MASTERARBEIT | MASTER'S THESIS

Titel | Title

Development of a new glycolipid database for high-resolution mass spectrometry workflows

verfasst von | submitted by

Amanda Pantucek BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Master of Science (MSc)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 066 862
Studienrichtung lt. Studienblatt | Degree programme as it appears on the student record sheet:	Masterstudium Chemie
Betreut von | Supervisor:	Ass.-Prof. Dipl.-Ing. Evelyn Rampler Privatdoz. Bakk. PhD

Für Kastalia

Acknowledgement

I would like to express my sincere gratitude to my supervisor, Evelyn Rampler, for her continuous guidance, encouragement, and support throughout the course of my master's thesis. Her expertise and dedication have been invaluable to my academic and personal growth.

My heartfelt thanks also go to my co-supervisor, Katharina Hohenwallner, for her insightful feedback, patience, and willingness to help whenever I needed direction. Her support has greatly enriched this work.

I am grateful to the Rampler Group and the Köllensperger Group for their constant readiness to help, the motivating discussions, and the countless lunch breaks that brightened even the busiest days. Being part of such supportive and welcoming teams has meant a great deal to me.

I would also like to thank my husband, Roland, for his unwavering love, patience, and encouragement. My deepest appreciation goes to my family—my mother, father, and brother—for always believing in me and supporting me in every step of my academic journey.

A special thank you goes to my university colleagues who became friends: Sebi, who has been around since the first year even though we eventually lost him to organic chemistry; Philipp, whose academic path often ran parallel to my own; and Philipp, with whom boat trips and Spritzerstände were always memorable and fun. I am also grateful to my friends outside the university for providing much-needed distractions, balance, and unwavering support throughout this journey.

Finally, I extend my gratitude to everyone who contributed to this journey, directly or indirectly. This thesis would not have been possible without your support

Abstract

Glycosphingolipids (GSLs) are structurally complex lipids with key biological functions in membranes and signaling. Their diversity along with the presence of numerous isomeric species, poses significant challenges for reliable identification by mass spectrometry (MS).

This thesis presents the systematic development of a spectral library specifically tailored for GSLs to support accurate and reproducible annotation in lipidomics. High-quality MS/MS spectra were collected, curated, and organized to cover a representative range of GSL classes.

The resulting library provides a robust foundation for rapid spectrum matching, enabling more consistent identification of known compounds in complex biological datasets. While limited by its current coverage and inability to capture novel or rare species, the curated spectral library addresses an important gap in lipidomics resources. Moreover, this work highlights the need for ongoing expansion using authentic standards and in silico-generated spectra to broaden coverage. Integrating spectral libraries with complementary annotation strategies, such as decision rule–based approaches, will be essential for achieving a more comprehensive understanding of GSL diversity.

Overall, this thesis establishes a dedicated GSL spectral library as a valuable resource for the lipidomics community and lays the groundwork for future methodological and biomedical applications.

Zusammenfassung

Glycosphingolipide (GSLe) bilden eine heterogene Klasse komplexer Membranlipide, die vielfältige biologische Funktionen übernehmen. Aufgrund ihrer strukturellen Vielfalt und zahlreicher isomerer Varianten stellt ihre eindeutige Identifizierung in massenspektrometrischen Analysen jedoch eine große Herausforderung dar.

Im Rahmen dieser Arbeit wurde erstmals eine Spektralbibliothek speziell für GSLe erstellt. Dazu wurden hochauflösende MS/MS-Spektren gesammelt, aufbereitet und in einer Datenbank zusammengeführt, die verschiedene zentrale GSL-Klassen abdeckt. Diese Bibliothek ermöglicht einen schnellen und verlässlichen Spektrenabgleich und trägt so zu einer konsistenteren Annotation bekannter Verbindungen in komplexen lipidomischen Datensätzen bei.

Trotz dieser Fortschritte bleibt der Ansatz durch die aktuelle Abdeckung der Bibliothek begrenzt: neuartige oder seltene GSL-Strukturen können mit reinen Bibliotheksmethoden nicht erfasst werden. Daher unterstreicht die Arbeit die Bedeutung einer kontinuierlichen Erweiterung der Datenbank durch authentische Standards sowie qualitativ hochwertige in silico generierte Spektren. Langfristig erscheint zudem die Kombination von Spektralbibliotheken mit regelbasierten Strategien besonders vielversprechend, um die strukturelle Diversität von GSLen umfassend abzubilden.

Mit der hier präsentierten GSL-Spektralbibliothek steht der Lipidomik-Gemeinschaft ein neues Werkzeug zur Verfügung, das nicht nur methodische Entwicklungen unterstützt, sondern auch die Grundlage für zukünftige biologische und biomedizinische Anwendungen schafft.

List of Abbreviations

ACN	acetonitrile
AGC	automatic gain control
BCA	bicinchoninic acid
CD	Compound Discoverer
CI	chemical ionization
CID	collision-induced dissociation
CRC	colorectal cancer
DDA	data-dependent acquisition
EI	electron impact ionisation
ESI	electrospray-ionization
FA	fatty acid
FAB	fast atom bombardment
GSL	glycosphingolipid
HESI	heated electrospray-ionization
HPLC	high-performance liquid chromatography
IPA	isopropanol
LC	liquid chromatography
LMWC	low molecular weight compound
MALDI	Matrix-Assisted Laser Desorption/Ionization
MIT	maximum injection time
MS	mass spectrometry
MS/MS	tandem mass spectrometry
MTBE	methyl-tert-butyl ether
NCE	normalized collision energy
nGSL	neutral glycosphingolipid mixture
PGS	pooled glycosphingolipid standard
RF	radio frequency
RT	retention time
SPB	sphingoid base
TGE	total ganglioside extract
TLC	thin-layer chromatography
TOF	Time-of-flight
TSI	thermospray ionization
UHPLC	ultra-high-performance liquid chromatography
UVPD	UV photodissociation

Contents

1 Introduction

Gangliosides play a crucial structural role in shaping the plasma membrane. They participate in a wide range of biological processes, including cell-cell communication, cellular growth, host-pathogen interactions, and signal transduction. In lipidomics mass spectrometry (MS) has become a central analytical tool. Libraries offer a straightforward solution for lipid identification; however, no MS libraries currently exist for glycosphingolipids. The aim of this thesis is to establish a high-quality spectral library for GSLs, using both UV photodissociation (UVPD), and collision-induced dissociation (CID) data. The following chapters provide background on lipids and lipidomics, explain the principles of mass spectrometry, and discuss the rationale for building a dedicated glycosphingolipid library.

1.1 Lipids

Lipids are water-insoluble biomolecules that dissolve readily in organic solvents such as chloroform. They fulfill a wide variety of biological roles, including serving as fuel molecules, concentrated energy stores, signaling mediators, and essential components of membranes. [1] Lipids are hydrophobic or amphiphatic small molecules [2] as seen in Figure 1.

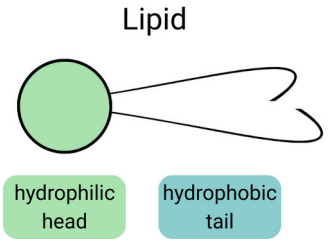

Figure 1: Schematic representation of a lipid to highlight its amphiphilic character.

Lipids can be classified on the basis of either structural or functional criteria. Structural classification reflects the fact that many lipids are composed of recurring building blocks, which define the corresponding lipid class. Following the recommendations of the International Committee for Lipid Classification and Nomenclature, a comprehensive classification system divides biologically relevant lipids into eight classes:
fatty acids (FA), glycerolipids (GL), glycerophospholipids (GP), sphingolipids (SP), sterols (ST), prenols (PR), saccharolipids (SL), and polyketides (PK).
While this system is extensive, it also has limitations: many lipid molecules can belong to multiple classes. For instance, glycerophospholipids form a subset of glycerolipids. An alternative classification, as seen in Figure2 groups lipids into three broad categories - fatty acid derivatives, isoprene derivatives, and polyketides - each of which can be further divided into subclasses. [8]

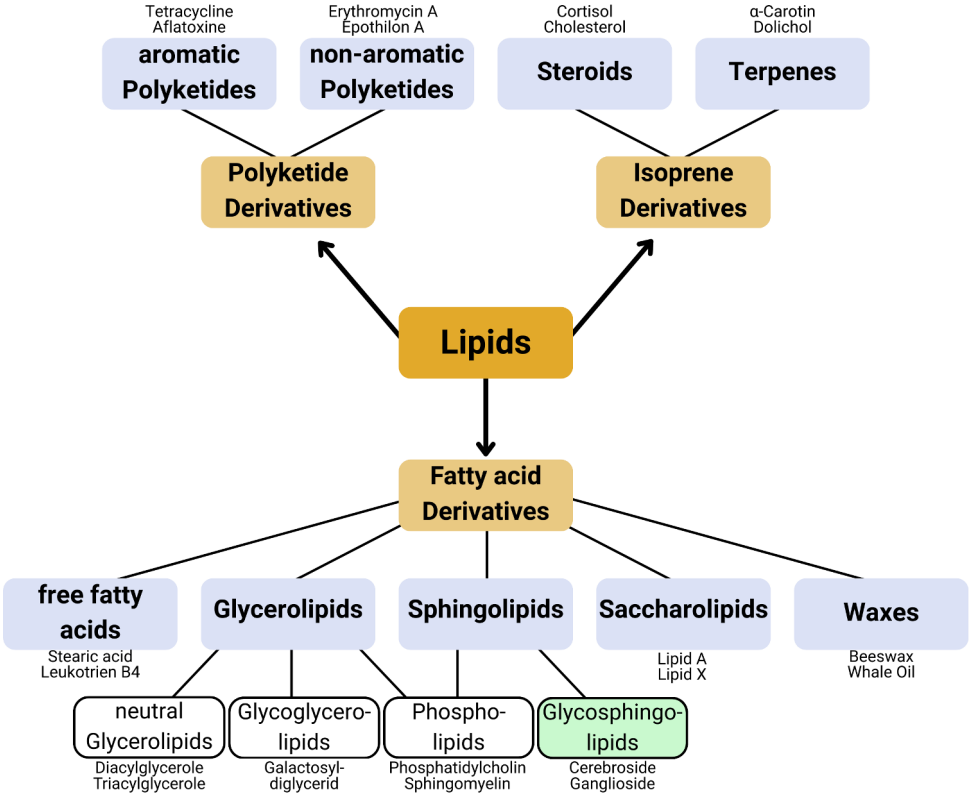

Figure 2: Classification of biologically relevant lipids [8], modified.

