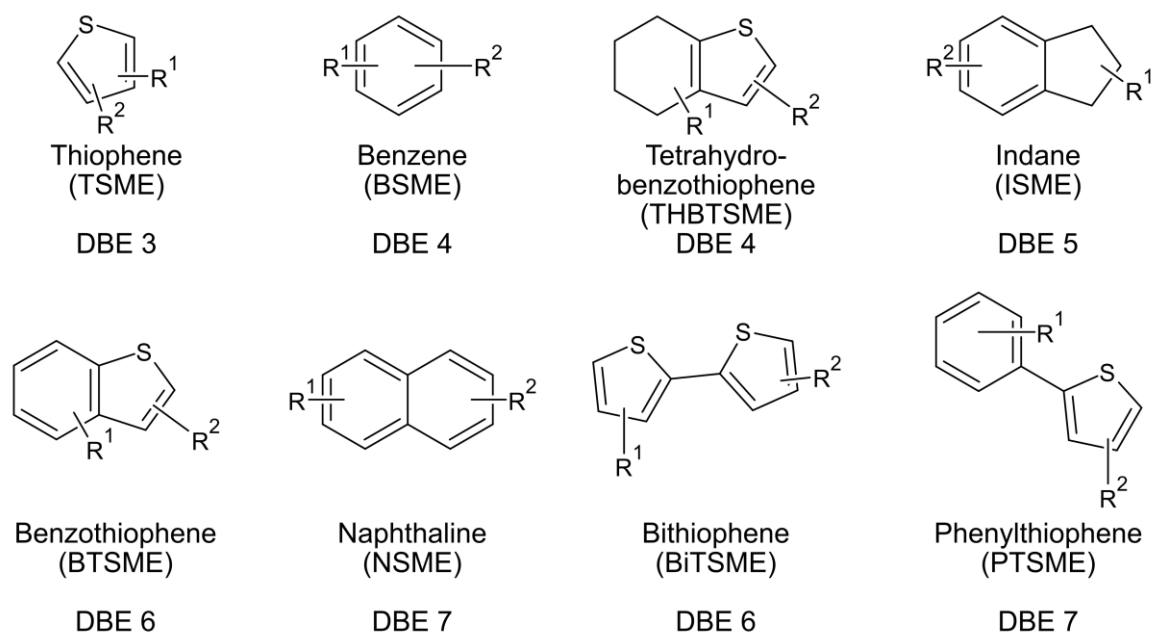


Figure 5.10 Core structures and derivatives as they occur in the manufacturing process and analysis. Including their double bond equivalent (DBE). R¹ and R²: H, core structural motive; R¹ and R²: H or C_nH_{n+2}, alkylated precursor compound (distillates); R¹: H or C_nH_{n+2} and R²: SO₃⁻Na⁺, sodium sulfonates (API); R¹: H or C_nH_{n+2} and R²: SO₃CH₃, sulfonate methyl ester (SME).

Because of the limited database entries and commercially available standard materials high mass accuracy and resolution were used for identification. Here, the stability of the molecular ions [M]⁺, which were detected within a mass error of +/- 5 ppm, the characteristic fragmentation (Figure 4.3), and chromatographic separation of isobaric structures allows the assignment to the eight SMEs [74].

The affinity of the sulfonation is visible in the carbon number distribution within the compound classes (Figure 5.9 C and D). Although the TSMEs in the API cover the same carbon number range as the thiophenes in the precursor distillate (C₄ to C₁₇), the focus shifts to lower carbon numbers, from C₁₀ in the precursor to C₇ in SBS. This is caused by the isomer distribution and alkylation pattern that cannot be sulfonated at carbon number > 7.

This method was further evaluated versus different kinds of extraction combined with liquid derivatization by TMSDAM. Here, extraction procedure via SPE with the different extraction sorbents (HLB, SAX, and WAX) and L/L extraction with acidified water and toluene followed by neutralization and methylation prior to injection with TMSDAM was performed (Figure 5.11).

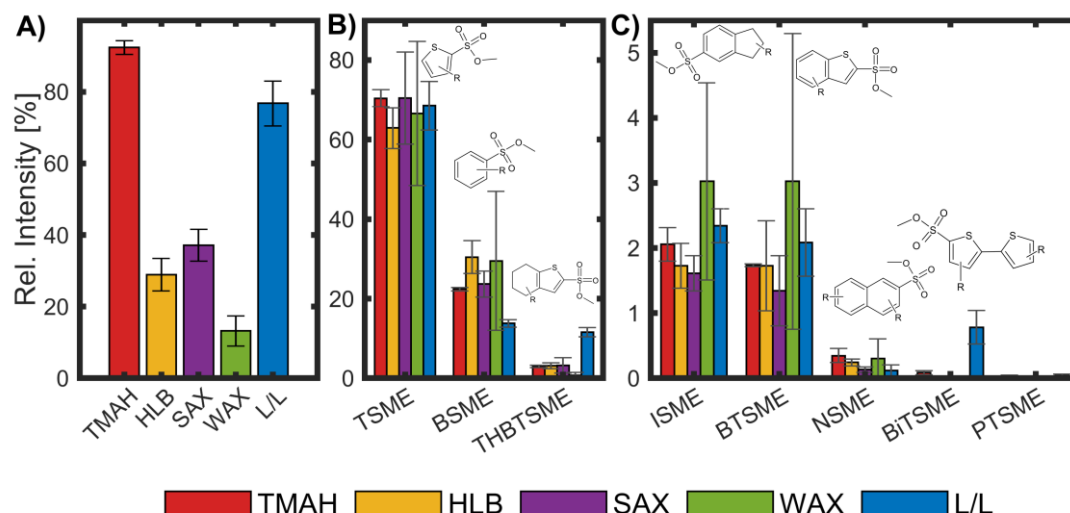


Figure 5.11 Comparison of the sulfonated species with different sample preparation methods [111]. Online derivatization by tetramethyl ammonium hydroxide (TMAH), offline derivatization with trimethylsilyl diazomethane after extraction via solid phase extraction (extraction sorbents: hydrophilic-lipophilic balance (HLB), strong anion exchanger (SAX), weak anion exchanger (WAX)) or liquid-liquid extraction (L/L). **A:** Bar plot of the relative abundance from the sulfonates to the overall ion current (excluding column bleed region). **B and C:** Bar plot of the mean value and the standard deviation ($n=3$) for the sulfonated compound class distributions of the different methods normalized for each method to the total ion current of the sulfonates.

The most significant difference between the sample preparation methods is the ratio between sulfonated and non-sulfonated compounds (Figure 5.11 A). In particular, the SPE approaches lead to a low relative abundance of the detected sulfonated compounds, while assessing $> 50\%$ of the relative abundance to non-sulfonated $\text{CHS}_{0.2}$ species. However, these non-sulfonated compounds are very reactive with sulfuric acid and consist of relatively volatile ($< 300\text{ }^\circ\text{C}$) thiophenes and benzothiophenes, which would be visible in the TGA-MS (Figure 5.7) [102, 119, 120]. However, the TGA-MS demonstrates a thermal instability of sulfonic acids and a minor mass loss for SBS close to the injection temperature ($350\text{ }^\circ\text{C}$). Here, the evolved gases showed m/z values mainly assignable to thiophene core structures. This also indicates a partial thermal degradation of the SBS. Therefore, in SPE approaches, predominant CHS species result from an insufficient extraction and/or derivatization. Combined with their absence or low abundance after the L/L extraction and online derivatization method, which should be more sensitive for non-sulfonated residuals or volatile compounds, this shows the unsuitability of the tested SPE approaches.

The comparison shows that the ratio of sulfonated to non-sulfonated species directly correlates with the derivatization efficiency, and only the L/L extraction approach showed a comparable derivatization efficiency to the online TMAH method. However, the effect of the extraction is visible in the compound class distribution (Figure 5.11 B and C). While the standard deviations of the relative abundances from the total sulfonates are comparable, they significantly differ for the individual compound classes. For seven of the eight compound classes, the online derivatization shows the highest repeatability, and although the distribution of the compound classes is comparable between the sample preparation methods, there are visible differences [111]. In particular, the abundance of BSME and THBTSME significantly differ with the L/L method from the other approaches and are closer to the ratio of the light

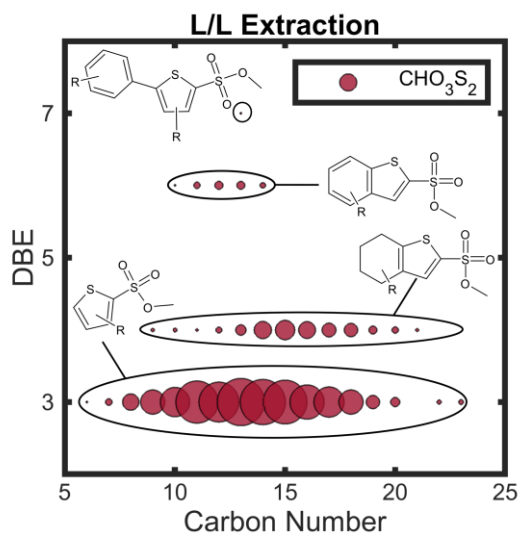


Figure 5.12 Double bound equivalent (DBE) vs. carbon number plot of the CHO_3S_2 species after the liquid-liquid extraction (L/L) and derivatization via trimethylsilyl diazomethane (TMSDAM) of sodium bituminosulfonate (SBS).

distillate (Figure 5.4). In addition, the L/L extraction method shows higher percentages of BiTSME. However, the similarity of the online derivatization to the remaining compound classes and the distribution of the SPE methods indicate the reliability of the online derivatization for the qualitative chemical profile. The shift to more non-polar compounds of the L/L extraction is also visible in the carbon number distribution. Here, the extraction shifts the profile to higher carbon numbers (Figure 5.12), focusing on C_{12} and reaching up to C_{23} for the TSMEs. This intensity maximum is even at a higher carbon number than the light distillate, where aromatic compounds with carbon numbers over 17 were not detected (Figure 5.5). Therefore, the shift in the carbon number and compound class distribution is an effect of the L/L extraction. This extraction targets the non-polar core

structural motive and is more affine for larger compounds. As a result, although this distribution most likely does not represent the quantitative composition of SBS, it complements the qualitative data of the online derivatization by selectively addressing the larger compounds.

