universität wien

MASTERARBEIT / MASTER'S THESIS

Titel der Masterarbeit / Title of the Master's Thesis

„Targeted isolation of undescribed flavonoids from *Matricaria chamomilla* using molecular networking"

verfasst von / submitted by

Anna Haberl, BSc

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

Magistra pharmaciae (Mag.pharm.)

Wien, 2022 / Vienna, 2022

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 066 605
Studienrichtung lt. Studienblatt / degree programme as it appears on the student record sheet:	Masterstudium Pharmazie
Betreut von / Supervisor:	Univ.-Prof. Dr. Judith M. Rollinger

TABLE OF CONTENT

1 ABSTRACT

Matricaria chamomilla L. (Asteraceae), also known as German chamomile, has been traditionally used for the treatment of various ailments including symptoms of gastro-intestinal complaints, wound healing, and common cold. It is one of the most researched medicinal plants and its main constituents are already well characterized.
The aim of this master thesis is the identification and subsequent isolation of so far undescribed minor compounds from *M. chamomilla* using molecular networking. Molecular networking is a dereplication strategy that enables the targeted search for undescribed or unknown natural products by the comparison of MS/MS fragmentation patterns of different compounds.

In this study, 27 microfractions of a chamomile flower crude extract were analyzed by molecular networking. Three undescribed flavonoid glycosides (compounds 1-3) have been identified and their putative structures were defined according to the network topology. Thereby, compounds 1 and 2 are potential derivatives of Quercetagetin-7-O-glucoside with a caffeoyl moiety and only differ in a methyl residue. Compound 3 was annotated as Spinacetin-7-O-glucoside.
For the targeted isolation of compounds 1-3, the chamomile flower crude extract was fractionated. Therefore, preparative liquid-liquid extraction followed by high-performance counter current chromatography (HPCCC) fractionation, size-exclusion chromatography and UPLC fractionation were deployed to isolate the compounds 1-3. Finally, 2.75 mg, 2.80 mg and 0.65 mg of compounds 1-3 were isolated, respectively, as well as five additional compounds. Purity > 95% was determined for all compounds, except for compound 2 (~94%) and two of the additionally isolated substances.

2 ZUSAMMENFASSUNG

Matricaria chamomilla L. (Asteraceae), auch bekannt als Echte Kamille, wird traditionell zur Behandlung verschiedener Beschwerden, wie gastrointestinalen Beschwerden, zur Wundheilung oder bei Erkältungskrankheiten, eingesetzt. Sie ist eine der am besten erforschten Heilpflanzen und ihre Hauptbestandteile wurden bereits umfassend charakterisiert.
Das Ziel dieser Masterarbeit ist die Identifizierung und anschließende Isolierung bisher unbeschriebener Nebenverbindungen aus *M. chamomilla* mittels molecular networking. Molecular networking ist eine Dereplikationsstrategie, die die gezielte Suche nach unbeschriebenen oder unbekannten Naturstoffen durch den Vergleich von MS/MS-Fragmentierungsmustern verschiedener Verbindungen ermöglicht.

In dieser Studie wurden 27 Mikrofraktionen eines Kamillenblüten-Rohextrakts mittels molecular networking analysiert. Drei unbeschriebene Flavonoidglykoside (Verbindungen 1–3) wurden dabei identifiziert und ihre mutmaßlichen Strukturen wurden gemäß der Topologie des Netzwerks definiert. Demnach handelt es sich bei den Verbindungen 1 und 2 um potentielle Derivate von Quercetagetin-7-O-glucosid mit einem Kaffeoylrest, die sich nur in einem Methylrest unterscheiden. Verbindung 3 wurde als Spinacetin-7-O-glucosid annotiert.
Für die gezielte Isolierung der Verbindungen 1-3 wurde der Kamillenblüten-Rohextrakt fraktioniert. Dabei wurden präparative Flüssig-Flüssig-Extraktion, Fraktionierung mittels Hochleistungsgegenstromchromatographie (HPCCC-Fraktionierung), Größenausschluss-chromatographie und eine UPLC-Fraktionierung eingesetzt, um die Verbindungen 1-3 zu isolieren. Schlussendlich konnten jeweils 2,75 mg, 2,80 mg und 0,65 mg der Verbindungen 1–3 isoliert werden, sowie fünf weitere bekannte Inhaltsstoffe. Für alle Verbindungen wurde eine Reinheit > 95% bestimmt, außer für Verbindung 2 (~94%) und zwei der zusätzlich isolierten Substanzen.

3 INTRODUCTION

3.1 Molecular networking

Molecular networking is a MS/MS-based dereplication method. Molecular networks (MNs) are constructed by comparison of MS/MS data of compounds within one or more samples (Aron et al., 2020). According to Yang et al. (2013), "*a molecular network is a visual representation of molecular relatedness (chemical similarity) of any given set of compounds*."

Figure 1 shows the basic principle of molecular networking. The first step is the generation of MS/MS spectra, which then have to be processed for further analysis.

All MS/MS scans are aligned by an algorithm which identifies identical spectra – and accordingly identical compounds. In the MN, each compound is represented as a single node (consensus cluster of identical MS/MS spectra). The MS/MS scans of all recorded molecules are compared to one another and cosine scores are calculated which represent the MS/MS structural similarity between nodes. Based on these cosine scores, detected spectrum-to-spectrum alignments are illustrated as connections – so-called edges – between the nodes. Nodes with structural similarity – and hence similar MS/MS fragmentation pattern – form clusters, which are referred to as molecular families. Nodes which are structurally unrelated to any other MS/MS spectra remain unconnected as single nodes.
The MS/MS scans are compared to MS spectral libraries, which suggest possible compound matches and/or analogues and provide further structural information helpful for dereplication and annotation. Besides the precursor masses, m/z differences between nodes are presented in the MN. These mass shifts can also contribute to the dereplication and annotation of compounds (Aron et al., 2020).

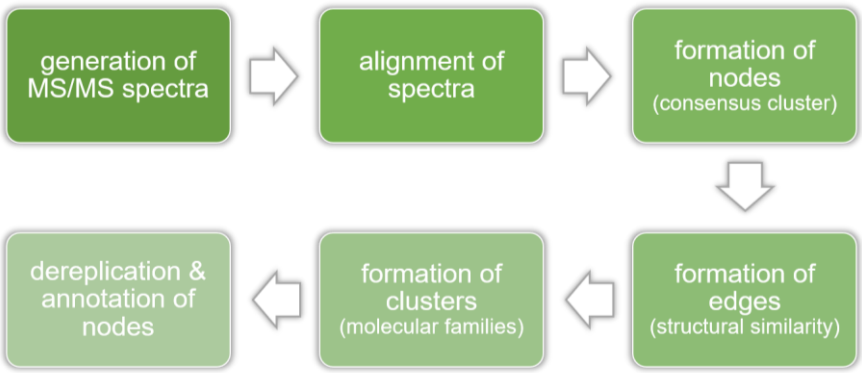


Figure 1: Basic principle of molecular networking

3.1.1 Molecular networking in the field of natural products research

A bioassay-guided fractionation often results in the isolation of already known natural products (NPs) (Olivon et al., 2017). Molecular networking enables the rapid dereplication of known compounds and potential analogues within complex extracts. It guides the isolation workflow towards unknown or undescribed constituents (Yang et al., 2013).

As the isolation of single bioactive NPs from their complex matrices is highly laborious, methods are required that indicate potentially bioactive constituents prior to any isolation steps. Molecular networking not only maps MS/MS spectral similarities of different compounds, but can also be combined with additional information like bioactivity or taxonomical data (Olivon et al., 2017).

Therefore, molecular networking is not only a strategy to find unknown or undescribed compounds, but also to distinguish between bioactive and non-active NPs. This contributes significantly to the optimization of the isolation process of targeted NPs.

3.2 Matricaria chamomilla

3.2.1 Therapeutic use of chamomile flowers

Matricaria chamomilla L., also known as German Chamomile, is a member of the Asteraceae family. It is one of the oldest and most commonly used medicinal herbs (Srivastava et al., 2010).

Chamomile and its preparations – mainly teas, extracts and essential oils – have been used for a variety of health issues for centuries. Its diverse medicinal properties have been studied extensively for years (El Joumaa & Borjac, 2022). The dried flowers of *M. chamomilla* (*Matricariae flos*) are traditionally used for their anti-inflammatory, antioxidant, sedative,

anxiolytic and spasmolytic effects. For the chamomile essential oil, an antimicrobial activity has been reported (Srivastava et al., 2010). *Matricariae flos* have been registered for "traditional use" by the European Medicines Agency (EMA) for following indications: (i) minor gastrointestinal complaints such as bloating and spasms, (ii) symptoms of common cold, (iii) minor ulcers and inflammations of the mouth and throat, (iv) irritations of skin and mucosae in the anal and genital region, (v) minor inflammation of the skin and superficial wounds (EMA/HMPC, 2015). Other medical applications of chamomile flowers and preparations thereof are insomnia, depression and anxiety. The essential oil of *M. chamomilla* is widely used in cosmetics and also popular in aromatherapy (Srivastava et al., 2010).
According to El Joumaa and Borjac (2022), the integration of chamomile preparations in the treatment of various medical conditions can be beneficial due to the diverse effects and relatively high safety in terms of toxicity.

3.2.2 Chemical composition of chamomile flowers

(i) Essential oil

The essential oil is primarily produced in the flowers of *M. chamomilla* (Schilcher et al., 2005) and mainly consists of sesquiterpenes such as azulenes (2-18%) – especially Chamazulene – α-Bisabolol (up to 50%), α-Bisabolol oxides A and B and trans-β-Farnesene (up to 45%). Other major constituents are spiroethers (cis- and trans-en-in-dicycloethers) with an amount of 20-30% (EMA/HMPC, 2015).

(ii) Phenolic compounds

Flavonoids:

Chamomile flowers contain several flavonoids, primarily Apigenin (~17%), Quercetin (~10%), Patuletin (~7%) and Luteolin (~2%) (McKay & Blumberg, 2006). The flavonoids are mainly present in the form of their glycosides (Avula et al., 2014; Tsivelika et al., 2021), whereas mainly multiply methoxylated flavonoids are present as aglycones (Avula et al., 2017; Xie et al., 2014). In addition, there are various esterified derivatives of Apigenin-7-O-glucoside within chamomile, such as derivatives with acetyl and/or malonyl residues (Avula et al., 2014; Švehlíková et al., 2004; Xie et al., 2014).

Phenolic acids:

Chlorogenic acid and its derivatives are found in chamomile (Zhao et al., 2019), as well as glycosylated phenolic acids such as cis- and trans-glucosyloxymethoxy cinammic acids (GMCAs) (Avula et al., 2014).

Coumarins:

Both coumarin glycosides and their aglycones are present within chamomile flowers (Avula et al., 2017; Petruľová-Poracká et al., 2013), with Umbelliferone and Herniarin being the main coumarins (McKay & Blumberg, 2006).

(iii) Other compounds

Other constituents of *M. chamomilla* are sesquiterpene lactones (Tschiggerl & Bucar, 2012), polyamines (Park et al., 2017) and polysaccharides (Slavov et al., 2019).

3.3 Aim of the work

The aim of this thesis is the identification and isolation of undescribed constituents from *M. chamomilla* using molecular networking. The research question is, if there are still undescribed minor compounds within this plant species, which has been extensively analyzed for decades. Therefore, molecular networking presents a suitable tool to answer this question.

In this study, 27 microfractions (MFs) from a chamomile flower crude extract – generated by Loidolt (2022) – were investigated. Based on these MFs a molecular network was generated to identify potential undescribed constituents. Fractionation of the chamomile flower crude extract was then performed for the targeted isolation of selected undescribed compounds. Figure 2 illustrates the basic workflow of this study.

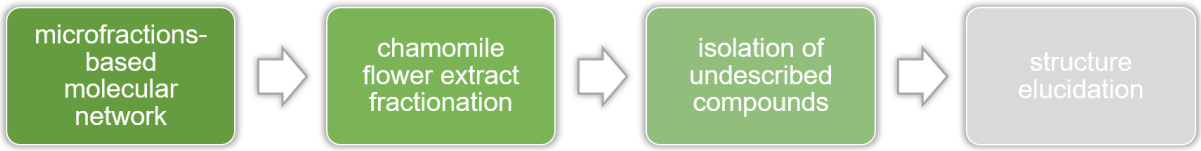


Figure 2: Basic workflow for identification and isolation of undescribed compounds from *M. chamomilla*

4 RESULTS AND DISCUSSION

4.1 Dereplication and annotation of chamomile microfractions

The 27 MFs generated by Loidolt (2022) and the chamomile flower crude extract (MatchaBDM_LS) were analyzed by Ultra performance liquid chromatography (UPLC), using the PDA, QDa and ELSD as detectors (see chapter 6.6.2; QDa and ELSD chromatograms see Appendix). Total ion chromatograms (TICs) in positive ionization mode were evaluated to assess the m/z values of supposed constituents. Most of the detected peaks in the crude extract were assigned to already well-known constituents of *M. chamomilla* that belong to the compound classes of flavonoids, sesquiterpenes, sesquiterpene lactones, phenolic acids and spiroethers. But in contrast to the crude extract, the TICs of the MFs revealed the occurrence of additional compounds, which could be annotated partly based on m/z data comparison with literature. Of 165 peaks detected, 31 peaks (~19%) could be assigned to already known substances from *M. chamomilla*. For some of the 134 remaining peaks, there are multiple possible annotations, while others could only be assigned to a certain compound class (e.g. flavonoid glycoside), but the majority remained entirely unidentified. (MFs annotation list see Appendix)

4.2 Molecular network generation from chamomile microfractions

4.2.1 Network generation

MS/MS data of the 27 MFs was recorded by Prof. Mehdi Beniddir (Université Paris-Scalay, France; see chapter 6.4.1). Using the software MZmine 2 (Pluskal et al., 2010), the data was processed and prepared for molecular networking. The MN was created by the web-based platform Global Natural Products Social Molecular Networking (GNPS) (Wang et al., 2016) and visualized with Cytoscape (ver. v3.9.1) (Shannon et al., 2003). The settings for the MN generation are described in detail in chapter 6.4. An overview of the MN derived from the 27 MFs is included in Figure 3.

4.2.2 Network analysis

In the course of molecular networking, five big clusters were generated. These so-called molecular families were then assigned to natural compound classes of *M. chamomilla* (see Figure 3). The first molecular family contains glycosylated phenolic compounds including

flavonoids and coumarins (1). The second cluster consists of chlorogenic acid derivatives (2). In the third family, coumaryl polyamines are clustered (3). The nodes in the fourth cluster represent sesquiterpenlactones (4). And the last molecular family contains spiroether compounds (5).

In addition, cis- and trans-GMCA were found as single nodes. Furthermore, a molecular family of flavonoid aglycones – including flavonoids like Apigenin (m/z = 271), Quercetin (m/z = 303) and Patuletin (m/z = 333) – was also formed. These nodes are "artefacts", because these aglycones are not present in the plant themselves, but are fragments resulting from the ionization of the glycosides. This was determined by comparing the retention times of the glycosides and the associated aglycones.

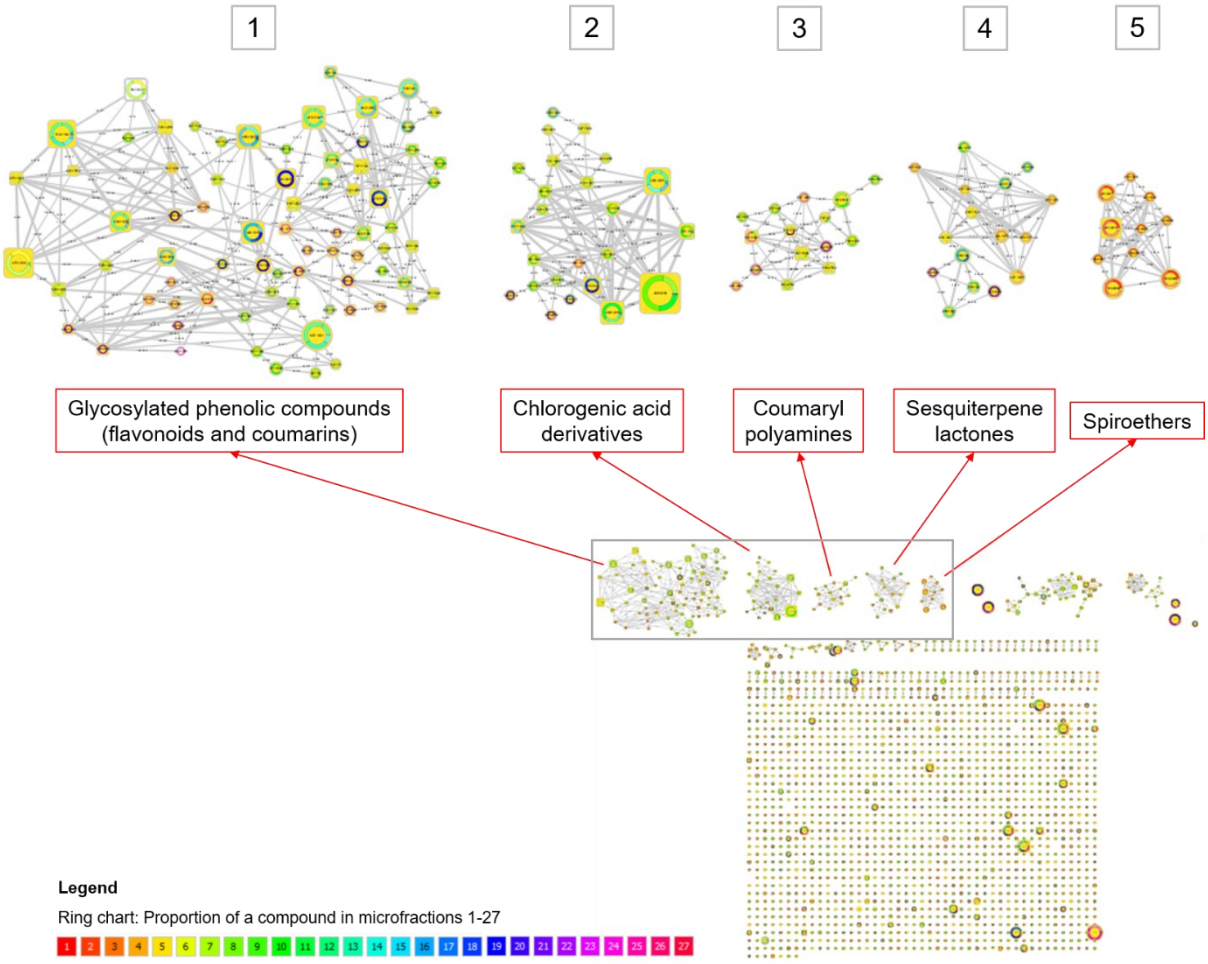


Figure 3: MN overview and molecular families of chamomile microfractions (MF1-27) with their assigned natural compound classes.

The largest cluster – referred to as molecular family 1 (Figure 4) – was selected for further in-depth analysis as a lot of the compounds had MS spectral library matches, which facilitated the dereplication and annotation of the nodes.

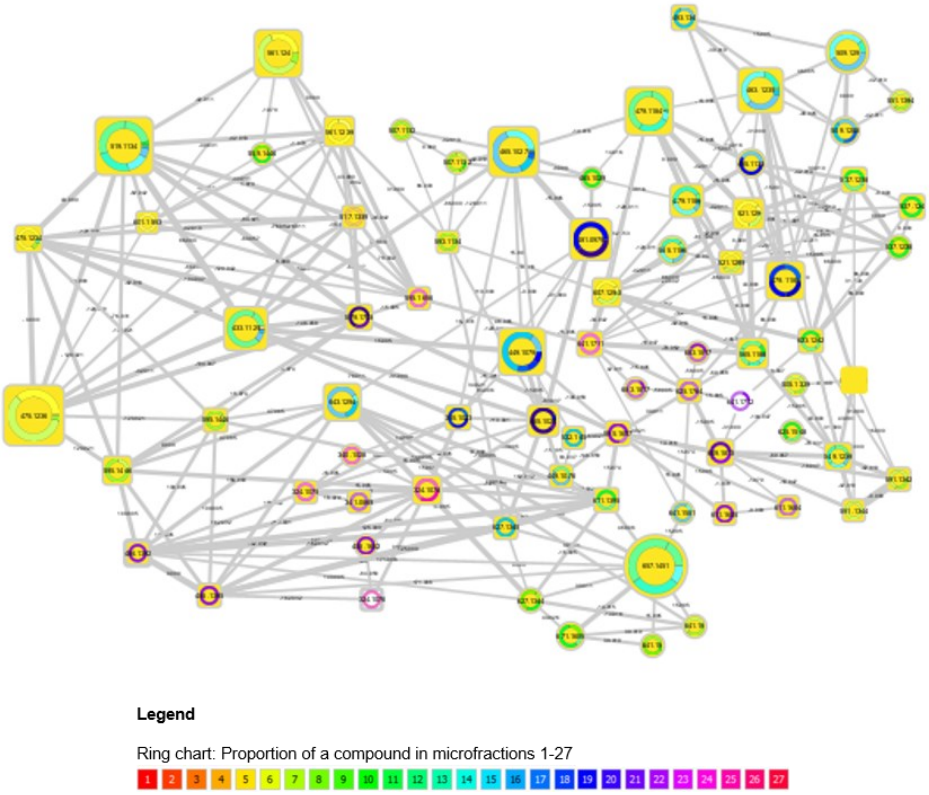


Figure 4: Molecular family 1 containing glycosylated flavonoids and coumarins

As mentioned above, molecular family 1 mainly consists of glycosylated flavonoids. Glycosylated coumarins are also present within this cluster. The analysis of molecular family 1 is given in Table 1.

Table 1: Analysis of molecular family 1; blue-colored nodes = dereplicated compounds from *M. chamomilla*; green-colored nodes = targeted undescribed glycosylated flavonoids; m/z = mass-to-charge ratio (positive mode); RT = retention time; MFs = chamomile microfractions (MF1-27)

node name	m/z	RT mean [min]	MFs	annotation	MS spectral library match	analog suggestion
1094	433.1129 [M+H]+	3.54	7 11 12	Apigenin-7-O-glucoside	Apigenin-7-O-glucoside	Aloe emodin
2584	475.1236 [M+H]+	4.06	5 6 7 8	Apigenin-7-O-?-acetyl-glucoside	6''-O-Acetyl-genistin	Apigenin-7-O-glucuronide
1374	475.1234 [M+H]+	4.37	6	Apigenin-7-O-?-acetyl-glucoside	6''-O-Acetyl-genistin	Apigenin-7-O-glucoside
804	519.1134 [M+H]+	3.94	10 11 12 13 14 15 16	Apigenin-7-O-malonyl-glucoside	6''-O-Malonylgenistin	Apigenin-7-O-glucuronide
2581	561.1240 [M+H]+	4.39	6 7 8	Apigenin-7-O-?-acetyl-malonyl-glucoside	-	Apigenin-7-O-glucuronide
3310	561.1239 [M+H]+	4.51	5 6	Apigenin-7-O-?-acetyl-malonyl-glucoside	-	Apigenin-7-O-glucoside
1701	559.1446 [M+H]+	4.25	8 9 10	Apigenin-7-O-?-acetyl-acetyl-acetyl-glucoside	-	-
4292	601.1553 [M+H]+	4.85	5	Apigenin-7-O-?-acetyl-acetyl-acetyl-acetyl-glucoside	-	Apigenin-7-O-glucuronide
3692	517.1339 [M+H]+	5.02	4	Apigenin-7-O-?-acetyl-acetyl-glucoside	-	Apigenin-7-O-glucoside
5502	579.1708 [M+H]+	3.34	21 22	Rhoifolin (= Apigenin-7-O-neohesperidoside)	5-Hydroxy-7-[3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxy-2-[4-(3,4,5-trihydroxy-6-methyloxan-2-yl)oxyphenyl]chromen-4-one	Apigenin-7-O-glucoside

6262	595.1656 [M+H]+	2.53	25	Apigenin-7,?-O-diglucoside	-	Pelargonin
2575	595.1446 [M+H]+	4.16	7 8	Apigenin-7-O-?-caffeoyl-glucoside	-	Apigenin-7-O-glucuronide
2587	595.1446 [M+H]+	4.40	7 8	Apigenin-7-O-?-caffeoyl-glucoside	-	Apigenin-7-O-glucuronide
5726	341.0869 [M+H]+	1.97	24	Aesculin	Aesculin	Esculetin (6,7-Dihydroxycoumarin)
6114	486.1392 [M+H]+	2.34	23	Caffeoyl-umbelliferon-glucoside	-	3-[(2S,3R,4S,5S,6R)-4,5-dihydroxy-6-(hydroxymethyl)-3-[(2S,3R,4S,5R,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxyoxan-2-yl]oxy-5-hydroxy-2-(4-hydroxyphenyl)-7-methoxychromen-4-one
6112	486.1395 [M+H]+	2.54	23	Caffeoyl-umbelliferon-glucoside	-	3-[(2S,3R,4S,5S,6R)-4,5-dihydroxy-6-(hydroxymethyl)-3-[(2S,3R,4S,5R,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxyoxan-2-yl]oxy-5-hydroxy-2-(4-hydroxyphenyl)-7-methoxychromen-4-one
6252	324.1075	1.73	25	Coumarin glycoside	-	Aesculin
2391	324.1076	0.61	26	Coumarin glycoside	-	Aesculin
582	324.1078	0.85	26 27	Coumarin glycoside	-	Aesculin
6118	408.1652	3.06	23	Coumarin glycoside	-	-
2284	355.1023	2.84	17 18 19	Coumarin glucoside	Undulatoside A	Aesculin
823	643.1294 [M+H]+	3.29	14 15 16	Quercetagetin-7-O-caffeoyl-?-glucoside	-	Myricetin 3-O-β-D-galactoside 6''-O-gallate
2498	465.1028 [M+H]+	2.79	20 21 22	Quercetin-?-O-glucoside	Quercetin-4'-O-glucoside	Aesculin

801	465.1027 [M+H]+	3.10	13 14 15 16 17	Quercetin-?-O-glycoside	Hyperoside	Quercetin
5429	481.0979 [M+H]+	2.74	19 21	Quercetagetin-7-O-glucoside	Gossypin	Gossypetin
1818	507.1132 [M+H]+	3.63	7 8 9 10	Quercetin-7-O-?-acetyl-glucoside	-	-
2620	507.1132 [M+H]+	3.93	8	Quercetin-7-O-?-acetyl-glucoside	-	-
2597	593.1134 [M+H]+	3.93	8	Quercetin-7-O-?-malonyl-acetyl-glucoside	-	Quercetin-4'-O-glucoside
1070	627.1345 [M+H]+	3.25	16	Quercetin-7-O-?-caffeoyl-glucoside	-	Quercetin-3-O-rhamnoside
1674	627.1344 [M+H]+	3.85	8 9 10	Quercetin-7-O-?-caffeoyl-glucoside	-	-
1061	449.1079 [M+H]+	3.14	15 16 17 18 19	Luteolin-7-O-glucoside	Glucoluteolin	Apigetrin
1188	449.1076 [M+H]+	3.54	15 16	Luteolin-4'-O-glucoside	Luteolin-4'-O-glucoside	Luteolin-7-O-glucuronide
1116	532.145	3.41	16	-	-	Luteolin-7-O-glucuronide
5745	595.1657 [M+H]+	3.04	23 24	Luteolin-?-O-neohesperidoside	Luteolin-7-O-rutinoside	Cyanidin-3-O-rutinoside
1659	611.1395 [M+H]+	3.82	10 11	Luteolin-?-O-?-caffeoyl-glucoside	-	[6-[2-(3,4-dihydroxyphenyl)-8-hydroxy-4-oxochromen-7-yl]oxy-3,4,5-trihydroxyoxan-2-yl]methyl (E)-3-(4-hydroxyphenyl)prop-2-enoate

807	657.1451 [M+H]+	3.67	11 12 13 14	caffeoylated flavonoid glycoside with 5 OH + 1 OCH3 → caffeoylated Quercetagetin-glucoside derivative?	-	-
1650	671.1605	4.02	9 10 11	caffeoylated flavonoid glycoside with 4 OH + 2 OCH3	-	-
1762	641.1500	4.06	8 9 10	caffeoylated flavonoid glycoside with 4 OH + 1 OCH3	-	-
2614	641.1500	4.22	7 8	caffeoylated flavonoid glycoside with 4 OH + 1 OCH3	-	-
884	641.1501	3.57	14 15	caffeoylated flavonoid glycoside with 4 OH + 1 OCH3	-	-
6120	611.1605 [M+H]+	2.98	23	Luteolin-?,?-O-diglucoside	-	Quercetin-3-O-glucosyl-rhamnosyl-glucoside
5737	611.1604 [M+H]+	2.74	24	Luteolin-?,?-O-diglucoside	-	Quercetin-3-O-glucosyl-rhamnosyl-glucoside
803	463.1235 [M+H]+	3.61	12 13 14 15 16	Chrysoeriol-7-O-glucoside	Hispidulin-4'-O-glucoside	Scutellarein-4'-methylether
5501	609.1813 [M+H]+	3.38	22 23	Chrysoeriol-7-O-neohesperidoside	Chrysoeriol-7-O-neohesperidoside	Diosmin
820	549.1239 [M+H]+	3.98	11 12 13 14 15	Chrysoeriol-7-O-?-malonyl-glucoside	-	Chrysoeriol-7-O-neohesperidoside
2956	591.1344 [M+H]+	4.58	7	Chrysoeriol-7-O-?-malonyl-acetyl-glucoside	-	Chrysoeriol-7-O-neohesperidoside
2601	591.1342 [M+H]+	4.46	7 8	Chrysoeriol-7-O-?-malonyl-acetyl-glucoside	-	Chrysoeriol-7-O-neohesperidoside