Because the goal of this thesis is to construct a mass spectrometry library for glycosphingolipids, the following sections focus on glycolipids, with particular emphasis on glycosphingolipids.

1.1.1 Glycolipids

A major class of membrane lipids is formed by carbohydrate-containing lipids, collectively known as glycolipids. In animal cells, glycolipids are derived from sphingosine, similar to sphingomyelin. In these molecules, the amino group of the sphingosine backbone is acylated with a fatty acid, while the primary hydroxyl group carries one or more carbohydrate residues instead of phosphorylcholine. Glycolipids are asymmetrically distributed within the plasma membrane: the carbohydrate moieties are always located on the extracellular side. [1]

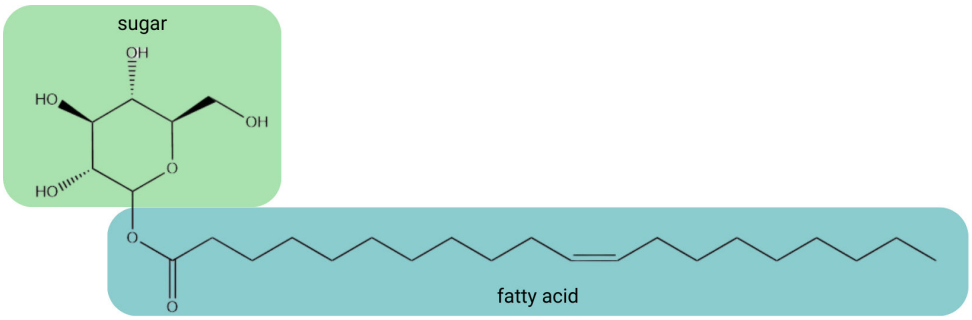

Figure 3: Exemplary representation of simple glycolipid.

1.1.2 Glycosphingolipids

Glycosphingolipids (GSLs) are a subclass of glycolipids with ceramide backbones, consisting of sphingoid bases N-acylated with long- or very-long-chain fatty acids. GSLs are involved in diverse cellular processes, including membrane organization and signaling. However, compared with other lipid classes, analytical tools for studying GSLs remain relatively underdeveloped, highlighting the need for advanced methodologies to characterize their structures, functions in health and disease.

Structure and Characterization

GSLs consist of a hydrophobic ceramide backbone linked to one or more hydrophilic glycan chains. This amphiphilic structure is essential for their role within membranes.[12] [9]

They display extensive structural variation, including differences in alkyl chain length, number and position of double bonds, and hydroxylation. The shorthand nomenclature for sphingoid bases indicates the number of hydroxyl groups (m = mono-, d = di-, t = tri-), the number of carbon atoms, and the number of double bonds (e.g., 18:1 for sphingosine). For example, GM1 36:1;O2 denotes a glycosphingolipid with 36 carbon atoms, one double bond, and two oxygen atoms.

Levels of structural resolution in GSL analysis are summarized in Table 1.

Table 1: Summary of structure resolution levels.

Structure resolution level	Evidence	Example
Unknown GSL fine structure	MS, Enzymatic assay	m/z 1565.90
Partially resolved GSL	—	C73H131N3O31; GM1 36:1;O2/18:0; GM1 36:1;O2(4E)/18:0
Assumed GSL	high mass accuracy MS1, MS2, CCS, pos/neg matching, RT dependencies	Gal(β1-3)GalNAc(β1-4)[Neu5Ac(α2-3)]Gal(β1-4)Glc(β1-1)Cer(18:1(4E)/18:0)

Ceramide fatty acids typically range from 14 to 26 carbons. The large number of possible combinations of sphingoid bases and fatty acids gives rise to substantial diversity among GSL species. Complex GSLs are generated by attaching carbohydrate head groups to carbon 1 of the ceramide. Major subcategories include glucosylceramides, galactosylceramides, and ceramide phosphoinositols. These lipids are arranged asymmetrically in membranes, with carbohydrate residues exposed extracellularly, where they contribute to cell interactions and membrane microdomain formation.

Function and Classification

GSLs carry out diverse functions through interactions with cellular and extracellular proteins, regulation of immune processes, and roles in intracellular trafficking. Some GSLs undergo metabolic processing, such as desialylation, to fulfill their biological functions. [10]

Based on their charge, GSLs are broadly divided into neutral and acidic species. Neutral GSLs contain carbohydrate residues but lack negatively charged groups. Examples include galactocerebroside, lactosylceramide, and globotriaosylceramide. According to their oligosaccharide cores, neu-

tral GSLs can be further grouped into ganglio-, globo-, and lacto-series (types 1 and 2). Acidic GSLs carry negatively charged substituents. They comprise two major groups: gangliosides, which contain sialic acid residues, and sulfatides, which contain sulfate groups.
Gangliosides are typically derived from the ganglio-series but can also arise from globo- or lacto-series precursors. This structural distinction underlies their diverse biological roles, including contributions to membrane organization, cell adhesion, receptor modulation, and signal transduction. [7]
The general structure of a ganglioside is shown in Figure 4.

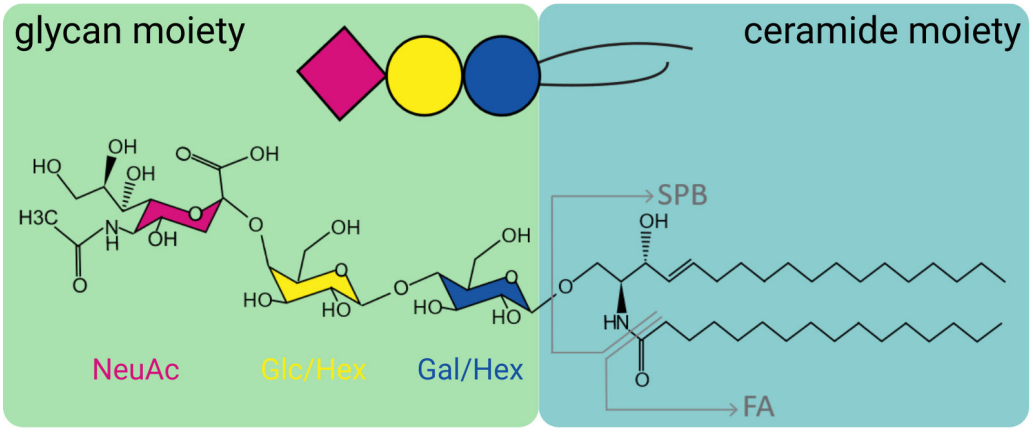

Figure 4: General structure of ganglioside GM3, consisting of one sialic acid (NeuAc), two hexoses (Gal and Glc), and a ceramide backbone (sphingoid base and fatty acid).

1.2 Lipidomics

The rise of lipidomics was driven by advances in genomics and proteomics, which emphasized the need to study metabolism at a systems level. Since its emergence in the late 1990s, lipidomics has focused on characterizing lipid molecular species and their roles as cellular messengers. The methodology relies primarily on mass spectrometry to profile, detect, and identify lipid species. With technological progress, lipidomics has expanded to cover a wide diversity of lipid structures. Applications include: determining phospholipid compositions in various organisms, studying lipid roles in diseases like HIV and Parkinson's, identifying biomarkers for hypercholesterolemia, and screening for enzyme inhibitors. [6]

1.3 Mass Spectrometry

Mass spectrometry (MS) converts analytes into ionized gaseous species, which are then separated according to their mass-to-charge (m/z) ratio. Unlike spectroscopic techniques based on electromagnetic radiation, MS generates ions through processes such as electron or ion bombardment. MS is widely applied to determine molecular structures, molecular masses, and isotopic ratios, as well as to perform qualitative and quantitative analyses of complex mixtures. In a mass spectrum, intensity reflects the relative abundance of ions at a given m/z. [11]

In mass spectrometry, similar to gas chromatography, lipids must first be vaporized, typically by increasing temperature and reducing pressure (high vacuum). Difficult-to-vaporize lipids can be desorbed using fast atom bombardment (FAB) or matrix-assisted laser desorption ionization (MALDI). For neutral lipids, ionization after vaporization is necessary, using methods like electron impact ionization (EI), chemical ionization (CI), electrospray ionization (ESI), or thermospray Ionization (TSI). EI involves bombarding lipid molecules with high-energy electrons, causing characteristic fragmentation. CI, a softer ionization method, creates mainly charged molecular ions with minimal fragmentation. Electron capture mass spectrometry, a CI variant used for eicosanoids, involves thermal electrons interacting with a reaction gas to form negatively charged ions without fragmentation. This technique is ideal for determining molecular weights. Unlike EI and CI, which operate in a vacuum, ESI functions at normal pressure, making it suitable for online coupling with liquid chromatography. ESI-MS is currently the primary method for characterizing cellular lipidomes. [8]

1.3.1 UV Photodissociation and Collision-Induced Dissociation

The generation of product ions for qualitative and quantitative analysis in mass spectrometry (MS) is commonly achieved through collision-induced dissociation (CID). This involves colliding selected precursor ions with a neutral gas, causing fragmentation dependent on the analyte and collision energy. Alternative dissociation methods, such as electron-based methods and UV photodissociation (UVPD), have been developed to enhance selectivity and provide complementary information. UVPD is widely used for analyzing biological molecules but is less common for singly charged low molecular weight compounds, where CID remains popular. UVPD was first applied to mass spectrometers in the mid-1980s and has been adapted for various instruments, such as ion traps and hybrid triple quadrupole linear ion traps. These setups facilitate UVPD due to their optical access and ability to trap selected ions. Time-of-flight (TOF) instruments offer high resolution and accurate mass for UVPD, but integration can be challenging. Recent advancements include implementing UVPD in various instruments and the development of UVPD-compatible Orbitrap mass spectrometers. UVPD uses various light sources, covering a wide wavelength range, and can be applied to intact molecules or derivatized analytes for improved detection selectivity. Despite technological advances, most UVPD applications focus on proteins, peptides, lipids, and carbohydrates, with limited work on low molecular weight compounds (LMWC). Recent studies demonstrated UVPD's ability to generate informative fragments for compounds that fragment poorly with CID. The implementation of a 266-nm laser on a triple quadrupole linear ion trap allowed the combination of UVPD and CID in the same LC-MS analysis. [3]

1.4 Libraries

1.4.1 Libraries for Glycosphingolipids

Currently, no publicly available MS/MS libraries exist for GSLs.

1.4.2 Challenges in Glycosphingolipid Analysis

Key challenges in GSL analysis include complex sample extraction, extreme structural diversity in both glycan and ceramide moieties, the scarcity of authentic standards, and limited software support.