Besides the derivatization efficiency and shifts in the chemical profile caused by the affinity of the derivatization agent, the derivatization could induce side reactions. However, none of the offline approaches detected potentially hydrolyzable compounds or missed potential products [111]. This limits the possibility of TMAH-induced side reactions and supports the reliability of this method.

The sample preparations differ in their derivatization efficiency, repeatability, compound class distributions, and carbon number profile. The previous analyses of the light distillate and distillate precursor evaluate the results of the different sample preparations. Here, the online derivatization aligns the most with the expectation of the sulfonation products of these starting materials. The chemical profile is further supported and complemented by the offline sample preparations.

Based on the combined TGA-MS and GC×GC results of the starting materials, process intermediates, and SBS with different sample preparations, multiple conclusions can be drawn:

- Sodium bituminosulfonate does not contain quantitative amounts of non-sulfonated organic compounds, and the presence of non-sulfonated organic compounds in SBS indicates thermal degradation and insufficient sample preparation.
- The abundance order of the sulfonate methyl esters and covered carbon numbers are similar to their precursors in the distillate, but the abundances are more focused on fewer isomers in the API.
- With the online derivatization method, most of the GC amenable compounds were detected.

The online derivatization method and its development met all three of the scope's objectives (S. I-III). While the GC×GC method elucidates the chemical profile to a molecular level (S. I), the comprehensive method development demonstrates the reliability of the qualitative results (S. II) and links them to the defining historic production process (S. III).

5.4 Transfer of the Online Derivatization Method to Ammonium Bituminosulfonates

After developing and verifying the online derivatization method for SBS, the method is transferred to both ABS matrices. While ABS (mix) is the sulfonated product of the combined light and middle distillate, ABS is only based on the middle distillate. Both contain higher boiling fractions and are directly sulfonated without a refining procedure and at higher temperatures than SBS. These factors can cause a more diverse chemical profile. In addition, the separation of the oil phase after the sulfonation and neutralization is only partially possible. Consequently, while non-sulfonated organic compounds indicate thermal degradation during the analysis for SBS, both ABS matrices could contain them as non-reacted residuals from the manufacturing process.

5.4.1 Chemical Description of Ammonium Bituminosulfonate (ABS)

The first analyzed ABS originates from the middle distillate. It should contain the molecules with lowest vapor pressure of the three APIs and tests the limits of the online derivatization method (Figure 5.13).

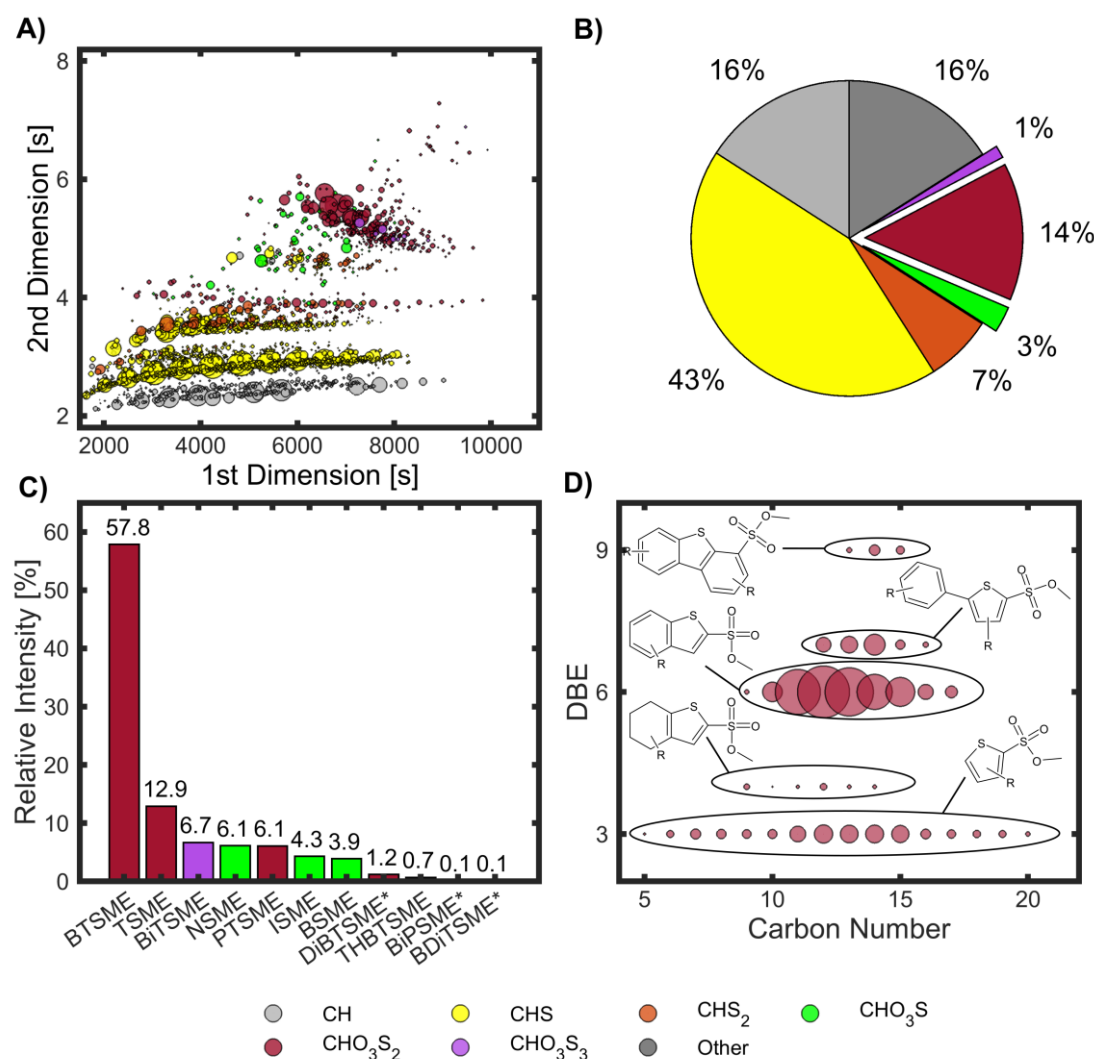


Figure 5.13 Chemical composition of ammonium bituminosulfonate of the middle distillate (ABS). A: Scatter plot of the classified sulfonated peaks CHO_3S , CHO_3S_2 , and CHO_3S_3 from the comprehensive two-dimensional gas chromatography ($\text{GC}\times\text{GC}$) species. The circle area representing the relative abundance of the peak (relative to

the most abundant isomer). **B**: Pie chart of the relative abundances of the classified compound classes with highlighted CHO_3S , CHO_3S_2 , and CHO_3S_3 species. **C**: Bar plot of the normalized (to their summed abundance) distribution of the sulfonated compound classes. These include the eight that were detected in SBS (thiophene (TSME), benzene (BSME), tetrahydrobenzothiophenes (THBTSME), indane (ISME), benzothiophenes (BTSME), naphthalene (NSME), bithiophene (BiTSME), and phenylthiophene sulfonate methyl ester (PTSME)), and three additional with * highlighted classes: dibenzothiophene (DiBTSME), biphenyl (BiPSME), and benzodithiophene sulfonate methyl ester (BDiTSME). **D**: double bound equivalent (DBE) vs. carbon number plots of the CHO_3S_2 species. The circle area represents the relative abundance (to the most abundant sum formula of this species) and the color the elemental composition.

Compared to SBS, the GC×GC results of ABS are more diverse. They comprise more classified peaks (> 2000) with a broad polarity range and elute along the entire space of the second dimension (Figure 5.13 A). While GC×GC measured almost only sulfonated compounds in SBS, the $\text{CHO}_3\text{S}_{1-3}$ species represent less than 20 % of the overall ion current for ABS (Figure 5.13 B). Most of the peaks and the relative abundance are assigned to non-sulfonated compounds, but two of the most abundant are present in the BTSME class ($\text{C}_{11}\text{H}_{12}\text{O}_3\text{S}_2$ and $\text{C}_{12}\text{H}_{14}\text{O}_3\text{S}_2$). Although compounds with other elemental compositions were detected, Ph. Eur. limited compounds like phenol and pyridine were not found. Compared to the precursors in the middle distillate and in contrast to SBS, ABS shows a higher aromaticity in the sulfonated products and a shift in their abundance order (Figure 5.13 C). Exemplary, while thiophenes are the most abundant compound class in the middle distillate, approx. two times more abundant than benzothiophenes, after the sulfonation, BTSME are more than four times more abundant than TSME. Besides the eight sulfonated compound classes that were measured in SBS, in ABS, three additional and more aromatic compound classes were detected: dibenzothiophene (DiBTSME), biphenyl (BiPSME), and benzodithiophene sulfonate methyl ester (BDiTSME). Besides the shift in the compound class distribution, the detected carbon numbers are also shifted (Figure 5.13 D). While the benzothiophenes class in the middle distillate covers C_8 to C_{20} with the maximum abundance at C_{13} , the maximum and the maximal abundant carbon number of BTSMEs in ABS is decreased to C_{17} and C_{11} , respectively.

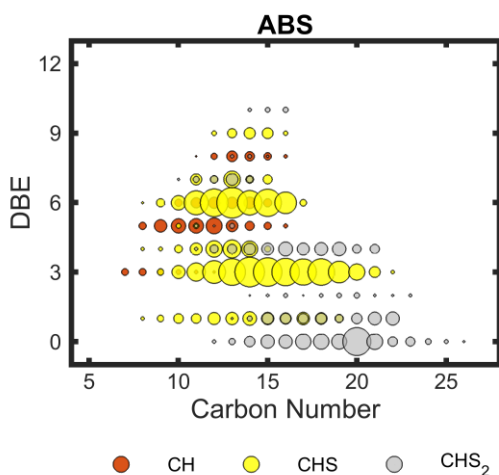


Figure 5.14 Double bound equivalent (DBE) vs carbon number plot of the detected CH, CHS, and CHS_2 classes in ammonium bituminosulfonate (ABS).

abundance due to an incomplete separation more likely.