1671	625.1548 [M+H]+	4.02	10 11	Chrysoeriol-7-O-?-malonyl-glucoside with 1 OH + 2 OCH3	-	-
2606	505.1341 [M+H]+	4.14	7 8	Chrysoeriol-7-O-?-acetyl-glucoside	-	Chrysoeriol-7-O-glucoside
2640	505.1339 [M+H]+	4.46	7 8	Chrysoeriol-7-O-?-acetyl-glucoside	-	-
1678	465.1029	3.74	10	flavonoid glucoside with 5 OH	-	-
815	479.1184 [M+H]+	3.57	10 11 12 13 14	Isorhamnetin-?-O-glucoside	-	Tamarixetin
806	479.1185 [M+H]+	3.45	12 13 14 15 16 17	Isorhamnetin-?-O-glucoside	-	Tamarixetin
1108	479.1185 [M+H]+	3.24	16 17 18 19 20 21	Isorhamnetin-7-O-glucoside	-	Jaceoside
2926	521.129 [M+H]+	4.15	5 6 7	Isorhamnetin-7-O-?-acetyl-glucoside	-	Isorhamnetin-3-O-glucoside
2953	521.1289 [M+H]+	4.44	5 6 7	Isorhamnetin-7-O-?-acetyl-glucoside	-	Isorhamnetin-3-O-glucoside
813	565.1186 [M+H]+	3.69	11 12 13 14 15	Isorhamnetin-?-O-?-malonyl-glucoside	-	-
845	565.1188 [M+H]+	3.96	10 11 12 13 14 15	Isorhamnetin-?-O-?-malonyl-glucoside	-	Isorhamnetin-3-O-glucoside

2643	607.1293 [M+H]+	4.44	5 6 7 8	Isorhamnetin-?-O-?-malonyl-acetyl-glucoside	-	Isorhamnetin-3-O-galactoside
6235	641.1711 [M+H]+	2.29	25	Isorhamnetin-?,?-O-diglucoside	Rhamnetin-3-O-sophoroside	3-[(2S,3R,4S,5S,6R)-4,5-dihydroxy-6-(hydroxymethyl)-3-[(2S,3R,4S,5R,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxyoxan-2-yl]oxy-5-hydroxy-2-(4-hydroxyphenyl)-7-methoxychromen-4-one
5721	625.1764	3.10	24	flavonoid neohesperidoside with 4 OH + 1 OCH3	-	Chrysoeriol-7-O-neohesperidoside
5725	683.1817 [M+H]+	2.98	23 24	Isorhamnetin-?,?-O-?-acetyl-diglucoside	-	-
5739	683.1817	2.82	23 24	-	-	-
5720	641.1712	3.10	23 24	flavonoid neohesperidoside with 4 OH + 2 OCH3	-	5-hydroxy-3-(4-methoxyphenyl)-7-[3,4,5-trihydroxy-6-[(3,4,5-trihydroxy-6-methyloxan-2-yl)oxymethyl]oxan-2-yl]oxychromen-4-one
1223	495.1133 [M+H]+	3.22	15 19	Patuletin-7-O-glucoside	-	-
1509	537.1238 [M+H]+	3.70	10 11 12	Patuletin-7-O-?-acetyl-glucoside	-	2-(3,4-dihydroxyphenyl)-3,5-dihydroxy-8-methoxy-7-[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxychromen-4-one
1653	537.1240 [M+H]+	3.62	10	Patuletin-7-O-?-acetyl-glucoside	-	2-(3,4-dihydroxyphenyl)-3,5-dihydroxy-8-methoxy-7-[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxychromen-4-one

1654	537.1238 [M+H]+	3.62	8 9 10	Patuletin-7-O-?-acetyl-glucoside	-	-
1506	623.1242 [M+H]+	3.94	9 10 11 12	Patuletin-7-O-?-malonyl-acetyl-glucoside	-	2-(3,4-dihydroxyphenyl)-3,5-dihydroxy-8-methoxy-7-[(2S,3R,4S,5S,6R)-3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]oxychromen-4-one
800	509.129 [M+H]+	3.62	13 14 15	flavonoid glucoside with 4 OH + 2 OCH3 → Spinacetin-7-O-glucoside?	-	-
897	509.1288	3.50	14 15 16 17	flavonoid glucoside with 4 OH + 2 OCH3	-	-
2609	551.1394	4.12	7 8	flavonoid acetyl-glucoside with 4 OH + 2 OCH3	-	-
1071	493.134	3.70	15 16	flavonoid glucoside with 3 OH + 2 OCH3	-	Hispidulin-4'-O-glucoside

The analysis of molecular family 1 revealed that the clustered flavonoid glycosides have following structural differences:

(i) the amount and position of their hydroxy and methoxy groups,
(ii) the amount and type of glycoside moieties
(iii) and the esterification of glycosides with acetyl, malonyl and/or caffeoyl moieties.

Starting from the dereplication of certain compounds like Apigenin-7-O-glucoside and Aesculin, it was possible to annotate the nodes based on:

(i) molecular weight,
(ii) m/z differences between connected compounds,
(iii) the dereplication and annotation of the chamomile microfractions (see chapter 4.1)

(iv) and literature research.

In the course of the dereplication process of molecular family 1, already known flavonoid glycosides and glycosylated coumarins from *M. chamomilla* were identified (listed in Table 2; see blue colored nodes in Table 1).

Table 2: Dereplicated flavonoid glycosides and glycosylated coumarins from *M. chamomilla* with their respective molecular weight.

known compounds	**MW [g/mol]**
Apigenin-7-O-glucoside	432
Luteolin-7-O-glucoside	448
Patuletin-7-O-glucoside	494
Quercetin-7-O-glucoside	464
Quercetagetin-7-O-glucoside	480
Isorhamnetin-3-O-glucosid	478
Isorhamnetin-7-O-glucosid	478
Chrysoeriol-7-O-glucosid	462
Apigenin-7-O-?-acetyl-glucoside	474
Apigenin-7-O-?-malonyl-glucoside	518
Apigenin-7-O-?-malonyl-?-acetyl-glucoside	561
Apigenin-7-O-neohesperidoside (Rhoifolin)	578
Apigenin-7-O-?-caffeoyl-glucoside	595
Aesculin (glucoside of Aesculetin)	340

In addition, flavonoid glycosides and glycosylated coumarins were found that could not be dereplicated for *M. chamomilla*. The presence of esterified derivatives was already known for Apigenin (e.g. caffeoylated Apigenin-7-O-glucoside), but presumably not for other contained flavonoids within chamomile. Not identified glycosylated coumarins as well as caffeoylated Umbelliferone glucosides are also present within molecular family 1.

Nodes which were found in molecular family 1 that could not be dereplicated for *M. chamomilla* were of high interest fur further analysis. For quantitative assessment of the compounds, the UPLC-ELSD chromatograms of the MFs were also considered (see

Appendix). Finally, three compounds which were annotated as undescribed compounds for *M. chamomilla* by molecular networking and were detected by ELSD within the MFs were selected for targeted isolation (see Figure 5).

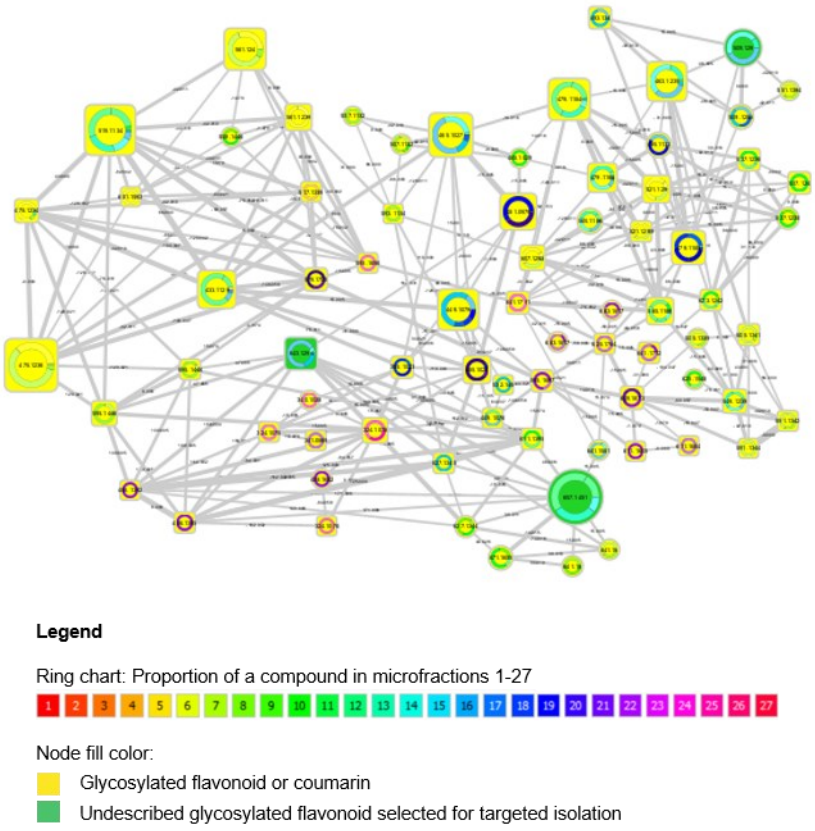


Figure 5: Undescribed glycosylated flavonoids (compounds 1-3; green-colored nodes) of molecular family 1 selected for targeted isolation.

Based on the MN analysis, the three nodes are glycosylated flavonoids (see green marked nodes in Table 1). Figure 6 shows the putative annotations and derived structures of the undescribed compounds 1-3 considering the MN topology and the chemical structures of known compounds from chamomile. It is assumed that all three compounds belong to the flavonoid class of flavonols. The hydroxyl and methoxy groups of the compounds were assigned to the flavonoid scaffold according to the putative structures of the neighbor nodes. A β-D-glucose is expected to be linked to the flavonoids in position 7 as described for most flavonoid glycosides found in chamomile flowers. Compounds 1 and 2 only differ in one methyl residue. Since compound 1 was in close proximity to nodes annotated as Isorhamnetin derivatives, it is expected that this compound possesses a methoxy group in position 3', while compound 2 has a hydroxyl group in this position. Compounds 1 and 2 both carry a caffeoyl moiety which is predicted to be in position 6''. Compounds 2 and 3 are

presumably compounds known from other species (Fursa et al., 1969), whereas compound 1 is apparently unknown.

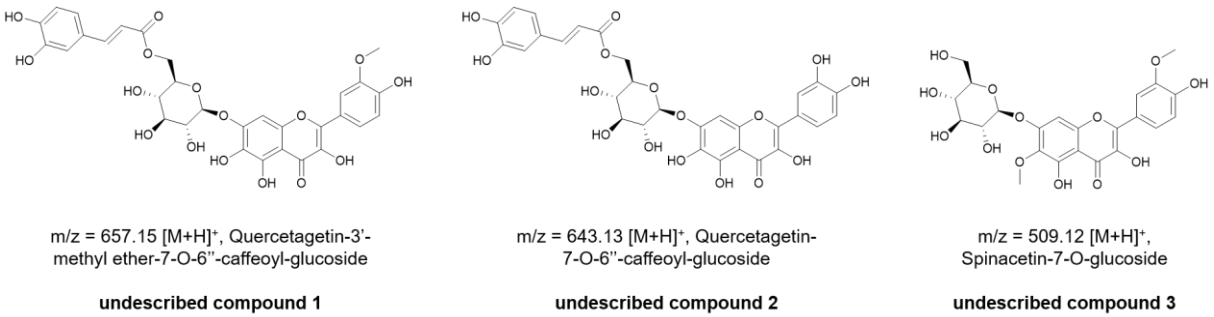


Figure 6: m/z ratios, putative annotations and derived structures of undescribed glycosylated flavonoids (compounds 1-3) of *M. chamomilla*

Since these three constituents could not be dereplicated for *M. chamomilla*, isolation was conducted to confirm the identity of the undescribed compounds 1-3. Therefore, it was decided to perform a separation and fractionation of a chamomile flower crude extract (MatchaBDM_LS) generated by Loidolt (2022). The aim was to isolate compounds 1-3 followed by NMR structure elucidation which will verify the correctness of the predicted structures.

Furthermore, several compounds that could not be identified in the annotation and dereplication process of the 27 MFs (described in chapter 4.1) were found when analyzing molecular family 1 (MFs annotation list see Appendix). The principle of molecular networking made it not only possible to assign compounds to the substance class of glycosylated flavonoids or coumarins, but also well-supported assumptions could be made about the putative structures.

4.3 Targeted isolation of undescribed compounds

In Figure 7, an overview of the MN-guided isolation workflow of undescribed compounds 1-3 is given. Each step is explained in detail in the following chapters.

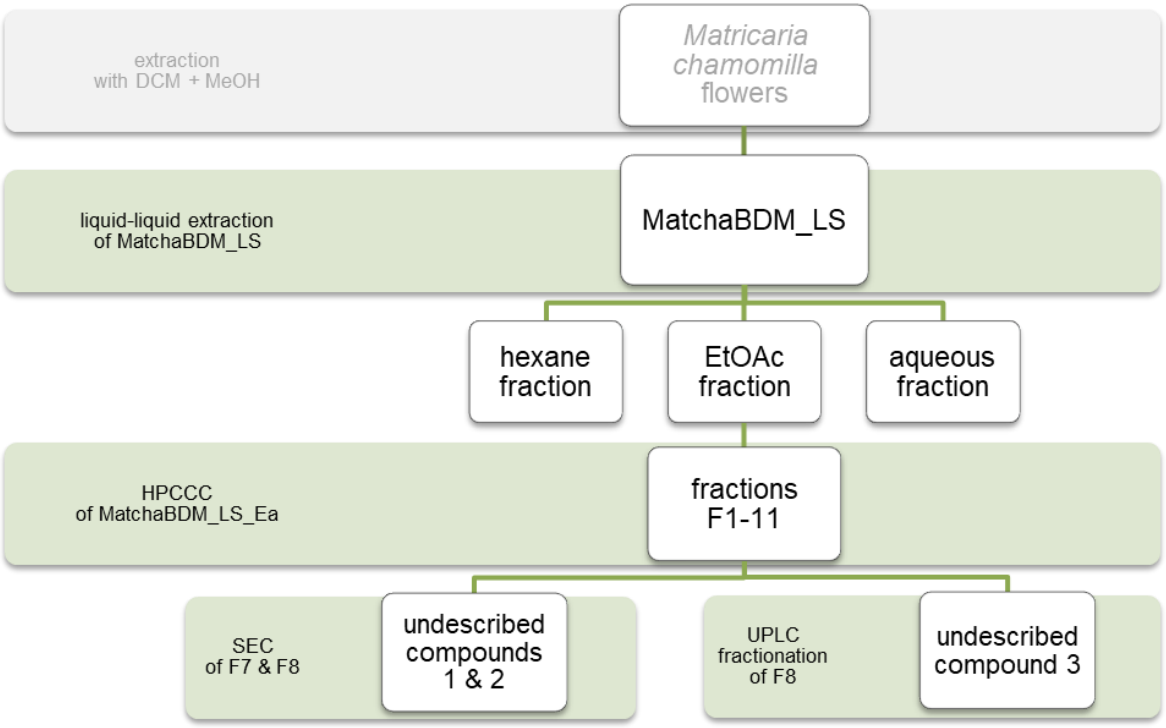


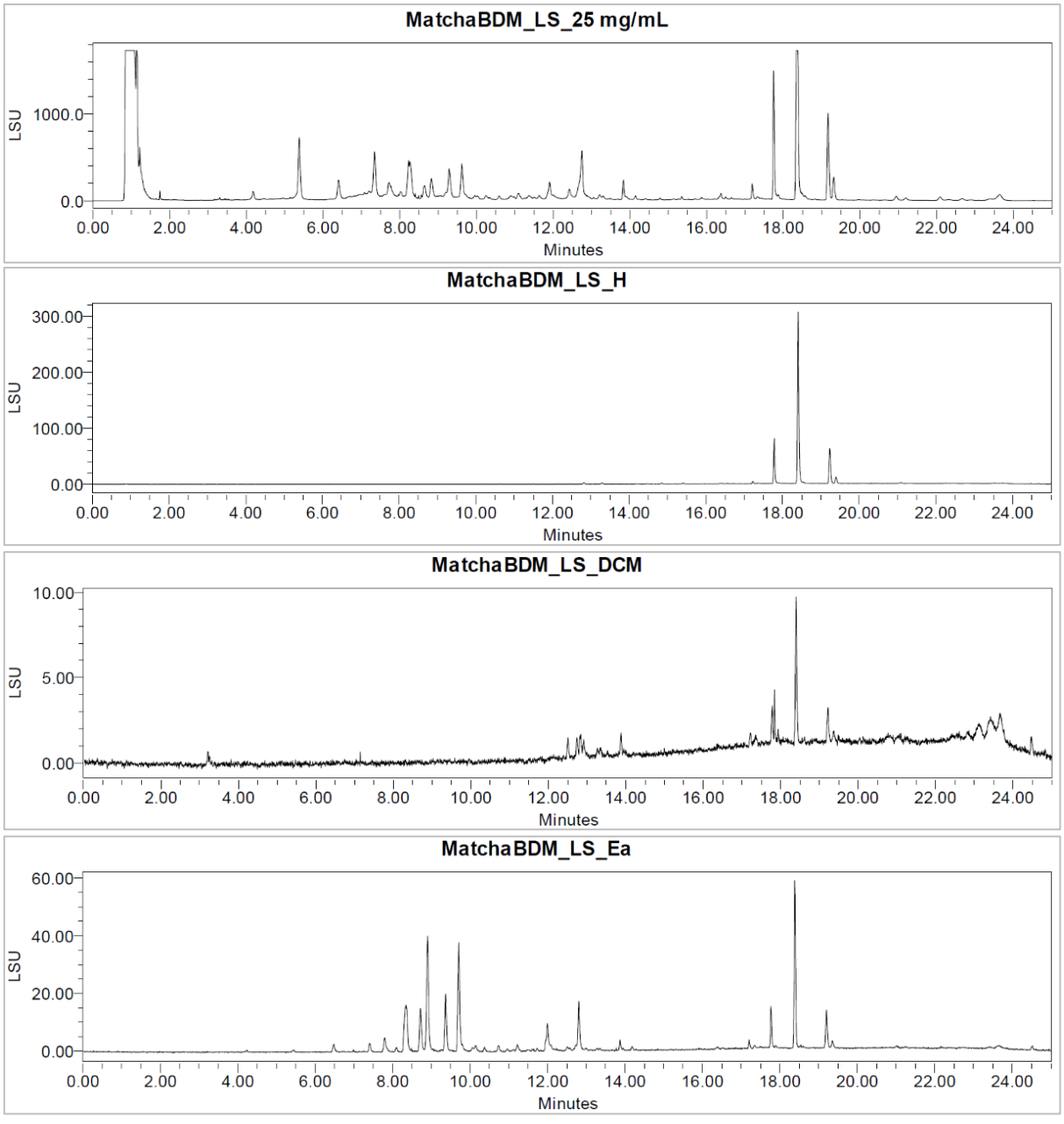
Figure 7: Phytochemical workflow for targeted isolation of undescribed compounds 1-3 from chamomile flower crude extract (MatchaBDM_LS); MatchaBDM_LS_Ea = ethyl acetate fraction of the crude extract; HPCCC = High-performance counter current chromatography; SEC = Size-exclusion chromatography; UPLC = Ultra performance liquid chromatography.

4.3.1 Liquid-liquid extraction

4.3.1.1 Analytical separation of extract components

A liquid-liquid extraction (LLE) of the chamomile flower crude extract (MatchaBDM_LS) was performed with several solvents that differ in their polarity. The aim was to separate the unidentified lipophilic and hydrophilic components of the extract, as well as to obtain an enriched fraction containing compounds 1-3, which were the components of interest. Therefore, the crude extract was separated with ddH_2O and the following organic solvents: hexane, dichloromethane, ethyl acetate and butanol. At the end of this separation process (see chapter 6.5.1), five LLE fractions were obtained.

The five LLE fractions were then measured using UPLC-PDA-ELSD (method see chapter 6.6.2). Figure 8 shows the ELSD chromatograms of the five fractions and the crude extract (MatchaBDM_LS 25 mg/mL). The lipophilic compounds (retention time span ~17.80-19.40 min), which are the predominant constituents quantitatively, were separated by hexane, but not completely. Therefore, it was concluded that more hexane extraction steps were required to improve the separation of the lipophilic compounds. The hydrophilic components (retention time around 0.90 min) mainly accumulated in ddH_2O. Ethyl acetate turned out to be the most suitable solvent to gain the components in the retention time range of interest (~6.00-14.00 min) from the extract. The results of this analytical separation process showed that the extraction steps with dichloromethane and butanol did not contribute significantly to the desired outcome and were neglected in subsequent extractions.



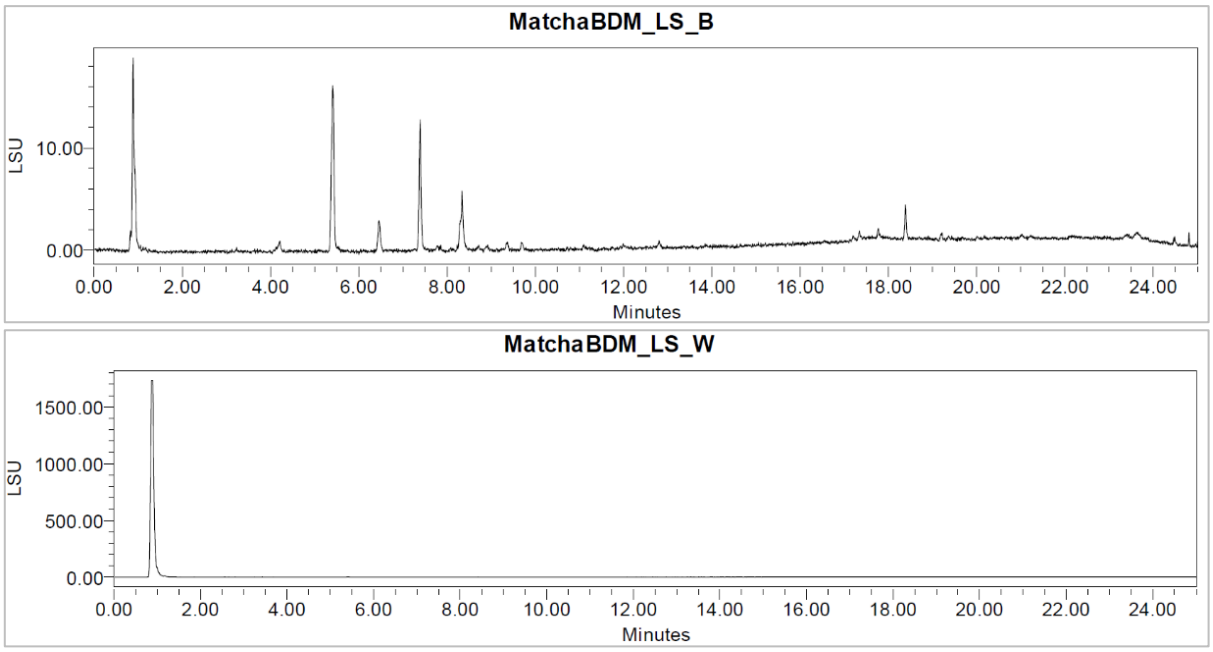


Figure 8: UPLC-ELSD chromatograms of chamomile flower crude extract (MatchaBDM_LS 25 mg/mL), hexane fraction (H), dichloromethane fraction (DCM), ethyl acetate fraction (Ea), butanol fraction (B), aqueous fraction (W) using method MC_032B90_PE_25_5.

A second analytical LLE was performed, this time using only hexane followed by ethyl acetate as organic solvents. In contrast to the first attempt, the extraction was now carried out five times instead of three times with the same amount of ddH_2O. In addition, twice the amount of hexane and ethyl acetate was used.

UPLC-PDA-ELSD measurements of the hexane and the ethyl acetate fraction were performed. The ELSD chromatograms (Figure 9) showed, that the refined separation process resulted in more lipophilic components being removed by hexane. However, in terms of quantity, the ethyl acetate fraction still mainly contained lipophilic components. It was concluded that the number of hexane extraction steps needed to be further increased to deplete these lipophilic compounds.

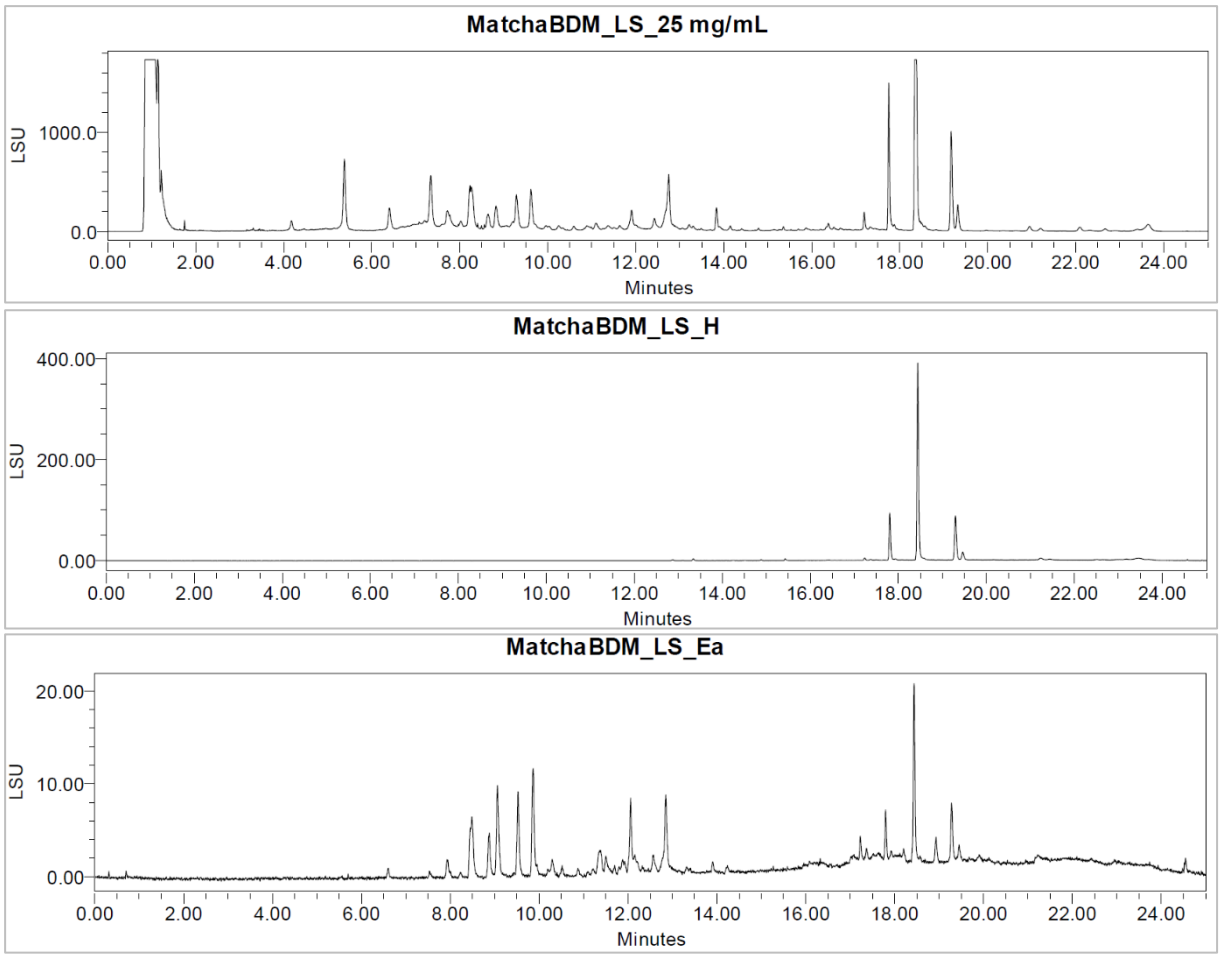


Figure 9: UPLC-ELSD chromatograms of chamomile flower crude extract (MatchaBDM_LS 25 mg/ml), hexane fraction (H), ethyl acetate fraction (Ea) using method MC_032B90_PE_25_5.