1.4.3 Evaluation Strategies

In mass spectrometry–based lipid analysis, two main strategies are commonly used for spectrum annotation: spectral library matching and rule-based approaches.
Spectral library matching works by comparing unknown MS/MS spectra with reference spectra stored in databases. These references usually come from authentic standards or are generated computationally. When the library is comprehensive and of high quality, this method is fast and effective. However, it is limited to the compounds present in the database, making it unsuitable for detecting novel or rare lipids. It can also lead to over-annotation and struggles with complex spectra containing signals from multiple compounds, which are common in lipidomics.[4]
Decision-rule-based annotation, on the other hand, uses defined rules that describe the typical fragmentation patterns of different lipid classes. This makes the approach more reliable and flexible, as it considers structural resolution and can distinguish between isomers or isobars. Crucially, it also allows the identification of previously unknown (de novo) compounds, something library matching cannot achieve. These rules are usually derived from experimental data or computational predictions and can be adapted by the user to specific MS setups. This strategy is especially useful for glycolipids, since both their glycan and ceramide components consist of well-defined building blocks that produce predictable fragmentation patterns.[4]

2 Material and Methods

2.1 Standards and Solvents

All solvents used were of LC-MS grade and sourced from Fisher Scientific (Vienna, Austria), VWR International (Vienna, Austria), Merck (Vienna, Austria), and Ajax Finechem (Melbourne, Australia). Glycolipid standards were obtained from Cayman Chemical, Avanti Polar Lipids, and Merck. These were diluted to final concentrations of approximately 5 µM or 10 µM in a 35%/65% (v/v) H_2O/IPA solution. Concentrations were calculated based on the 36:1;O species in each mixture. All standards utilized in this study are detailed in Table 2. The deuterated d5 GM1 18:1;O2/18:0 standard, purchased from Avanti Polar Lipids, Inc. (Alabaster, Alabama, USA), was employed as an internal standard for subsequent analyses. BOLD-100 was synthesized at the Institute of Inorganic Chemistry, University of Vienna, and dissolved in dimethyl sulfoxide (DMSO) to prepare a 20 mM stock solution. Nintedanib was obtained from MedChemExpress (Monmouth Junction, NJ, USA).

Table 2: Standards used in this work.

No.	Name	Vendor	Catalogue No.
1	GD1a	Avanti	860055P
2	GD1b	Avanti	860056P
3	GD2	Cayman	25487
4	GD3	Merck	345752
5	GM1	Cayman	19579
6	GM3	Avanti	860058P
7	GM4	Cayman	24851
8	GQ1b	Merck	345754
9	Lactosylceramide and Ganglioside GM3 and GD3 Mixture (LacCer)	Cayman	29361
10	Neutral Glycosphingolipid Mixture (nGSL)	Cayman	29360
11	Disialoganglioside Mixture (disG)	Cayman	29358
12	Ganglioside Total (TGE)	Avanti	860053P
13	Pooled GSL standard (PGS)	Pool 1-12	
14	Colon cancer (HCT116, 5 conditions with 5 replicates each) - Vienna only		

Standards 1-13 were measured in both Melbourne (Lumos, UVPD) and Vienna (IQ-X ,HCD) in positive and negative ion modes. Colon cancer samples (14) were measured only in Vienna (IQ-X, HCD).

2.2 Sample Preparation

To extract gangliosides an adapted SIMPLEX protocol was used. The cancer cell samples were quenched directly on the 6-well plates with N2, transported on dry ice, and extracted on the same day. 50 µL of an internal ganglioside standard (GM3-d5 in MeOH) were added directly to the samples to compensate for losses during extraction. 500 µL of cold MeOH were added, and the cells were detached using a cell scraper. Each well was properly scratched for 1 minute while still being cooled on dry ice. The cell suspension was then transferred to a 5 mL Eppendorf tube, and the well plate was rewashed with 500 µL of cold MeOH, which was then also added to the Eppendorf tube. The samples remained cooled on dry ice. 3190 µL of cold methyl-tert-butyl ether (MTBE) were added and vortexed. The samples were incubated for 1 hour on a Thermo shaker at 4 °C and 280 rpm. To induce phase separation, 800 µL H_2O were added, and the samples were centrifuged for 5 minutes at 1580 rcf. The lipid fraction (upper phase) was transferred into new 5 mL Eppendorf tubes and dried in a GenVac overnight. 2700 µL MeOH were added to the remaining lower phase and incubated for 1 hour in a shaker at room temperature. Afterwards, the tubes were centrifuged for 12 minutes at 4 °C and 4500 rpm. The metabolite fraction (supernatant) was transferred into a new 5 mL Eppendorf tube, dried in a GenVac and stored at -80 °C. The remaining pellet is the protein fraction, which was stored at -80 °C. The dried lipid fraction was resolved in 100 µL IPA (65%), vortexed, and transferred into 250 µL inlets, which were then brought back into glass vials. For a QC sample pool, 5 µL of each sample were used. For condition pools, 15 µL of each sample of the corresponding condition were used. Five replicates per condition were prepared. To determine the protein concentration, the BCA assay (Micro BCA™ Protein Assay Kit, Thermo Scientific) was used. Four different conditions were used as seen in Table 3.

Table 3: Colorectal cancer cell conditions used in this work.

Name	**Cells**	**Entity**	**sensitive (S) / resistent (R)**	**Selection**	**Biological background**
HCT116	HCT116	hCRC	S	—	—
HCT116/Nin	HCT116	hCRC	S	—	—
HCTR	HCTR	hCRC	R	BOLD-100	loss of CD147 in R compared to S
HCT116OxR	HCT116OxR	hCRC	R	Oxaliplatin	—

The cell line HCT116/Nin was treated with Nintedanib for 4 hours with a concentration of 7.5 µM.

2.3 LCMS Method

2.3.1 IQ-X (HCD)

A Vanquish Horizon UHPLC system, coupled via heated electrospray-ionization (HESI) to a high-field Orbitrap ID-X™ Tribrid™ mass spectrometer (both from Thermo Fisher Scientific) was used. Reverse-phase chromatography was performed using an Acquity HSS T3 column (2.1 mm × 150 mm, 1.8 μm, Waters) paired with a VanGuard pre-column (2.1 mm × 5 mm, 100 Å, 1.8 μm). The column temperature was maintained at 40 °C, with a flow rate of 0.25 mL/min, and an injection volume of 5 µL. Between injections, the injector needle was rinsed with a solution of 75% isopropanol (IPA) including 0.1% formic acid (FA). The mobile phases consisted of Solvent A (acetonitrile (ACN)/H_2O, 3:2, v/v) and Solvent B (IPA/ACN, 9:1, v/v) both containing 0.1% formic acid and 10 mM ammonium formate. A 30-minute gradient was applied under the following conditions: 0.0–2.0 min 30% B, 2.0–3.0 min ramp to 55% B, 3.0–17.0 min ramp to 67% B, 17.0–22.0 min ramp to 100%, 22.0–26.0 min 100% B, 26.0 min fast switch to 30% B, and 26.0–30.0 equilibration at starting conditions (30% B).
The ESI source parameters were set at 3.5 kV (positive ion mode) and 3.0 kV (negative ion mode). The sheath gas flow was set to 40, auxiliary gas to 8, and sweep gas to 1. The capillary temperature (ion transfer tube temperature) was 275 °C, and the auxiliary gas heater (vaporizer temperature) was 300 °C. The radio frequency (RF) level was adjusted to 60%. Positive and negative modes were measured consecutively. Spectral data were acquired using profile mode, generating 3 consecutive scans per cycle, and an intensity threshold of 1.0x10^3 for the precursor was applied. A mass range of m/z 400-2,000 at a resolution of 120,000 was selected for full MS runs. The automatic gain control (AGC) target was set to Standard, with a maximum injection time (MIT) of 100 ms, and 1 Microscan. A dynamic exclusion period of 3 seconds was applied for triggered m/z. Furthermore, isotopes were excluded. MS2 and MS3 scans were performed using data-dependent acquisition (DDA). For MS2 scanning, Quadrupole isolation with an Orbitrap resolution of 30,000 was applied. These experiments were executed in an unscheduled mode using a ganglioside-specific inclusion list with a mass tolerance of 5 ppm and an exclusion list including certain m/z values with a mass tolerance of 10 ppm. The Scan Range was m/z 150-2,000. The normalized collision energy (NCE) was set to 27/30/33% (positive ion mode) and 27/30/33% (negative ion mode), with an isolation window of m/z 1.5. The AGC target was set to Standard, and MIT was set to 300 ms. For MS3 spectra, the mass range was m/z 400-800. In negative mode, a targeted mass trigger was set to m/z 290.0876, with a mass tolerance of 10 ppm. The ion trap was used as the mass analyzer with a fixed collision energy of 33% (CID activation), an activation time of 10 ms, and an activation Q of 0.25. The Ion Trap scan rate was set to Rapid, with isolation windows of m/z 1.5 for MS1 and 2.0 for MS2. The AGC target was set to Standard, and MIT was set to Automatic, and 1 Microscan. An unscheduled targeted mass exclusion list (containing m/z 524.265 and 581.1827) was applied. MS3 scans in the Ion Trap were parallelized with MS1 and MS2 scans in the Orbitrap, increasing the amount of information obtained and saving time.

2.3.2 Fusion Lumos (UVPD)

A Vanquish Horizon UHPLC system, coupled via heated electrospray-ionization (HESI) to an Orbitrap Fusion Lumos mass spectrometer (Thermo Fisher, San Jose, CA, USA) was used. The mass spectrometer (MS) was expanded with a user-installed 213 nm laser (Ekspla NL204, 0.2 mJ, 7–10 ns pulse duration, 1 kHz repetition rate, Vilnius, Lithuania) for UVPD fragmentation. Most parameters used were set as described above. Differences include ESI parameters set to 2.9 kV for negative ion mode, a sheath gas flow of 50, a vaporizer temperature of 350 °C, and a RF level of 45%. Furthermore, the cycle time was set to 0.6 seconds, the intensity threshold was set to 5.0x10^3 and a mass range of m/z 250-2,000 was selected for full MS runs. A dynamic exclusion period of 5 seconds was applied for triggered m/z. Only a ganglioside-specific inclusion list was used. The normalized collision energy (NCE) was set to 27%. No MS3 spectra were generated.

3 Data Processing

3.1 Compound Discoverer

Data processing was performed with Compound Discoverer (CD) 3.3 SP2 using raw files from both Fusion Lumos and IQ-X measurements. An untargeted lipidomics workflow was adapted from the local database setup. The workflow included the following nodes as seen in Figure 5: Input Files, Select Spectra, Align Retention Times (ChromAlign), Detect Compounds, Group Compounds, Assign Compound Annotations, Predict Compositions, and Search Mass Lists.