In contrast to SBS, many non-sulfonated compounds were detected (Figure 5.14), and the incomplete phase separation is visible in presence of non-reactive aliphatic CH and CHS species. Although they were also present in the light distillate, they were separated from the SBS. Besides non- and low-reactive compounds, nearly all the precursors of the middle distillate are visible in ABS, including generally reactive compound classes like thiophenes and benzenes. However, most of these reactive species include isomers that are completely alkylated or stereochemically hindered and cannot form sulfonates. Although similar compounds could also be created by thermal degradation during the injection, their absence in the measurement of SBS indicates a low degradation tendency with the online derivatization method and makes their presence and

abundance due to an incomplete separation more likely.

In summary, the higher distillation phase of the middle distillate is less reactive and results in a high percentage of non-sulfonated compounds in ABS. Although the ratio of non-sulfonated to sulfonated

compounds may be affected by thermal decomposition and a non-uniform detector response of the compound classes, the presence and high abundances of these partially non- or low-reactive compounds are plausible but they do not reflect the quantitative amount. The distribution of the sulfonates themselves is also shifted to more aromatic core structural motives. These results align with the differences in the manufacturing process of SBS and ABS.

5.4.2 Chemical Description of Ammonium Bituminosulfonate (ABS (mix))

In contrast to the other ABS, ABS (mix) is based on the sulfonation of the combined light and middle distillate phases but is otherwise produced similarly. Therefore, the chemical composition of this API should be closer to ABS than SBS. To show this, ABS (mix) is measured with the same online derivatization method as the other two bituminosulfonates (Figure 5.15).

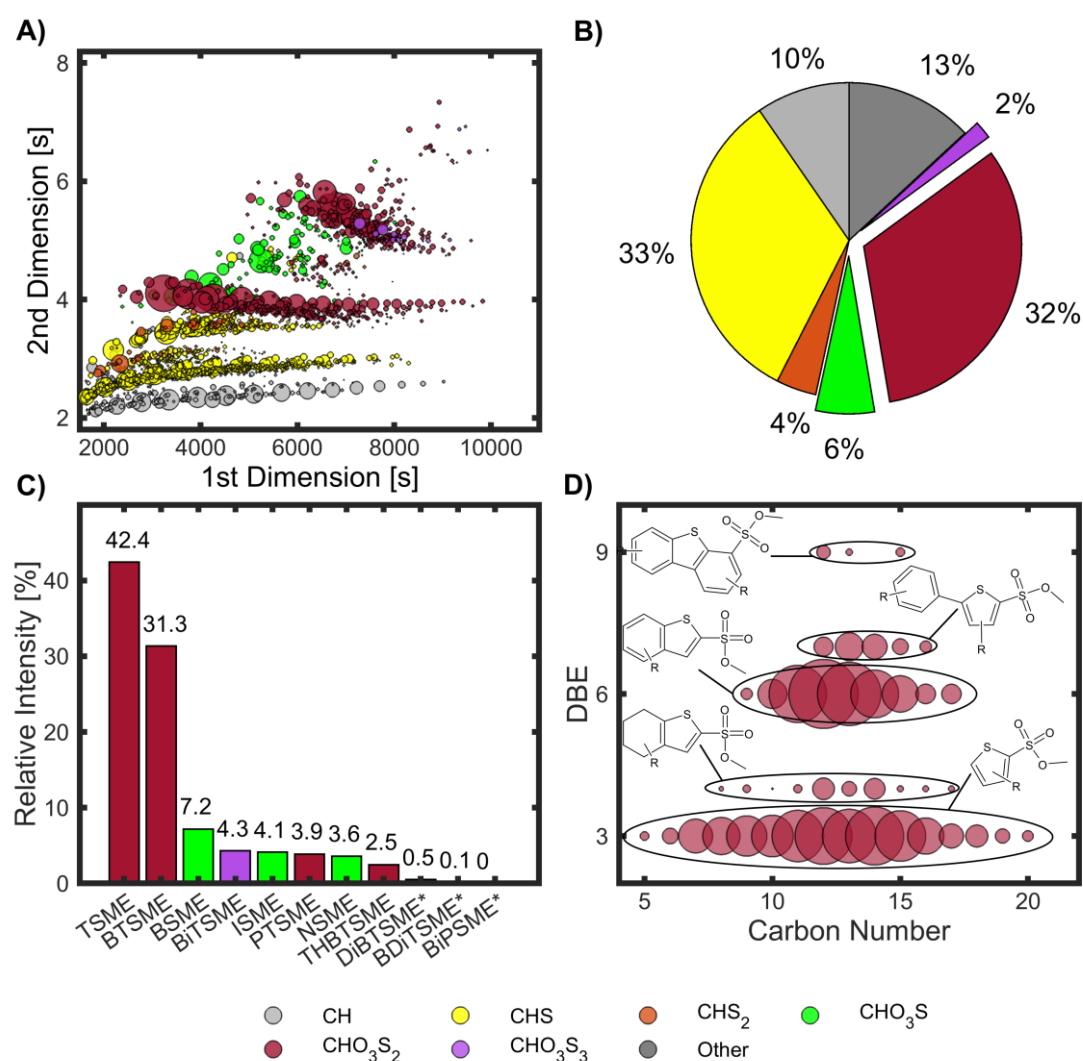


Figure 5.15 Chemical composition of ammonium bituminosulfonate of the combined distillation phases (ABS (mix)). **A:** Scatter plot of the classified sulfonated peaks CHO₃S, CHO₃S₂, and CHO₃S₃ from the comprehensive two-dimensional gas chromatography (GC×GC) species. The circle area representing the relative abundance of the peak (relative to the most abundant isomer). **B:** Pie chart of the relative abundances of the classified compound classes with highlighted CHO₃S, CHO₃S₂, and CHO₃S₃ species. **C:** Bar plot of the normalized (to their summed abundance) distribution of the sulfonated compound classes. These include the eight that were detected in SBS (thiophene (TSMES), benzene (BSME), tetrahydrobenzothiophenes (THBITSME), indane (ISME), benzothiophenes

(BTSME), naphthalene (NSME), bithiophene (BiTSME), and phenylthiophene sulfonate methyl ester (PTSME)), and three additional with * highlighted classes: dibenzothiophene (DiBTSME), biphenyl (BiPSME), and benzodithiophene sulfonate methyl ester (BDiTSME). **D**: double bound equivalent (DBE) vs. carbon number plots of the CHO₃S₂ species. The circle area represents the relative abundance (to the most abundant sum formula of this species) and the color the elemental composition.

Like the other ABS matrix, the classified peaks (>1800) in the GC×GC of ABS (mix) elute along the entire space of the second dimension (Figure 5.15 A). Here, the combined distillation phase is visible with the two most abundant peaks from SBS and ABS also present within the five most abundant signals in ABS (mix) (1. C₇H₁₀O₃S₂, 2. C₈H₁₂O₃S₂, 5. C₁₁H₁₂O₃S₂, and 4. C₁₂H₁₄O₃S₂). In addition, ABS (mix) shows a ratio of sulfonated to non-sulfonated compounds (close to 1:1) (Figure 5.15 B). This indicates a better reactivity of the light distillation phase, which is missing in ABS. However, the incomplete separation of the starting material prevents a fully sulfonated API.

In the comparison of the sulfonated compound class, the broader distillation range of the starting material is apparent. The distribution of the sulfonates is between SBS and ABS, with TSME and BTSME being the most abundant compound classes (Figure 5.15 C). In addition, all eleven compound classes of ABS are also detected in ABS (mix). The carbon number distribution of the sulfonated compounds resembles the aromatic non-sulfonated compounds in a mixture of light and middle distillate (Figure 5.5) and covers nearly the entire range detected in SBS and ABS (Figure 5.9 D, Figure 5.13 D, and Figure 5.15 D). Exemplarily shown for TSME, which is covered from C₄ in SBS to C₂₀ in ABS and from C₄ to C₂₀ in ABS (mix).

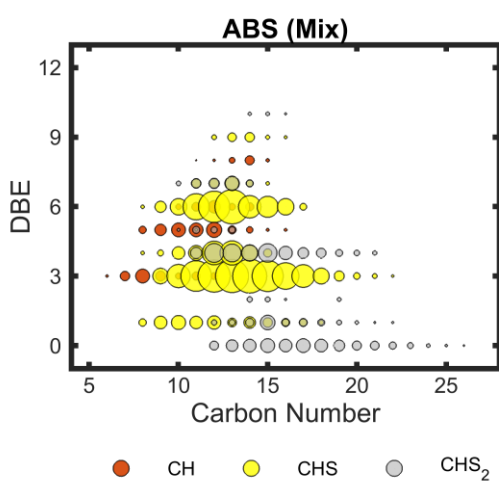


Figure 5.16 Double bond equivalent (DBE) vs carbon number plot of the GC×GC detected CH, CHS, and CHS₂ classes in ammonium bituminosulfonate of the combined distillation phases (ABS (mix)).

The non-sulfonated compounds are comparable to those detected in the ABS (Figure 5.16). Here, similar non- and low-reactive compounds are present. However, even the light distillate contains alkylated thiophene and benzene isomers that cannot be sulfonated (Figure 5.5). The similarity of the non-sulfonated species between both bituminosulfonates is caused by the overlap of the non-reactive species and the higher reactivity of the low alkylated compounds in the light distillate fraction.

Alternative L/L extraction with toluene and water was not possible. Adding both phases to ABS caused the formation of one mixed phase, which did not separate. Also, other organic phases (e.g., DCM, n-, and cyclohexane) did not sufficiently separate the sulfonates from the non-sulfonated species, even though some formed a phase separation. Here, the amphiphilic properties of the sulfonates in ABS prevent their sufficient separation and isolated measurements.

In conclusion, the online derivatization is not only the best-tested method for measuring SBS but is also suitable, in contrast to the tested L/L extractions, to address the qualitative chemical composition of both ABS matrices. However, these matrices' polarity range may reach the limits of GC×GC, and more aromatic sulfonates or disulfonated compounds could be undetectable. Therefore, an orthogonal method without thermal injection is applied to cover the limitations and cross-verify the GC×GC results.