4.3.1.2 Targeted enrichment and depletion of extract components

To gain sufficient amounts of compounds 1-3, a preparative LLE of the crude extract was performed (see chapter 6.5.2). The extract was processed a total of ten times with hexane and then eight times with ethyl acetate.

After separation and evaporation, a UPLC-PDA-ELSD measurement of the three LLE fractions obtained was performed. The ELSD chromatograms in Figure 10 show that the lipophilic components were significantly more depleted than in the previous analytical tests due to the increased extraction steps with hexane. The hydrophilic compounds accumulated in ddH_2O. The ethyl acetate fraction predominantly contained those components within the retention time range of interest (~6.00-14.00 min), including compounds 1-3.

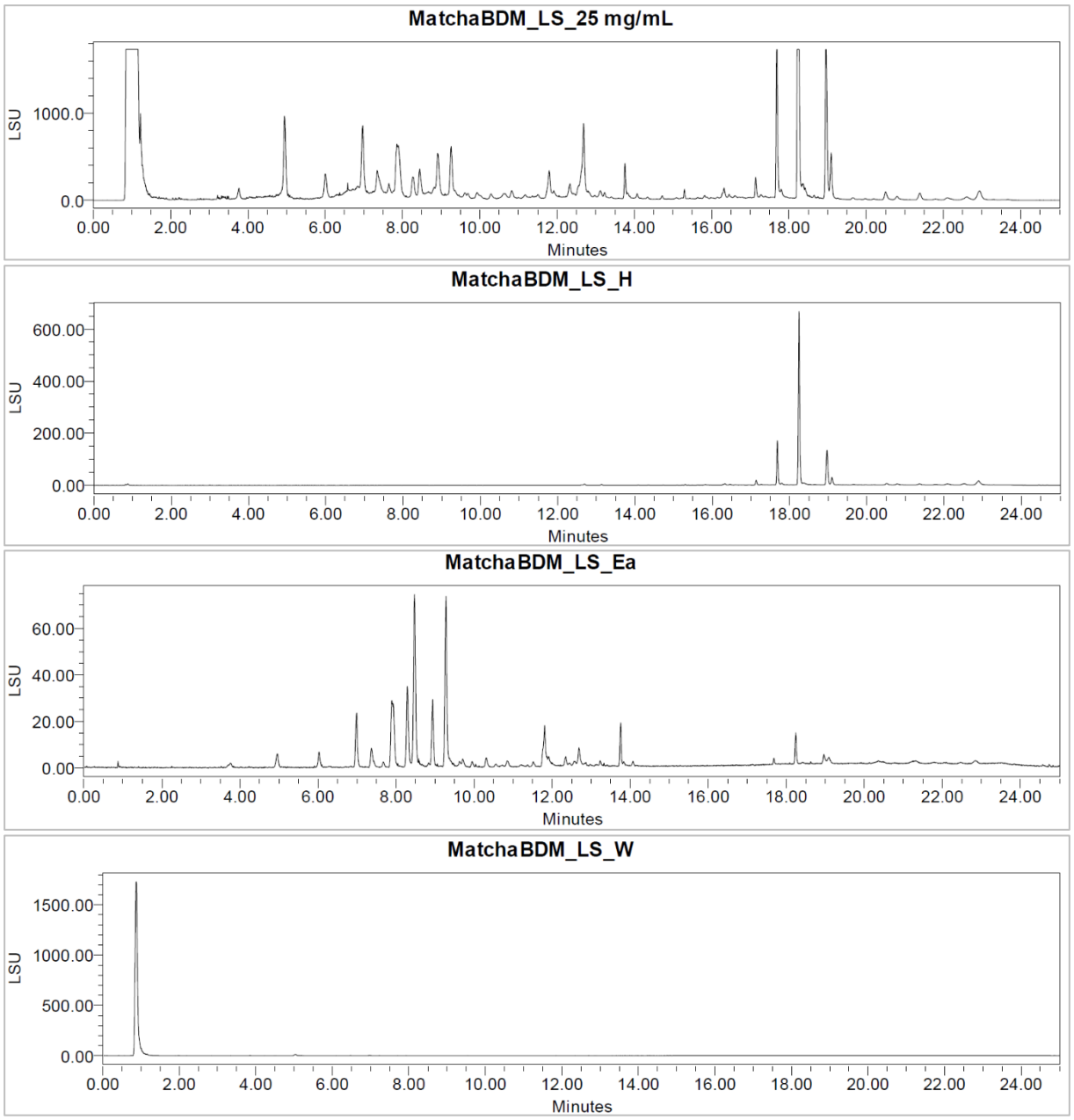


Figure 10: UPLC-ELSD chromatograms of chamomile flower crude extract (MatchaBDM_LS 25 mg/mL), hexane fraction (H), ethyl acetate fraction (Ea), aqueous fraction (W) using method MC_032B90_PE_25_5.

Thin layer chromatography (TLC) of the three LLE fractions and the crude extract (MatchaBDM_LS 10 mg/mL) was performed (Figure 11) as complementary method (TLC system described in 6.6.1). In addition, Apigenin-7-O-glucoside (1 mg/mL) was added as a reference substance for analysis. The lipophilic constituents of the hexane fraction were located near the solvent front (R_f = 0.9). The constituents from the ethyl acetate fraction covered the whole TLC plate including Apigenin-7-O-glucoside with an R_f value of 0.7. The hydrophilic components from the aqueous fraction were found at R_f < 0.3 expect one spot located at R_f = 0.55.

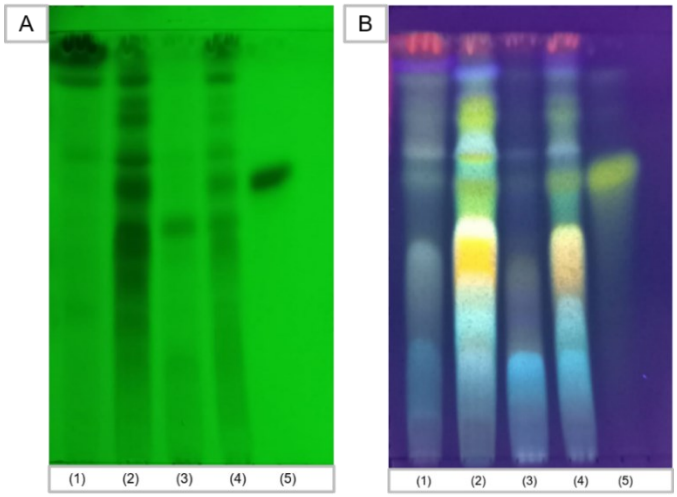


Figure 11: TLC of (1) hexane fraction, (2) ethyl acetate fraction, (3) aqueous fraction, (4) chamomile flower crude extract (MatchaBDM_LS 10 mg/mL), (5) Apigenin-7-O-glucoside (1 mg/mL); detected at (A) UV_{254}, (B) UV_{366} after spraying with nature substance reagent A/PEG400

4.3.2 HPCCC fractionation

In order to separate the ethyl acetate LLE fraction (MatchaBDM_LS_Ea), high-performance counter current chromatography (HPCCC) was used hyphenated with an Interchim PuriFlash® 4250 device (see chapter 6.6.3). Two semi-preparative runs were performed. The ELSD chromatograms of the two runs (see Appendix) recorded by the Interchim device were comparable to a large extent, which shows the reproducibility of this separation method.

185 fractions were collected in the first run and 175 fractions in the second run. TLC of both runs was performed to analyze the collected fractions (TLC see Appendix). TLC confirmed that the separation process was comparable for both runs. The collected fractions of both runs were pooled to 12 fractions (HPCCC fractions) according to the TLC analysis.

UPLC-PDA-QDa and UPLC-PDA-ELSD measurements were performed to determine in which HPCCC fractions compounds 1-3 were located (all chromatograms and scans see Appendix). In addition, no differences between fractions 9 and 10 were observed. They were subsequently pooled to one fraction which resulted in 11 final HPCCC fractions.

The annotation of the QDa scans (negative mode) and the ELSD chromatograms of the fractions showed, that fractions 7 and 8 (F7 and F8; Figure 12 and Figure 13) contained the undescribed compounds 1-3.

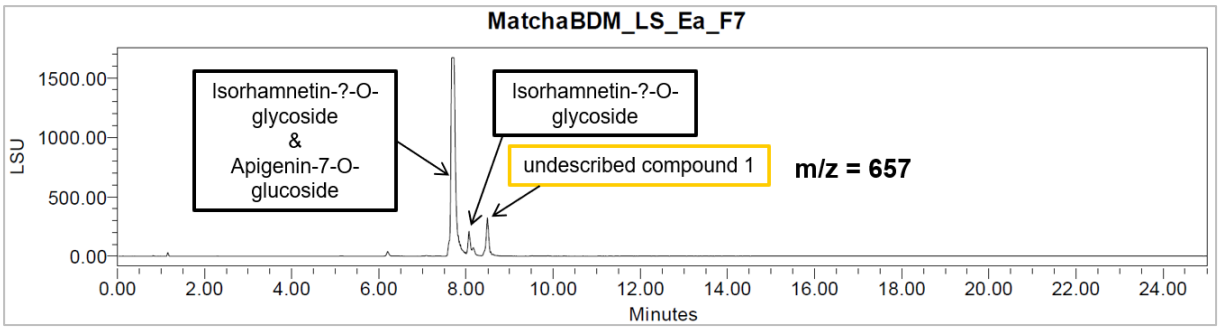


Figure 12: UPLC-ELSD chromatogram of HPCCC fraction F7 with annotated compounds using method MC_032B90_PE_25_5

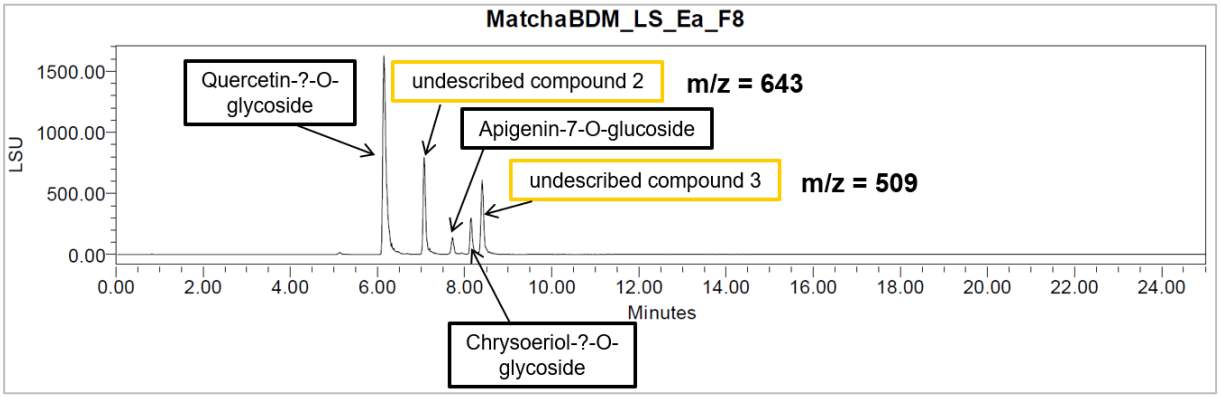


Figure 13: UPLC-ELSD chromatogram of HPCCC fraction F8 with annotated compounds using method MC_032B90_PE_25_5

In addition, a collective TLC-analysis of the 11 HPCCC fractions was performed (Figure 14). This analysis showed, that the fractionation by HPCCC was efficient in separating the constituents according to their polarity.

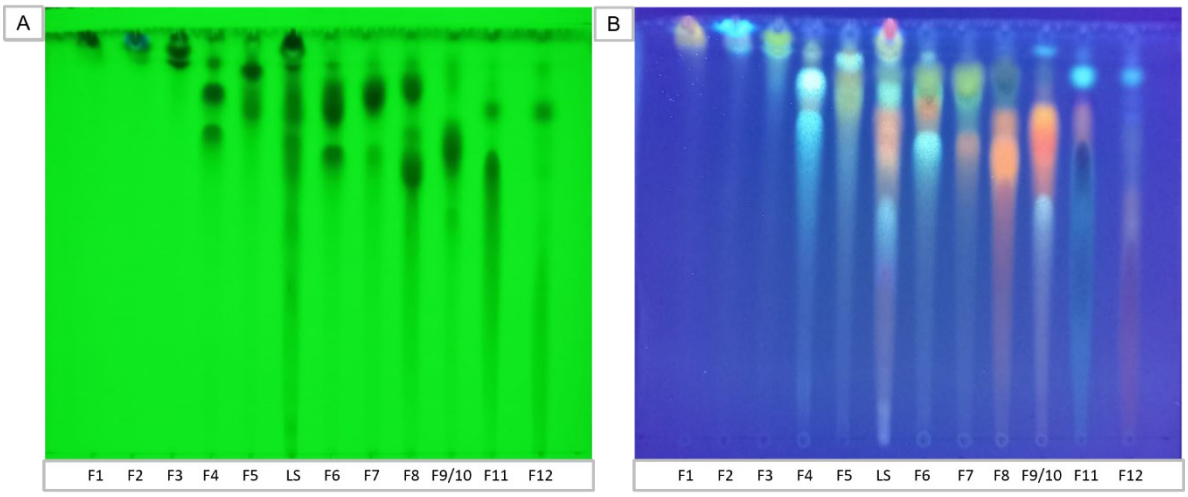


Figure 14: Collective TLC of generated HPCCC fractions (F1-F12) and chamomile flower crude extract (LS = MatchaBDM_LS 25 mg/mL); detected at (A) UV_{254}, (B) UV_{366} after spraying with nature substance reagent A/PEG400.

4.3.3 Isolation using size-exclusion chromatography

Size-exclusion chromatography (SEC) was subsequently used for the isolation of compounds 1-3 (see chapters 6.6.4 and 6.8.1.1).
In this chromatographic method, molecules are separated according to their size (= molecular weight). The pore volume of the packing medium is essential for the separation process. When a sample passes through the SEC column, larger molecules are not able to enter the pores and elute faster. Smaller molecules can penetrate the pores, remaining within the pores due to certain mechanisms and interactions and therefore elute later on. (Hunt & Holding, 2013)

4.3.3.1 Fraction 9

Since SEC had not yet been used in the study of the chamomile flower crude extract, fraction 9 (F9) served as an example for the separation process and was applied first onto a SEC column (length: 0.5 m; diameter: 2.0 cm). Figure 15 shows the annotated main compounds of F9.

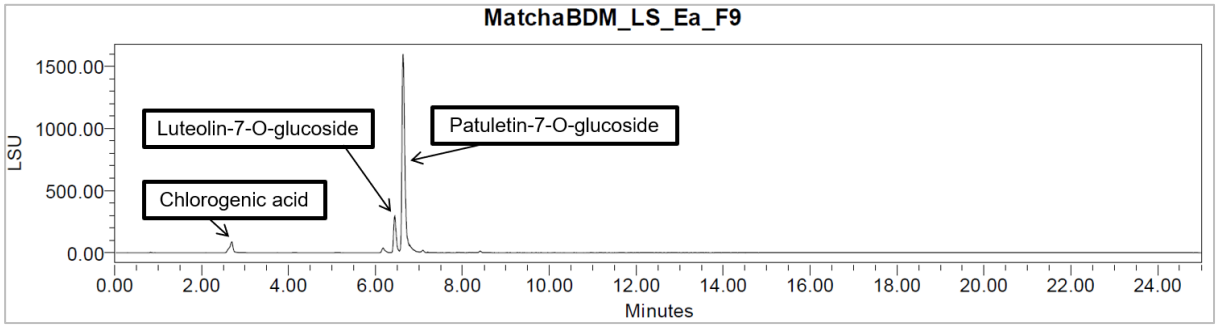


Figure 15: UPLC-ELSD chromatogram of HPCCC fraction F9 with annotated compounds using method MC_032B90_PE_25_5

After separation of F9, TLC of the collected fractions (SEC fractions) was performed. As shown in Figure 16, the two main flavonoids contained in F9, annotated as Luteolin-7-O-glucoside and Patuletin-7-O-glucoside (identified by UPLC-QDa measurements, QDa scans not shown; method described in chapter 6.6.2), were not sufficiently separated. The molecular weight difference between the two flavonoids is probably not significant enough for an effective separation by the 0.5 m column. Chlorogenic acid, the third major component of F9, was successfully separated from the other components.

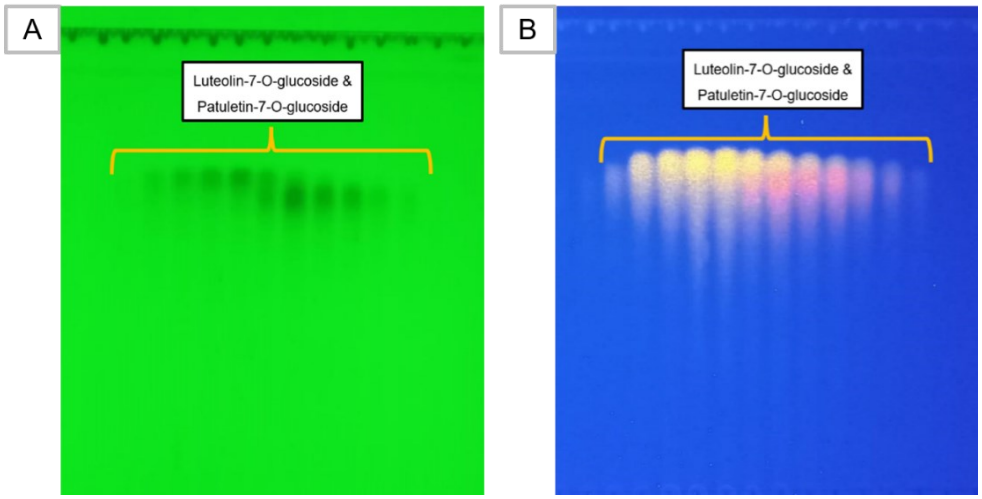


Figure 16: TLC of F9 SEC fractions containing not sufficiently separated flavonoids (Luteolin-7-O-glucoside and Patuletin-7-O-glucoside); detection at (A) UV_{254}, (B) UV_{366} after spraying with nature substance reagent A/PEG400

The SEC fractions containing the two flavonoids were combined and then once again separated by SEC, this time using a column with a length of 1.0 m and 2.0 cm in diameter. The second isolation attempt using the longer column showed better results, as can be seen in Figure 17. Therefore, it was decided to use the 1.0 m SEC column for the isolation attempts of compounds 1-3 in F7 and F8.

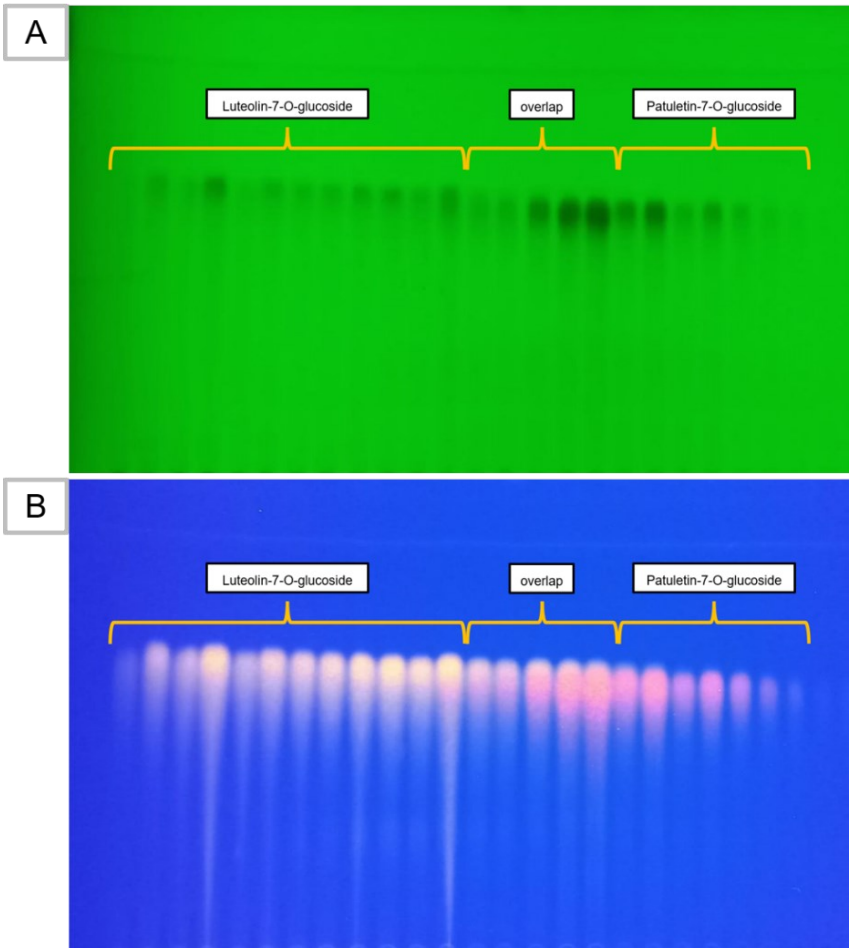


Figure 17: TLC of F9 SEC fractions containing partly separated flavonoids (Luteolin-7-O-glucoside and Patuletin-7-O-glucoside); detection at (A) UV_{254}, (B) UV_{366} after spraying with nature substance reagent A/PEG400.

Another conclusion that was drawn from the chromatographic separation of F9 was, that the components were not eluting in decreasing order of their molecular weight, as was initially expected. According to that assumption, Patuletin-7-O-glucoside (MW = 494.4 g/mol) should have eluted prior to Luteolin-7-O-glucoside (MW = 448.4 g/mol), which was not the case. Therefore, the elution sequence of the components cannot be predicted and has to be determined by TLC and additional UPLC measurements.

4.3.3.2 Fraction 8

F8 was applied next onto the SEC column (length: 1.0 m; diameter: 2.0 cm). As shown in Figure 13, this fraction mainly consists of five components, including the undescribed compounds 2 (m/z = 643) and 3 (m/z = 509). The chromatographic method was successful in isolating compound 2 (m/z = 643) based on the TLC analysis of the collected SEC fractions (Figure 20 A and B). In addition, Quercetin-?-O-glucoside could also be separated from the other components of F8 (Figure 19). In contrast, compound 3 (m/z = 509) could not be

adequately isolated. It co-eluted with two other components (annotated as Apigenin-7-O-glucoside and Chrysoeriol-?-O-glycoside; see Figure 18 A and B). Therefore, a different method was required for isolating compound 3 (see chapter 4.3.4).

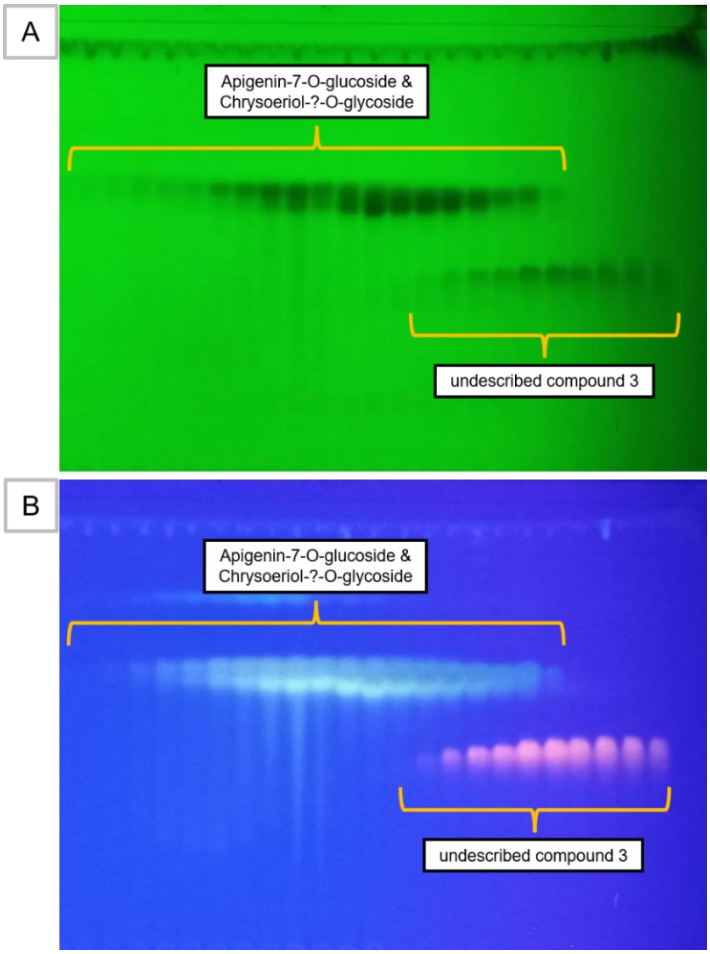


Figure 18: TLC of F8 SEC fractions containing Apigenin-7-O-glucoside, Chrysoeriol-?-O-glycoside and undescribed compound 3 (m/z = 509); detection at (A) UV_{254}, (B) UV_{366} after spraying with nature substance reagent A/PEG400.

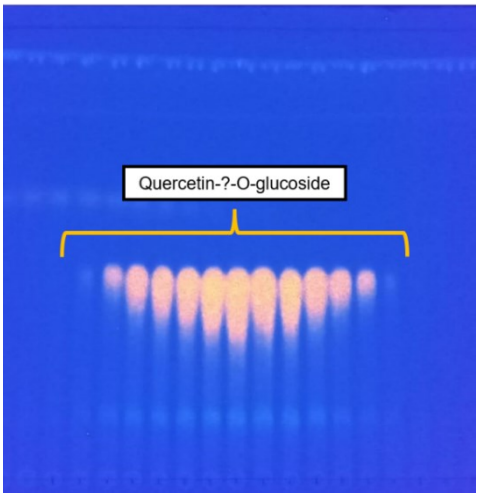


Figure 19: TLC of F8 SEC fractions containing isolated Quercetin-?-O-glucoside; detection at UV_{366} after spraying with nature substance reagent A/PEG400 (UV_{254} not documented).

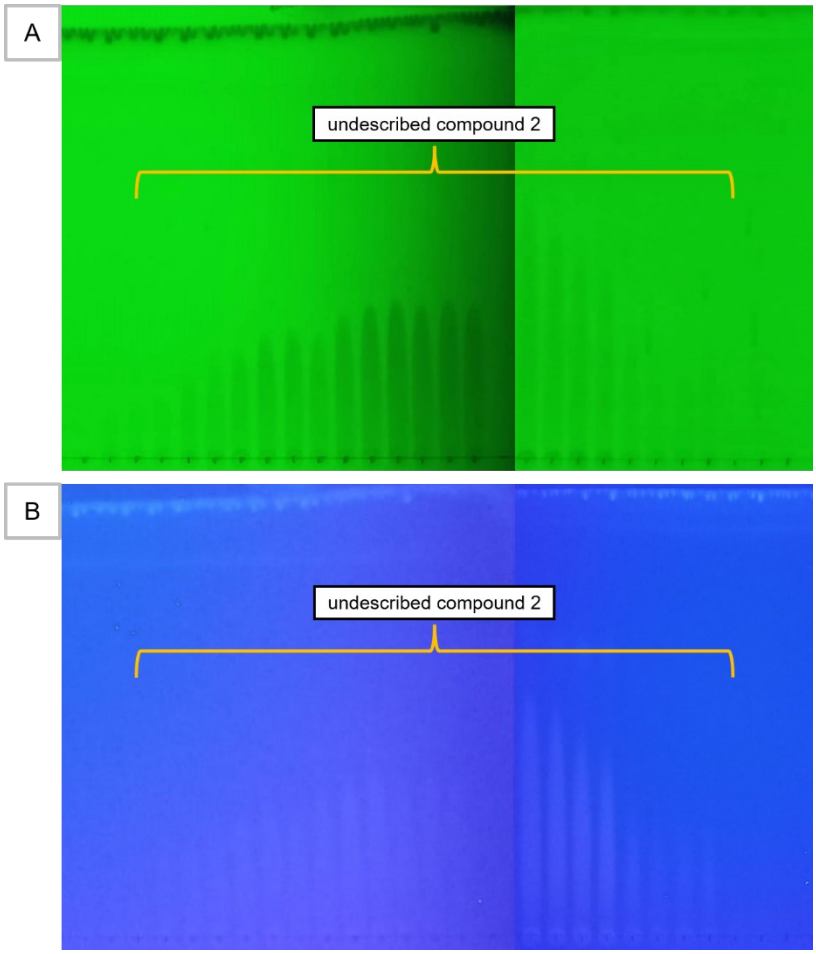


Figure 20:TLC of F8 SEC fractions containing isolated undescribed compound 2 (m/z = 643); detection at (A) UV_{254}, (B) UV_{366} after spraying with nature substance reagent A/PEG400.