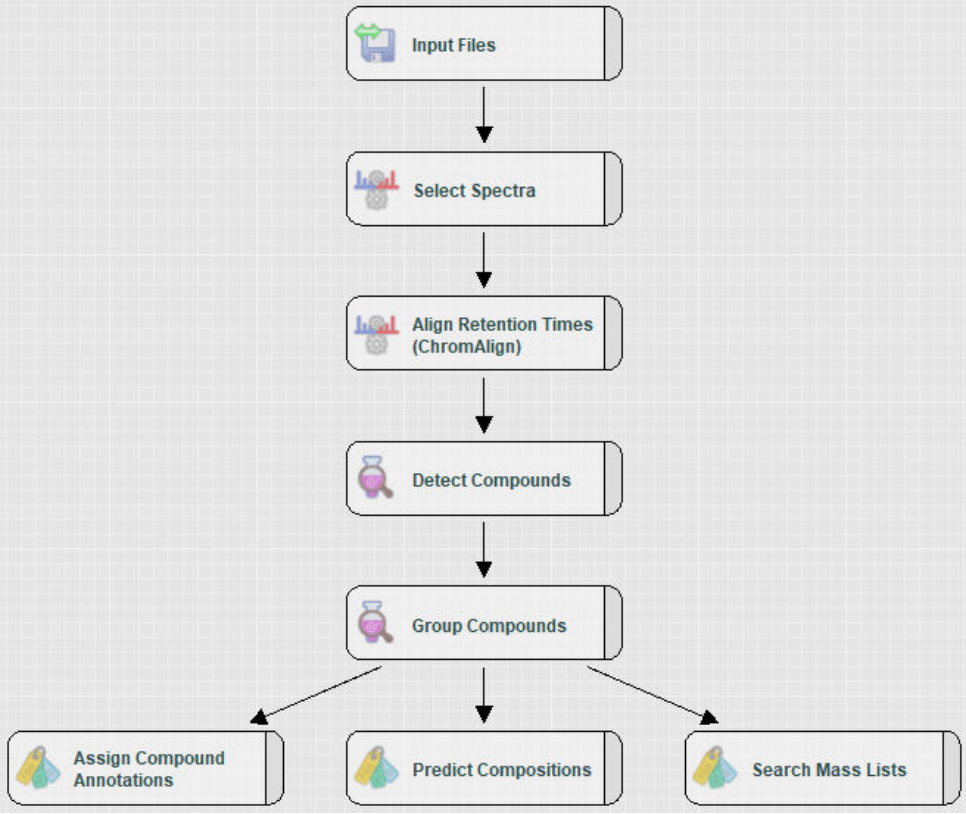

Figure 5: Workflow Compound Discoverer depicting the used nodes.

For optimization, only six lipid classes (GD1, GD2, GD3, GM1, GM3, and GQ1) and two species (36:1;O2 and 38:1;O2) were included, using raw standard measurements from Melbourne. Different input mass lists were compared as seen in Table 4.

Table 4: Information included in the different mass lists tested.

Masslist	Species Name	Molecular Form	Molecular Weight	Retention Time
Inputlist	✓	✓	—	—
Inputlist molweight	✓	✓	✓	—
Inputlist RT	✓	✓	—	✓

All input lists produced satisfactory results. Since molecular weight may vary depending on adducts,

and retention time is highly useful for identification, further CD runs used the mass list including species name, molecular form, and retention times. The complete list can be found in chapter 6 in table 6.

Different settings and filters were tried. The final settings were as follows:

The following parameters were configured for the "Select Spectra" node: In the Spectrum Properties Filter, the lower retention time was set to 5, and the upper limit to 25. The precursor mass range was defined with a minimum of 200 Da and a maximum of 2000 Da, with a minimum peak count of 1. All other parameters were set to 0. In the Scan Event Filters, Mass Analyzer, MS Order, Activation Type, Scan Type, and Polarity Mode were set to "Any." The collision energy was set to a maximum of 1000 and a minimum of 0. MS1 Mass Range and FAIMS CV were not specified. The S/N Threshold (FT-only) in the Peak Filters was set to 1.5. Under Replacements for Unrecognized Properties, Charge Replacements were set to 1, Mass Analyzer to ITMS, MS Order to MS2, Activation Type to CID, and Polarity to positive. The MS resolution at 200 was configured to 60,000 and the MSn resolution at 200 to 30,000. In the general settings, precursor selection was set to "Use MS(n-1) Precursor," and the isotope pattern was considered in precursor reevaluation. Profile spectra were automatically provided, and chromatogram storage was disabled. For the "Align Retention Times (ChromAlign)" node no reference file was used. In the "Detect Compounds" node the general settings were set to a mass tolerance of 5 ppm, a minimal peak intensity of 30,000 and a minimum of 7 scans per peak. Only the most intense isotope was used. For Trace detection the maximum number of gaps to correct and the minimal number of adjacent non/zeros were both set to 2. Under Peak Detection the chromatographic S/N threshold was 1.5, the baseline was not removed and the threshold for the gap ratio was 0.35. The maximum peak width was 1 minute and the relative valley depth had a minimum of 0.1. Under Isotope Pattern Detection Isotopes were grouped for Br and Cl. The thresholds for the Zig-Zag Index was set to 0.2, for Jaggedness to 0.4 and for Modality to 0.9. The peak quality was used for isotope grouping , filtering out of features with bad peaks only was done as well as removal of potentially false positive isotopes. For Compound detection the following ions were used: $[2M+FA-H]^{-1}$; $[2M+Na]^{+1}$; $[2M+NH_4]^{+1}$; $[M+2H]^{+2}$; $[M+3H]^{+3}$; $[M+H]^{+1}$; $[M+H+Na]^{+2}$; $[M+H+NH_4]^{+2}$; $[M+H-H_2O]^{+1}$; $[M+Na]^{+1}$; $[M+NH_4]^{+1}$; $[M-2H]^{-2}$; $[M-H]^{-1}$; $[M-H+Hac]^{-1}$; $[M-H-H_2O]^{-1}$, as well as these base ions: $[M+2H]^{+2}$; $[M+H]^{+1}$; $[M-2H]^{-2}$; $[M-H]^{-1}$. Singlets were removed. Under AquireX settings the detection of persistent background ions was not used. For the “predict compositions” node under prediction settings the mass tolerance was set to 5 ppm. The element counts were defined with a minimum of C and H and a maximum of C_{90} H_{190} N_{10} O_{18}. RDBE had a minimum of 0 and a maximum of 40, while the H/C range 0.1 to 4 was. A maximum number of candidates was set to 10 and of internal candidates to 200. Under pattern matching the intensity tolerance was 30% with a threshold of 0.1%. The S/N threshold was 3, the minimum spectral fit was 30% and the minimum pattern cov. Was 90%. Dynamic recalibration was used. Under fragments matching a mass tolerance of 5 ppm with a S/N threshold of 3 was used. Under general settings in the node “group compounds” the mass tolerance was 5 ppm and the RT tolerance was 0.3 minutes. Peak alignment was used and $[M+2H]^{+2}$; $[M+H]^{+1}$; $[M+3H]^{+3}$; $[M-2H]^{-2}$; $[M-H]^{-1}$ were the preferred

ions. Area integration was done for the most common ion. In the peak rating contributions the Area contribution was 3 and the CV contribution was 10. The contributions for FWHM to base, jaggedness, modality and zig-zag index were 5. In the peak rating filter the peak rating threshold was 4 and the number of files was 0. The parameters of the “Assign Compound Annotations” node were set to a mass tolerance of 5 ppm under general settings. For data sources MassList Search and Predicted Compositions were used. Under scoring rules mzLogic and spectral distance were used and SFit threshold and SFit range were set to 20. Under reprocessing the clear names were set to false. For the node “Search mass lists” under search settings the mass list was imported and retention times were used with a RT tolerance of 0.5 minutes and a mass tolerance of 5 ppm.
Separate runs were performed for the positive and negative mode and the HCD and UVPD data, using standard data from 8 lipid classes and 5 standard mixtures as listed in 2. The run in positive mode generated 122 hits for the HCD data and 87 hits for the UVPD data. The run in negative mode generated 110 hits for the HCD data and 79 hits for the UVPD data.

3.2 mzVault

Data from CD were exported into mzVault (v2.3 SP1). Separate libraries were created for positive and negative ionization modes, each including HCD and UVPD spectra. The library in negative mode had 71 entries including 634 spectra. 10 lipid classes were added by hand. All spectra were manually checked for the presence of diagnostic fragments and background noise. After applying those quality criteria, the library included 81 entries for lipid species with 114 spectra including 8 GSL classes. The lipid species included in the library are listed in Table 5.

Table 5: GSL classes in the library for negative ionization mode.

lipid class	**new nomenclature**
GM1	Hex3_HexNAc_NeuAc_Cer
GM2	Hex2_HexNAc_NeuAc_Cer
GM3	Hex2_NeuAc_Cer
GD1	Hex3_HexNAc_NeuAc2_Cer
GD2	Hex2_HexNAc_NeuAc2_Cer
GD3	Hex2_NeuAc2_Cer
GT1	Hex3_HexNAc_NeuAc3_Cer
GQ1	Hex3_HexNAc_NeuAc4_Cer

The library in positive mode had 71 entries including 608 spectra. 11 lipid classes were added by hand. All spectra were manually checked for the presence of diagnostic fragments and background noise. After applying those quality criteria, the library included 82 entries for lipid species with 116 spectra including 10 GSL classes. The lipid species included in the library are listed in Table 6.

Table 6: GSL classes in the library for positive ionization mode.

lipid class	new nomenclature
GM1	Hex3_HexNAc_NeuAc_Cer
GM2	Hex2_HexNAc_NeuAc_Cer
GM3	Hex2_NeuAc_Cer
GD1	Hex3_HexNAc_NeuAc2_Cer
GD2	Hex2_HexNAc_NeuAc2_Cer
GD3	Hex2_NeuAc2_Cer
GT1	Hex3_HexNAc_NeuAc3_Cer
GQ1	Hex3_HexNAc_NeuAc4_Cer
LacCer	Hex2_Cer
Gb3	Hex3_Cer

All spectra were manually curated for quality, ensuring reproducibility of the spectral library.

4 Results and Discussion

4.1 Creating a Library for Glycosphingolipids

A dedicated MS/MS spectral library for glycosphingolipids (GSLs) was successfully developed. The final library includes 82 species with 116 spectra in positive ionization mode and 81 species with 114 spectra in negative ionization mode. A complete list of entries is provided in table 6 in chapter 6. The distribution of entries by lipid class is summarized in Table 7.