5.5 Application of an Orthogonal Technique to Cross-Verify Complex Pharmaceuticals

The primary objective of the cross-verification via ESI(-) FT-ICR MS is the evaluation of the detected sulfonates. Here, the sensitivity for anions, minimal thermal stress for the sample, minimal size limitation, and high mass resolution and accuracy make this technique ideal for an orthogonal analysis of the APIs. In addition, while APIs had to be derivatized to measure the sulfonates (RSO_3CH_3) via GC×GC, they can [121] and were directly measured (RSO_3^-) via the ESI(-) FT-ICR MS (Figure 5.17).

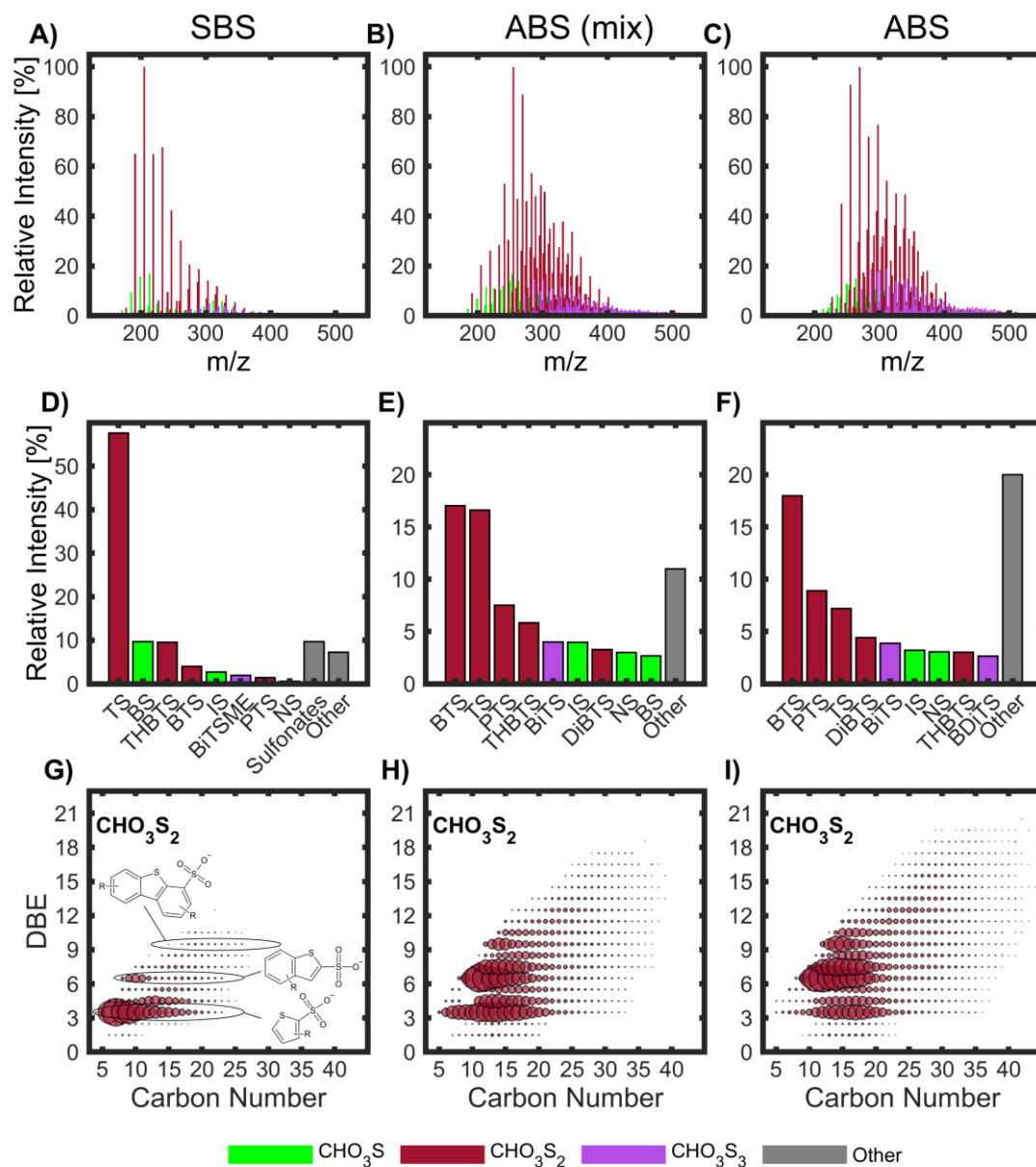


Figure 5.17 Chemical composition of bituminosulfonates [116]. From left to right, SBS, ABS (mix), ABS [111]. A-C: Color-coded mass spectra from the direct infusion negative electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry (DI ESI(-) FT-ICR MS) with the relative abundance normalized to the most abundant m/z value. D-F: Bar plots of the nine most abundant compound classes (thiophenesulfonates (TS), benzenesulfonate (BS), tetrahydrobenzothiophenesulfonate (THBTs), bithiophenesulfonate (BiTS), benzothiophenesulfonate (BTS), indanesulfonate (IS), phenylthiophenesulfonate (PTS), naphthalinesulfonate (NS), dibenzothiophenesulfonate (DiBTs), benzodithiophenesulfonate (BDiTS)) and the summed abundance non-sulfonated rest (other). G-I: Double bond equivalent (DBE) vs. carbon number plot of the CHO_3S_2 species (measured and depicted in the deprotonated form $[M]^-$).

The mass spectra of the FT-ICR MS show homolog rows of m/z values and their calculated sum formulas ($[M]^-$) that are similar to aromatic sulfonates (Figure 5.17 A-C). They include predominately CHO_3S , CHO_3S_2 , and CHO_3S_3 species, representing more than 80 % of the total ion current in all APIs. In addition, they show a high overlap with the corresponding methyl esters measured via GC×GC HR-ToF-MS. This enables the assignment to their core structural motive and compound class.

Here, the eleven compound classes that were found via GC×GC HR-ToF-MS are also the most abundant compound classes of the ESI(-) FT-ICR MS (Figure 5.17 D-F). However, even if they generally show the same most abundant compound classes, the measurements also reflect the discrimination caused by either of the techniques. The FT-ICR MS measurements show higher relative abundances for all three APIs for more aromatic and alkylated compounds, and more unique sum formulas were assigned (Figure 5.17 G-I). This is a result of the high selectivity and sensitivity of the ESI(-) for sulfonates combined with the scan time of the FT-ICR MS, which also increases the sensitivity. However, as a result of the high overlap from the remaining more abundant sum formulas and the comparability of the compound class distributions, the ESI(-) FT-ICR MS confirms the suitability of the GC×GC HR-ToF-MS to cover the majority of the sulfonates in the APIs.

As the second objective, the comparison should confirm the overall measured chemical profile. For SBS, both techniques show almost only sulfonated species (Figure 5.9 and Figure 5.17 A and D), comparable chemical composition, and, therefore, show the validity of the results [111]. This is not the case for both ABS APIs. While the GC×GC shows high abundances of non-sulfonated compounds, the FT-ICR MS detects CH, CHS, and CHS_2 species only with minimal abundances [116]. This is also a result of the selectivity of the ESI(-) that underrepresents non-polar compounds in a polar matrix [89]. Therefore, this ionization technique is unable to address the entire matrix. As a consequence, APPI was used as a second ionization technique for the FT-ICR MS to also compare the non-sulfonated compounds. With the APPI, the mass spectra show comparable abundance ratios of sulfonates and non-sulfonated species to the GC×GC, as well as similar sum formulas. However, the APPI detected m/z ratios above 500 with relatively high abundances in both ABS matrices, which were neither detected via GC×GC HR-ToF-MS nor ESI(-) [116]. They indicate potential side reactions that cannot be addressed via GC×GC. Therefore, the FT-ICR MS only partially supports the GC×GC results and shows its limitation for both ABS APIs.

The cross-verification with the orthogonal method shows the capacity of the online derivatization GC×GC HR-ToF-MS to detect and identify the vast majority of sulfonates in the APIs. It also covers non-sulfonated residuals from the starting materials to a limited extent or without thermal degradation. Furthermore, it resolves the complexity of the APIs on another dimension, combining the isomeric resolution power of the GC×GC with the ultrahigh mass resolution and accuracy of the FT-ICR MS. However, it also demonstrates the limitation regarding the size and vaporability of the GC×GC for the analysis of both ABS matrices.

Combined with the complementary sample preparations and TGA-MS experiments, the reliability of the measured chemical profile of SBS could be shown. Therefore, the complementary risk-based and the holistic cross-verification via FT-ICR MS validate the GC×GC HR-ToF-MS elucidated chemical profile and the discovered effects from the manufacturing process for SBS and show the possibilities but also the limitations for the analysis of the two ABS matrices.

6 Conclusion & Outlook

Bituminosulfonates, as an example of numerous APIs of complex drugs, have been used in medical treatments for more than 140 years, and regulatory challenges and their chemical complexity have prevented their detailed and reliable chemical description. Using a comprehensive analytical approach, this study designed a holistic analytical workflow for method development and addressed these challenges: The developed online derivatization GC×GC HR-ToF-MS method comprehensively elucidates the qualitative chemical composition of APIs allocable to NBCDs for the first time (S. I); The reliability of the non-targeted analysis is addressed and shown in the method development and in a cross-verification via ESI(-) and APPI FT-ICR MS (S. II); And through measurements of the starting materials and process intermediates, the chemical profile of the APIs was linked to the defining manufacturing process (S. III).

With this approach, the possibly critical parameters of sample preparation, volatility, and thermal stability were identified beforehand. Due to the limited available literature, they were also implemented in the method development. An evaluation of these parameters with different sample preparation methods and TGA-MS shows their minimal effect on the developed online derivatization GC×GC HR-ToF-MS method. In addition to this risk-based approach, the holistic cross-verification with different ionization techniques for FT-ICR MS complements the evaluation. Both approaches significantly improve the confidence of the results and compensate for limited previous studies, similar matrices, and commercially available standard materials.