4.3.3.3 Fraction 7

F7 containing the undescribed compound 1 (m/z = 657) was also fractionated by SEC. Due to the suspected structural similarity of compounds 1 and 2 (m/z = 643), it was assumed that compound 1 would be the component eluting last as well. TLC of the last collected colored SEC fractions – likely containing compound 1 – was performed. The suspected compound 1 was isolated from the other components in this fraction (Figure 21). The comparison of Figure 20 and Figure 21 shows that the suspected compound 1 has similar properties in TLC analysis as compound 2. The R_f value of both compounds is approximately the same. Both compounds are not strongly colored even after spraying with nature substance reagent A/PEG400, as is usually the case with other flavonoids. These results support the assumption that the isolated substance is compound 1.

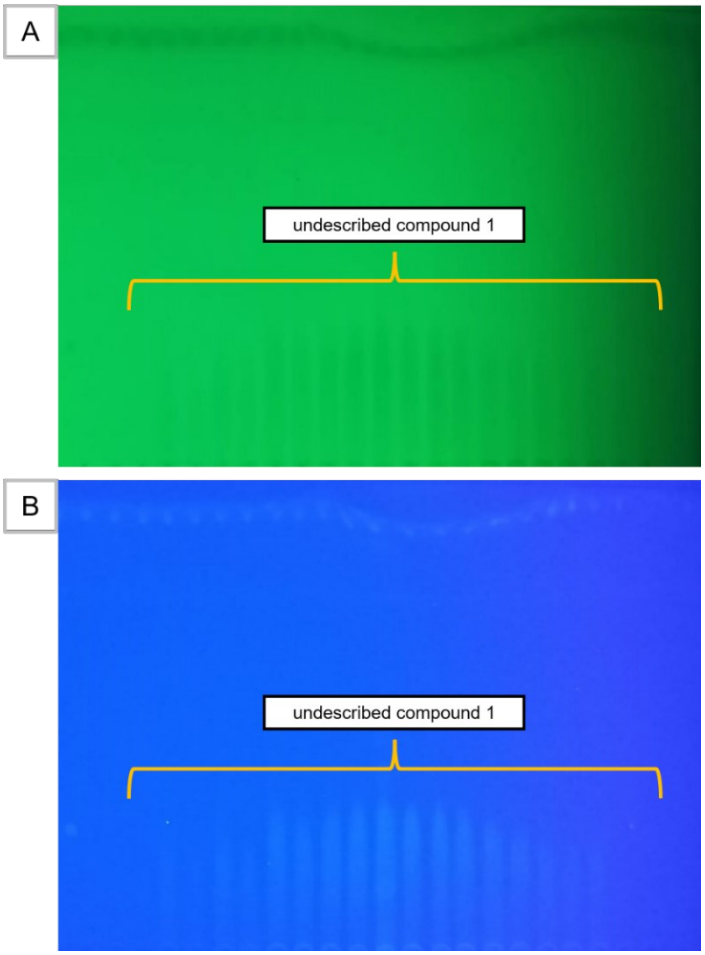


Figure 21: TLC of F7 SEC fractions containing isolated undescribed compound 1 (m/z = 657); detection at (A) UV_{254}, (B) UV_{366} after spraying with nature substance reagent A/PEG400.

4.3.4 UPLC isolation

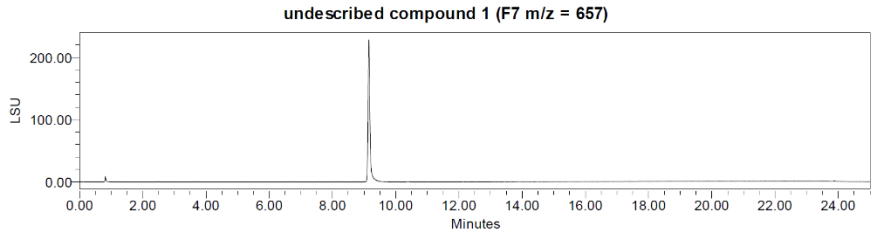
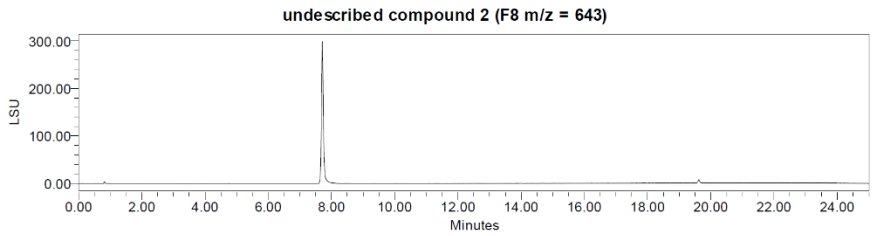
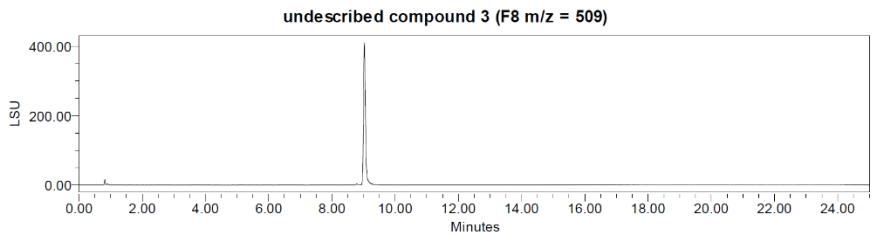
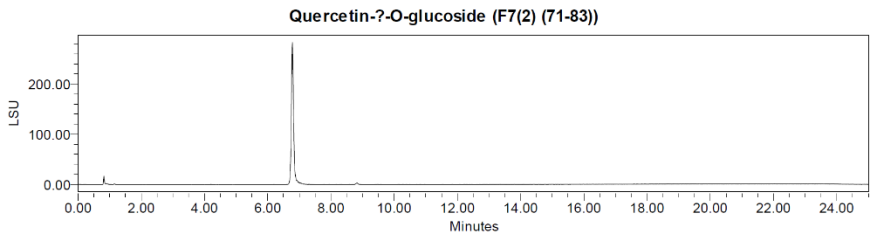
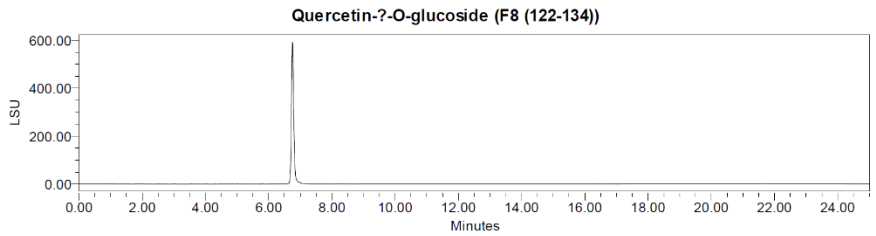
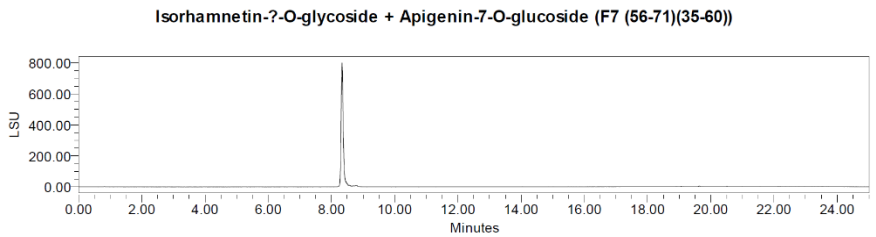
Since the isolation of compound 3 (m/z = 509) from F8 could not be accomplished by SEC, a UPLC method (see chapters 6.6.2 Table 14 and 6.8.1.2) was developed for the fractionation of pooled SEC fractions containing compound 3 as well as Apigenin-7-O-glucoside and Chrysoeriol-?-O-glycoside (in the following referred to as F8*). Waters Fraction Manager – Analytical (WFM-A) was connected to the UPLC for the collection of fractions. PDA chromatograms were recorded for setting, monitoring and optimization of the process.
A maximum of 10 µL of F8* could be injected per run. Therefore, a large number of injections had to be carried out in order to isolate appropriate amounts of compound 3 for subsequent NMR measurement.
Since the retention times of the peaks were constantly shifting, the collection time windows had to be continuously adjusted. Also, recurring error messages hindered the fractionation. Due to these problems, only compound 3 was collected at the end of the process. Other initially collected components were discarded, as they were low in yields anyway.

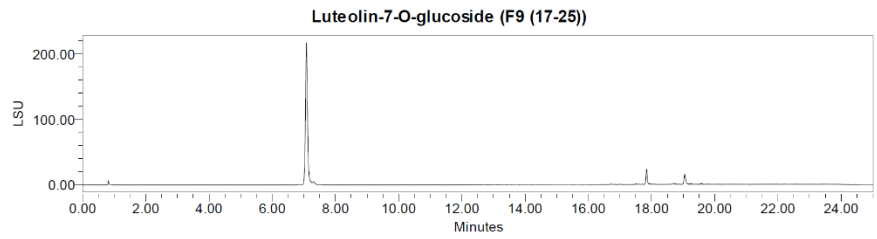
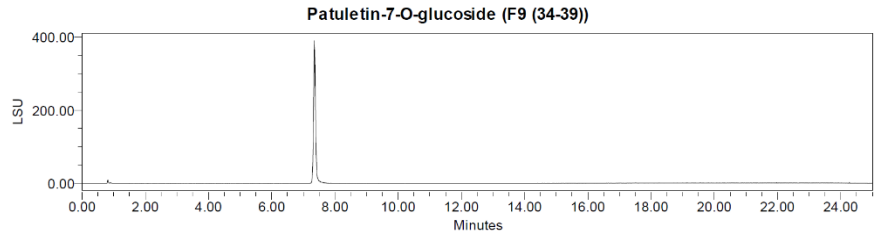
4.4 Purity determination of isolated compounds

UPLC-PDA-ELSD measurements were performed to determine the purity of the isolated substances (method see chapter 6.6.2). In Table 3, the ELSD chromatograms and the evaluated purity values in percent are given.

Out of the eight isolated compounds, six substances have a purity ≥ 95%, including compounds 1 (m/z = 657 $[M+H]^+$; purity: ~97%) and 3 (m/z = 509 $[M+H]^+$; purity: ~96%). Compound 2 (m/z = 643 $[M+H]^+$) has an estimated purity about 94%, which is still acceptable. Luteolin-7-O-glucoside only has a purity about 86%. The relatively low purity is probably due to traces of other components in F9 that could not be completely separated by SEC.

Table 3: ELSD chromatograms and evaluated purity values in percent of isolated substances

isolated substance	ELSD chromatogram	purity [%]
undescribed compound 1 (m/z = 657)		97.07
undescribed compound 2 (m/z = 643)		94.23
undescribed compound 3 (m/z = 509)		95.68
Quercetin-?-O-glucoside (F7)		94.70
Quercetin-?-O-glucoside (F8)		99.40
Isorhamnetin-?-O-glycoside + Apigenin-7-O-glucoside		97.73

Luteolin-7-O-glucoside		85.53
Patuletin-7-O-glucoside		98.11

In addition, TLC was performed to examine the purity of the isolated substances (Figure 22). For comparison, the chamomile flower crude extract and the HPCCC fractions F7, F8 and F9 were applied as well.

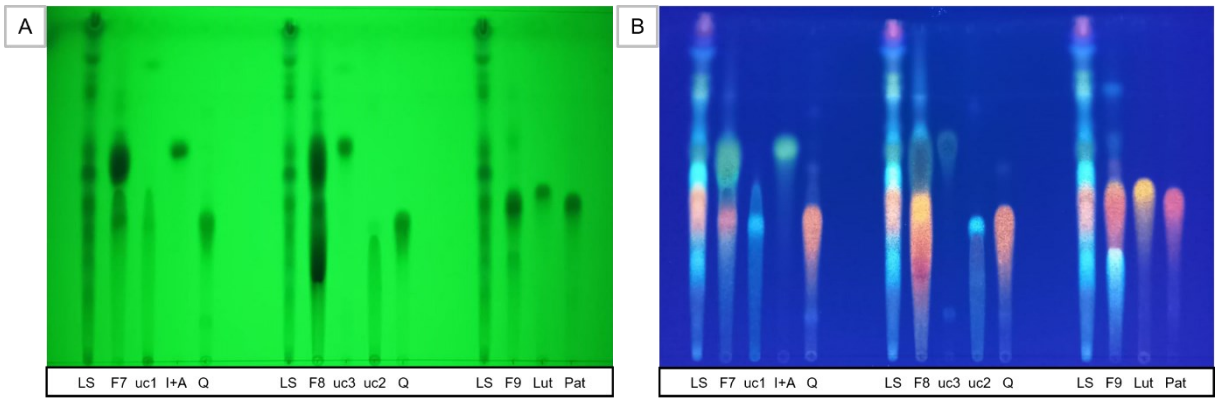


Figure 22: Collective TLC of chamomile flower crude extract (LS = MatchaBDM_LS 25 mg/mL), HPCCC fractions F7, F8 and F9 and isolated substances (uc1-3 = undescribed compounds 1-3; I+A = Isorhamnetin-?-O-glycoside + Apigenin-7-O-glucoside; Q = Quercetin-?-O-glucoside; Lut = Luteolin-7-O-glucoside; Pat = Patuletin-7-O-glucoside); detected at (A) UV_{254}, (B) UV_{366} after spraying with nature substance reagent A/PEG400.

Compound 3 (m/z = 509) as well as Patuletin-7-O-glucoside show relatively sharply defined single spots, which indicates high purity.

As mentioned before, compounds 1 (m/z = 657) and 2 (m/z = 643) have very similar properties in TLC. In both cases, not a single well-separated spot is visible, but sort of a “streak” is created during elution. Nevertheless, both substances are high in purity.

Quercetin-?-O-glucoside isolated from F8 is visibly more pure than Quercetin-?-O-glucoside from F7, as the later more prominently shows a second spot in the lower section. This confirms the evaluated purity values.

Luteolin-7-O-glucoside looks quite pure on TLC. As mentioned above, the evaluation of the ELSD chromatogram showed a relatively low purity of about 86% likely caused by impurities with other components of F9.

Table 4 summarizes the yields and purity values of the isolated undescribed compounds 1-3. Both purity and quantity of the three substances are sufficient for the performance of NMR measurements.

Table 4: Yields and purity of isolated undescribed compounds 1-3 from HPCCC fractions F7 and F8

isolated substance	**origin fraction**	**yield [mg]**	**purity [%]**
undescribed compound 1 (m/z = 657)	F7	2.75	97
undescribed compound 2 (m/z = 643)	F8	2.80	94
undescribed compound 3 (m/z = 509)	F8	0.65	96

5 CONCLUSION AND OUTLOOK

Starting point of this master thesis have been the 27 previously generated microfractions of a chamomile flower crude extract, which were here analyzed in detail by using molecular networking. The investigation of the largest formed cluster – i.e. molecular family 1 – not only revealed already known flavonoid glycosides as well as glycosylated coumarins of *M. chamomilla*, but also indicated a large number of nodes that could not be dereplicated. The principle of molecular networking made it possible to make well-supported assumptions about the putative annotations and structures of these undescribed constituents. Three of these nodes were selected for their targeted isolation, as they have been shown to be relatively high in quantity within the 27 chamomile microfractions. Analysis showed that these undescribed compounds are all glycosylated flavonoids. Compounds 1 (m/z = 657 $[M+H]^+$) and 2 (m/z = 643 $[M+H]^+$) are predicted to be caffeoylated Quercetagetin-7-O-glucoside derivatives that only differ in one methyl residue. Compound 3 (m/z = 509 $[M+H]^+$) is presumably Spinacetin-7-O-glucoside.

For the targeted isolation, different separation and fractionation techniques were applied on a chamomile flower crude extract. LLE was performed followed by HPCCC, SEC and UPLC fractionation, which finally led to the isolation of the targeted compounds. NMR structure elucidation of compounds 1-3 will be subject of further research in the Phytochemistry and Biodiscovery group of the Division of Pharmacognosy.

For the analysis of multi-component mixtures, it is advantageous to use modern, computer-aided methods such as molecular networking to examine the spectrum of constituents and to identify bioactive compounds. The results of this thesis support the use of molecular networking as a suitable dereplication tool for the identification of unknown compounds. Although German chamomile is a well-studied natural remedy, it was possible to identify undescribed components that may additionally contribute to the pharmacological effects.

6 MATERIALS AND METHODS

6.1 Chamomile flower crude extract and microfractions

The 27 chamomile microfractions that were analyzed by UPLC-QDa measurements and molecular networking were generated by Loidolt (2022). The chamomile flower crude extract (MatchaBDM_LS) used in the separation and fractionation process to isolate the undescribed compounds 1-3 was obtained from Loidolt (2022) as well.

6.2 Labeling of extracts and fractions

The label of the chamomile flower crude extract, “MatchaBDM_LS”, resulted from the plant name *Matricaria chamomilla* in addition to the abbreviation “LS” for large-scale extract, the initial of the plant organ – “B” for blossom – and the abbreviation “DM” for dichloromethane and methanol as extraction solvents (Loidolt, 2022). The abbreviation “Ea” for ethyl acetate was added to describe the ethyl acetate fraction obtained from the preparative liquid-liquid extraction, resulting in “MatchaBDM_LS_Ea”. The fractions gained from subsequent fractionation steps were named after the respective method, resulting in the labels “LLE fractions”, “HPCCC fractions”, “SEC fractions” and “UPLC fractions”.

6.3 Chamomile microfractions dereplication and annotation process

In Table 13 (see chapter 6.6.2), the method for UPLC-PDA-QDa analysis of the 27 chamomile microfractions (MFs) is given. RT and m/z values of supposed constituents were determined using the software Empower®. Total ion chromatograms (TIC) in positive ionization mode were analyzed. An arbitrary threshold for intensity of 7.0E6 was set. To identify the constituents within the 27 MFs, the documented RT and m/z values were compared to those of the chamomile flower crude extract (MatchaBDM_LS), which were evaluated by (Loidolt, 2022). Also, literature research was performed to assign peaks to previously published constituents of *M. chamomilla* based on m/z and molecular weight values. Therefore, SciFinder® (scifinder.cas.org/) was used to gather required data. The occurrence of components within different MFs and their concentration profile were determined by comparing RT values and the intensity of respective peaks in consecutive MFs.

6.4 Molecular network generation

6.4.1 MS/MS data generation

MS/MS (= MS2) data of the 27 MFs was generated by Prof. Mehdi Beniddir (Université Paris-Scalay, France).

UPLC-ESI-HRMS2 analyses were achieved by coupling the UPLC system to a hybrid quadrupole time of flight mass spectrometer Agilent 6546 (Agilent Technologies, Massy, France) equipped with an ESI source, operating in both positive and negative ion mode. A BEH Waters ACQUITY C18 UPLC column (2.1 × 150 mm; i.d. 1.8 µm, Agilent) was used, with a flow rate of 0.5 $mL \cdot min^{-1}$ and a linear gradient from 5% B (A: H_2O + 0.1% formic acid, B: Acetonitrile + 0.1% formic acid) to 100% B over 12 minutes. Source parameters were set as followed: capillary temperature at 320°C, source voltage at 3500 V, sheath gas flow rate at 11 $L \cdot min^{-1}$. The divert valve was set to waste for the first 3 minutes. MS scans were operated in full-scan mode from m/z 100 to 1200 (0.1 s scan time) with a mass resolution of 67.000 at m/z = 922. A MS1 scan was followed by MS2 scans of the four most intense ions above an absolute threshold of 3000 counts. Selected parent ions were fragmented at a collision energy fixed at 45 eV and an isolation window of 1.3 amu. In the positive ion mode, purine $C_5H_4N_4$ $[M+H]^+$ ion (m/z = 121.050873) and the hexakis (1H,1H,3H-tetrafluoropropoxy)-phosphazene $C_{18}H_{18}F_{24}N_3O_6P_3$ $[M+H]^+$ ion (m/z = 922.009798) were used as internal lock masses. In the negative ion mode, trifluoroacetic acid (CF_3CO_2H, m/z = 112.98559) and the trifluoroacetate adduct with m/z = 1033.988109 were used. A permanent MS/MS exclusion list criterion was set to prevent oversampling of the internal calibrant. LC-UV and MS data acquisition and processing were performed using MassHunter® Workstation software (Agilent Technologies, Massy, France).

Using the software MSConvert (Chambers et al., 2012), the generated MS/MS data files were converted from the ".d" to the ".mzXML" format.

6.4.2 MZmine – data-processing

MS/MS ".mzXML" data files were processed using the software MZmine 2 (Pluskal et al., 2010). In Table 5, the settings for each MS/MS data-processing step are given in chronological order.

Table 5: Settings for MS/MS (= MS2) data-processing using MZmine 2

mass detection	
parameter	**value**
MS1 noise level	1.0E4
MS2 noise level	1.0E1
otherwise default settings	
ADAP chromatogram builder (Myers et al., 2017) (MS1 level selected)	
parameter	**value**
minimum group size of scans	4
group intensity threshold	1.0E4
minimum highest intensity	1.0E4
m/z tolerance	10 ppm
chromatogram deconvolution (Myers et al., 2017)	
a) ADAP wavelets algorithm	
parameter	**value**
S/N threshold	23
minimum feature height	1.0E4
coefficient/area threshold	10
Peak duration length	0.1-0.5 min
RT wavelet range	0.0-0.08
b) m/z center calculation	
AUTO	
c) m/z tolerance range for MS2 scan pairing	
0.02 Da	
d) RT tolerance range for MS2 scan pairing	
0.15 min	

Isotopic peaks grouper algorithm	
parameter	**value**
m/z tolerance	10 ppm
RT tolerance	0.3 min
maximum charge	1
representative isotope	most intense
Peak alignment	
parameter	**value**
m/z tolerance	10 ppm
weight for m/z	50
weight for RT	50
absolute RT tolerance	0.12 min
Feature list rows filter	
parameter	**value**
m/z	100.00-1200.00
RT	0.04-18.0 min

Finally, ".mgf" files (containing preclustered spectral data) and ".csv" metadata files (containing quantitative information) were generated.
Using Excel, a metadata (sample information) table was generated to link the MS files of the measured samples to the corresponding microfractions (MF1-27).

6.4.3 GNPS – molecular network generation

The ".mgf" preclustered data file, the ".csv" metadata file as well as the metadata table were exported to GNPS (Wang et al., 2016). With the provided online workflow, the molecular network was generated using the settings detailed in Table 6.

Table 6: Settings for molecular network generation using GNPS

basic options	
parameter	**value**
quantification table source	MZmine
precursor ion mass tolerance	0.02 Da
MS2 fragment ion mass tolerance	0.02 Da
advanced network options	
parameter	**value**
minimum pairs cosine score	0.7
network TopK	10
minimum matched fragment ions	6
maximum connected component size	100
maximum shift between precursors	500
advanced library search options	
parameter	**value**
library search minimum matched peaks	6
search analogs	do search
top results to report per query	3
score threshold	0.7
maximum analog search mass difference	200 Da
advanced filtering options	
parameter	**value**
for all filters	don’t filter
advanced quantification options	
parameter	**value**
normalization per file	row sum normalization
aggregation method for peak abundances per group	sum

advanced multivariate statistics options	
parameter	**value**
PCoA distance metric	cosine
advanced univariate statistics options	
parameter	**value**
not selected	-
advanced external tools	
parameter	**value**
dereplicator	don't run

6.4.4 Cytoscape – network visualization and analysis

The generated MN was then visualized and analyzed using Cytoscape (ver. v3.9.1) (Shannon et al., 2003). Various style settings can be used to display different properties of the clustered compounds. In this network, substance properties were distinguished by the size, shape, and color scheme of the nodes as well as the thickness of the edges (described in Table 7). The m/z differences of connected compounds are also presented.

Table 7: Molecular network visualization settings

setting	**property**
node size	relative quantity of a compound (based on ionization)
node shape	square: spectral library match □ circle: no spectral library match ○
color palette	MF that contains a compound → 27 different colors representing one MF each
ring chart	proportion of a compound in consecutive MFs
edge thickness	similarity of two connected compounds (cosine-score)

The components of the largest molecular family (molecular family 1) were dereplicated and annotated based on their molecular weights, m/z differences, GNPS spectral library matches

and analogues, as well as literature research and chamomile microfractions dereplication and annotation results.

6.5 Liquid-liquid extraction

6.5.1 Analytical extraction

Approximately 100 mg chamomile flower crude extract (MatchaBDM_LS) was dissolved in 20 mL of ddH_2O and 40 mL of hexane in an ultrasonic bath.
LLE was performed using a separatory funnel. After separation of the two phases the aqueous phase was again blended with 40 mL of fresh hexane. In total, this separation process was performed three times and the three collected hexane fractions were pooled. The aqueous phase was then extracted three times with 40 mL of dichloromethane, three times with 40 mL of ethyl acetate and then three more times with 40 mL of water-saturated butanol. The three collected fractions of dichloromethane, ethyl acetate and butanol were pooled, respectively. The aqueous fraction was evaporated to dryness on a rotary evaporator, while the organic fractions were evaporated using the GeneVac® evaporation device.
In Table 8, the respective yields of the five LLE fractions gained are shown.

Table 8: Yields of LLE fractions (1)

fraction	tare [g]	weight [g]	yield [mg]
hexane	97.9869	97.9960	9.10
dichloromethane	95.6330	95.6426	9.60
ethyl acetate	99.8157	99.8363	20.60
butanol	101.0023	101.0512	48.90
ddH_2O	58.7637	58.8131	49.40

To optimize the outcome, a second analytical extraction attempt was carried out. 100 mg crude extract were extracted with 40 mL ddH_2O and five times with 80 mL hexane. The aqueous phase was then extracted three times with 40 mL of ethyl acetate.
Table 9 shows the yields of the three LLE fractions.

Table 9: Yields of LLE fractions (2)

fraction	tare [g]	weight [g]	yield [mg]
hexane	11.96569	11.97420	8.60
ethyl acetate	17.60177	17.62060	18.83
ddH_2O	47.24140	47.30170	60.30

For UPLC measurements, samples of the LLE fractions were prepared containing a concentration of 5 mg/mL. The chamomile flower crude extract with a concentration of 25 mg/mL was also measured. The method used was MC_032B90_PE_25_5, which is described in chapter 6.6.2.

6.5.2 Preparative extraction

For the preparative separation, about 5 g of chamomile flower crude extract was dissolved in 700 mL ddH_2O and 700 mL hexane in an ultrasonic bath.
A separatory funnel was used for LLE. Extraction of the aqueous phase was repeated, until the color of the respective organic phase remained colorless. In total, the extraction with hexane was carried out ten times, the subsequent extraction with 700 mL ethyl acetate eight times. The collected hexane and ethyl acetate fractions were pooled, respectively. During the extraction process, the hexane and ethyl acetate were each recovered on a rotary evaporator and reused in subsequent extraction steps. The hexane, ethyl acetate and aqueous fractions were evaporated to dryness.
In Table 10, the yields of the LLE fractions obtained are presented.

Table 10: Yields of preparative LLE fractions

fraction	tare [g]	weight [g]	yield [mg]
hexane	10.22976	11.00395	774.19
ethyl acetate	9.53831	10.47473	936.42
ddH_2O	48.11650	50.67230	2557.80

For UPLC-ELSD measurement of the hexane and ethyl acetate fractions, 1 mg each of the dry extract was dissolved in 200 µL methanol (5 mg/mL). The method used was MC_032B90_PE_25_5, which is described in chapter 6.6.2.

In addition, TLC was performed as described in 6.6.1 using the hexane fraction, the ethyl acetate fraction, the aqueous fraction, the crude extract (MatchaBDM_LS 10 mg/mL) as well as 6 µL of Apigenin-7-O-glucoside (1 mg/mL) as a reference substance for comparison.

6.6 Chromatographic methods

6.6.1 TLC system

The TLC system for *M. chamomilla* was developed by Loidolt (2022). For the stationary phase, silica gel 60 F_{254} plates were used. The composition of the mobile phase as well as the TLC system parameters are given in Table 11. Chamber saturation was awaited before TLC was performed. After completion of the chromatographic process, the TLC was analyzed under ultraviolet light (UV_{254}). For the detection of flavonoids and other secondary metabolites, the TLC was derivatized with the spraying reagent natural product reagent A (1%) and subsequently with polyethylene glycol 400 solution (PEG400). After spraying, UV_{366} was used for the visualization of the separated constituents of the applied samples.

Table 11: TLC system parameters

stationary phase	TLC silica gel 60 F_{254}
mobile Phase	$CHCl_3$: MeOH : ddH_2O : FA = 67 : 30 : 2 : 1
spraying reagents	natural substance reagent A 1% in methanol + polyethylene glycol 400 in 96% ethanol
detection	UV_{254}, UV_{366}

6.6.2 UPLC methods

The UPLC methods used for the analyses of the chamomile microfractions as well as the generated fractions and isolates were developed by Loidolt (2022). In Table 12 and Table 13, the specific parameters are shown.