Table 7: Number of entries per lipid class.

lipid class	entries in positive ionization mode	entries in negative ionization mode
GM1	9	9
GM2	2	2
GM3	21	25
GD1	10	12
GD2	10	11
GD3	14	15
GT1	2	4
GQ1	2	3
LacCer	6	—
Gb3	6	—

Most spectra were obtained using authentic standards specific to each class. For GT1 and Gb3, where dedicated standards were unavailable, spectra were generated from the total ganglioside standard and the neutral glycosphingolipid standard, respectively.
Due to the absence of molecular species–specific standards, identifications were reported at the lipid species level (e.g., 36:1;O2) rather than the molecular lipid species level. While fragmentation data can sometimes suggest more detailed molecular structures, reporting at the species level avoids over-annotation.
Representative examples of successful library-based annotation are shown in Figure 6 and 7, which present four GSL species (two acidic and two neutral) identified in positive ion mode with HCD fragmentation.

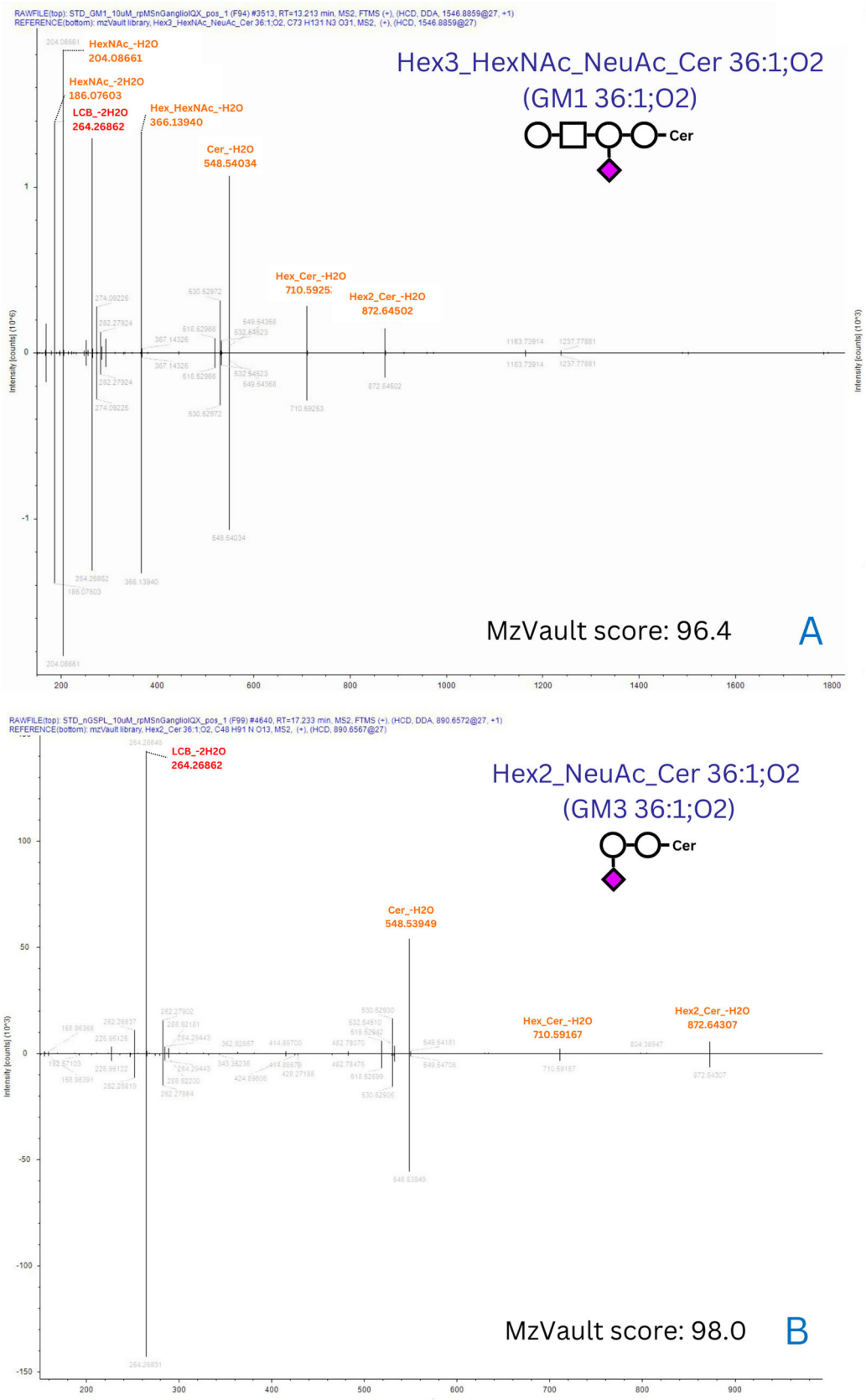

Figure 6: Mirror plots of two selected GSL species obtained from library search in positive ionization mode with HCD fragmentation, based on standards and cancer samples: Two potential acidic ganglioside classes (A) GM1 36:1;O2 and (B) GM3 36:1:O2.
Figure by Amirezza Dowlati Beirami

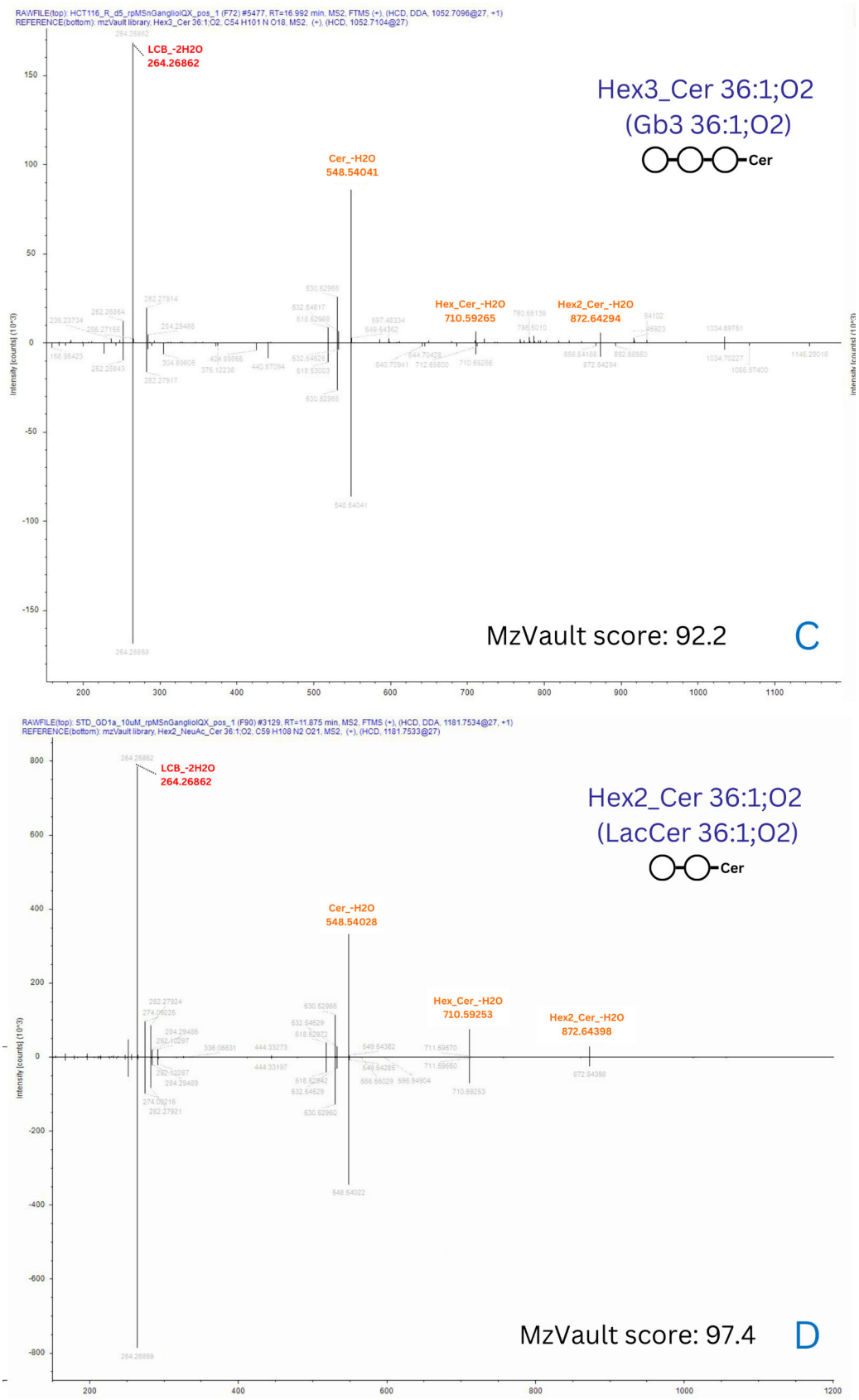

Figure 7: Mirror plots of two selected GSL species obtained from library search in positive ionization mode with HCD fragmentation, based on standards and cancer samples: Two potential neutral species including a globoside (C) Gb3 36:1;O2 and a lactosyl ceramide (D) LacCer 36:1:O2.
Figure by Amirezza Dowlati Beirami

These mirror plots confirm high-quality matches, with mzVault scores above 90. In the annotated GSL (glycosphingolipid) examples shown in Figure 6 and Figure 7, all species share a common structure composed of the same sphingoid base and ceramide, specifically 36:1;O2. This is confirmed by two key fragment ions: Firstly the sphingoid base fragment at m/z 264, which results from the loss of water and indicates a dihydroxylated 18:1;O2 structure, and secondly the ceramide fragment at m/z 548, which corresponds to Cer36:1 after water loss. The consistent presence of the m/z 264 ion in all examples allows the assignment of the molecular structure as 18:1;O2/18:0, showing that positive ion mode is particularly useful for revealing ceramide structure, unlike negative ion mode, which tends to highlight glycan fragments instead. Additionally, the ion at m/z 872.645 (Hex2Cer-H2O) is present in all spectra and serves as a precursor for more complex GSLs. This includes the parent ion for Hex2-Cer 36:1;O2 (Lactosylceramide, or LacCer) and a major fragment for larger GSLs like GM1, GM3, and Gb3, all with the same ceramide base. A key distinction between acidic GSLs (Figure 6A and 6B) and neutral GSLs (Figure 7C and 7D) is the presence of a prominent m/z 292 fragment in the acidic species, corresponding to the loss of sialic acid.