Based on the chromatographic and mass spectrometric information of the GC×GC HR-ToF-MS and the established reliability of its analytical results, this study elucidates the chemical composition of the bituminosulfonates to a molecular level. Here, not only by their sum formulas but also by their core structural motives were identified. In addition, this study shows the consistency of these core structural motives across the manufacturing process and the influences of the reactivity and steric effects on the sulfonation of the complex mixtures. The depth of the chemical description is the starting point for a chemical evaluation of pharmaceutical effects, which require a chemical description for more than 30 years. Furthermore, with the obtained fingerprint, copycat products can be chemically distinguished. As a consequence, the shown depth of the chemical analyses is need to sufficiently characterize APIs of complex drugs.

However, there are open, analytical challenges and follow-up questions to address: the exact chemical stereochemical identification of isolated isomeric marker compounds, the transfer to a representative QC method with commonly used techniques, and an evaluation of the transferability of this approach and GC×GC method to other complex drugs (NBCDs and biologicals).

Numerous other APIs allocable to NBCDs or biologicals face similar challenges and altered regulatory requirements, similar to REACH, which would require a comparable in-depth chemical description. So far, this study is pioneering, but as regulations continue to change, it has the potential to lead the way and be an example for others to follow.

7 References

1. Groten JP, Feron VJ, Sühnel J. Toxicology of simple and complex mixtures. *Trends Pharmacol Sci.* 2001;22:316–22. doi:10.1016/s0165-6147(00)01720-x.
 2. Groten JP. Mixtures and interactions. *Food Chem Toxicol.* 2000;38:S65-71. doi:10.1016/s0278-6915(99)00135-0.
 3. Manz KE, Feerick A, Braun JM, Feng Y-L, Hall A, Koelmel J, et al. Non-targeted analysis (NTA) and suspect screening analysis (SSA): a review of examining the chemical exposome. *J Expo Sci Environ Epidemiol.* 2023;33:524–36. doi:10.1038/s41370-023-00574-6.
 4. Pomati F, Castiglioni S, Zuccato E, Fanelli R, Vigetti D, Rossetti C, Calamari D. Effects of a complex mixture of therapeutic drugs at environmental levels on human embryonic cells. *Environ Sci Technol.* 2006;40:2442–7. doi:10.1021/es051715a.
 5. Leonti M, Verpoorte R. Traditional Mediterranean and European herbal medicines. *J Ethnopharmacol.* 2017;199:161–7. doi:10.1016/j.jep.2017.01.052.
 6. Tian S, Wang J, Li Y, Xu X, Hou T. Drug-likeness analysis of traditional Chinese medicines: prediction of drug-likeness using machine learning approaches. *Mol Pharm.* 2012;9:2875–86. doi:10.1021/mp300198d.
 7. Crommelin DJA, Shah VP, Klebovich I, McNeil SE, Weinstein V, Flühmann B, et al. The similarity question for biologicals and non-biological complex drugs. *Eur J Pharm Sci.* 2015;76:10–7. doi:10.1016/j.ejps.2015.04.010.
 8. Klein K, Borchard G, Shah VP, Flühmann B, McNeil SE, Vlieger JSB de. A pragmatic regulatory approach for complex generics through the U.S. FDA 505(j) or 505(b)(2) approval pathways. *Ann N Y Acad Sci.* 2021;1502:5–13. doi:10.1111/nyas.14662.
 9. Crommelin DJA, Vlieger JSB de, Weinstein V, Mühlebach S, Shah VP, Schellekens H. Different pharmaceutical products need similar terminology. *AAPS J.* 2014;16:11–4. doi:10.1208/s12248-013-9532-0.
 10. Schellekens H, Klinger E, Mühlebach S, Brin J-F, Storm G, Crommelin DJA. The therapeutic equivalence of complex drugs. *Regul Toxicol Pharmacol.* 2011;59:176–83. doi:10.1016/j.yrtph.2010.09.021.
 11. Sahoo N, Manchikanti P, Dey S. Herbal drugs: standards and regulation. *Fitoterapia.* 2010;81:462–71. doi:10.1016/j.fitote.2010.02.001.
 12. Siddiqui MR, AlOthman ZA, Rahman N. Analytical techniques in pharmaceutical analysis: A review. *Arabian Journal of Chemistry.* 2017;10:S1409-S1421. doi:10.1016/j.arabjc.2013.04.016.
 13. La Torre BG de, Albericio F. The Pharmaceutical Industry in 2019. An Analysis of FDA Drug Approvals from the Perspective of Molecules. *Molecules* 2020. doi:10.3390/molecules25030745.
 14. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). Specification: Test Procedures and Acceptance Criteria for new Drug Products: Chemical Substances Q6A. ICH Quality Guidelines.
 15. Kumar M, Bhatia R, Rawal RK. Applications of various analytical techniques in quality control of pharmaceutical excipients. *J Pharm Biomed Anal.* 2018;157:122–36. doi:10.1016/j.jpba.2018.05.023.
 16. Khagga B, Kaitha MV, Dammu R, Mogili S. ICH guidelines – “Q” series (quality guidelines) - A review. *GSC Biol. and Pharm. Sci.* 2019;6:89–106. doi:10.30574/gscbps.2019.6.3.0034.
 17. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). Validation of Analytical Procedures: Text and Methodology Q2(R2) (Draft Version). ICH Quality Guidelines.
-

References

18. Hussaarts L, Mühlebach S, Shah VP, McNeil S, Borchard G, Flühmann B, et al. Equivalence of complex drug products: advances in and challenges for current regulatory frameworks. *Ann N Y Acad Sci.* 2017;1407:39–49. doi:10.1111/nyas.13347.
 19. Kinch MS. An overview of FDA-approved biologics medicines. *Drug Discov Today.* 2015;20:393–8. doi:10.1016/j.drudis.2014.09.003.
 20. Mühlebach S. Regulatory challenges of nanomedicines and their follow-on versions: A generic or similar approach? *Adv Drug Deliv Rev.* 2018;131:122–31. doi:10.1016/j.addr.2018.06.024.
 21. Gaspar RS, Silva-Lima B, Magro F, Alcobia A, Da Costa FL, Feio J. Non-biological Complex Drugs (NBCDs): Complex Pharmaceuticals in Need of Individual Robust Clinical Assessment Before Any Therapeutic Equivalence Decision. *Front Med (Lausanne).* 2020;7:590527. doi:10.3389/fmed.2020.590527.
 22. Zagalo DM, Simões S, Sousa J. Regulatory Science Approach in Pharmaceutical Development of Follow-on Versions of Non-Biological Complex Drug Products. *J Pharm Sci.* 2022;111:2687–713. doi:10.1016/j.xphs.2022.07.015.
 23. Liu Y-H, Chen Y-S, Tseng T, Jiang M-L, Gau C-S, Chang L-C. Regulatory considerations for generic products of non-biological complex drugs. *J Food Drug Anal.* 2023;31:20–31. doi:10.38212/2224-6614.3441.
 24. Blackman LD, Sutherland TD, Barro PJ de, Thissen H, Locock KES. Addressing a future pandemic: how can non-biological complex drugs prepare us for antimicrobial resistance threats? *Mater Horiz.* 2022;9:2076–96. doi:10.1039/D2MH00254J.
 25. Cauchon NS, Oghamian S, Hassanpour S, Abernathy M. Innovation in Chemistry, Manufacturing, and Controls—A Regulatory Perspective From Industry. *J Pharm Sci.* 2019;108:2207–37. doi:10.1016/j.xphs.2019.02.007.
 26. Griffin JP. A History of Drug Regulation in the UK. In: Griffin JP, Posner J, Barker GR, editors. *The Textbook of Pharmaceutical Medicine*: Wiley; 2013. p. 317–346. doi:10.1002/9781118532331.ch17.
 27. Malheiro V, Duarte J, Veiga F, Mascarenhas-Melo F. Exploiting Pharma 4.0 Technologies in the Non-Biological Complex Drugs Manufacturing: Innovations and Implications. *Pharmaceutics* 2023. doi:10.3390/pharmaceutics15112545.
 28. Zagalo DM, Sousa J, Simões S. Quality by design (QbD) approach in marketing authorization procedures of Non-Biological Complex Drugs: A critical evaluation. *Eur J Pharm Biopharm.* 2022;178:1–24. doi:10.1016/j.ejpb.2022.07.014.
 29. Ballin NZ, Laursen KH. To target or not to target? Definitions and nomenclature for targeted versus non-targeted analytical food authentication. *Trends in Food Science & Technology.* 2019;86:537–43. doi:10.1016/j.tifs.2018.09.025.
 30. Brodie RR, Hill HM. Validation issues arising from the new FDA guidance for industry on bioanalytical method validation. *Chromatographia.* 2002;55:S91-S94. doi:10.1007/BF02493361.
 31. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). Analytical Procedure Development Q14. ICH Quality Guidelines.
 32. Rocco P, Musazzi UM, Franzè S, Minghetti P. Copies of nonbiological complex drugs: generic, hybrid or biosimilar? *Drug Discov Today.* 2019;24:250–5. doi:10.1016/j.drudis.2018.08.003.
 33. Schellekens H, Stegemann S, Weinstein V, Vlieger JSB de, Flühmann B, Mühlebach S, et al. How to regulate nonbiological complex drugs (NBCD) and their follow-on versions: points to consider. *AAPS J.* 2014;16:15–21. doi:10.1208/s12248-013-9533-z.
 34. Koch J, Moser R, Demel J. Analyse des Ammoniumbituminosulfonates Ichthyol®. *Arch. Pharm. Pharm. Med. Chem.* 1985;318:198–206. doi:10.1002/ardp.19853180303.
-