Table 12: UPLC-PDA-ELSD method for (i) identifying and analyzing compounds within fractions, (ii) estimating the quantity of compounds within fractions, (iii) determining the purity of isolated substances

instrument	ACQUITY UPLC H-Class System by Waters		
column	ACQUITY BEH C18 (2.1 x 100 mm, 1.7 µm)		
flow rate [mL/min]	0.3		
temperature [°C]	40		
detectors	PDA (205, 220, 254, 360 nm) ELSD		
mobile phase A	H_2O + 0.1% FA		
mobile phase B	ACN + 0.1% FA		
gradient	time [min]	A [%]	B [%]
	0	90	10
	9	72	28
	13	40	60
	16	2	98
	22	2	98
	22.1	90	10
	25	90	10

Table 13: UPLC-PDA-QDa method for identifying and analyzing compounds within (micro-)fractions

instrument	ACQUITY UPLC H-Class System by Waters		
column	ACQUITY BEH C18 (2.1 x 100 mm, 1.7 µm)		
flow rate [mL/min]	0.3		
temperature [°C]	40		
PDA [nm]	205, 220 254, 360		
ISM	H_2O/MeOH 9:1 + 0.1% FA		
QDa	cone voltage: pos = 15V / neg = 30V		
mass range [Da]	100-900		
mobile phase A	H_2O + 0.1% FA		
mobile phase B	ACN + 0.1% FA		
gradient	time [min]	A [%]	B [%]
	0	90	10
	9	72	28
	13	40	60
	16	2	98
	22	2	98
	22.1	90	10
	25	90	10

UPLC-settings for the isolation of undescribed compound 3 (m/z = 509) from HPCCC fraction 8 (F8) are given in Table 14.

Table 14: UPLC-PDA-WFM-A method for isolating undescribed compound 3 (m/z = 509) from F8

instrument	ACQUITY UPLC H-Class System by Waters				
column	ACQUITY BEH C18 (2.1 x 100 mm, 1.7 µm)				
flow rate [mL/min]	0.3				
temperature [°C]	40				
PDA [nm]	205, 220, 254, 360				
WFM-A	collection flow rate [mL/min]	default vessel fill [%]	default waste vessel fill [%]	needle dispense position	valve state between vessels
	0.50	100	100	above vessel	waste
mobile phase A	H_2O				
mobile phase B	ACN				
gradient	time [min]	A [%]		B [%]	
	0	82		18	
	6	78.5		21.5	
	6.50	2		98	
	8.60	82		18	
	11.60	82		18	

6.6.3 HPCCC system

6.6.3.1 Solvent system and sample preparation

For an efficient HPCCC separation of the ethyl acetate fraction, HEMWat solvent systems were used. 28 HEMWat solvent systems are described that consist of n-hexane, ethyl acetate, methanol, water and in some cases also butanol. The systems differ in the volume ratio of the solvents mentioned. The polarity decreases in ascending order, meaning that lower system numbers are higher in polarity.

In the course of a previous master thesis on *M. chamomilla*, a HEMWat system for the fractionation of the chamomile flower crude extract (MatchaBDM_LS) was developed (Loidolt, 2022). This system was applied for the fractionation of the ethyl acetate LLE fraction of the crude extract (MatchaBDM_LS_Ea).
HPCCC was performed on a Dynamic Extractions Ltd HPCCC instrument using normal phase mode and gradient elution. The lower layer (LL) of HEMWat system 13 was selected as the stationary phase. The upper layers (UL) of HEMWat systems 13, 11, 9 and 7 were applied successively as mobile phase. Table 15 shows the compositions of the used HEMWat systems, which were prepared in advance of the HPCCC runs.

Table 15: Composition of used HEMWat systems in percent; LL = lower layer, UL = upper layer

phase	n-hexane [%]	EtOAc [%]	MeOH [%]	ddH_2O [%]
LL 13	0	10.54	25.33	64.23
UL 13	30.69	65.40	3.14	0.77
UL 11	20.76	75.77	2.40	1.07
UL 9	15.08	81.93	1.81	1.18
UL 7	5.11	92.05	0.76	2.08

For sample preparation, approximately 400 mg of the ethyl acetate fraction were weighed into a centrifuge tube and dissolved in 5 mL of the starting UL 13 and 5 mL of the LL 13. The sample was sonicated in an ultrasonic bath until complete dissolution and then centrifuged (3500 rpm for 10 min).

6.6.3.2 Semi-preparative HPCCC fractionation

The HPCCC was hyphenated to the Interchim PuriFlash® 4250, which enables the use of the PuriFlash®'s fraction collector and its integrated detectors. In this case, the ELSD detector was used for monitoring the process (chromatograms see Appendix).
To prepare the HPCCC for the chromatographic process, the semi-preparative column was filled with stationary phase (LL 13) with a flow rate of 10 mL/min and rotation speed of 200 rpm. Then, the first mobile phase (UL 13) was pumped into the column at a rotation speed of 1600 rpm and a flow rate of 6 mL/min. When the hydrodynamic equilibrium between lower and upper layer occurred, the sample was injected into the HPCCC system. At this point, fraction collection started (6 mL per tube). 270 mL of the four ULs were added to the system in descending order (UL 13 – 11 – 9 – 7) using gradient elution. After the consumption of 210 mL of an UL, 60 mL of the next UL was added. When 60 mL of this mixture was consumed, the remaining volume of the next UL was added.
At the end of the fractionation process, elution extrusion was performed for 30 minutes. Therefore, MeOH/ddH_2O (1:1) was pumped through the system, while the rotation level was decreased to 200 rpm and the flow rate was increased to 10 mL/min. Fractions were collected until no more eluting compounds were detected by ELSD.

The fractionation process was performed twice using the same settings (see Table 16 and Table 17). In the first run, 185 fractions were collected by the Interchim, 175 fractions in the second run. The fractions were then evaporated to dryness using the GeneVac® evaporation device.

Table 16: Parameters for semi-preparative HPCCC fractionation of the ethyl acetate LLE fraction (MatchaBDM_LS_Ea)

instruments	Dynamic Extractions Spectrum HPCCC Interchim PuriFlash® 4250
mode	normal phase / preparative
total column volume [mL]	136
flow rate [mL/min]	6 / 10
rotation speed [rpm]	1600
maximum pressure [bar]	13
collected volume per fraction [mL]	6
detector	ELSD

Table 17: HEMWat solvent system for HPCCC fractionation of the ethyl acetate LLE fraction (MatchaBDM_LS_Ea) with gradient elution; LL = lower layer, UL = upper layer; *before hydrodynamic equilibrium

phase	used layer	solvent ratio				volume [mL]
		n-hexane	EtOAc	MeOH	ddH_2O	
HEMWat 13	LL UL	2	5	2	5	300 180* + 270
HEMWat 11	UL	1	4	1	4	270
HEMWat 9	UL	1	6	1	6	270
HEMWat 7	UL	1	19	1	19	270
elution extrusion	-	0	0	1	1	300

6.6.4 Size-exclusion chromatography

For SEC, the samples (HPCCC fractions F7, F8 and F9) were dissolved in 2 mL methanol. (As F7 required 5 mL MeOH to be solved, this fraction was divided into two samples and applied twice onto the SEC column.) The respective fraction was applied onto the SEC column, packed with Sephadex® LH-20.
As flavonoids tend to be yellow in color, collection of fractions – using Spectra/Chrom® CF-2 Fraction Collector – was started when yellow colored solvent approached the end of the SEC column. The separation process was monitored by TLC. The chromatography was stopped when no more eluents were detected by TLC.

6.7 Pooling of HPCCC fractions

For pooling of the collected HPCCC fractions, TLC was performed.

Based on the TLC results (see Appendix), the fractions of both HPCCC runs were pooled according to their apparent constituents (see Table 18). 12 preliminary HPCCC factions were generated due to this pooling. UPLC-ELSD and UPLC-QDa measurements showed that fractions 9 and 10 (F9 and F10) are similar to a large extent. Therefore, F9 and F10 were pooled to one fraction, meaning that 11 final HPCCC fractions (F1-F11) resulted.

Table 18: Combination of collected fractions of two semi-preparative HPCCC runs and yields of generated HPCCC fractions; *fractions F9 and F10 are considered to be one fraction due to their constituents

HPCCC fraction	collected tubes of 1st run	collected tubes of 2nd run	yield [mg]
F1	1-4	1-4	49.80
F2	5-13	5-13	26.47
F3	14-34	14-39	52.45
F4	35-46	40-50	61.57
F5	47-55	51-60	18.91
F6	56-74	61-78	45.37
F7	75-88	79-90	28.43
F8	89-103	91-100	23.98
***F9**	104-122	101-117	20.11
***F10**	123-140	118-134	25.96
F11	141-165	135-154	38.82
F12	166-185	155-175	75.27

6.8 Isolation of pure substances

6.8.1 Isolation methods

6.8.1.1 SEC

The method of isolating substances from HPCCC fractions F7, F8 and F9 using SEC is described in chapter 6.6.4.

SEC fractions containing relevant components were analyzed by TLC to determine the separation pattern and to assess if compounds might already be isolated. SEC fractions containing isolated undescribed compounds 1 and 2 were transferred into vials and placed under the sample concentrator for solvent evaporation.

6.8.1.2 UPLC

The UPLC method (MC_031B18_PF_11) for the isolation of undescribed compound 3 is given in chapter 6.6.2 Table 14.

For the fractionation of F8*, UPLC-PDA connected to Waters Fraction Manager – Analytical (WFM-A) was used and 10 µL of sample were injected per run. Time-based peak collection was performed, where collection windows are configurated based on the recorded PDA chromatogram of the sample.
The collected fractions (UPLC fractions) containing compound 3 were pooled, transferred into a vial and then placed under the sample concentrator.

6.8.2 Yields of isolated compounds

In Table 19, the yields of the substances isolated from HPCCC fractions F7, F8 and F9 are given.

Table 19: Yields of isolated substances from HPCCC fractions F7, F8, F9

isolated substance	**origin fraction**	**yield [mg]**
undescribed compound 1 (m/z = 657)	F7	2.75
undescribed compound 2 (m/z = 643)	F8	2.80
undescribed compound 3 (m/z = 509)	F8	0.65
Quercetin-?-O-glucoside	F8	5.93
Quercetin-?-O-glucoside	F7	0.70
Isorhamnetin-?-O-glycoside + Apigenin-7-O-glucoside	F7	10.45
Luteolin-7-O-glucoside	F9	1.81
Patuletin-7-O-glucoside	F9	1.99

6.8.3 Purity measurements

For determining the purity of the isolated substances, UPLC-ELSD measurements were performed using method MC_032B90_PE_25_5 (see chapter 6.6.2 Table 12). Purity of the compounds was determined by integration of the peak areas of the ELSD chromatograms.

6.9 Instruments, solvents and reagents

6.9.1 Instruments and lab ware

device	description
analytical balance	Sartorius BP210D
balance	Sartorius LC4801P
centrifugal evaporator	GeneVac® EZ-2 plus
centrifuge	Laborfuge 400 Function Line, Heraeus
concentrator	Sample concentrator FSC400D, Techne
flash chromatography	PuriFlash® 4250 by Interchim
fraction collector	Spectra/Chrom® CF-2 Fraction Collector
heat gun	Steinel HG 2000 E
HPCCC	Dynamic Extractions Spectrum
HPCCC chiller	Accel 500 LC, Thermo Scientific
pipettes	Eppendorf Research (10 µL, 100 µL, 1000 µL)
pipette tips	Eppendorf; Sarstedt; StarLab
rotary evaporator	R-210 and Heating Bath B-491, Büchi laboratory technology AG
SEC column	OMNI glass columns (50 cm x 2 cm; 100 cm x 2 cm)
SEC packing material	Sephadex® LH-20
TLC capillaries	Micropipettes 1/2/3/4/5* µL, BLAUBRAND®intraMARK, BRAND®
TLC plates	TLC Silica gel 60 F_{254}, Aluminium sheets 20x20 cm, Merck
ultrasonic bath	Transsonic T 460, Elma
UPLC	ACQUITY UPLC H-Class System by Waters modules: SM, QSM, PDA, ELSD, ISM, QDa, WFM-A
UPLC column	ACQUITY BEH C18 (2.1 x 100 mm, 1.7 µm)
UV-lamp	CAMAG®
vacuum pump	V-710, Büchi laboratory technology AG
vortex shaker	Genius 3, IKA

6.9.2 Solvents and reagents

solvent/reagent	description
Acetonitrile	HiPerSolv Chromanorm ≥ 99.9%, gradient grade for HPLC
Butanol	AnalaR Normapur, VWR Chemicals®, 99.9%
Chloroform	AnalaR Normapur, stabilized with about 0.6% ethanol, VWR Chemicals®, BDH®
Dichloromethane	Rectapur, distilled according to ÖAB
Ethyl acetate	Rectapur, distilled according to ÖAB
Formic acid 99%	≥ 98%, p.a., ACS, Carl Roth GmbH
Methanol	HiPerSolv Chromanorm ≥ 99.9%, for HPLC Rectapur (distilled according to ÖAB)
n-Hexane	Rectapur, distilled according to ÖAB
Nature substance reagent A	≥98%, p. a., Carl Roth GmbH
PEG400	Rotipuran® Ph. Eu., Carl Roth GmbH + Co KG

7 LIST OF ABBREVIATIONS

ACN	Acetonitrile
B	Butanol
$CHCl_3$	Chloroform
DCM	Dichloromethane
ddH_2O	Double-distilled water
ELSD	Evaporative light scattering detector
EtOAc, Ea	Ethyl acetate
F	Fraction
FA	Formic acid
H	n-Hexane
HEMWat	n-Hexane, Ethyl acetate, Methanol, Water
HPCCC	High-performance counter current chromatography
LL	Lower layer
LLE	Liquid-liquid extraction
LS	Large-scale
MeOH	Methanol
MF(s)	Microfraction(s)
MN	Molecular network
MS	Mass spectrometry
MS/MS = MS2	Tandem mass spectrometry
m/z	Mass-to-charge ratio
NMR	Nuclear magnetic resonance
NP(s)	Natural product(s)
PDA	Photo diode array
PEG400	Polyethylene glycol 400
QDa	Quadrupole Dalton mass detector
rpm	Revolutions per minute
RT	Retention time
SEC	Size-exclusion chromatography
TIC(s)	Total ion chromatogram(s)
TLC	Thin layer chromatography
UL	Upper layer
UPLC	Ultra performance liquid chromatography
UV	Ultraviolet
WFM-A	Waters fraction manager – analytical

8 REFERENCES

Aron, A. T., Gentry, E. C., McPhail, K. L., Nothias, L.-F., Nothias-Esposito, M., Bouslimani, A., Petras, D., Gauglitz, J. M., Sikora, N., Vargas, F., van der Hooft, J. J. J., Ernst, M., Kang, K. B., Aceves, C. M., Caraballo-Rodríguez, A. M., Koester, I., Weldon, K. C., Bertrand, S., Roullier, C., Sun, K., Tehan, R. M., Boya P., C. A., Christian, M. H., Gutiérrez, M., Ulloa, A. M., Tejeda Mora, J. A., Mojica-Flores, R., Lakey-Beitia, J., Vásquez-Chaves, V., Zhang, Y., Calderón, A. I., Tayler, N., Keyzers, R. A., Tugizimana, F., Ndlovu, N., Aksenov, A. A., Jarmusch, A. K., Schmid, R., Truman, A. W., Bandeira, N., Wang, M., & Dorrestein, P. C. (2020). Reproducible molecular networking of untargeted mass spectrometry data using GNPS. *Nature Protocols*, *15*(6), 1954–1991. https://doi.org/10.1038/s41596-020-0317-5

Avula, B., Wang, Y.-H., Isaac, G., Yuk, J., Wrona, M., Yu, K., & Khan, I. A. (2017). Metabolic profiling of Hoodia, chamomile, Terminalia Species and evaluation of commercial preparations using ultrahigh-performance liquid chromatography quadrupole-time-of-flight mass spectrometry. *Planta Medica*, *83*(16), 1297–1308. https://doi.org/10.1055/s-0043-109239

Avula, B., Wang, Y.-H., Wang, M., Avonto, C., Zhao, J., Smillie, T. J., Rua, D., & Khan, I. A. (2014). Quantitative determination of phenolic compounds by UHPLC-UV–MS and use of partial least-square discriminant analysis to differentiate chemo-types of chamomile/Chrysanthemum flower heads. *Journal of Pharmaceutical and Biomedical Analysis*, *88*, 278–288. https://doi.org/10.1016/j.jpba.2013.08.037

Chambers, M. C., Maclean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T. A., Brusniak, M.-Y., Paulse, C., Creasy, D., Flashner, L., Kani, K., Moulding, C., Seymour, S. L., Nuwaysir, L. M., Lefebvre, B., Kuhlmann, F., Roark, J., Rainer, P., Suckau, D., Hemenway, T., Huhmer, A., Langridge, J., Connolly, B., Chadick, T., Holly, K., Eckels, J., Deutsch, E. W., Moritz, R. L., Katz, J. E., Agus, D. B., MacCoss, M., Tabb, D. L., & Mallick, P. (2012). A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology*, *30*(10), 918–920. https://doi.org/10.1038/nbt.2377

El Joumaa, M. M., & Borjac, J. M. (2022). Matricaria chamomilla: A valuable insight into recent advances in medicinal uses and pharmacological activities. *Phytochemistry Reviews*. https://doi.org/10.1007/s11101-022-09817-0

EMA/HMPC. (2015). *Assessment report on Matricaria recutita L., flos and Matricaria recutita L., aetheroleum* (No. 55837/2011). European Medicines Agency (EMA). https://www.ema.europa.eu/en/documents/herbal-report/final-assessment-report-matricaria-recutita-l-flos-matricaria-recutita-l-aetheroleum-first-version_en-0.pdf, accessed on 09/21/22

Fursa, N. S., Litvinenko, V. I., & Krivenchuk, P. E. (1969). Flavonoids from Armoracia rusticana and Barbarea arcuata. *Chimija Prirodnych Soedinenij*, *5*(4), 320.

Hunt, B. J., & Holding, S. R. (Eds.). (2013). *Size Exclusion Chromatography* (1st ed.). Springer New York, NY.

Loidolt, M. (2022). *Phytochemical investigation of extracts against SARS-CoV-2* [Master Thesis]. University of Vienna.

McKay, D. L., & Blumberg, J. B. (2006). A review of the bioactivity and potential health benefits of chamomile tea (Matricaria recutita L.). *Phytotherapy Research*, *20*(7), 519–530. https://doi.org/10.1002/ptr.1900

Myers, O. D., Sumner, S. J., Li, S., Barnes, S., & Du, X. (2017). One step forward for reducing false positive and false negative compound identifications from mass spectrometry metabolomics data: New algorithms for constructing extracted ion chromatograms and detecting chromatographic peaks. *Analytical Chemistry*, *89*(17), 8696–8703. https://doi.org/10.1021/acs.analchem.7b00947

Olivon, F., Allard, P.-M., Koval, A., Righi, D., Genta-Jouve, G., Neyts, J., Apel, C., Pannecouque, C., Nothias, L.-F., Cachet, X., Marcourt, L., Roussi, F., Katanaev, V. L., Touboul, D., Wolfender, J.-L., & Litaudon, M. (2017). Bioactive natural products prioritization using massive multi-informational molecular networks. *ACS Chemical Biology*, *12*(10), 2644–2651. https://doi.org/10.1021/acschembio.7b00413

Park, S.-B., Song, K., & Kim, Y. (2017). Tetra-cis/trans-coumaroyl polyamines as NK1 receptor antagonists from Matricaria chamomilla. *Planta Medica International Open*, *4*, e43–e51. https://doi.org/10.1055/s-0043-106742

Petruľová-Poracká, V., Repčák, M., Vilková, M., & Imrich, J. (2013). Coumarins of Matricaria chamomilla L.: Aglycones and glycosides. *Food Chemistry*, *141*(1), 54–59. https://doi.org/10.1016/j.foodchem.2013.03.004

Pluskal, T., Castillo, S., Villar-Briones, A., & Orešič, M. (2010). MZmine 2: Modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics*, *11*(1), 395. https://doi.org/10.1186/1471-2105-11-395

Schilcher, H., Imming, P., & Goeters, S. (2005). Active chemical constituents of Matricaria chamomilla L. syn. Chamomilla recutita (L.) Rauschert. In H. Schilcher & R. Franke (Eds.), *Chamomile: Industrial Profiles*. CRC Press. https://books.google.at/books?id=wXmmDwAAQBAJ

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, *13*(11), 2498–2504. https://doi.org/10.1101/gr.1239303

Slavov, A., Yantcheva, N., & Vasileva, I. (2019). Chamomile wastes (Matricaria chamomilla): New source of polysaccharides. *Waste and Biomass Valorization*, *10*(9), 2583–2594. https://doi.org/10.1007/s12649-018-0270-5

Srivastava, J. K., Shankar, E., & Gupta, S. (2010). Chamomile: A herbal medicine of the past with a bright future (review). *Molecular Medicine Reports*, *3*(6), 895–901. https://doi.org/10.3892/mmr.2010.377

Švehlíková, V., Bennett, R. N., Mellon, F. A., Needs, P. W., Piacente, S., Kroon, P. A., & Bao, Y. (2004). Isolation, identification and stability of acylated derivatives of apigenin 7-O-glucoside from chamomile (Chamomilla recutita [L.] Rauschert). *Phytochemistry*, *65*(16), 2323–2332. https://doi.org/10.1016/j.phytochem.2004.07.011

Tschiggerl, C., & Bucar, F. (2012). Guaianolides and volatile compounds in chamomile tea. *Plant Foods for Human Nutrition (New York, NY, United States)*, *67*(2), 129–135. https://doi.org/10.1007/s11130-012-0277-1

Tsivelika, N., Irakli, M., Mavromatis, A., Chatzopoulou, P., & Karioti, A. (2021). Phenolic profile by HPLC-PDA-MS of greek chamomile populations and commercial varieties and their antioxidant activity. *Foods*, *10*(10). https://doi.org/10.3390/foods10102345

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

Dorrestein, P. C., & Bandeira, N. (2016). Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnology*, *34*(8), 828–837. https://doi.org/10.1038/nbt.3597

Xie, X.-Y., Chen, F.-F., & Shi, Y.-P. (2014). Simultaneous determination of eight flavonoids in the flowers of Matricaria chamomilla by high performance liquid chromatography. *Journal of AOAC International*, *97*(3), 778–783. https://doi.org/10.5740/jaoacint.13-029

Yang, J. Y., Sanchez, L. M., Rath, C. M., Liu, X., Boudreau, P. D., Bruns, N., Glukhov, E., Wodtke, A., de Felicio, R., Fenner, A., Wong, W. R., Linington, R. G., Zhang, L., Debonsi, H. M., Gerwick, W. H., & Dorrestein, P. C. (2013). Molecular networking as a dereplication strategy. *Journal of Natural Products*, *76*(9), 1686–1699. https://doi.org/10.1021/np400413s

Zhao, Y., Sun, P., Ma, Y., Wang, K., Chang, X., Bai, Y., Zhang, D., & Yang, L. (2019). Simultaneous quantitative determination of six caffeoylquinic acids in Matricaria chamomilla L. with high-performance liquid chromatography. *Journal of Chemistry*, *2019*, 4352832. https://doi.org/10.1155/2019/4352832

9 ACKNOWLEDGEMENTS

First and foremost, I would like to thank **Univ.-Prof. Dr. Judith M. Rollinger** for giving me the opportunity to write my master thesis at the Division of Pharmacognosy. I have always been very interested in this field and I was very excited when I got accepted to do my thesis in the Phytochemistry and Biodiscovery group. Prof. Rollinger was always interested in the progress of my work. Whenever there were difficulties, she tried to help with her impressive expertise. I highly appreciated her efforts.

Furthermore, many thanks to **Mag. pharm. Andreas Wasilewicz** for his supervision in the course of my thesis. Andi was always eager to help me with any problem I had during my work. Despite his busy work, he always took the time to answer my questions. As Andi was noticeably very invested in the project, he was very patient whenever giving me instructions and his suggestions were always well thought-out. Nevertheless, I was able to work very independently in the laboratory, which I really appreciated. I wish him all the best for his future career.

Special thanks to **Ruzi** for her guidance and support in the laboratory. Whenever you couldn't find something, all you had to do was turn to her and she helped you straight away. Thank you Ruzi for patiently explaining how to perform size-exclusion chromatography and for helping to solve the mystery of how to operate the fraction collector.

Many thanks to all the **members of the Division of Pharmacognosy**, especially the team of the Phytochemistry and Biodiscovery group. All of you made my time at the division very enjoyable. I also want to thank my fellow **master students** for the great time we had working together in lab and the countless laughs we shared during lunch breaks. I enjoyed coming to university every day.

Finally, I would like to thank my **family** for the endless encouragement and support during my studies.