The libraries are dominated by HCD spectra. In negative ion mode, 78 out of 81 entries have HCD spectra, while only 28 include UVPD spectra. In positive ion mode, 80 out of 82 entries have HCD spectra, with 34 entries including UVPD spectra. The distribution is shown in Table 8.

Table 8: Number of compounds per fragmentation strategy.

Lipid class	**Compounds in positive ionization mode**		**Compounds in negative ionization mode**	
	HCD	**UVPD**	**HCD**	**UVPD**
GM1	9	1	9	—
GM2	2	—	2	—
GM3	20	15	25	1
GD1	10	8	12	4
GD2	10	2	11	8
GD3	14	4	15	10
GT1	1	2	4	2
GQ1	—	2	2	3
LacCer	6	—	—	—
Gb3	6	—	—	—

UVPD generates more fragments overall, including glycan-specific ions useful for structural characterization. It can also produce unique diagnostic fragments, such as those representing neutral fatty acid losses, which are absent in HCD spectra. This makes UVPD particularly valuable for identifying molecular lipid species of abundant GSLs. HCD, by contrast, produces consistent and easily

interpretable spectra, making it the backbone of the library.[5]

Because the library is built on fragment pattern matching, manual inspection of fragment assignments remains essential. Without molecular standards, structural assignment is limited to the lipid species level to avoid misannotation.
The spectral library was successfully extended to public repositories.
This library is intended as quick-check tools to assess whether GSLs are present and worth further investigation, rather than as replacements for comprehensive structural analysis.

4.2 Proof of Principle with Cancer Samples

A proof of principle analysis was conducted using the HCT116 colorectal cancer (CRC) cell line. Glycolipids such as GD1, GM1, GM3, Gb3, Gb4, and GalCer have previously been associated with CRC. [13] In this study, several of these glycolipids, including LacCer, GM3, GM2, and Gb3, were confirmed at the MS1 level. Other species, such as HexCer, GM4, GD3, and GM1, were detected in very low abundance, suggesting potential relevance but requiring further validation. The presence of GD1 and Gb4 could not be definitively established in this dataset and requires additional analysis for confirmation. Because the dataset included two sensitive conditions (HCT116 and HCT116 treated with Nintedanib) and two resistant conditions (HCTR treated with BOLD-100 and HCT116OxR treated with oxaliplatin), differences in GSL profiles were investigated.

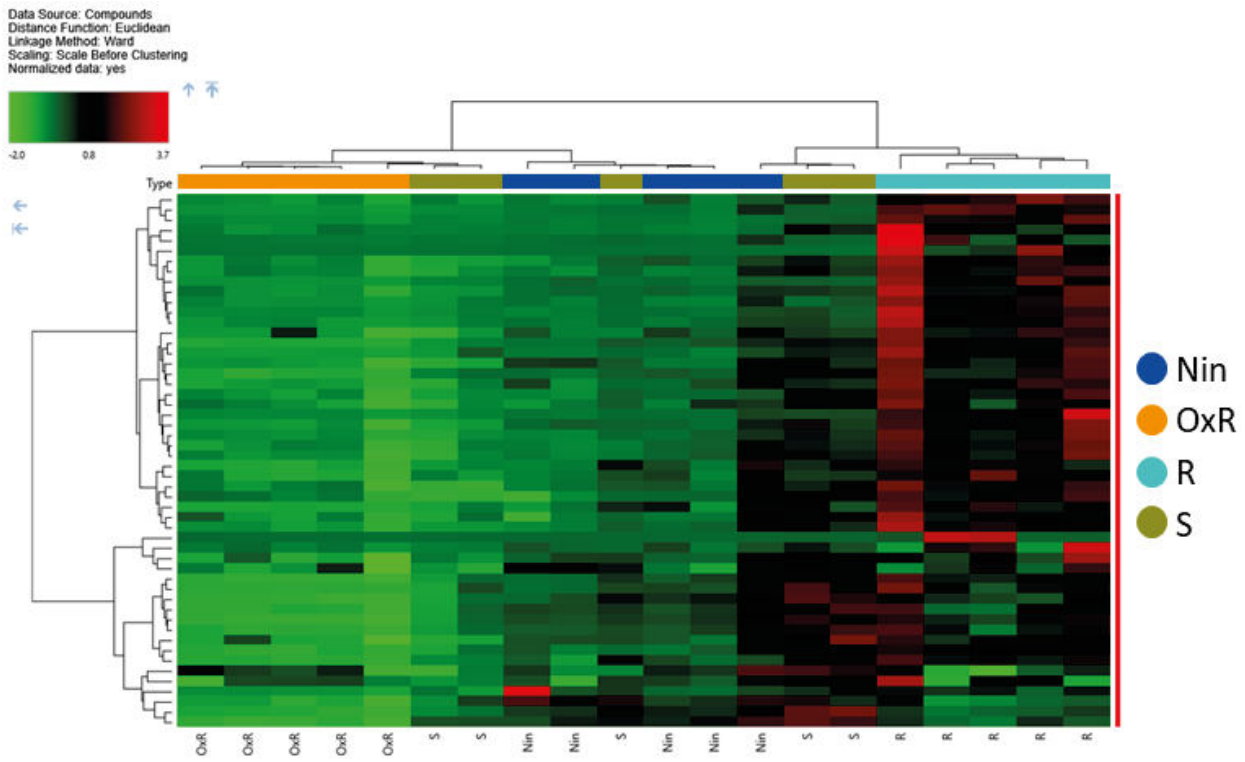

Figure 8: Heatmap of normalized GSL profiles across resistant and sensitive cell lines, based on 30 GSL species identified with the Compound Discoverer library search. Five replicates per condition were included. Distinct clustering patterns separate BOLD-100–resistant (R, right) and oxaliplatin-resistant (OxR, left) cell lines. Sensitive cell lines (HCT116 and HCT116 + Nintedanib) cluster together, suggesting minimal differences between them. Figure by Amirezza Dowlati Beirami

The heatmap analysis shown in Figure 8 revealed distinct clustering of resistant versus sensitive phenotypes. BOLD-100–resistant cells showed elevated GSL levels. Oxaliplatin-resistant cells displayed reduced GSL levels for several species. The two sensitive conditions clustered together, indicating minimal variation. These findings demonstrate that the spectral library enables GSL profiling capable of differentiating cancer cell phenotypes, underscoring the potential of GSL signatures as biomarkers of therapeutic response.

5 Conclusion and Outlook

In this thesis, a dedicated spectral library for glycosphingolipids (GSLs) was successfully generated, providing a novel resource for the lipidomics community. Using Compound Discoverer and mzVault, approximately 80 GSL species were compiled at the lipid species level, covering both positive and negative ionization modes. This work represents one of the first systematic efforts to assemble a spectral library specifically tailored to GSLs. The resulting collection enables rapid and reproducible spectral matching and thus offers a valuable tool for the annotation and evaluation of GSLs in complex biological samples.

The practical utility of the library was demonstrated by its application to cancer-related datasets. The results confirmed that profiling GSLs in cancer samples is feasible, underscoring the potential of this approach as a quick-check for the presence and distribution of GSLs. Importantly, it was shown that such profiling can distinguish between sensitive and resistant cancer cells, highlighting the possible relevance of GSL signatures as indicators of therapeutic response. The integration of the spectral library into public repositories further extends its accessibility, ensuring that it can be readily applied and expanded by other researchers.

Looking ahead, the continued expansion of the library will be crucial to broaden its coverage across different GSL classes and to include additional reference spectra from authentic standards and high-quality in silico predictions. Moreover, future work should consider comparative approaches, such as rule-based annotation, which could complement library-based matching and facilitate the detection of novel or rare compounds. Ultimately, the combination of these strategies holds promise for establishing more comprehensive lipidomics pipelines.

In conclusion, this work delivers a spectral library that not only supports reliable annotation of glycosphingolipids but also demonstrates clear applicability in biomedical research. The ability to use GSL profiling to separate cancer cell phenotypes emphasizes the translational potential of the approach. With further development and integration into broader analytical frameworks, the library provides a solid foundation for advancing both fundamental and applied studies of glycosphingolipid biology.

References

[1] J. M. Berg, J. L. Tymoczko, G. J. Gatto jr., and L. Stryer. *Stryer Biochemie*. Springer Spektrum, Berlin, Heidelberg, 8 edition, 2018.

[2] E. Fahy, S. Subramaniam, R. C. Murphy, M. Nishijima, C. R. Raetz, T. Shimizu, F. Spener, G. Van Meer, M. J. Wakelam, and E. A. Dennis. Update of the LIPID MAPS comprehensive classification system for lipids, 4 2009.

[3] R. Giraud, Y. J. C. Le Blanc, G. Mircea, and G. Hopfgartner. Ultraviolet photodissociation and collision-induced dissociation for qualitative/quantitative analysis of low molecular weight compounds by liquid chromatography-mass spectrometry. 415:7117–7126, 2023.

[4] K. Hohenwallner. *Development of high-resolution mass spectrometry and automated annotation strategies for comprehensive glycosphingolipid profiling and structural characterization - Universität Wien*. PhD thesis, 2024.

[5] K. Hohenwallner, L. M. Lamp, L. Peng, M. Nuske, J. Hartler, G. E. Reid, and E. Rampler. FAIMS Shotgun Lipidomics for Enhanced HILIC-like Separation and Automated Annotation of Gangliosides. 3 2024.

[6] P. T. Ivanova, S. B. Milne, D. S. Myers, and H. A. Brown. Lipidomics: a mass spectrometry based systems level analysis of cellular lipids. *Current Opinion in Chemical Biology*, 13(5-6):526–531, 12 2009.

[7] F. L. Kiechle and X. Zhang. Review: Glycosphingolipids in Health and Disease. Technical report, 2004.

[8] J. Kurreck, J. W. Engels, and F. Lottspeich. *Bioanalytik*. Springer Spektrum, Berlin, Heidelberg, 4 edition, 2022.

[9] D. H. Kwak, J. S. Ryu, C. H. Kim, K. Ko, J. Y. Ma, K. A. Hwang, and Y. K. Choo. Relationship between ganglioside expression and anti-cancer effects of the monoclonal antibody against epithelial cell adhesion molecule in colon cancer. *Experimental and Molecular Medicine*, 43(12):693–701, 2011.

[10] A. H. Merrill and M. N. Vu. Glycolipids. In *Encyclopedia of Cell Biology*, volume 1, pages 180–193. Elsevier Inc., 2016.

[11] M. Otto. *Analytische Chemie*. Wiley-VCH, Weinheim, 5 edition, 2019.