References

35. Sethmann A. Update zur antibiotischen Behandlung unkomplizierter Hautabszesse. *Evid Self Med* 2022. doi:10.52778/efsm.22.0036.
36. Gugenheim PG, McLachlan T, Jones JM. Ichthammol and its pharmacopoeial use. *J. Chem. Technol. Biotechnol.* 1950;69:107–9. doi:10.1002/jctb.5000690406.
37. Unna PG. Aphorismen über Schwefeltherapie und Schwefelpräparate. IV. *Ichthyol. Mon. Prakt. Dermatol.* 1882;1:328–33.
38. Boyd AS. Ichthammol revisited. *Int J Dermatology.* 2010;49:757–60. doi:10.1111/j.1365-4632.2010.04551.x.
39. Warnecke J, Wendt A. Anti-inflammatory action of pale sulfonated shale oil (ICHTHYOL pale) in UVB erythema test. *Inflamm Res.* 1998;47:75–8. doi:10.1007/s000110050282.
40. Schewe C, Schewe T, Rohde E, Diezel W, Czarnetzki BM. Inhibitory effects of sulfonated shale oils (ammonium bituminosulphonates, Ichthyols) on enzymes of polyenoic fatty acid metabolism. *Arch Dermatol Res.* 1994;286:137–41. doi:10.1007/BF00374208.
41. Heuser E, Becker K, Idelevich EA. Bactericidal Activity of Sodium Bituminosulfonate against *Staphylococcus aureus*. *Antibiotics (Basel)* 2022. doi:10.3390/antibiotics11070896.
42. Idelevich EA, Becker K. In Vitro Activity of Sodium Bituminosulfonate: Susceptibility Data for the Revival of an Old Antimicrobial. *Microb Drug Resist.* 2020;26:1405–9. doi:10.1089/mdr.2019.0390.
43. Blisse M, Idelevich EA, Becker K. Investigation of In-Vitro Adaptation toward Sodium Bituminosulfonate in *Staphylococcus aureus*. *Microorganisms* 2020. doi:10.3390/microorganisms8121962.
44. Fink S, Sethmann A, Hipler U-C, Wiegand C. In vitro investigation of the principle of action of ammonium bituminosulfonate ointments on a 3D skin model. *Eur J Pharm Sci.* 2022;172:106152. doi:10.1016/j.ejps.2022.106152.
45. Cholcha W, Leuschner J, Leuschner F. Untersuchungen zur Verträglichkeit von dunklem sulfoniertem Schieferöl nach lokaler und systemischer Applikation. [The adverse effects of dark sulfonated shale oil following local and systemic administration]. *Arzneimittelforschung.* 1994;44:844–9.
46. Cholcha W, Leuschner J, Leuschner F. Untersuchungen zur Verträglichkeit von hellem sulfoniertem Schieferöl nach lokaler und systemischer Applikation. [The tolerability of pale sulfonated shale oil following local and systemic administration]. *Arzneimittelforschung.* 1994;44:170–7.
47. Korting HC, Schöllmann C, Cholcha W, Wolff L. Efficacy and tolerability of pale sulfonated shale oil cream 4% in the treatment of mild to moderate atopic eczema in children: a multicentre, randomized vehicle-controlled trial. *J Eur Acad Dermatol Venereol.* 2010;24:1176–82. doi:10.1111/j.1468-3083.2010.03616.x.
48. Schiffmann S, Gunne S, Henke M, Ulshöfer T, Steinhilber D, Sethmann A, Parnham MJ. Sodium Bituminosulfonate Used to Treat Rosacea Modulates Generation of Inflammatory Mediators by Primary Human Neutrophils. *J Inflamm Res.* 2021;14:2569–82. doi:10.2147/jir.S313636.
49. Beckurts H, Frerichs H. Comparative Investigation of the Composition of Ichthyol-Ammonium and Some of Its Substitutes**Translated in full from *Arch. d. Pharmacie*, 1912, NO. 6 and 7. From the Pharmaceutical Institute of the Herzoglichen Technischen Hochschule in Braunschweig. *The Journal of the American Pharmaceutical Association* (1912). 1913;2:335–43. doi:10.1002/jps.3080020318.
50. Scheibler H. Über die chemischen Bestandteile der schwefelreichen, bituminösen Teeröle (Ichthyolöle). *I. Ber. Dtsch. Chem. Ges.* 1915;48:1815–26. doi:10.1002/cber.19150480276.

References

51. Pailer M, Begutter H. Untersuchung stark schwefelhaltiger Schieferle, 4. Mitt. Monatshefte für Chemie. 1973;104:297–311. doi:10.1007/BF00911171.
52. Pailer M, Grnhaus H. Untersuchung stark schwefelhaltiger Schieferle, 5. Mitt. Monatshefte für Chemie. 1973;104:312–37. doi:10.1007/BF00911172.
53. Pailer M, Berner-Fenz L. Untersuchung stark schwefelhaltiger Schieferle, 6. Mitt. Monatshefte für Chemie. 1973;104:339–51. doi:10.1007/BF00903097.
54. Pailer M, Jiresch W. Über die Basenfraktion eines stark schwefelhaltigen Schieferles. Monatshefte für Chemie. 1969;100:121–31. doi:10.1007/BF00938246.
55. Pailer M, Oesterreicher W, Simonitsch E. Untersuchungen über die Geruchsträger des Seefelder Schieferles. Monatshefte für Chemie. 1965;96:1377–406. doi:10.1007/BF00904289.
56. Pailer M, Simonitsch E. über die Zusammensetzung einer mittleren Siedefraktion des Seefelder Schieferles. Monatshefte für Chemie. 1967;98:1477–91. doi:10.1007/BF00909018.
57. Pailer M, Hlozek V. Untersuchung stark schwefelhaltiger Schieferle, 7. Mitt. Monatshefte für Chemie. 1975;106:1259–84. doi:10.1007/BF00913600.
58. Kostopoulou M, Nikolaou A. Analytical problems and the need for sample preparation in the determination of pharmaceuticals and their metabolites in aqueous environmental matrices. *TrAC Trends in Analytical Chemistry*. 2008;27:1023–35. doi:10.1016/j.trac.2008.09.011.
59. Heywood A, Mathias A, Williams AE. Identification of sulfonic acids and sulfonates by mass spectrometry. *Anal. Chem*. 1970;42:1272–3. doi:10.1021/ac60293a031.
60. Reemtsma T. Methods of analysis of polar aromatic sulfonates from aquatic environments. *Journal of Chromatography A*. 1996;733:473–89. doi:10.1016/0021-9673(95)00738-5.
61. Tegelaar E, Leeuw J de, Derenne S, Largeau C. A reappraisal of kerogen formation. *Geochimica et Cosmochimica Acta*. 1989;53:3103–6. doi:10.1016/0016-7037(89)90191-9.
62. Vandenbroucke M, Largeau C. Kerogen origin, evolution and structure. *Organic Geochemistry*. 2007;38:719–833. doi:10.1016/j.orggeochem.2007.01.001.
63. Sibus JR, Wambaugh JF, Isaacs KK, Williams AJ, McEachran AD, Richard AM, et al. Integrating tools for non-targeted analysis research and chemical safety evaluations at the US EPA. *J Expo Sci Environ Epidemiol*. 2018;28:411–26. doi:10.1038/s41370-017-0012-y.
64. Hartner E, Paul A, Käfer U, Czech H, Hohaus T, Gröger T, et al. On the Complementarity and Informative Value of Different Electron Ionization Mass Spectrometric Techniques for the Chemical Analysis of Secondary Organic Aerosols. *ACS Earth Space Chem*. 2022;6:1358–74. doi:10.1021/acsearthspacechem.2c00039.
65. Dudzik D, Barbas-Bernardos C, García A, Barbas C. Quality assurance procedures for mass spectrometry untargeted metabolomics. a review. *J Pharm Biomed Anal*. 2018;147:149–73. doi:10.1016/j.jpba.2017.07.044.
66. Tistaert C, Dejaegher B, Vander Heyden Y. Chromatographic separation techniques and data handling methods for herbal fingerprints: a review. *Anal Chim Acta*. 2011;690:148–61. doi:10.1016/j.aca.2011.02.023.
67. Kharbach M, Marmouzi I, El Jemli M, Bouklouze A, Vander Heyden Y. Recent advances in untargeted and targeted approaches applied in herbal-extracts and essential-oils fingerprinting - A review. *J Pharm Biomed Anal*. 2020;177:112849. doi:10.1016/j.jpba.2019.112849.
68. Hulleman T, Turkina V, O'Brien JW, Chojnacka A, Thomas KV, Samanipour S. Critical Assessment of the Chemical Space Covered by LC-HRMS Non-Targeted Analysis. *Environ Sci Technol*. 2023;57:14101–12. doi:10.1021/acs.est.3c03606.