10 APPENDIX

Chamomile microfractions (MF1-27) annotation table

RT [min]	m/z (positive mode)	proposed substance(s)	MW [g/mol]	occurrence in microfractions
0.744	126.04	-	-	*5-6~, 6-7↑, 7-8↑, 8-9↑, 9-10~, 10-11↓, 11-12~, 12-13~, 13-14↑, 14-15↓, 15-16↑, 16-17~, 17-18~, 18-19~, 19-20~, 20-21↓, 21-22↑, 22-23~
0.862	158.17; 268.19; 144.12; 107.08; 124.08; 144.12; 123.09; 162.09; 377.18; 379.19	-	-	*5-6↓, 6-7↑, 7-8↑, 8-9↓, 9-10↑, 10-11↑, 11-12↓, 12-13~, 13-14↑, 14-15↓,15-16↑, 16-17↑, 17-18?, 18-19~, 19-20↓, 20-21↑, 21-22↓, 22-23↑, 23-24?, 24-25↑, 25-26↑, 26-27↓
1.035	268.19; 144.14	-	-	*24-25↑, 25-26↑, 26-27↑
1.098	268.19; 144.10	-	-	*24-25↑, 25-26↑, 26-27↓
1.100	123.10	-	-	*16-17↑, 17-18↓
1.152	123.10; 187.15	-	-	*16-17↑, 17-18↓
1.153	124.07	-	-	*9-10↑, 10-11↑, 11-12↓
1.153	132.16; 324.14; 294.24; 276.24	-	-	*26-27↓
1.156	154.13	-	-	*21
1.157	115.06; 161.06	-	-	*13-14↑, 14-15↓
1.157	268.21	-	-	*24-25↑
1.420	144.04	-	-	*9
1.520	141.09; 261.12	-	-	*7-8↑, 8-9↓
1.613	166.13	-	-	*26
1.681	162.09	-	-	*18-19↑
1.769	22.25; 242.25; 166.10	-	-	*21
2.815	146.10	-	-	*14-15↓
2.894	183.19	-	-	*7
2.992	347.10; 325.15; 163.08	-	-	*24

3.553	377.19 [M+Na]	chlorogenic acid	354.3	*17-18↑, 18-19↓, 19-20↓, 20-21↑
3.946	397,30	-	-	*25
4.518	307,16	-	-	*18-19↑
4.645	379.21; 177.11 [M+H-glucose]; 195.15 [methoxy caffeic acid+H]	(Z)-2-β-D-Glucopyranosyloxy-4-methoxycinnamic acid (cis-GMCA)	356.3	*19-20↓, 20-21↑, 21-22↑, 22-23↑, 23-24↓
5.079	363.20; 358.28	-	-	*16
5.280	457.35; 237.28; 563.35	-	-	*25
5.518	457.35; 237.28	-	-	*25
5.553	481.20 [M+H]	Quercetagetin-?-O-glucoside	480.4	*19-20↑, 20-21↓
5.706	457.36	-	-	*6-7↓
5.808	465.21 [M+H]	Quercetin-?-O-glycoside	464.4	*21
6.051	363.18; 358.26	-	-	*14
6.082	391.20; 177.12	-	-	*15-16↓
6.103	363.27; 167.08	-	-	*9-10↑, 10-11↓
6.233	163.10 [M+H]	Umbelliferon	162.1	*4-5↑
6.260	207.22	-	-	*7
6.379	285.20	-	-	*17-18↑
6.419	347.13; 218.26	-	-	*12-13↓
6.460	237.33	-	-	*9-10↑, 10-11↓
6.557	165.09; 171.18; 197.20	-	-	*5
6.563	379.19 [M+Na]; 177.12 [M+H-glucose]; 195.14 [methoxy caffeic acid+H]	(E)-2-β-D-Glucopyranosyloxy-4-methoxycinnamic acid (trans-GMCA)	356.1	*18-19↑, 19-20↑, 20-21↓, 21-22↓, 22-23↑, 23-24↓
6.675	347.12	Eupatoletin? Spinacetin? 3-O-Methylpatuletin?	-	*11-12↑, 12-13↓, 13-14↑, 14-15↓, 15-16↓
6.716	363.19; 358.26	-	-	*12-13↓
6.719	165.10	-	-	*5

6.772	351.20	-	-	*21
6.830	197.22	-	-	*7
6.929	351.24	-	-	*21
7.030	237.30; 295.25	-	-	*8-9↓
7.056	465.24 [M+H]	Quercetin-?-O-glycoside	464.4	*14-15↑
7.255	361.18 [M+H]; 487.20	Jaceidin? Chrysoplenol?	-	*11-12↑, 12-13↓, 13-14↑, 14-15~
7.367	449.21 [M+H]	Luteolin-7-O-glucoside	448.4	*15-16↑
7.422	225.22; 167.18	-	-	*7
7.507	197.24	-	-	*5
7.609	495.21 [M+H]	Patuletin-7-O-glucoside	494.4	*15-16↑, 16-17~, 17-18↓, 18-19↓
7.995	643.32	-	-	*15
8.150	499.20; 539.20 [M+Na]	Dicaffeoylquinic acid	516.4	8-9↑, 9-10↓, 10-11↓, 11-12↓, 12-13↑, 13-14↑
8.217	471.18; 287.12	-	-	*12
8.254	455.36	-	-	*5-6↑
8.321	393.33	-	-	*21-22↓
8.451	295.24	-	-	*11
8.462	311.26	-	-	*7
8.532	263.25	-	-	*5
8.534	499.24; 539.17 [M+Na]	Dicaffeoylquinic acid	516.4	*8-9↑, 9-10↓
8.574	501.17; 317.10	-	-	*15
8.705	433.25 [M+H]	Apigenin-7-O-glucoside	432.4	*11-12↑, 12-13↑, 13-14↑
8.834	293.23	-	-	*5
8.890	531.20	-	-	*15
8.900	455.36	-	-	*21-22↓, 22-23↑
8.960	185.17; 211.30	-	-	*4-5↑
8.985	539.20 [M+Na]; 499.25	Dicaffeoylquinic acid	516.4	*9-10↑, 10-11↑

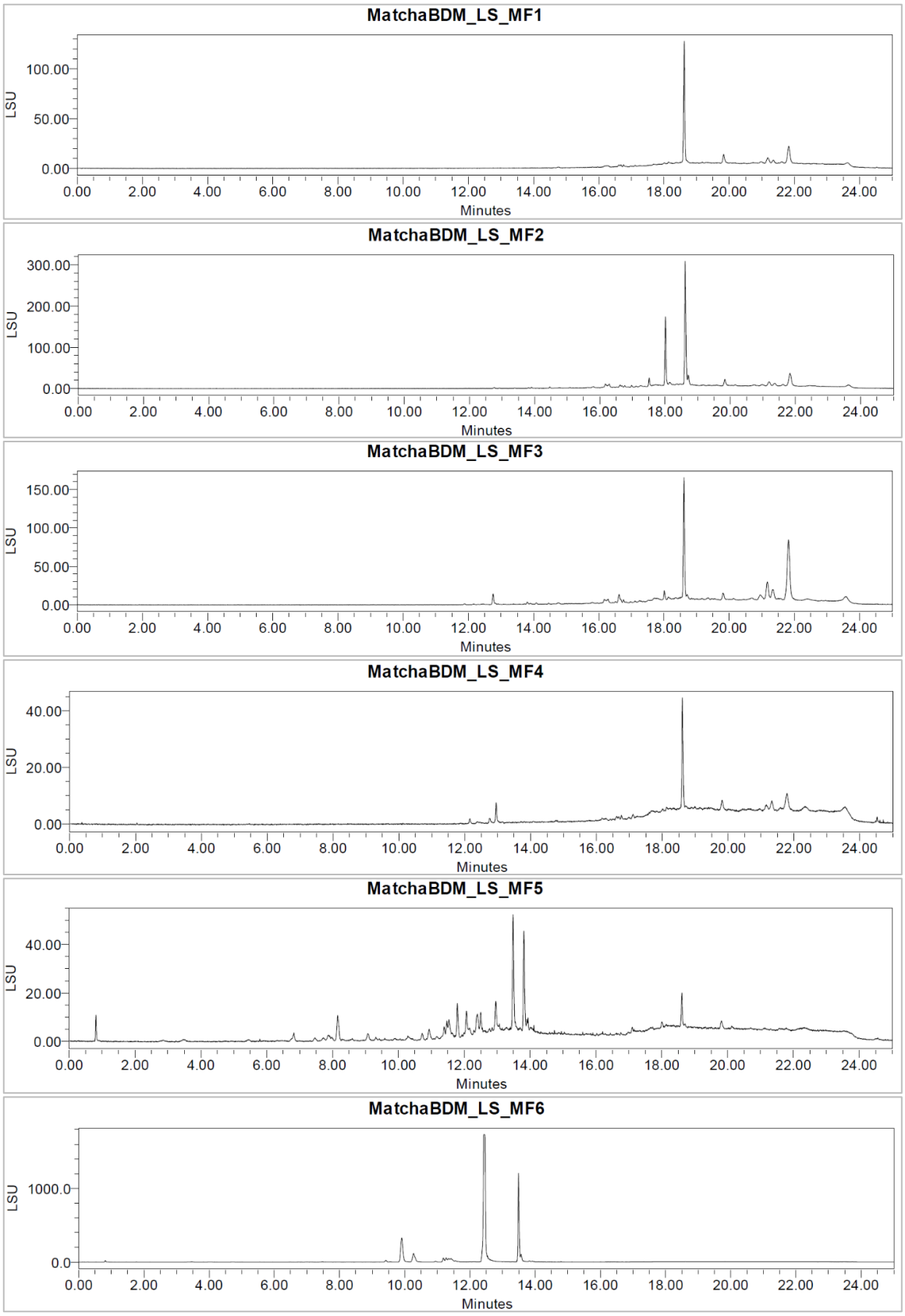
8.971	309.25	-	-	*7
9.040	479.20	Isorhamnetin-7-O-glucoside? Isorhamnetin-3-O-glucoside? Quercetin-3-O-glucuronide?	-	*11-12↑, 12-13↓, 13-14↑
9.072	463.25; 531.15	-	-	*14-15↓
9.240	309.24	-	-	*7
9.347	509.21 [M+H]; 531.19 [M+Na]	Flavonoid glycoside	-	*14-15↓
9.492	439.33 [M+Na]	Hydroxybisabolol oxide glucoside?	-	*20-21↑, 21-22↓
9.718	657.36 [M+H]	Flavonoid glycoside	-	*11-12↑, 12-13↓
9.848	349.22	-	-	*14
9.889	541.26; 307.20	-	-	*5
9.892	541.25	-	-	*11
10.046	541.24; 429.26; 627.24; 641.50 [M+H]; 321.42	Flavonoid glycoside	-	*9-10↑, 10-11↑, 11-12↓ *23-24↑
10.217	295.26	-	-	*5-6↑
10.325	519.15; 520.20	Apigenin-7-O-malonyl-glucoside?	-	*11-12↓, 12-13↓, 13-14↑
10.469	539.17; 163.05	-	-	*9-10↑
10.417	419.30	-	-	*13-14↑
10.471	421.30; 177.10; 537.20	-	-	*11-12↓
10.641	439.37	-	-	*20-21↑
10.645	475.25 [M+H]	Apigenin-7-O-?-acetyl-glucoside	474.1	*6-7↑, 7-8↓
10.910	177.11 [M+H]	Herniarin	176.2	*3-4↓
10.977	541.22	Flavonoid glycoside	-	*11
11.029	521.21; 543.20	-	-	*6-7↓
11.057	505.21	-	-	*7
11.250	595.38	-	-	*8

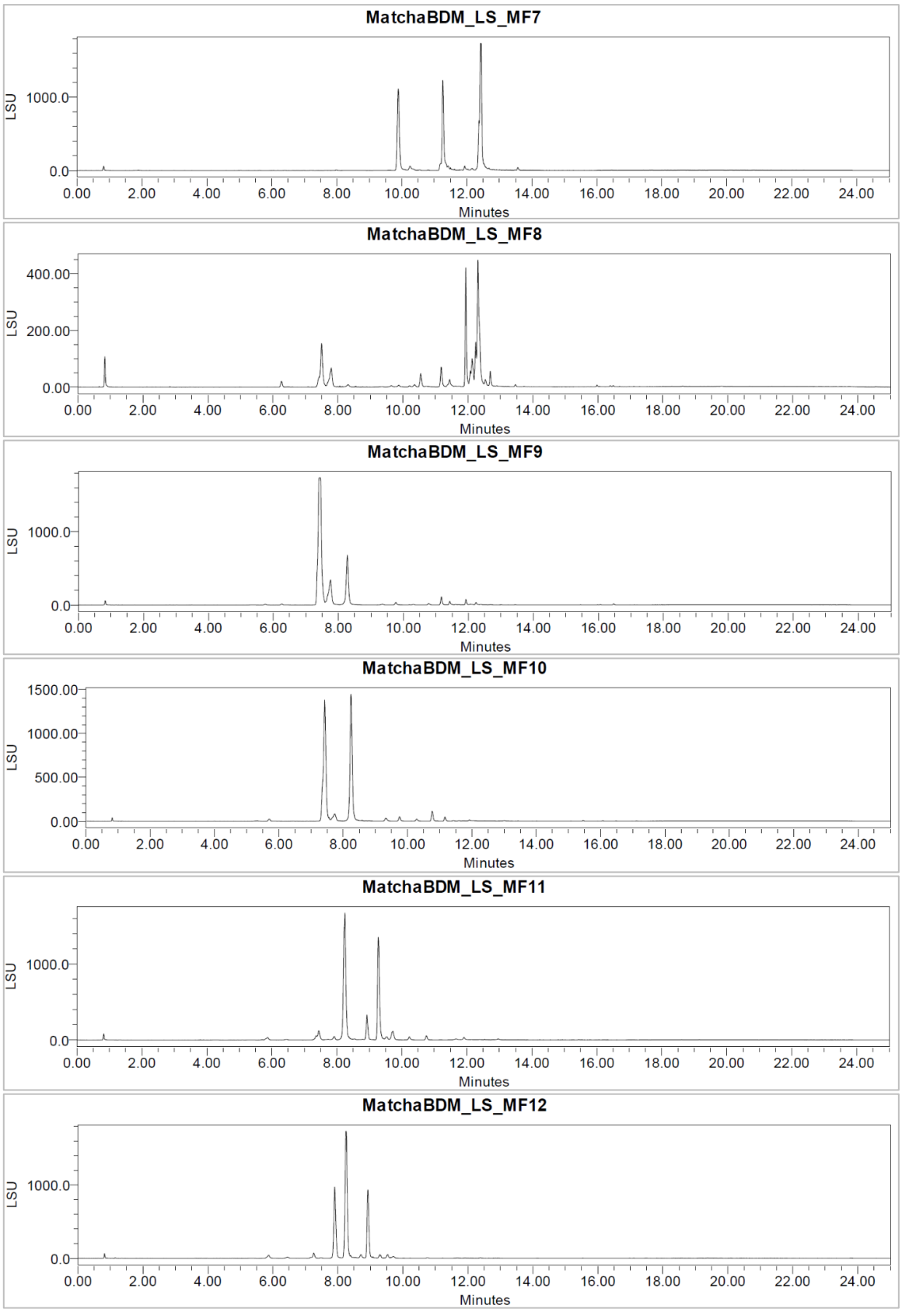
11.350	293.25	-	-	*4-5↑
11.359	387.28; 185.19	-	-	*12-13↓
11.367	541.24	Flavonoid glycoside	-	*9-10↑, 10-11↓
11.540	441.35	-	-	*17
11.570	455.36	-	-	*5-6↑
11.684	331.19	Eupaletin?	-	*8-9↑, 9-10↓, 10-11↓
11.685	389.29	-	-	*13-14↑
11.750	475.26 [M+H]; 561.26 [M+H]	Apigenin-7-O-?-acetyl-glucoside (474.1)? Apigenin-7-O-?-acetyl-malonyl-glucoside (~560)?	-	*6-7↑
11.776	369.33; 269.29	-	-	*5
11.895	333.20	Patuletin?	-	*8-9↑, 9-10↓
11.947	455.34; 235.33	-	-	*5-6↑
12.015	333.21; 149.10	-	-	*9
12.086	308.31; 455.35	-	-	*5-6↓
12.095	401.28; 473.24; 199.21	-	-	*11-12↓
12.096	335.18; 313.24	Hydroxybisabolol oxide glucoside?	-	*8
12.186	541.30; 525.33	-	-	*5
12.265	335.19	-	-	*7-8↑, 8-9↓
12.291	439.37 [M+Na]; 167.13; 473.27	Hydroxybisabolol oxide glucoside?	-	*10-11↑
12.296	271.19 [M+H]	Apigenin	270.2	*3
12.386	291.22; 251.28	-	-	*1-2↑, 2-3↑
12.410	417.31; 412.35; 804.56; 803.54	-	-	*8-9↓
12.422	541.33	-	-	*5
12.606	247.26 [M+H]	Achillin? Leucodin?	246.3	*3-4↑
12.630	487.24	-	-	*9
12.638	555.41	-	-	*11-12↑, 12-13↓, 13-14↑, 14-15↓

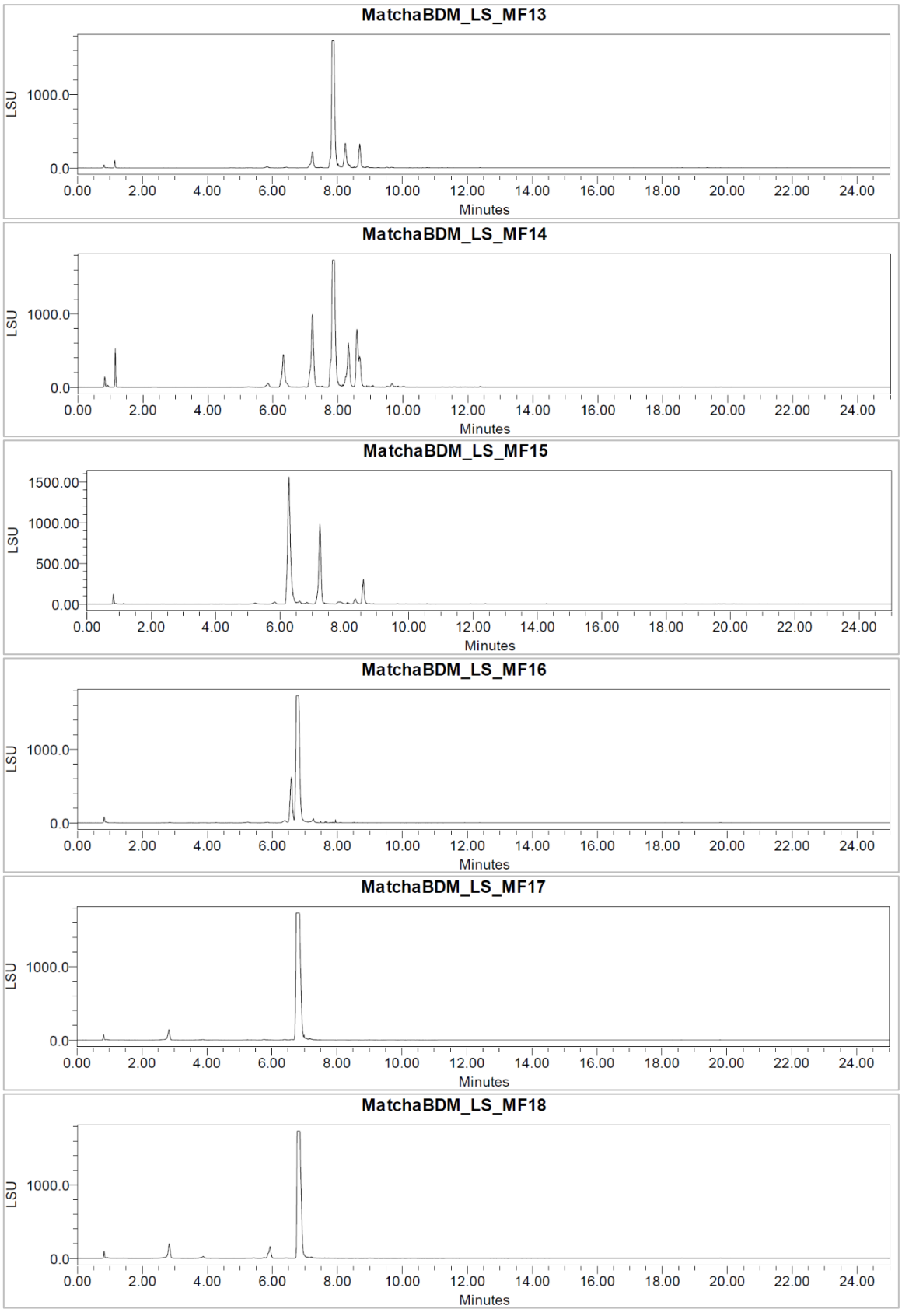
12.695	313.31; 353.34; 291.25	-	-	*5
12.776	275.26; 235.34	-	-	*4
12.790	419.31	-	-	*8
12.791	277.27; 219.27	-	-	*2-3↑
12.700	787.50 [M+H]; 788.48; 809.51	Polyamines	~786	*6-7↑, 7-8↑, 8-9↓
12.795	416.35; 421.32	-	-	*7-8↓
12.817	351.30; 293.32; 275.30	-	-	*5
12.853	305.25	-	-	*3-4↓
12.873	339.30; 337.23; 305.22	4-O-p-Coumaroylquinic acid? 3-O-p-Coumaroylquinic acid?	-	*2-3↑,3-4↓
12.906	155,15; 241,28	Achillin? Leucodin?	-	*1-2↑, 2-3↑, 3-4↓
12.939	419.28; 369.14	-	-	*7-8↑, 8-9↓
13.110	305.20 [M+H]; 235.36; 253.31; 293.27	Matricarin	304.3	*1-2↑, 2-3↑, 3-4↓
13.233	353.32 [M+Na]	Eupaletin?	-	*4-5↑
13.260	293.26; 235.36	-	-	*2-3↑, 3-4↓
13.416	253.35; 311.34	-	-	*4
13.601	307.26 [M+H]	Matricin	306.4	*3
13.694	201.21	-	-	*2-3↓
13.726	431.38; 317.29	-	-	*8
13.764	423.39 [M+Na]; 221.32 [M-glucose]	Bisabolol oxide A glucoside	-	*5-6↑, 6-7↓
14.020	275.26	-	-	*4-5↑
14.074	509.33	-	-	*5-6↓
14.106	375.20 [M+H]	Chrysosplenetin	374.3	*2-3↑
14.190	509.36	-	-	*6
14.196	201.23	-	-	*2-3↓

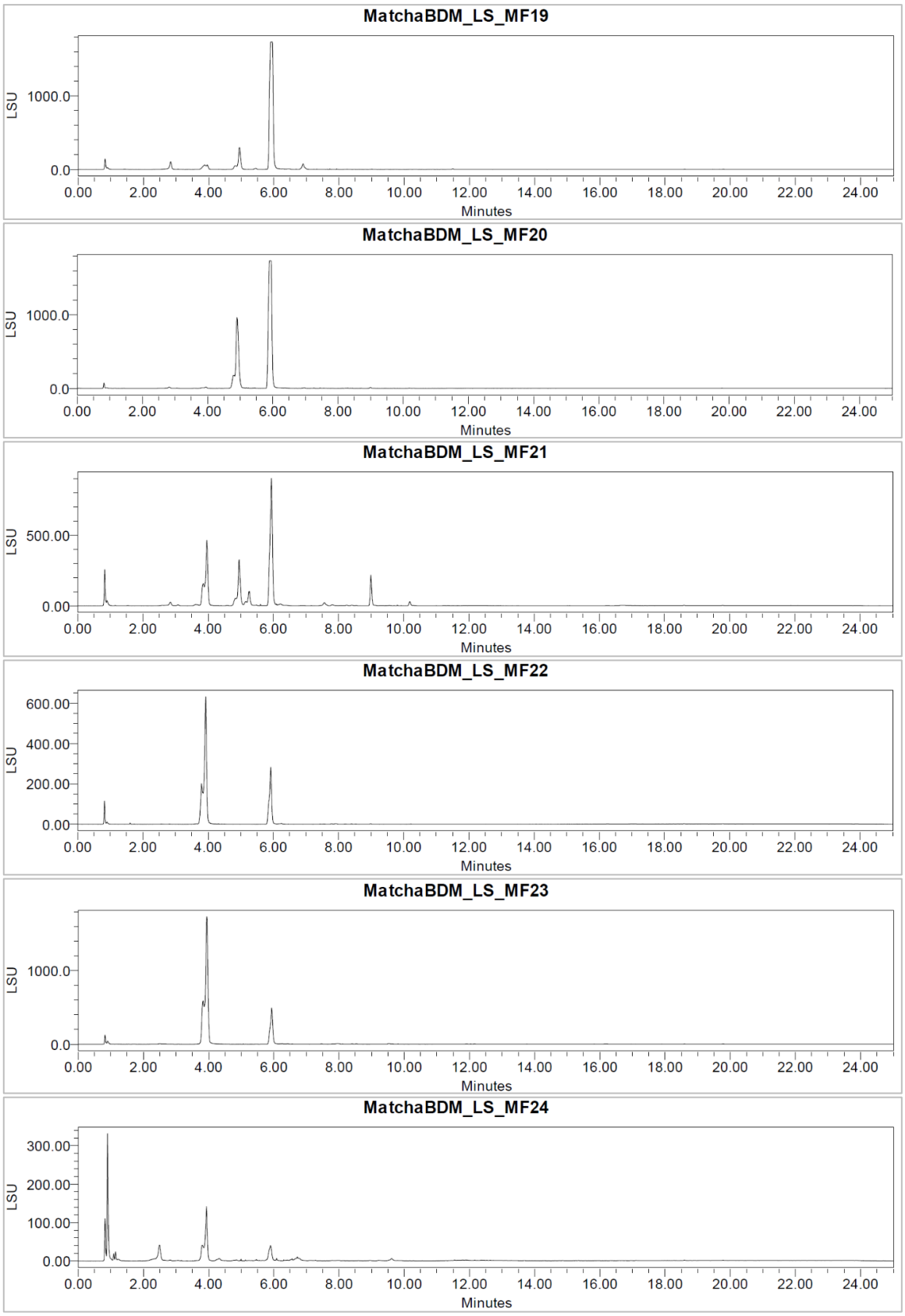
14.365	429.37	-	-	*3
14.745	201.13 [M+H]	(E)-En-yn-dicycloether	200.2	*1-2↑, 2-3↓, 3-4↓
14.965	367.29	-	-	*1-2↑, 2-3~
15.117	319.28	Bisabololoxide?	-	*2-3~
15.323	201.12 [M+H]	(Z)-En-yn-dicycloether	200.2	*1-2↑, 2-3↓, 3-4↓
15.609	699.55; 353.39	-	-	*10-11↓
16.012	221.31 [M-OH]; 277.34	Bisabololoxide? Farnesen?		*2-3↓
16.136	215.25	-	-	*2-3↓
16.201	520.46	-	-	*9-10↑
16.270	221.32 [M-OH]	Bisabololoxide? Farnesen?		*2
16.383	279.28 [M+H]	Fatty acid	-	*2-3↓
16.503	279.28 [M+H]; 277.29; 280.30	Fatty acid	-	*2-3↓
16.638	295.35 [M+H]; 277.30	Fatty acid	-	*2
16.673	496.45	-	-	*8-9↑
16.726	301.22	-	-	*9
17.513	-	-	-	*2
18.012	-	-	-	*2, 3
18.319	281.36; 263.39	-	-	*2
18.639	-	-	-	*1, 2, 3, 4
18.716	-	-	-	*2, 3
19.866	-	-	-	*1, 2, 3, 4
20.183	391.41; 413.42	-	-	*25
21.229	-	-	-	*1, 2, 3, 4
21.405	-	-	-	*1, 2, 3, 4
21.878	-	-	-	*1, 2, 3, 4

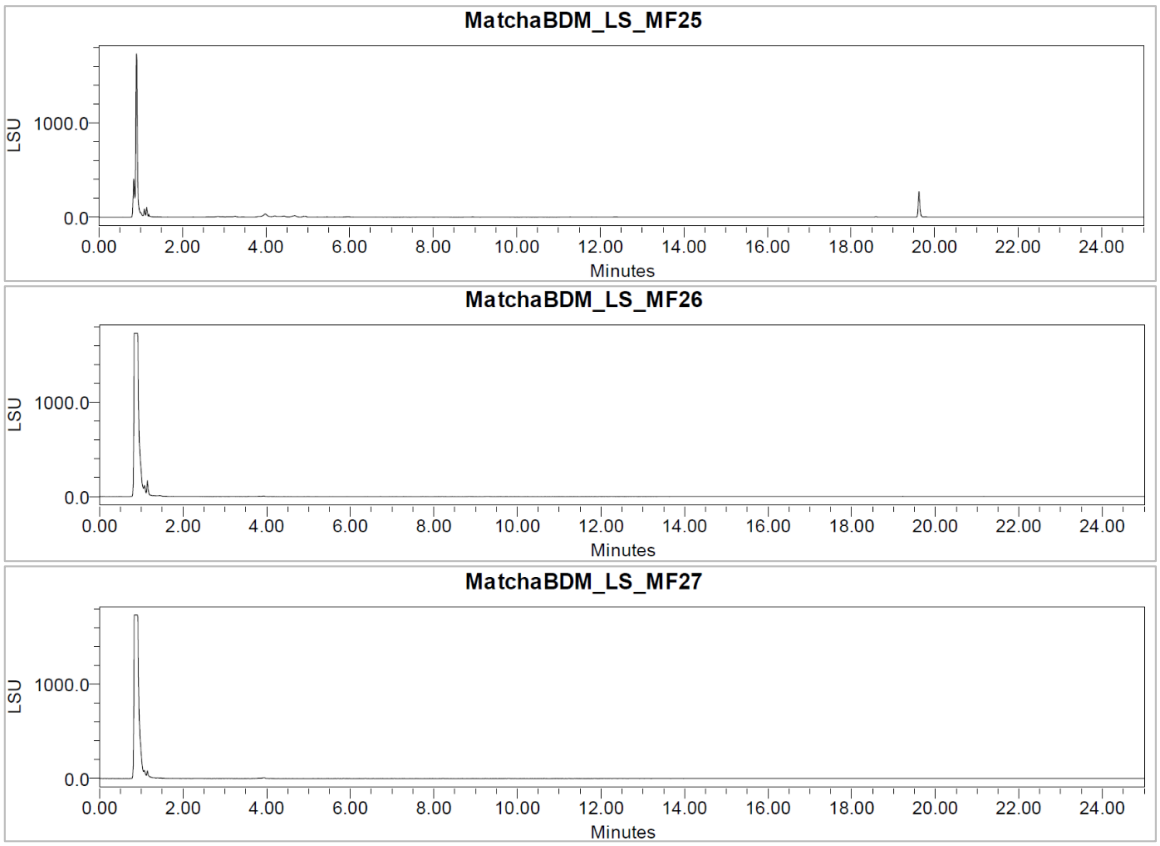
Chamomile microfractions (MF1-27) ELSD chromatograms



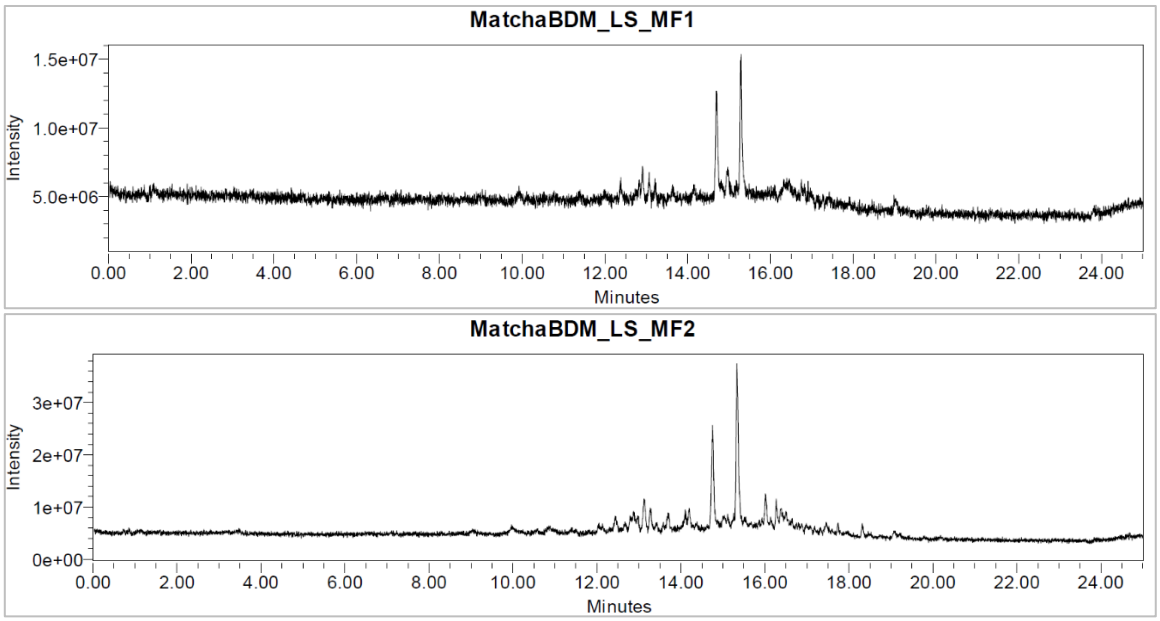

MatchaBDM_LS_MF7
LSU
1000.0
0.0
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes
MatchaBDM_LS_MF8
LSU
400.00
200.00
0.00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes
MatchaBDM_LS_MF9
LSU
1000.0
0.0
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes
MatchaBDM_LS_MF10
LSU
1500.00
1000.00
500.00
0.00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes
MatchaBDM_LS_MF11
LSU
1000.0
0.0
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes
MatchaBDM_LS_MF12
LSU
1000.0
0.0
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