[12] K. L. Wormwood Moser, G. Van Aken, D. DeBord, N. G. Hatcher, L. Maxon, M. Sherman, L. Yao, and K. Ekroos. High-defined quantitative snapshots of the ganglioside lipidome using high resolution ion mobility SLIM assisted shotgun lipidomics. *Analytica Chimica Acta*, 1146:77–87, 2 2021.

[13] D. Zhuo, X. Li, and F. Guan. Biological roles of aberrantly expressed glycosphingolipids and related enzymes in human cancer development and progression, 5 2018.

List of Figures

List of Tables

6 Attachment

Table 9: Complete input list used with Compound Discoverer.

Class	Species	Sum Formula	Retention Time [min]
GM1	GM1 34:1;O2	C71H127N3O31	11.1
GM1	GM1 35:1;O2	C72H129N3O31	12
GM1	GM1 36:1;O2	C73H131N3O31	13.2
GM1	GM1 37:1;O2	C74H133N3O31	14.4
GM1	GM1 38:1;O2	C75H135N3O31	15.8
GM1	GM1 39:1;O2	C76H137N3O31	17.1
GM1	GM1 40:1;O2	C77H139N3O31	18.6
GM1	GM1 41:1;O2	C78H141N3O31	19.9
GM1	GM1 42:1;O2	C79H143N3O31	20.6
GM1	GM1 42:2;O2	C79H141N3O31	19.6
GM1	GM1 43:2;O2	C80H143N3O31	20
GM1	GM1 44:1;O2	C81H147N3O31	21.5
GM1	GM1 44:2;O2	C81H145N3O31	20.7
GM1	GM1 34:1;O3	C71H127N3O32	10.7
GM2	GM2 32:1;O2	C63H113N3O26	9.4
GM2	GM2 34:0;O2	C65H119N3O26	12.2
GM2	GM2 34:1;O2	C65H117N3O26	11.4
GM2	GM2 34:2;O2	C65H115N3O26	9.7
GM2	GM2 35:1;O2	C66H119N3O26	12.5
GM2	GM2 36:1;O2	C67H121N3O26	13.4
GM2	GM2 38:1;O2	C69H125N3O26	16.1
GM2	GM2 40:1;O2	C71H129N3O26	19.3
GM2	GM2 40:2;O2	C71H127N3O26	16.4
GM2	GM2 40:2;O2	C71H127N3O26	16.9
GM2	GM2 42:1;O2	C73H133N3O26	20.9

Class	Species	Sum Formula	Retention Time [min]
GM2	GM2 42:2;O2	C73H131N3O26	19.1
GM2	GM2 42:2;O2	C73H131N3O26	19.8
GM2	GM2 43:2;O2	C74H133N3O26	20.1
GM2	GM2 44:2;O2	C75H135N3O26	20.8
GM2	GM2 34:0;O3	C65H119N3O27	11.5
GM2	GM2 34:1;O3	C65H117N3O27	10.9
GM2	GM2 34:2;O3	C65H115N3O27	9.3
GM2	GM2 35:1;O3	C66H119N3O27	11.6
GM2	GM2 42:2;O3	C73H131N3O27	18.7
GM2	GM2 43:2;O3	C74H133N3O27	20
GM2	GM2 44:2;O3	C75H135N3O27	20.7
GM3	GM3 32:0;O2	C55H102N2O21	10.3
GM3	GM3 32:1;O2	C55H100N2O21	9.7
GM3	GM3 33:1;O2	C56H102N2O21	10.6
GM3	GM3 34:0;O2	C57H106N2O21	12.4
GM3	GM3 34:1;O2	C57H104N2O21	11.4
GM3	GM3 34:2;O2	C57H102N2O21	9.9
GM3	GM3 35:0;O2	C58H108N2O21	14
GM3	GM3 35:1;O2	C58H106N2O21	12.4
GM3	GM3 36:0;O2	C59H110N2O21	15
GM3	GM3 36:1;O2	C59H108N2O21	13.8
GM3	GM3 36:1;O2	C59H108N2O21	13
GM3	GM3 36:2;O2	C59H106N2O21	11.9
GM3	GM3 36:2;O2	C59H108N2O21	12
GM3	GM3 37:1;O2	C60H110N2O21	15
GM3	GM3 38:0;O2	C61H114N2O21	17.8
GM3	GM3 38:1;O2	C61H112N2O21	15.3
GM3	GM3 38:1;O2	C61H112N2O21	16.7

Class	Species	Sum Formula	Retention Time [min]
GM3	GM3 38:2;O2	C61H110N2O21	14.1
GM3	GM3 39:1;O2	C62H114N2O21	18
GM3	GM3 40:0;O2	C63H118N2O21	20.1
GM3	GM3 40:1;O2	C63H116N2O21	19.3
GM3	GM3 40:2;O2	C63H114N2O21	16.7
GM3	GM3 40:2;O2	C63H114N2O21	17.3
GM3	GM3 41:1;O2	C64H118N2O21	20.2
GM3	GM3 41:2;O2	C64H116N2O21	18
GM3	GM3 41:2;O2	C64H116N2O21	18
GM3	GM3 41:2;O2	C64H116N2O21	18.6
GM3	GM3 42:1;O2	C65H120N2O21	21
GM3	GM3 42:2;O2	C65H118N2O21	19.4
GM3	GM3 42:2;O2	C65H118N2O21	20
GM3	GM3 43:1;O2	C66H122N2O21	21.4
GM3	GM3 43:1;O2	C66H122N2O21	21.4
GM3	GM3 43:2;O2	C66H120N2O21	19.9
GM3	GM3 43:2;O2	C66H120N2O21	20.3
GM3	GM3 43:2;O2	C66H120N2O21	20.3
GM3	GM3 43:2;O2	C66H120N2O21	20.3
GM3	GM3 44:1;O2	C67H124N2O21	21.7
GM3	GM3 44:2;O2	C67H122N2O21	20.6
GM3	GM3 44:2;O2	C67H122N2O21	20.9
GM3	GM3 45:2;O2	C68H124N2O21	21.2
GM3	GM3 46:2;O2	C69H126N2O21	21.6
GM3	GM3 32:1;O3	C55H100N2O22	9.2
GM3	GM3 33:1;O3	C56H102N2O22	10.1
GM3	GM3 34:0;O3	C57H106N2O22	11.8
GM3	GM3 34:1;O3	C57H104N2O22	11.1

Class	Species	Sum Formula	Retention Time [min]
GM3	GM3 34:2;O3	C57H102N2O22	9.4
GM3	GM3 35:1;O3	C58H106N2O22	12.2
GM3	GM3 36:0;O3	C59H110N2O22	12.9
GM3	GM3 36:1;O3	C59H108N2O22	13.4
GM3	GM3 36:2;O3	C59H106N2O22	11.3
GM3	GM3 38:2;O3	C61H110N2O22	13.8
GM3	GM3 40:0;O3	C63H118N2O22	17.9
GM3	GM3 40:0;O3	C63H118N2O22	18.4
GM3	GM3 40:1;O3	C63H116N2O22	15.6
GM3	GM3 40:2;O3	C63H114N2O22	12.9
GM3	GM3 41:0;O3	C64H120N2O22	19.7
GM3	GM3 41:1;O3	C64H118N2O22	16.9
GM3	GM3 42:1;O3	C65H120N2O22	18.2
GM3	GM3 42:1;O3	C65H120N2O22	18.2
GM3	GM3 42:2;O3	C65H118N2O22	15.4
GM3	GM3 42:2;O3	C65H118N2O22	15.8
GM3	GM3 44:1;O3	C67H124N2O22	20.4
GM3	GM3 44:2;O3	C67H122N2O22	18.2
GM4	GM4 34:1;O2	C51H95N2O16	11.2
GM4	GM4 36:1;O2	C53H99N2O16	13.5
GM4	GM4 36:2;O2	C53H97N2O16	13.5
GM4	GM4 38:1;O2	C55H103N2O16	16.1
GM4	GM4 38:2;O2	C55H101N2O16	16.1
GM4	GM4 39:1;O2	C56H105N2O16	17.5
GM4	GM4 40:1;O2	C57H107N2O16	19.1
GM4	GM4 41:1;O2	C58H109N2O16	20.1
GD1	GD1 34:1;O2	C82H144N4O39	9.9
GD1	GD1 35:1;O2	C83H146N4O39	10.8

Class	Species	Sum Formula	Retention Time [min]
GD1	GD1 36:0;O2	C84H150N4O39	12.7
GD1	GD1 36:1;O2	C84H148N4O39	11.9
GD1	GD1 36:2;O2	C84H146N4O39	10.2
GD1	GD1 37:1;O2	C85H150N4O39	13.1
GD1	GD1 38:1;O2	C86H152N4O39	14.3
GD1	GD1 39:1;O2	C87H154N4O39	15.6
GD1	GD1 40:1;O2	C88H156N4O39	16.8
GD1	GD1 41:1;O2	C89H158N4O39	18.4
GD1	GD1 42:2;O2	C90H158N4O39	17
GD2	GD2 34:1;O2	C76H134N4O34	10.1
GD2	GD2 35:1;O2	C77H136N4O34	11.1
GD2	GD2 36:1;O2	C78H138N4O34	12.2
GD2	GD2 37:1;O2	C79H140N4O34	13.4
GD2	GD2 38:0;O2	C80H144N4O34	15.4
GD2	GD2 38:1;O2	C80H142N4O34	14.6
GD2	GD2 38:2;O2	C80H140N4O34	12.6
GD2	GD2 39:1;O2	C81H144N4O34	16
GD2	GD2 40:1;O2	C82H146N4O34	17.3
GD2	GD2 41:1;O2	C83H148N4O34	18.8
GD2	GD2 42:2;O2	C84H148N4O34	17.6
GD3	GD3 34:1;O2	C68H121N3O29	10.8
GD3	GD3 35:1;O2	C69H123N3O29	11.8
GD3	GD3 36:0;O2	C70H127N3O29	14.3
GD3	GD3 36:1;O2	C70H125N3O29	13.1
GD3	GD3 36:2;O2	C70H123N3O29	11.4
GD3	GD3 37:1;O2	C71H127N3O29	14.3
GD3	GD3 38:1;O2	C72H129N3O29	15.8
GD3	GD3 39:1;O2	C73H131N3O29	17.2