References

69. Gao B, Holroyd SE, Moore JC, Laurvick K, Gendel SM, Xie Z. Opportunities and Challenges Using Non-targeted Methods for Food Fraud Detection. *J Agric Food Chem.* 2019;67:8425–30. doi:10.1021/acs.jafc.9b03085.
70. Blomberg J, Schoenmakers PJ, Beens J, Tijssen R. Comprehensive two-dimensional gas chromatography (GC×GC) and its applicability to the characterization of complex (petrochemical) mixtures. *J. High Resol. Chromatogr.* 1997;20:539–44. doi:10.1002/jhrc.1240201005.
71. Górecki T, Harynuk J, Panić O. The evolution of comprehensive two-dimensional gas chromatography (GC x GC). *J Sep Sci.* 2004;27:359–79. doi:10.1002/jssc.200301650.
72. Tranchida PQ, Franchina FA, Dugo P, Mondello L. Comprehensive two-dimensional gas chromatography-mass spectrometry: Recent evolution and current trends. *Mass Spectrom Rev.* 2016;35:524–34. doi:10.1002/mas.21443.
73. Byer JD, Siek K, Jobst K. Distinguishing the C3 vs SH4 Mass Split by Comprehensive Two-Dimensional Gas Chromatography-High Resolution Time-of-Flight Mass Spectrometry. *Anal. Chem.* 2016;88:6101–4. doi:10.1021/acs.analchem.6b01137.
74. Schwalb L, Tiemann O, Käfer U, Gröger T, Rüger CP, Gayko G, Zimmermann R. Analysis of complex drugs by comprehensive two-dimensional gas chromatography and high-resolution mass spectrometry: detailed chemical description of the active pharmaceutical ingredient sodium bituminosulfonate and its process intermediates. *Anal Bioanal Chem.* 2023;415:2471–81. doi:10.1007/s00216-022-04393-w.
75. Ng B, Quinete N, Gardinali PR. Assessing accuracy, precision and selectivity using quality controls for non-targeted analysis. *Sci Total Environ.* 2020;713:136568. doi:10.1016/j.scitotenv.2020.136568.
76. Riedl J, Esslinger S, Faulstich C. Review of validation and reporting of non-targeted fingerprinting approaches for food authentication. *Anal Chim Acta.* 2015;885:17–32. doi:10.1016/j.aca.2015.06.003.
77. Hollender J, Schymanski EL, Ahrens L, Alygizakis N, Béen F, Bijlsma L, et al. NORMAN guidance on suspect and non-target screening in environmental monitoring. *Environ Sci Eur* 2023. doi:10.1186/s12302-023-00779-4.
78. American Society for Testing and Materials. ASTM D8396-22: Standard Test Method for Group Types Quantification of Hydrocarbons in Hydrocarbon Liquids with a Boiling Point between 36 °C and 343 °C by Flow Modulated GCxGC-FID. 2022. <https://www.astm.org/d8396-22.html>. Accessed 23 Nov 2023.
79. Basics of Mass Spectrometry. In: Dass C, editor. *Fundamentals of Contemporary Mass Spectrometry*; Wiley; 2007. p. 1–14. doi:10.1002/9780470118498.ch1.
80. Glish GL, Vachet RW. The basics of mass spectrometry in the twenty-first century. *Nat Rev Drug Discov.* 2003;2:140–50. doi:10.1038/nrd1011.
81. Marshall AG, Hendrickson CL. High-resolution mass spectrometers. *Annu Rev Anal Chem (Palo Alto Calif).* 2008;1:579–99. doi:10.1146/annurev.anchem.1.031207.112945.
82. Reemtsma T. Determination of molecular formulas of natural organic matter molecules by (ultra-) high-resolution mass spectrometry: status and needs. *J Chromatogr A.* 2009;1216:3687–701. doi:10.1016/j.chroma.2009.02.033.
83. Xian F, Hendrickson CL, Marshall AG. High resolution mass spectrometry. *Anal. Chem.* 2012;84:708–19. doi:10.1021/ac203191t.
84. Awad H, Khamis MM, El-Aneed A. Mass Spectrometry, Review of the Basics: Ionization. *Applied Spectroscopy Reviews.* 2015;50:158–75. doi:10.1080/05704928.2014.954046.

References

85. Dawson JHJ, Guilhaus M. Orthogonal-acceleration time-of-flight mass spectrometer. *Rapid Comm Mass Spectrometry*. 1989;3:155–9. doi:10.1002/rcm.1290030511.
86. Coles J, Guilhaus M. Orthogonal acceleration — a new direction for time-of-flight mass spectrometry: Fast, sensitive mass analysis for continuous ion sources. *TrAC Trends in Analytical Chemistry*. 1993;12:203–13. doi:10.1016/0165-9936(93)80021-B.
87. Plaß WR, Dickel T, Scheidenberger C. Multiple-reflection time-of-flight mass spectrometry. *International Journal of Mass Spectrometry*. 2013;349-350:134–44. doi:10.1016/j.ijms.2013.06.005.
88. Gröger TM, Käfer U, Zimmermann R. Gas chromatography in combination with fast high-resolution time-of-flight mass spectrometry: Technical overview and perspectives for data visualization. *TrAC Trends in Analytical Chemistry*. 2020;122:115677. doi:10.1016/j.trac.2019.115677.
89. Modes of Ionization. In: Dass C, editor. *Fundamentals of Contemporary Mass Spectrometry*: Wiley; 2007. p. 15–65. doi:10.1002/9780470118498.ch2.
90. Maciel EVS, Pereira Dos Santos NG, Vargas Medina DA, Lanças FM. Electron ionization mass spectrometry: Quo vadis? *Electrophoresis*. 2022;43:1587–600. doi:10.1002/elps.202100392.
91. Cho Y, Ahmed A, Islam A, Kim S. Developments in FT-ICR MS instrumentation, ionization techniques, and data interpretation methods for petroleomics. *Mass Spectrom Rev*. 2015;34:248–63. doi:10.1002/mas.21438.
92. Sneddon J, Masuram S, Richert JC. *Gas Chromatography-Mass Spectrometry-Basic Principles, Instrumentation and Selected Applications for Detection of Organic Compounds*. Analytical Letters. 2007;40:1003–12. doi:10.1080/00032710701300648.
93. Ristic ND, Djokic MR, Konist A, van Geem KM, Marin GB. Quantitative compositional analysis of Estonian shale oil using comprehensive two dimensional gas chromatography. *Fuel Processing Technology*. 2017;167:241–9. doi:10.1016/j.fuproc.2017.07.008.
94. Lorentz C, Laurenti D, Zotin JL, Geantet C. Comprehensive GC × GC chromatography for the characterization of sulfur compound in fuels: A review. *Catalysis Today*. 2017;292:26–37. doi:10.1016/j.cattod.2017.04.052.
95. Bahaghighat HD, Freye CE, Synovec RE. Recent advances in modulator technology for comprehensive two dimensional gas chromatography. *TrAC Trends in Analytical Chemistry*. 2019;113:379–91. doi:10.1016/j.trac.2018.04.016.
96. Marshall AG, Rodgers RP. Petroleomics: chemistry of the underworld. *Proc Natl Acad Sci U S A*. 2008;105:18090–5. doi:10.1073/pnas.0805069105.
97. Jennerwein MK, Eschner M, Gröger T, Wilharm T, Zimmermann R. Complete Group-Type Quantification of Petroleum Middle Distillates Based on Comprehensive Two-Dimensional Gas Chromatography Time-of-Flight Mass Spectrometry (GC×GC-TOFMS) and Visual Basic Scripting. *Energy Fuels*. 2014;28:5670–81. doi:10.1021/ef501247h.
98. Coats AW, Redfern JP. Thermogravimetric analysis. A review. *Analyst*. 1963;88:906–24.
99. Blasio C de, editor. *Fundamentals of Biofuels Engineering and Technology*. Cham: Springer International Publishing; 2019.
100. Moldoveanu SC. Pyrolysis of Other Compounds With Heteroatoms. In: *Pyrolysis of Organic Molecules*: Elsevier; 2019. p. 763–780. doi:10.1016/B978-0-444-64000-0.00017-2.
101. Kaisersberger E, Post E. Practical aspects for the coupling of gas analytical methods with thermal-analysis instruments. *Thermochimica Acta*. 1997;295:73–93. doi:10.1016/S0040-6031(97)00099-3.

References

102. Käfer U, Gröger T, Rüger CP, Czech H, Saraji-Bozorgzad M, Wilharm T, Zimmermann R. Direct inlet probe - High-resolution time-of-flight mass spectrometry as fast technique for the chemical description of complex high-boiling samples. *Talanta*. 2019;202:308–16. doi:10.1016/j.talanta.2019.05.030.
103. Demeestere K, Dewulf J, Witte B de, van Langenhove H. Sample preparation for the analysis of volatile organic compounds in air and water matrices. *Journal of Chromatography A*. 2007;1153:130–44. doi:10.1016/j.chroma.2007.01.012.
104. Andrade-Eiroa A, Canle M, Leroy-Cancellieri V, Cerdà V. Solid-phase extraction of organic compounds: A critical review (Part I). *TrAC Trends in Analytical Chemistry*. 2016;80:641–54. doi:10.1016/j.trac.2015.08.015.
105. Andrade-Eiroa A, Canle M, Leroy-Cancellieri V, Cerdà V. Solid-phase extraction of organic compounds: A critical review. part ii. *TrAC Trends in Analytical Chemistry*. 2016;80:655–67. doi:10.1016/j.trac.2015.08.014.
106. Lin D-L, Wang S-M, Wu C-H, Chen B-G, Liu RH. Chemical derivatization for the analysis of drugs by GC-MS - A conceptual review. *J Food Drug Anal* 2008. doi:10.38212/2224-6614.2373.
107. Kühnel E, Laffan DDP, Lloyd-Jones GC, Martínez Del Campo T, Shepperson IR, Slaughter JL. Mechanism of methyl esterification of carboxylic acids by trimethylsilyldiazomethane. *Angew Chem Int Ed Engl*. 2007;46:7075–8. doi:10.1002/anie.200702131.
108. He Y, Buch A, Szopa C, Williams AJ, Millan M, Guzman M, et al. The search for organic compounds with TMAH thermochemolysis: From Earth analyses to space exploration experiments. *TrAC Trends in Analytical Chemistry*. 2020;127:115896. doi:10.1016/j.trac.2020.115896.
109. International Organisation for Standardisation. ISO 9001:2015(en) - Quality management systems — Requirements. ISO Standards.
110. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). Quality Risk Management Q9(R1). ICH Quality Guidelines.
111. Schwalb L, Tiemann O, Käfer U, Rüger CP, Gröger T, Zimmermann R. Applying a risk assessment guided evaluation for verifying comprehensive two-dimensional gas chromatography to analyse complex pharmaceuticals. *Anal Bioanal Chem*. 2024;416:1033–45. doi:10.1007/s00216-023-05093-9.
112. Lu H, Shi Q, Lu J, Sheng G, Peng P, Hsu CS. Petroleum Sulfur Biomarkers Analyzed by Comprehensive Two-Dimensional Gas Chromatography Sulfur-Specific Detection and Mass Spectrometry. *Energy Fuels*. 2013;27:7245–51. doi:10.1021/ef401239u.
113. Käfer U, Gröger T, Rohbogner CJ, Struckmeier D, Saraji-Bozorgzad MR, Wilharm T, Zimmermann R. Detailed Chemical Characterization of Bunker Fuels by High-Resolution Time-of-Flight Mass Spectrometry Hyphenated to GC × GC and Thermal Analysis. *Energy Fuels*. 2019;33:10745–55. doi:10.1021/acs.energyfuels.9b02626.
114. Nehlsen J, Benziger J, Kevrekidis I. Oxidation of Aliphatic and Aromatic Sulfides Using Sulfuric Acid. *Ind. Eng. Chem. Res*. 2006;45:518–24. doi:10.1021/ie050901m.
115. Shoji H, Majima K. Investigations on the side reactions in sulfonation reactions-hydrophobic compounds in the unsulfonated oil fraction. *J Americ Oil Chem Soc*. 1963;40:179–83. doi:10.1007/BF02632576.
116. Tiemann O, Rüger CP, Schwalb L, Chacón-Patiño ML, Gröger T, Zimmermann R. Rock2Skin: Characterization of Shale Oil-Based Non-Biological Complex Drugs Along the Production Process by High-Resolution Mass Spectrometry. *Analytical Chemistry* (submitted).
-