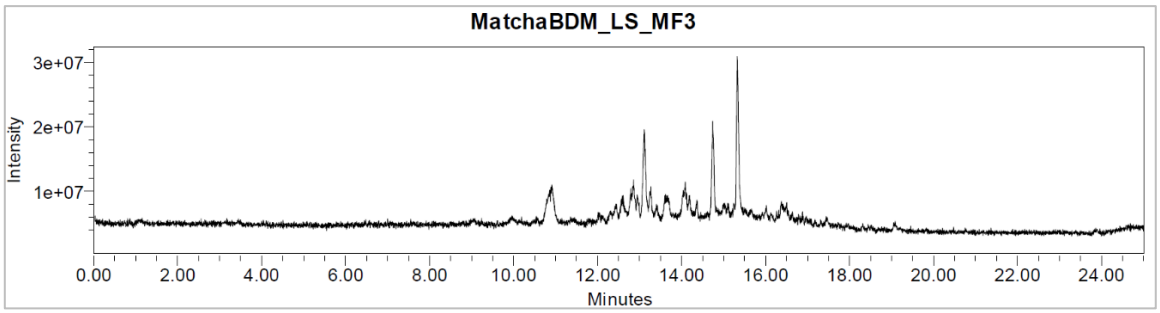

MatchaBDM_LS_MF13
LSU
1000.0
0.0
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes
MatchaBDM_LS_MF14
LSU
1000.0
0.0
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes
MatchaBDM_LS_MF15
LSU
1500.00
1000.00
500.00
0.00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes
MatchaBDM_LS_MF16
LSU
1000.0
0.0
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes
MatchaBDM_LS_MF17
LSU
1000.0
0.0
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes
MatchaBDM_LS_MF18
LSU
1000.0
0.0
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

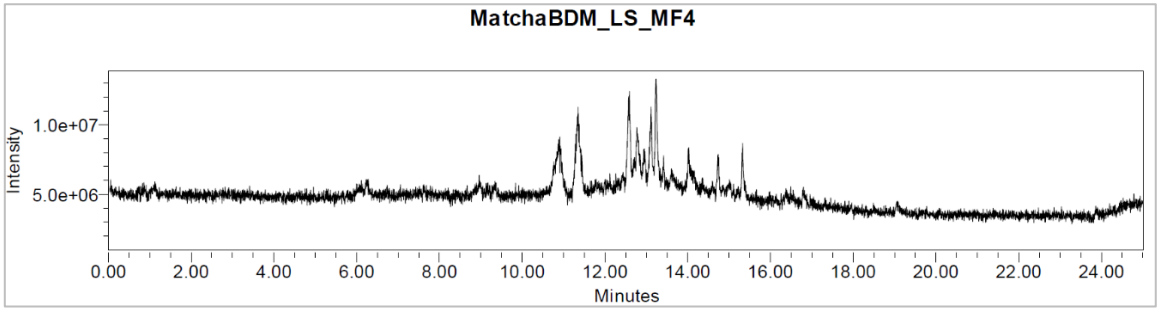

MatchaBDM_LS_MF19
LSU
1000.0
0.0
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes
MatchaBDM_LS_MF20
LSU
1000.0
0.0
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes
MatchaBDM_LS_MF21
LSU
500.00
0.00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes
MatchaBDM_LS_MF22
LSU
600.00
400.00
200.00
0.00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes
MatchaBDM_LS_MF23
LSU
1000.0
0.0
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes
MatchaBDM_LS_MF24
LSU
300.00
200.00
100.00
0.00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

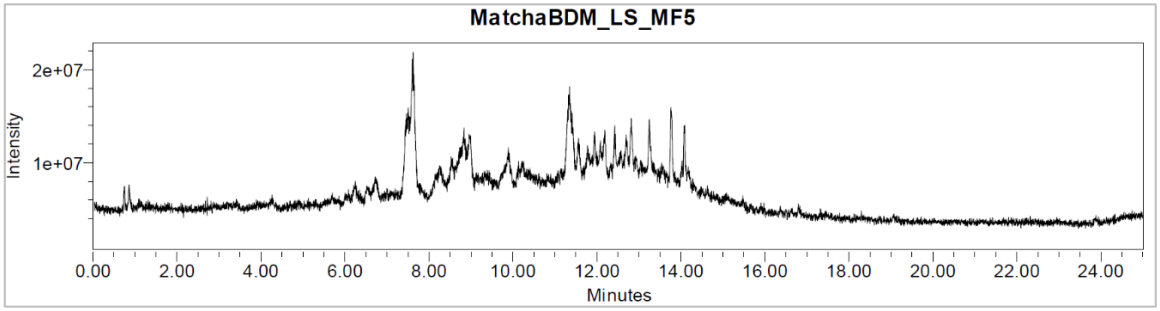


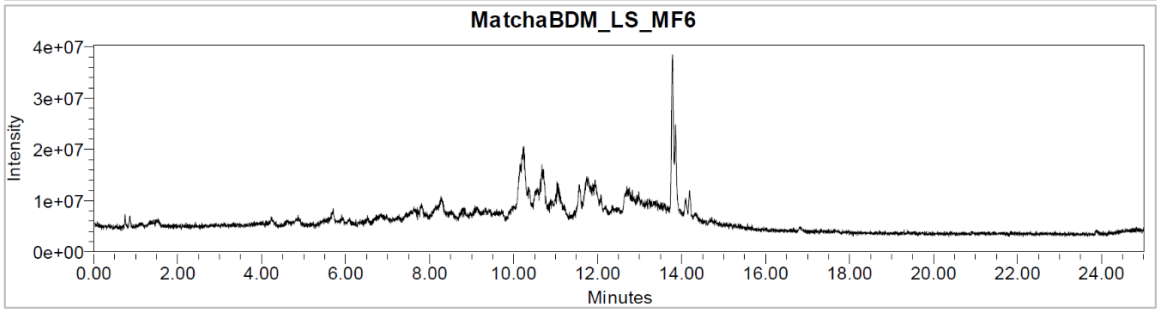
Chamomile microfractions (MF1-27) QDa scans (positive mode)

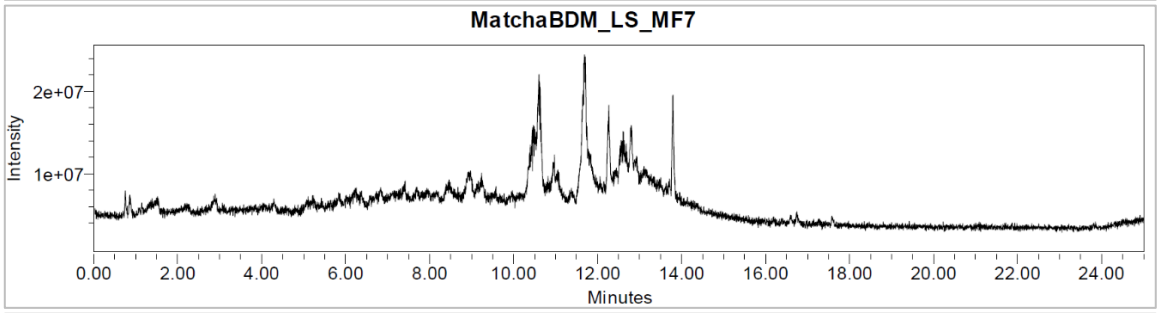


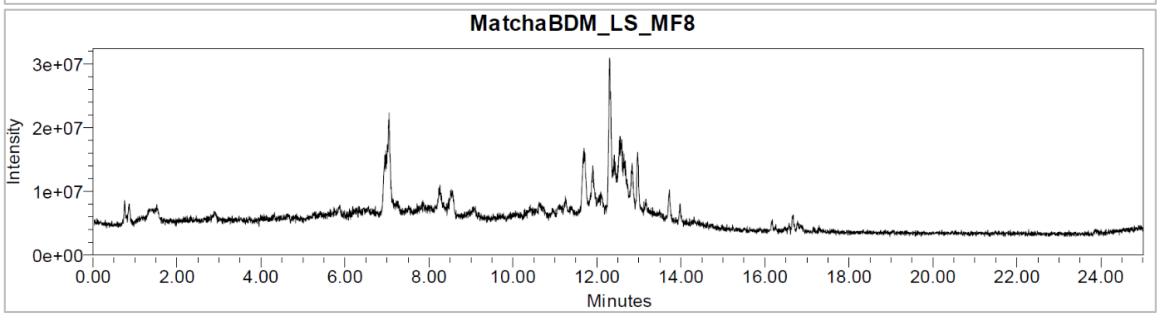

MatchaBDM_LS_MF3
Intensity
3e+07
2e+07
1e+07
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

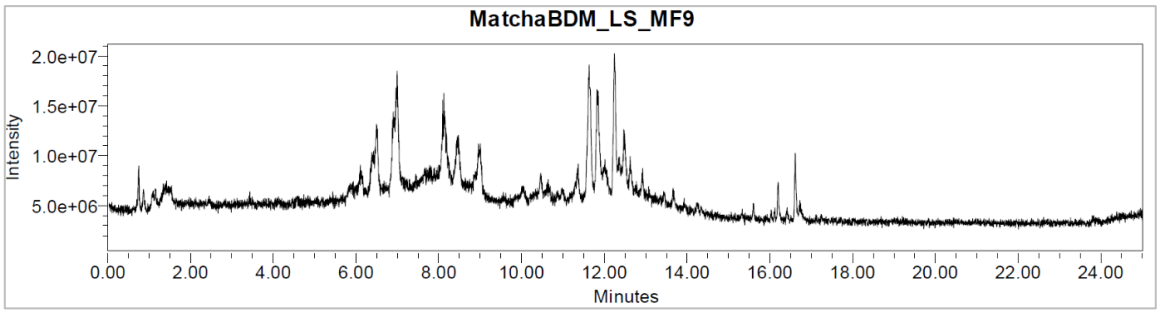

MatchaBDM_LS_MF4
Intensity
1.0e+07
5.0e+06
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

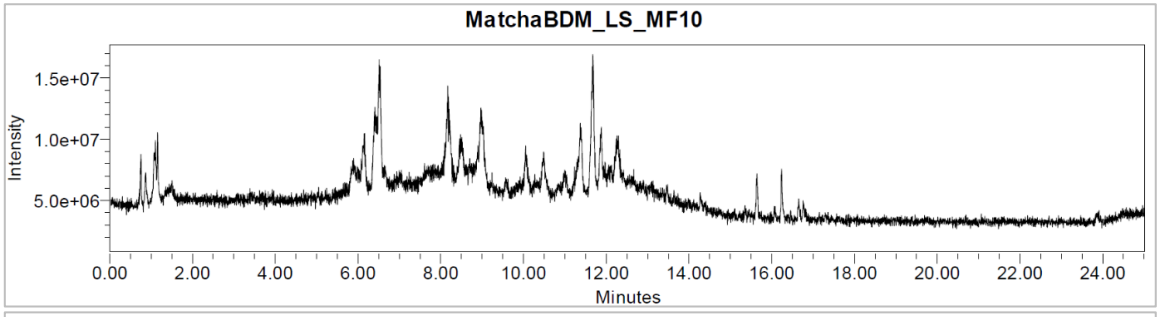

MatchaBDM_LS_MF5
Intensity
2e+07
1e+07
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

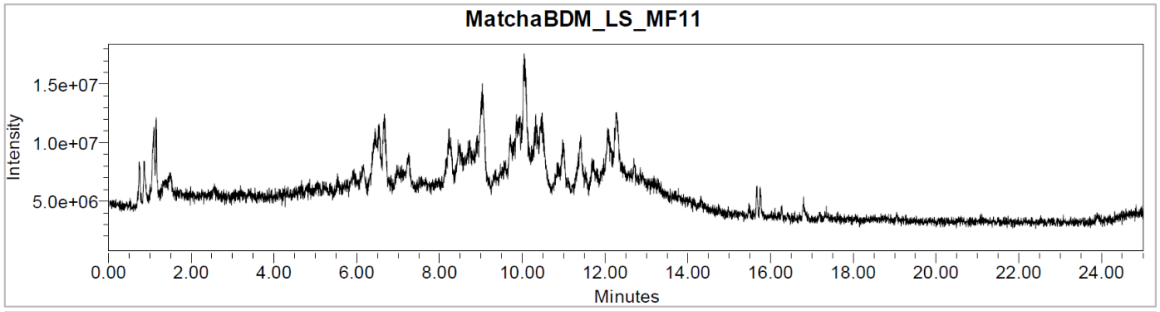

MatchaBDM_LS_MF6
Intensity
4e+07
3e+07
2e+07
1e+07
0e+00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

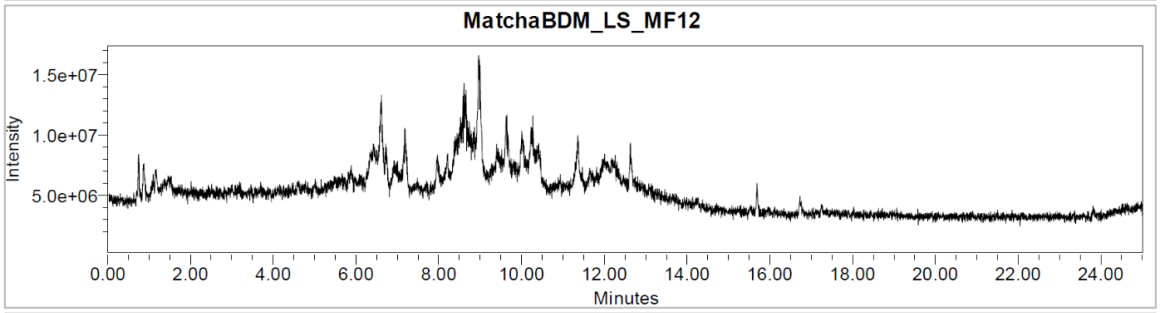

MatchaBDM_LS_MF7
Intensity
2e+07
1e+07
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

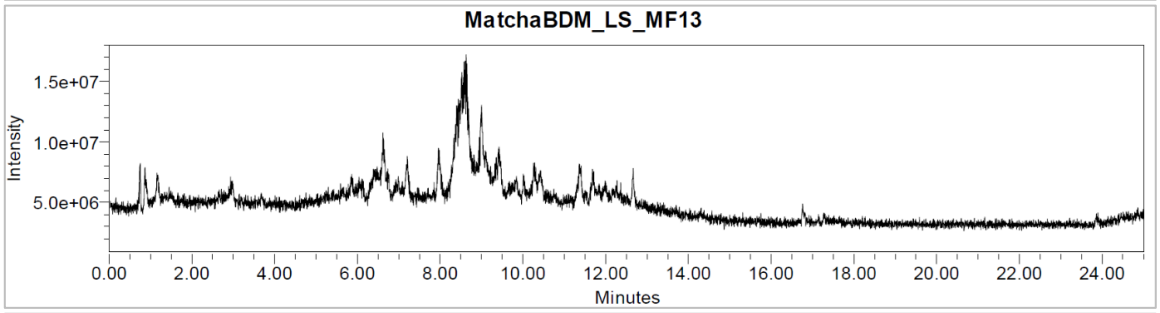

MatchaBDM_LS_MF8
Intensity
3e+07
2e+07
1e+07
0e+00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

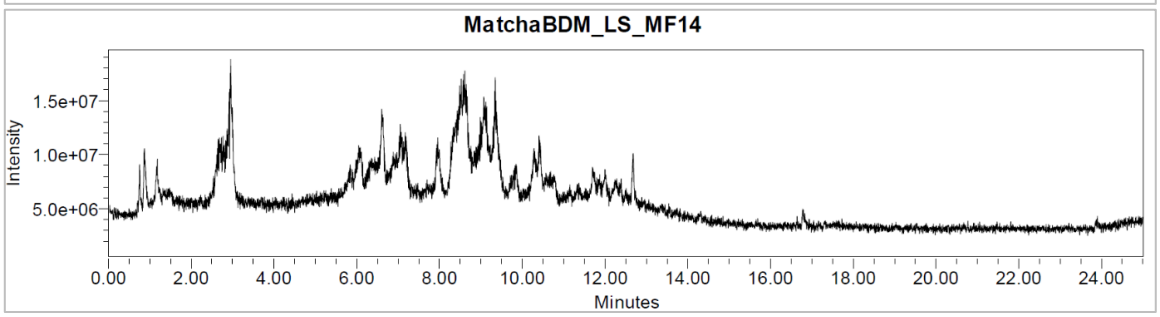

MatchaBDM_LS_MF9
Intensity
2.0e+07
1.5e+07
1.0e+07
5.0e+06
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

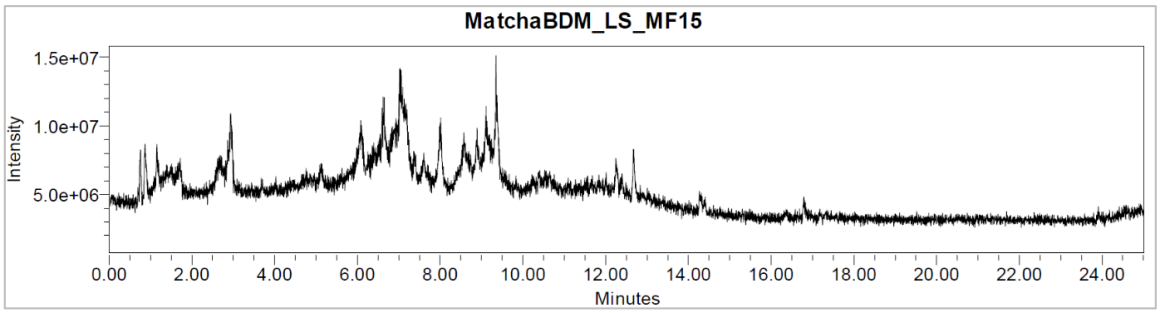

MatchaBDM_LS_MF10
Intensity
1.5e+07
1.0e+07
5.0e+06
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

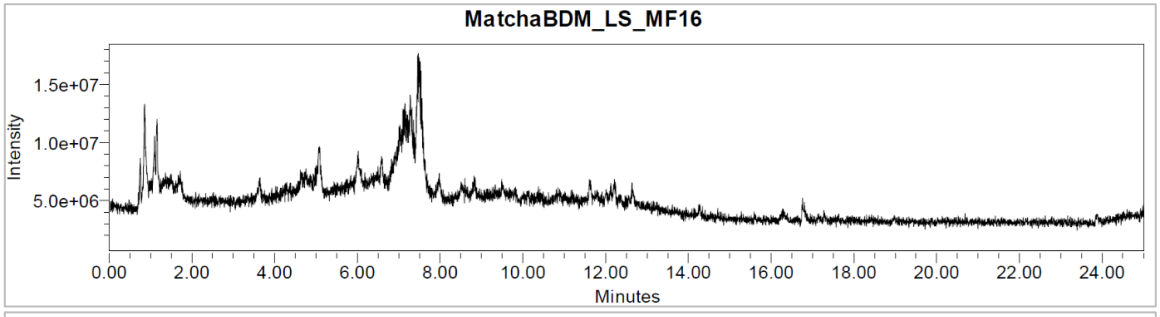

MatchaBDM_LS_MF11
Intensity
1.5e+07
1.0e+07
5.0e+06
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

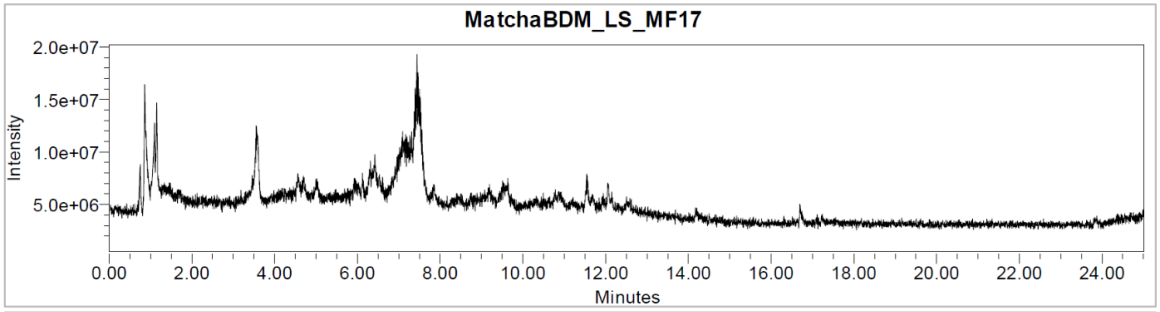

MatchaBDM_LS_MF12
Intensity
1.5e+07
1.0e+07
5.0e+06
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

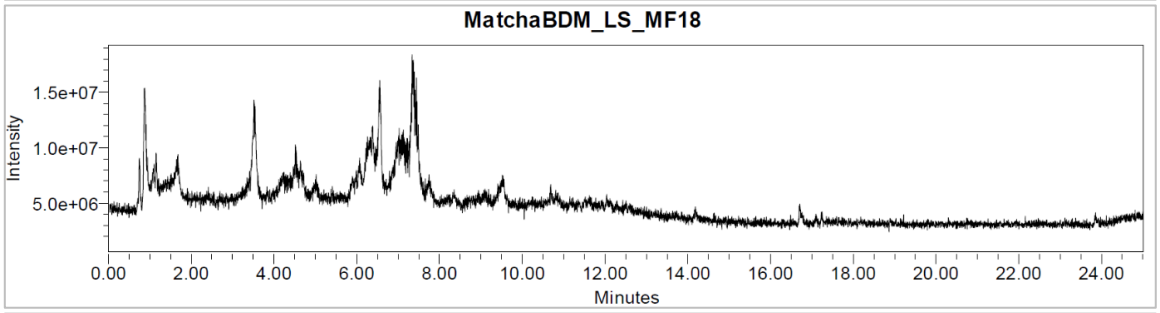

MatchaBDM_LS_MF13
Intensity
1.5e+07
1.0e+07
5.0e+06
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

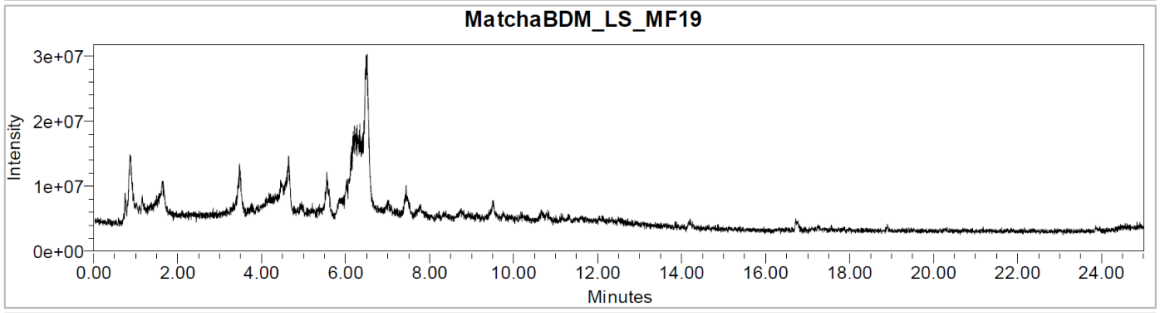

MatchaBDM_LS_MF14
Intensity
1.5e+07
1.0e+07
5.0e+06
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

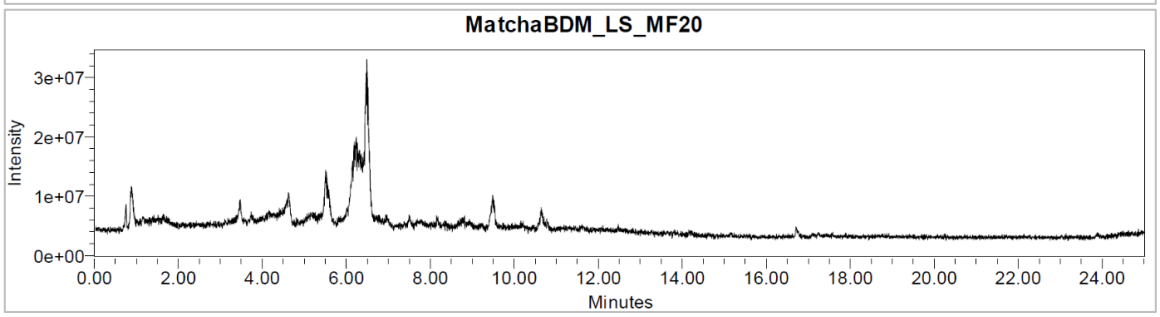

MatchaBDM_LS_MF15
Intensity
1.5e+07
1.0e+07
5.0e+06
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

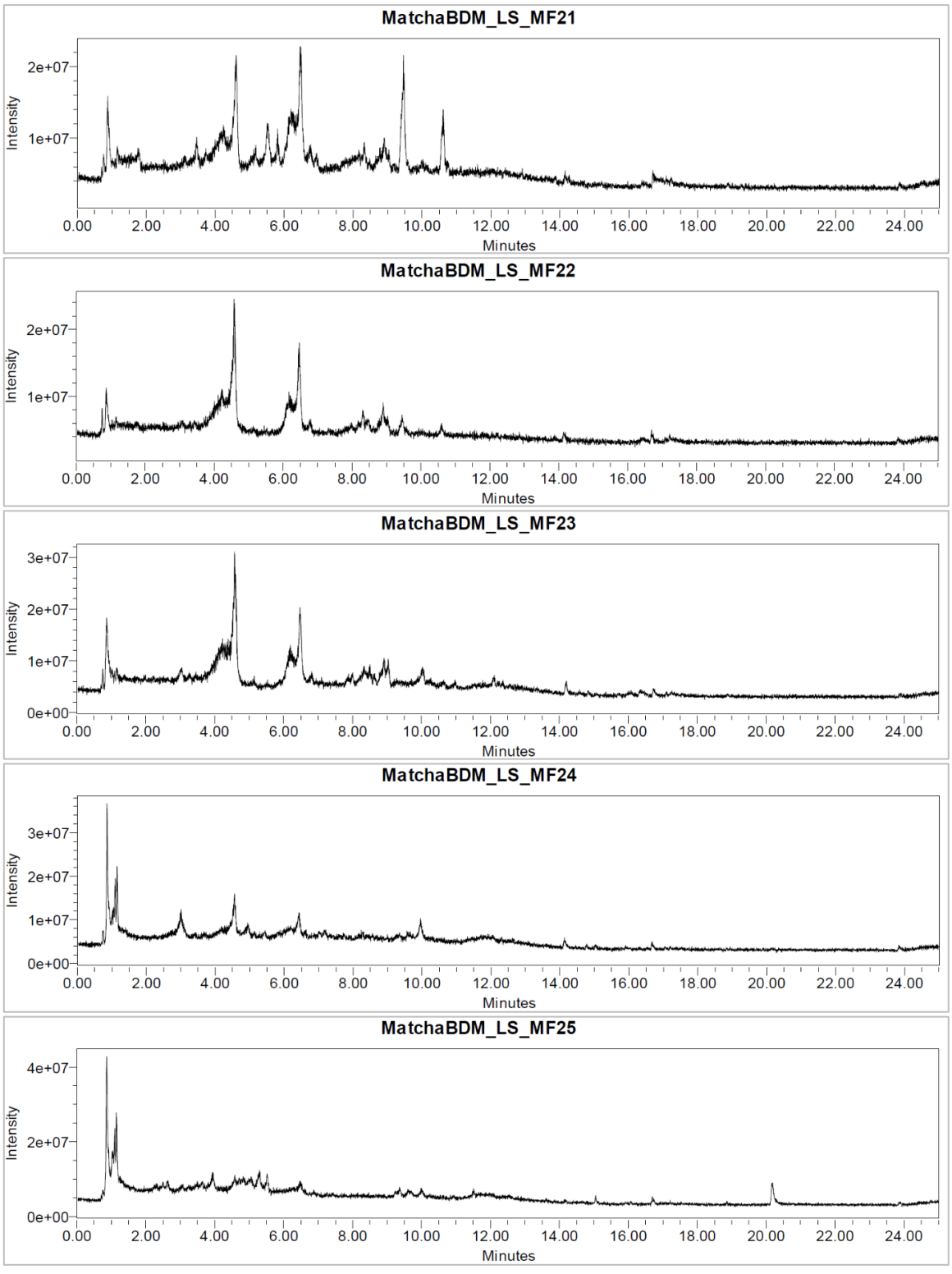

MatchaBDM_LS_MF16
Intensity
1.5e+07
1.0e+07
5.0e+06
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