Class	Species	Sum Formula	Retention Time [min]
GD3	GD3 40:1;O2	C74H133N3O29	18.5
GD3	GD3 41:1;O2	C75H135N3O29	19.6
GD3	GD3 41:2;O2	C75H133N3O29	17.1
GD3	GD3 42:1;O2	C76H137N3O29	20.7
GD3	GD3 42:2;O2	C76H135N3O29	18.5
GD3	GD3 42:2;O2	C76H135N3O29	19.3
GD3	GD3 43:2;O2	C77H137N3O29	19.8
GD3	GD3 44:1;O2	C78H141N3O29	21.5
GD3	GD3 44:2;O2	C78H139N3O29	20.5
GD3	GD3 46:2;O2	C80H143N3O29	20.7
GD3	GD3 34:1;O3	C68H121N3O30	10.5
GD3	GD3 36:1;O3	C70H125N3O30	10.7
GD3	GD3 42:1;O3	C76H137N3O30	17.4
GD3	GD3 42:2;O3	C76H135N3O30	18.2
GD3	GD3 42:2;O3	C76H135N3O30	18.6
GD3	GD3 43:2;O3	C77H137N3O30	19.9
GD3	GD3 44:2;O3	C78H139N3O30	20.6
GT1	GT1 34:1;O2	C93H161N5O47	9.2
GT1	GT1 36:0;O2	C95H167N5O47	11.2
GT1	GT1 36:1;O2	C95H165N5O47	11.1
GT1	GT1 37:1;O2	C96H167N5O47	12.1
GT1	GT1 38:1;O2	C97H169N5O47	13.2
GT1	GT1 39:1;O2	C98H171N5O47	14.5
GT1	GT1 40:1;O2	C99H173N5O47	15.7
GT2	GT2 34:1;O2	C87H151N5O42	—
GT2	GT2 36:1;O2	C89H155N5O42	—
GT2	GT2 38:1;O2	C91H159N5O42	—
GT3	GT3 36:1;O2	C81H142N4O37	—

Class	Species	Sum Formula	Retention Time [min]
GT3	GT3 38:1;O2	C83H146N4O37	—
GT3	GT3 40:1;O2	C85H150N4O37	—
GQ1	GQ1 34:1;O2	C104H178N6O55	8.6
GQ1	GQ1 35:1;O2	C105H180N6O55	9.6
GQ1	GQ1 36:0;O2	C106H184N6O55	10.5
GQ1	GQ1 36:1;O2	C106H182N6O55	10.4
GQ1	GQ1 36:2;O2	C106H180N6O55	9
GQ1	GQ1 37:1;O2	C107H184N6O55	11.1
GQ1	GQ1 38:1;O2	C108H186N6O55	12.4
GQ1	GQ1 39:1;O2	C109H188N6O55	13.2
GQ1	GQ1 40:1;O2	C110H190N6O55	15.1
GQ1	GQ1 41:1;O2	C111H192N6O55	16.5
GP1	GP1 36:1;O2	C117H199N7O63	10.1
GP1	GP1 38:1;O2	C119H203N7O63	12.2
GP1	GP1 40:1;O2	C121H207N7O63	14.6

Table 10: Library entries in positive and negative ionization mode.

Compound	Formula	positive ionization mode	negative ionization mode
GM1 34:1;O2	C71 H127 N3 O31	✓	✓
GM1 35:1;O2	C72 H129 N3 O31	✓	✓
GM1 36:1;O2	C73 H131 N3 O31	✓	✓
GM1 37:1;O2	C74 H133 N3 O31	✓	✓
GM1 38:1;O2	C75 H135 N3 O31	✓	✓
GM1 39:1;O2	C76 H137 N3 O31	✓	✓
GM1 40:1;O2	C77 H139 N3 O31	✓	✓
GM1 41:1;O2	C78 H141 N3 O31	✓	✓
GM1 42:1;O2	C79 H143 N3 O31	✓	✓
GM2 36:1;O2	C67 H121 N3 O26	✓	✓

Compound	Formula	positive ionization mode	negative ionization mode
GM2 38:1;O2	C69 H125 N3 O26	✓	✓
GM3 32:0;O2	C55 H102 N2 O21	—	✓
GM3 32:1;O2	C55 H100 N2 O21	—	✓
GM3 34:0;O2	C57 H106 N2 O21	✓	✓
GM3 34:1;O2	C57 H104 N2 O21	✓	✓
GM3 35:0;O2	C58 H108 N2 O21	✓	✓
GM3 35:1;O2	C58 H106 N2 O21	✓	✓
GM3 36:0;O2	C59 H110 N2 O21	✓	✓
GM3 36:1;O2	C59 H108 N2 O21	✓	✓
GM3 36:2;O2	C59 H106 N2 O21	—	✓
GM3 37:1;O2	C60 H110 N2 O21	✓	✓
GM3 38:0;O2	C61 H114 N2 O21	✓	✓
GM3 38:1;O2	C61 H112 N2 O21	✓	✓
GM3 38:2;O2	C61 H110 N2 O21	✓	✓
GM3 39:1;O2	C62 H114 N2 O21	✓	✓
GM3 40:0;O2	C63 H118 N2 O21	✓	✓
GM3 40:1;O2	C63 H116 N2 O21	✓	✓
GM3 40:2;O2	C63 H114 N2 O21	✓	✓
GM3 41:0;O3	C64 H120 N2 O22	—	✓
GM3 41:1;O2	C64 H118 N2 O21	✓	✓
GM3 41:2;O2	C64 H116 N2 O21	✓	✓
GM3 42:1;O2	C65 H120 N2 O21	✓	✓
GM3 42:2;O2	C65 H118 N2 O21	✓	✓
GM3 43:1;O2	C66 H122 N2 O21	✓	✓
GM3 43:2;O2	C66 H120 N2 O21	✓	✓
GM3 44:1;O2	C67 H124 N2 O21	✓	✓
GD1 34:1;O2	C82 H144 N4 O39	✓	✓
GD1 34:1;O3	C82 H144 N4 O40	—	✓

Compound	Formula	positive ionization mode	negative ionization mode
GD1 35:1;O2	C83 H146 N4 O39	✓	✓
GD1 36:0;O2	C84 H150 N4 O39	✓	✓
GD1 36:1;O2	C84 H148 N4 O39	✓	✓
GD1 36:2;O2	C84 H146 N4 O39	✓	✓
GD1 37:1;O2	C85 H150 N4 O39	✓	✓
GD1 38:1;O2	C86 H152 N4 O39	✓	✓
GD1 39:1;O2	C87 H154 N4 O39	✓	✓
GD1 40:1;O2	C88 H156 N4 O39	✓	✓
GD1 41:1;O2	C89 H158 N4 O39	—	✓
GD1 42:2;O2	C90 H158 N4 O39	✓	✓
GD2 34:1;O2	C76 H134 N4 O34	✓	✓
GD2 34:1;O3	C76 H134 N4 O35	—	✓
GD2 35:1;O2	C77 H136 N4 O34	✓	✓
GD2 36:1;O2	C78 H138 N4 O34	✓	✓
GD2 37:1;O2	C79 H140 N4 O34	✓	✓
GD2 38:0;O2	C80 H144 N4 O34	✓	✓
GD2 38:1;O2	C80 H142 N4 O34	✓	✓
GD2 38:2;O2	C80 H140 N4 O34	✓	✓
GD2 39:1;O2	C81 H144 N4 O34	✓	✓
GD2 40:1;O2	C82 H146 N4 O34	✓	✓
GD2 41:1;O2	C83 H148 N4 O34	✓	✓
GD3 34:1;O2	C68 H121 N3 O29	✓	✓
GD3 35:1;O2	C69 H123 N3 O29	✓	✓
GD3 36:0;O2	C70 H127 N3 O29	✓	✓
GD3 36:1;O2	C70 H125 N3 O29	✓	✓
GD3 36:2;O2	C70 H123 N3 O29	✓	✓
GD3 37:1;O2	C71 H127 N3 O29	✓	✓
GD3 38:1;O2	C72 H129 N3 O29	✓	✓

Compound	Formula	positive ionization mode	negative ionization mode
GD3 39:1;O2	C73 H131 N3 O29	✓	✓
GD3 40:1;O2	C74 H133 N3 O29	✓	✓
GD3 41:1;O2	C75 H135 N3 O29	✓	✓
GD3 41:2;O2	C75 H133 N3 O29	✓	✓
GD3 42:1;O2	C76 H137 N3 O29	—	✓
GD3 42:2;O2	C76 H135 N3 O29	✓	✓
GD3 43:2;O2	C77 H137 N3 O29	✓	✓
GD3 44:1;O2	C78 H141 N3 O29	✓	✓
GT1 36:1;O2	C95 H165 N5 O47	✓	✓
GT1 37:1;O2	C96 H167 N5 O47	—	✓
GT1 38:1;O2	C97 H169 N5 O47	✓	✓
GT1 40:1;O2	C99 H173 N5 O47	—	✓
GQ1 36:1;O2	C106 H182 N6 O55	✓	✓
GQ1 38:0;O2	C108 H188 N6 O55	—	✓
GQ1 38:1;O2	C108 H186 N6 O55	✓	✓
LacCer 32:1;O2	C44 H83 N O13	✓	—
LacCer 34:1;O2	C46 H87 N O13	✓	—
LacCer 36:1;O2	C48 H91 N O13	✓	—
LacCer 38:1;O2	C50 H95 N O13	✓	—
LacCer 40:1;O2	C52 H99 N O13	✓	—
LacCer 42:1;O2	C54 H103 N O13	✓	—
Gb3 34:1;O2	C52 H97 N O18	✓	—
Gb3 36:1;O2	C54 H101 N O18	✓	—
Gb3 38:1;O2	C56 H105 N O18	✓	—
Gb3 40:1;O2	C58 H109 N O18	✓	—
Gb3 42:1;O2	C60 H113 N O18	✓	—
Gb3 44:1;O2	C62 H117 N O18	✓	—

7 Declaration of Software and Tools Used

In preparing this thesis, a range of software tools and platforms were utilized to facilitate literature review, figure development, data analysis, and manuscript writing.

For literature review the following platforms were used: Google Scholar and the online library of the University of Vienna (u:search) for accessing and searching publications. Mendeley was employed for library management and to ensure consistent citation formatting throughout the manuscript.

For figure design and visualization, Canva and Adobe Illustrator were used to generate the graphics included in the thesis. Chemical structures were prepared using ChemDraw.

For data evaluation, MS/MS library generation and organization, as well as statistical analyses, the following software tools and databases were utilized: Compound Discoverer, LipidSearch, mzVault, Lipid Maps, SwissLipids.

For manuscript preparation, a range of tools supported drafting and refinement, including Overleaf, Microsoft Office applications (PowerPoint, Excel) for content organization, as well as DeepL and ChatGPT for grammar checks and phrasal adjustments.
Importantly, no content was generated solely by AI; all scientific interpretations, conclusions, and original writing remain entirely the author's own.