References

117. Diya'uddeen BH, Daud WMAW, Abdul Aziz AR. Treatment technologies for petroleum refinery effluents: A review. *Process Safety and Environmental Protection*. 2011;89:95–105. doi:10.1016/j.psep.2010.11.003.
118. Hsu CS, Robinson PR. Petroleum Distillation. In: Hsu CS, Robinson PR, editors. *Springer Handbook of Petroleum Technology*. Cham: Springer International Publishing; 2017. p. 533–550. doi:10.1007/978-3-319-49347-3_16.
119. Katritzky AR, Kim MS, Fedoseyenko D, Widyan K, Siskin M, Francisco M. The sulfonation of aromatic and heteroaromatic polycyclic compounds. *Tetrahedron*. 2009;65:1111–4. doi:10.1016/j.tet.2008.11.023.
120. Basu P. *Biomass gasification, pyrolysis and torrefaction: Practical design and theory*. 3rd ed. London, United Kingdom: Academic Press imprint of Elsevier; 2018.
121. Rojas-Ruiz FA, Gómez-Escudero A, Pachón-Contreras Z, Villar-García Á, Orrego-Ruiz JA. Detailed Characterization of Petroleum Sulfonates by Fourier Transform Ion Cyclotron Resonance Mass Spectrometry. *Energy Fuels*. 2016;30:2714–20. doi:10.1021/acs.energyfuels.5b02923.

List of Figures/Tables

Introduction:

Figure 2.1 Classification of drugs and their active pharmaceutical ingredients (APIs).....	3
Figure 2.2 Production process of the Bituminosulfonates.	5
Figure 2.3 Key dates for the development of bituminosulfonates.	6
Figure 2. 4 Enclosure of sulfur-rich sediments during the formation of the Alps.	7

Methods & Instrumentations:

Figure 4.1 Basic composition of a mass spectrometer.....	10
Figure 4.2 Schematic illustration of high-resolution multi-reflecting time-of-flight mass spectrometer with an electron ionization source.	12
Figure 4.3 Measured electron ionization fragmentation pattern of ortho-toluenesulfonate methyl ester.	13
Figure 4.4 Illustration of the instrumental setup of a thermally modulated comprehensive two-dimensional gas chromatograph (GC×GC).....	14
Figure 4.5 Comparison of a gas chromatography (GC) separation (left) and a comprehensive two-dimensional gas chromatograph (GC×GC) (right).	15
Figure 4.6 Illustration of the operation principle of a quad jet cryogenic modulator.	15
Figure 4.7 Data visualization of a comprehensive two-dimensional gas chromatography high-resolution time-of-flight mass spectrometry (GC×GC-HR-ToF-MS) data set.	16
Figure 4.8 Illustration of a comprehensive two-dimensional gas chromatograph (GC×GC) and thermobalance (see 4.2.2) hyphenated with an electron ionization high-resolution time-of-flight mass spectrometer (EI-HR-ToF-MS).....	17
Figure 4.9 Setup and data visualization in thermogravimetric analysis (TGA).....	18
Figure 4.10 Chemical equation of the reaction of sulfonic acids with trimethylsilyl diazomethane (TMSDAM).	20
Figure 4.11 Chemical equation of the reaction of sulfonates with tetramethylammonium hydroxide (TMAH).	20

Results & Discussion

Figure 5.1 Ishikawa diagram of influencing parameters for a qualitative analysis of bituminosulfonates via comprehensive two-dimensional gas chromatograph (GC×GC-HR-ToF-MS).	22
Figure 5.2 Analytical workflow for the development and verification of a comprehensive two-dimensional gas chromatograph (GC×GC) analysis.	23
Figure 5.3 Measuring plan derived from the analytical workflow.....	24
Figure 5.4 Chemical composition of the shale oil and its distillation products.	25
Figure 5.5 Double bound equivalent (DBE) vs. Carbon number plots of the shale oil and its distillation products.....	26
Figure 5.6 Statistical comparison of the light distillate and precursor distillate.	27
Figure 5.7 Thermogram of the measurement from the sodium bituminosulfonate (SBS).....	28
Figure 5.8 Thermogram of the Benzenesulfonates.	29

References

Figure 5.9 Chemical composition of sodium bituminosulfonate (SBS).	30
Figure 5.10 Core structures and derivatives as they occur in the manufacturing process and analysis	31
Figure 5.11 Comparison of the sulfonated species with different sample preparation methods.....	32
Figure 5.12 Double bound equivalent (DBE) vs. carbon number plot of the CHO_3S_2 species after the liquid-liquid extraction (L/L) and derivatization via trimethylsilyl diazomethane (TMSDAM) of sodium bituminosulfonate (SBS).....	33
Figure 5.13 Chemical composition of ammonium bituminosulfonate of the middle distillate (ABS).	34
Figure 5.14 Double bound equivalent (DBE) vs carbon number plot of the detected CH, CHS, and CHS_2 classes in ammonium bituminosulfonate (ABS).....	35
Figure 5.15 Chemical composition of ammonium bituminosulfonate of the combined distillation phases (ABS (mix)).	36
Figure 5.16 Double bound equivalent (DBE) vs carbon number plot of the GC×GC detected CH, CHS, and CHS_2 classes in ammonium bituminosulfonate of the combined distillation phases (ABS (mix)).	37
Figure 5.17 Chemical composition of bituminosulfonates.	38

Tables

Table 1 Typical properties of active pharmaceutical ingredients (APIs).....	3
Table 2 List of the isotopes	10

-

Curriculum Vitae

LUKAS FELIX SCHWALB

Personal data

Date of birth: 24.08.1994

Place of birth: Bad Godesberg now Bonn

Nationality: German

Contact details

Address: Mauthäuslstr. 9, 81379 München

Email: lukas.schwalb@helmholtz-munich.de

Education

Since 10/2020	PhD student, Chair for Analytical Chemistry, University of Rostock
10/2018 – 09/2020	Master of Science, Analytical Chemistry and Quality Assurance, University of Applied Science Bonn-Rhein-Sieg
09/2013 – 07/2018	Bachelor of Science Pharmaceutical Chemistry, Technical University of Applied Science Cologne
08/2005 – 06/2013	Abitur, Friedrich-Ebert-Gymnasium, Bonn

Employment History

Since 10/2020	Scientist, University of Rostock, Chair of Analytical Chemistry, Joint Mass Spectrometry Center (University Rostock & Helmholtz Munich), Munich
03/2020 – 10/2020	Master Thesis, Merck KGaA, Darmstadt
03/2018 – 10/2018	Chemist at Advanced Accelerator Applications, Bonn
09/2017 – 03/2018	Bachelor Thesis, Federal Institute for Drugs and Medical Devices, Bonn

Lukas Felix Schwalb

Rostock,

Conference Participations

Oral Presentation, 55th Annual DGMS Conference 2024, Freising, Germany

Targeting Complex Drugs: Complementary High-Resolution Mass Spectrometric Techniques for Detailed Chemical Description of Active Pharmaceutical Ingredients

Oral Presentation, DGMS interest group FT-MS and High-Resolution MS 2023, Rostock, Germany

Addressing the challenging alignment of two orthogonal high-resolution mass spectrometry techniques for the complementary chemical characterization of complex pharmaceuticals

Oral Presentation, ANAKON 2023, Vienna, Austria

A new approach for the comprehensive chemical description of complex pharmaceutical products

Oral Presentation, 33 Doktorandenseminar des Arbeitskreises Separation Science der GDCh-Fachgruppe Analytische Chemie, Hohenroda, Germany

Targeting the Analytical Gap in the Chemical Description of Complex Drugs

Poster Presentation, Multidimensional Chromatography Workshop 2023, Liege, Belgium

Adding discrimination dimensions with the application of selective ionization for two-dimensional comprehensive gas chromatography in the analysis of complex Drugs

Poster with Flash Presentation, GC×GC Symposium 2022, Online

Overcoming the analytical gap in the chemical description of complex drugs

Flash Presentation, Multidimensional Chromatography Workshop 2022, Online

Sulfonated shale oil distillate: A versatile API for pharmaceutical products but a challenging matrix for a comprehensive chemical characterization

Publications

Publication 1

Analysis of complex drugs by comprehensive two-dimensional gas chromatography and high-resolution mass spectrometry: detailed chemical description of the active pharmaceutical ingredient sodium bituminosulfonate and its process intermediates

Schwalb, L., Tiemann, O., Käfer, U., Gröger, T., Rüger, C. P., Gayko, G. and Zimmermann, R.

Anal Bioanal Chem (2023). <https://doi.org/10.1007/s00216-022-04393w>

Publication 2

Applying a risk assessment guided evaluation for verifying comprehensive two-dimensional gas chromatography to analyse complex pharmaceuticals

Schwalb, L., Tiemann, O., Käfer, U., Rüger, C. P., Gröger, T. and Zimmermann, R.

Anal Bioanal Chem (2023). <https://doi.org/10.1007/s00216-023-05093-9>

Publication 3 (Submitted)

Rock-to-Pharma: Characterization of Shale Oil-Based Non-Biological Complex Drugs Along the Production Process by High-Resolution Mass Spectrometry

Tiemann, O.; Rüger, C. P.; Schwalb, L.; Chacón-Patiño, M. L.; Gröger, T.; Zimmermann, R.

submitted to ACS Analytical Chemistry (08.03.2024)