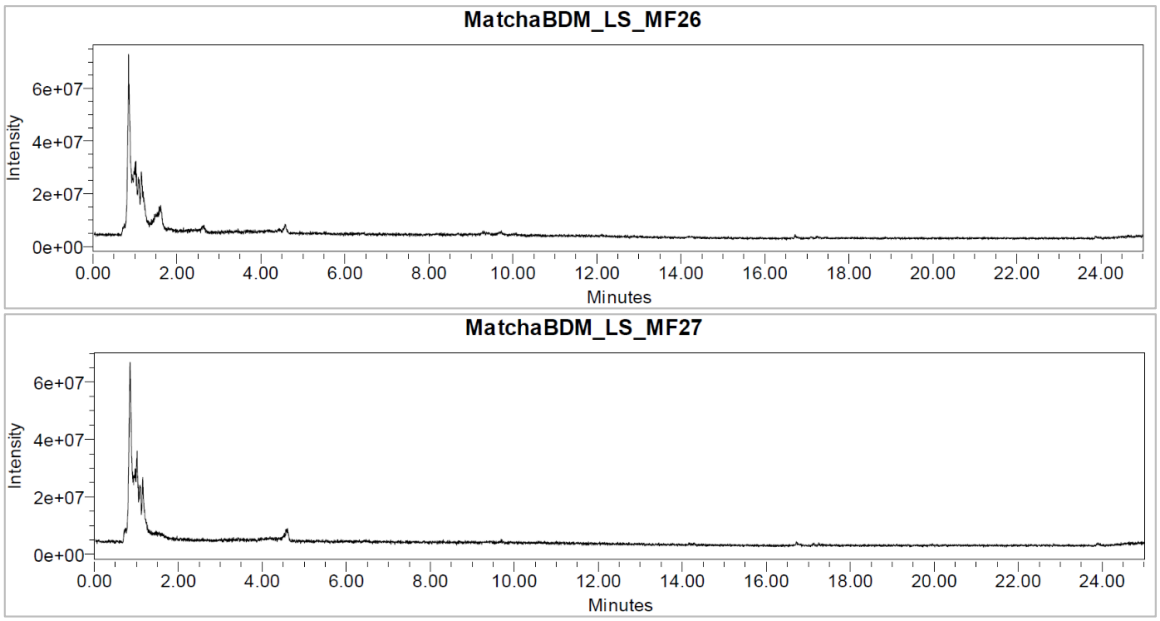

MatchaBDM_LS_MF17
Intensity
2.0e+07
1.5e+07
1.0e+07
5.0e+06
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes


MatchaBDM_LS_MF18
Intensity
1.5e+07
1.0e+07
5.0e+06
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes


MatchaBDM_LS_MF19
Intensity
3e+07
2e+07
1e+07
0e+00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

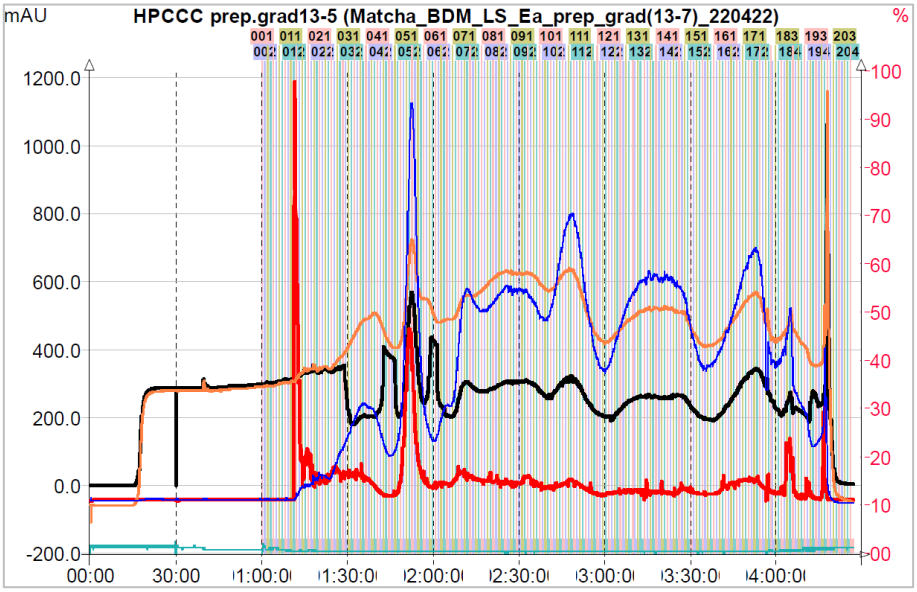

MatchaBDM_LS_MF20
Intensity
3e+07
2e+07
1e+07
0e+00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes


MatchaBDM_LS_MF21
Intensity
2e+07
1e+07
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes
MatchaBDM_LS_MF22
Intensity
2e+07
1e+07
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes
MatchaBDM_LS_MF23
Intensity
3e+07
2e+07
1e+07
0e+00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes
MatchaBDM_LS_MF24
Intensity
3e+07
2e+07
1e+07
0e+00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes
MatchaBDM_LS_MF25
Intensity
4e+07
2e+07
0e+00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

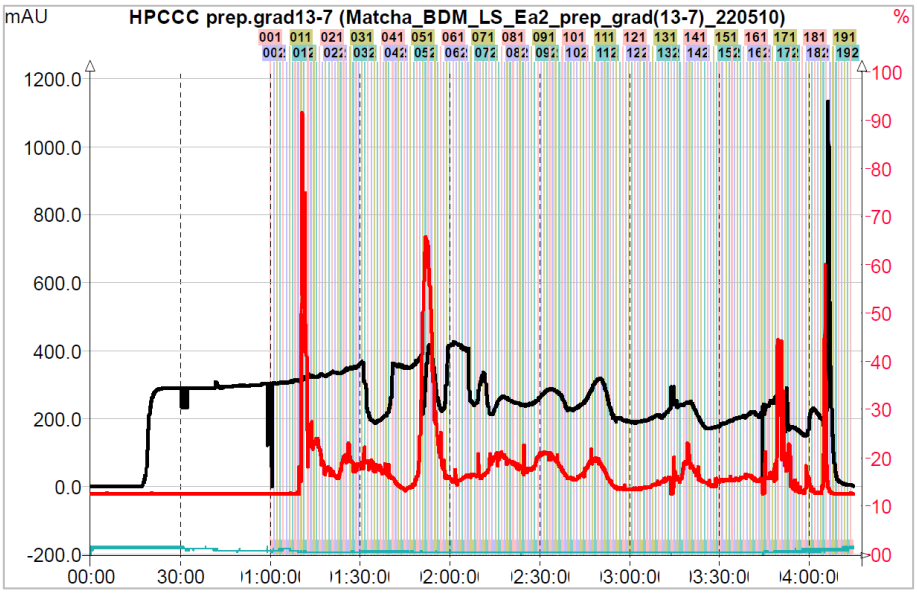


HPCCC runs ELSD chromatograms (red spectrums)

1st run:



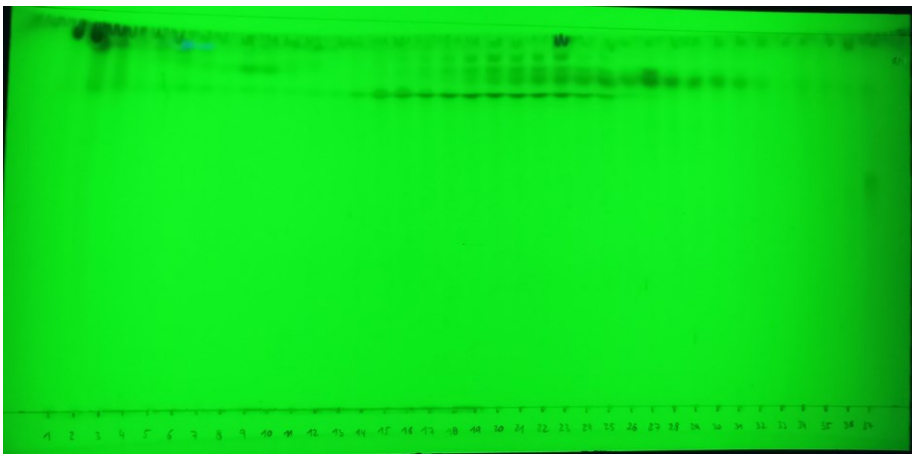
2nd run:

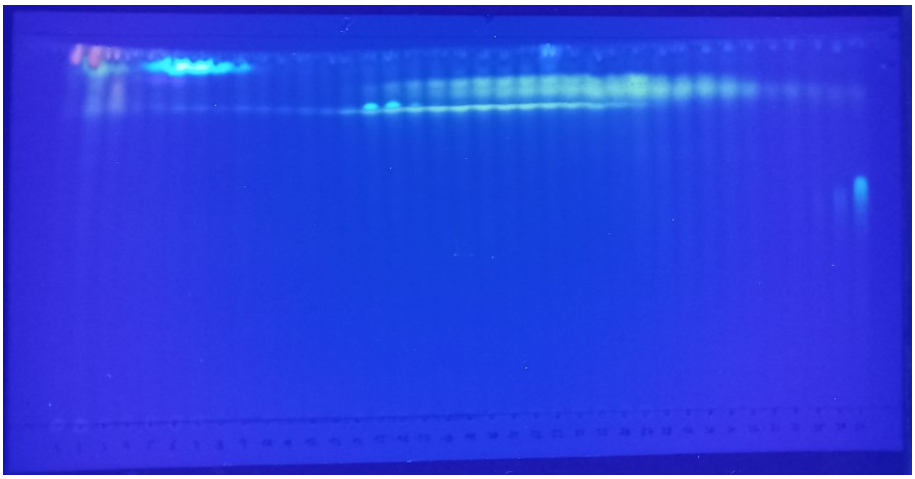


HPCCC runs TLCs

detection at UV_{254} and UV_{366} after spraying with nature substance reagent/PEG400

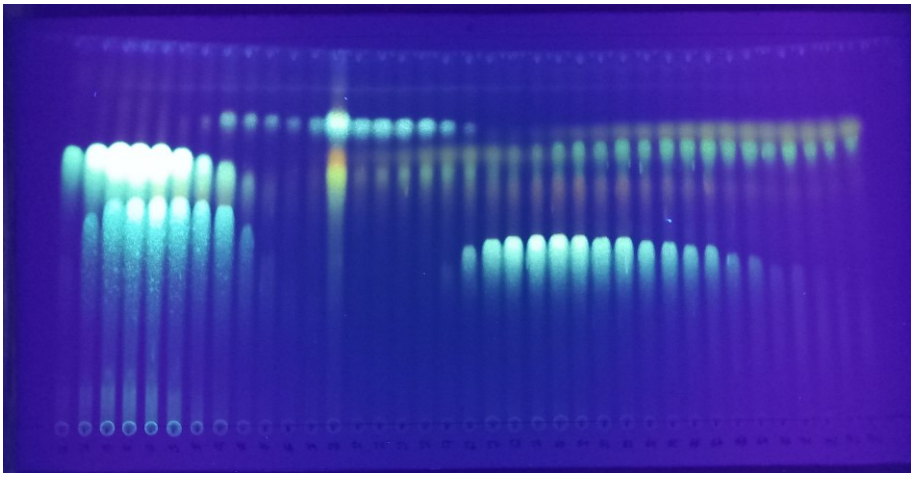
- 1st run, fractions 1-37:



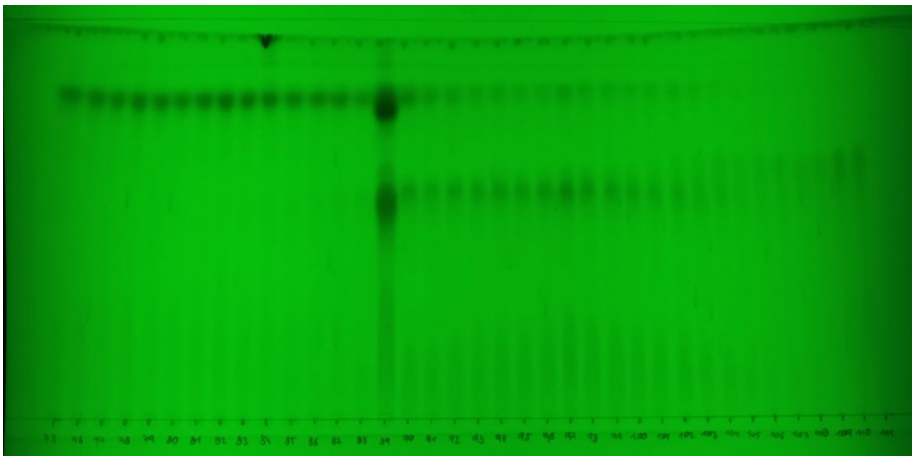


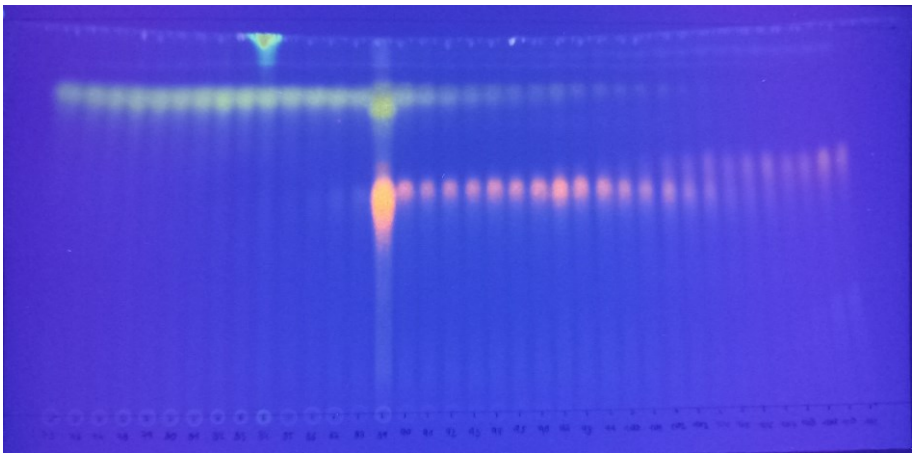
- 1st run, fractions 38-74:

UV_{254} not documented

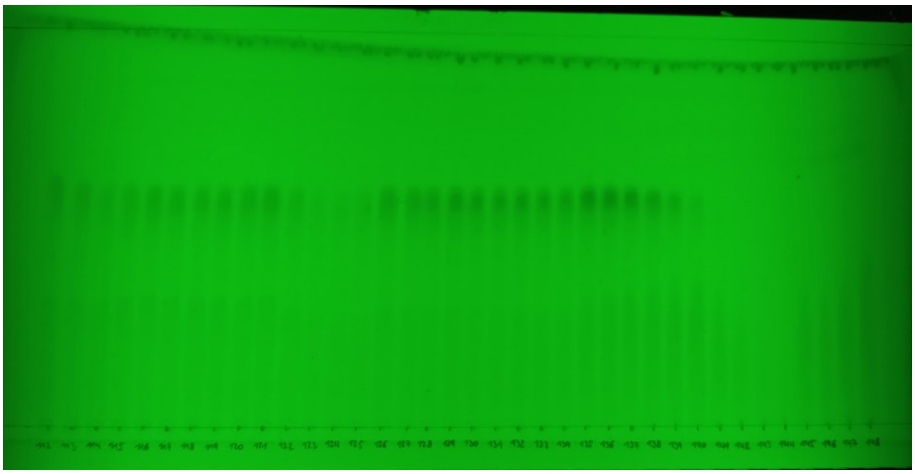


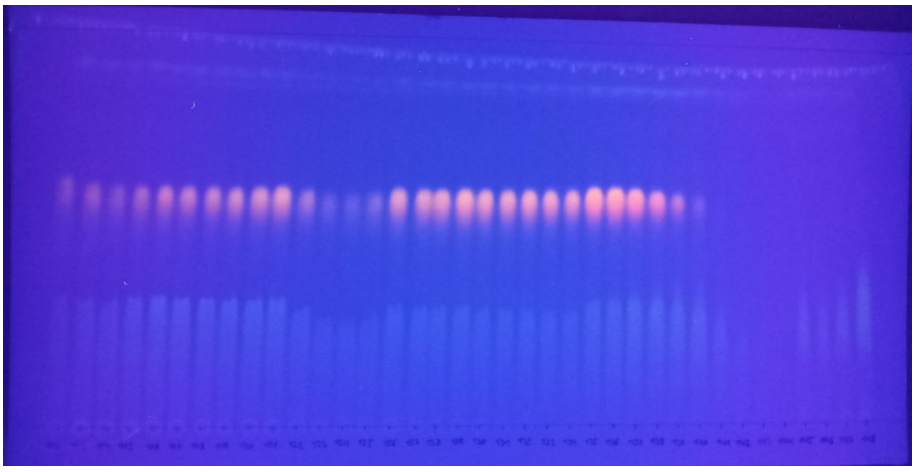
- 1st run, fractions 75-111:



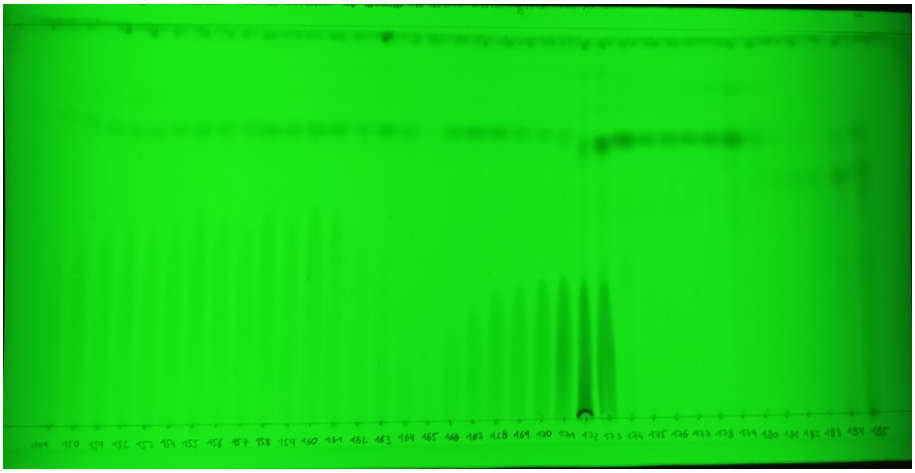


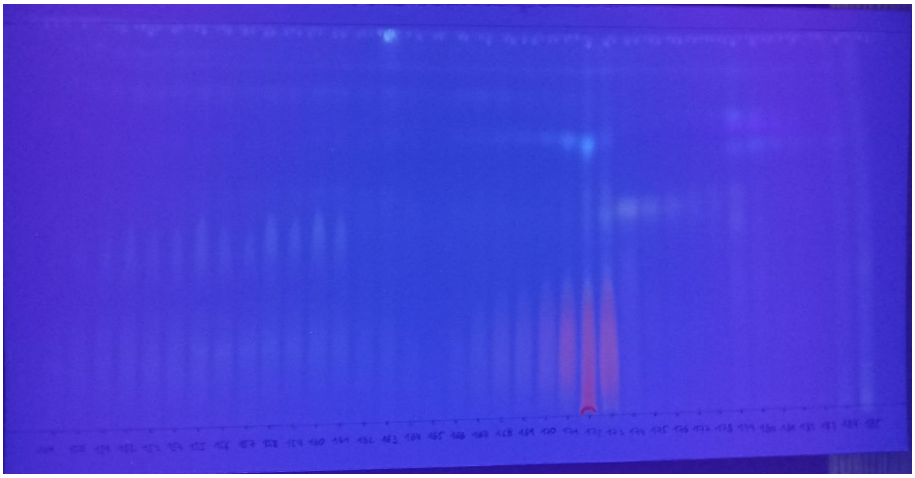
- 1st run, fractions 112-148:



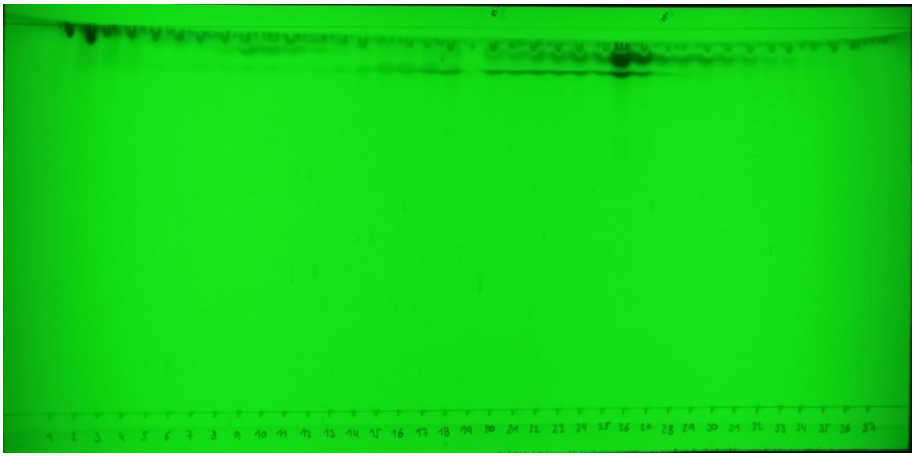


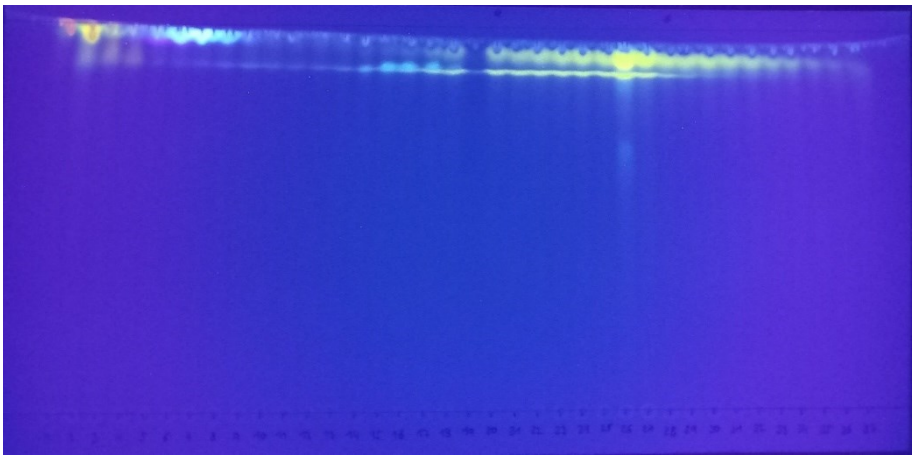
- 1st run, fractions 149-185:



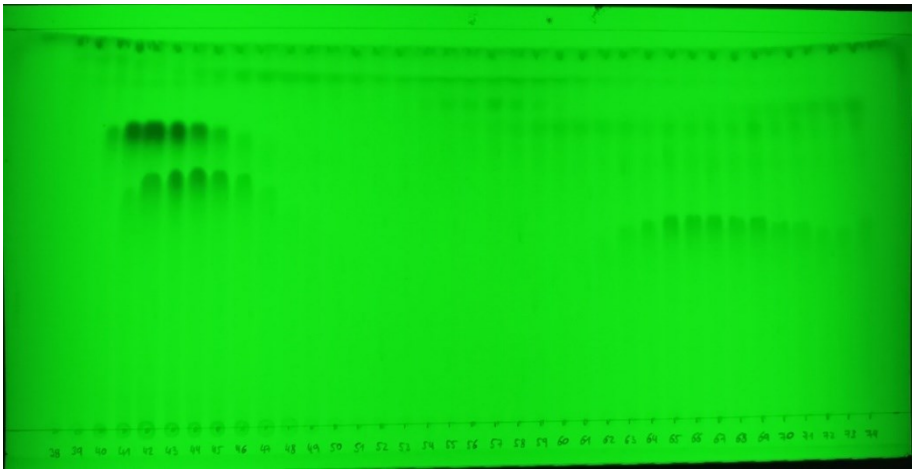


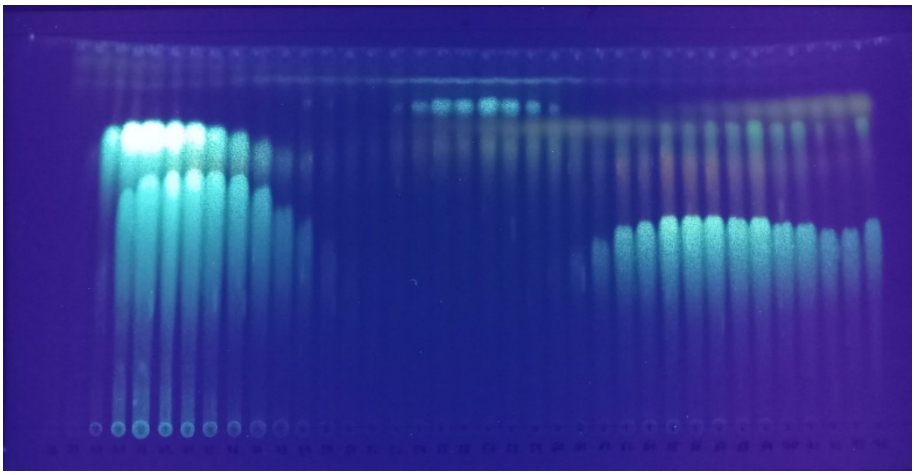
- 2nd run, fractions 1-37:



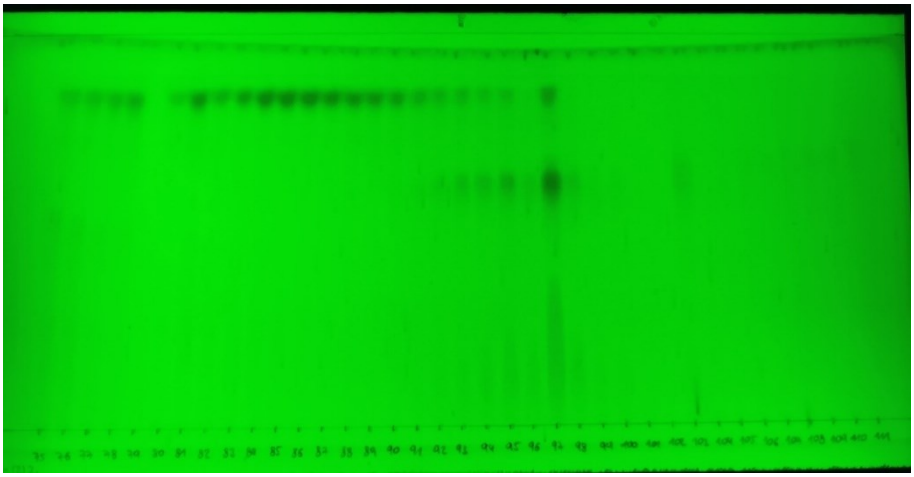


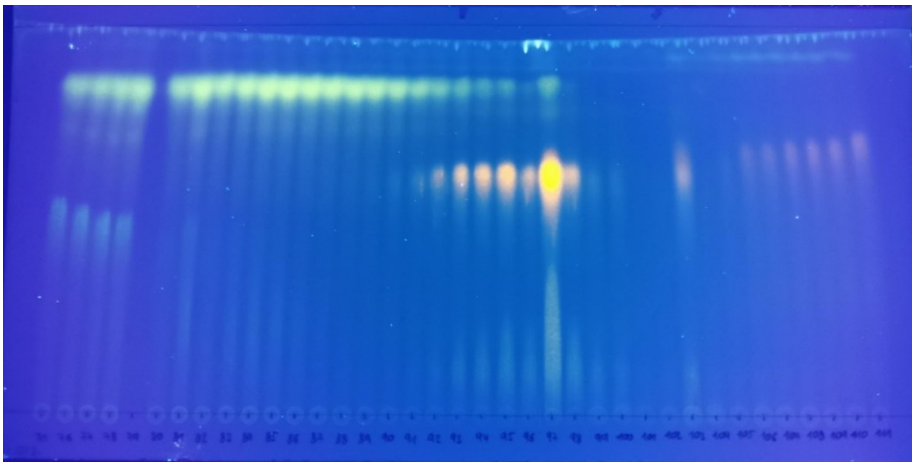
- 2nd run, fractions 38-74:



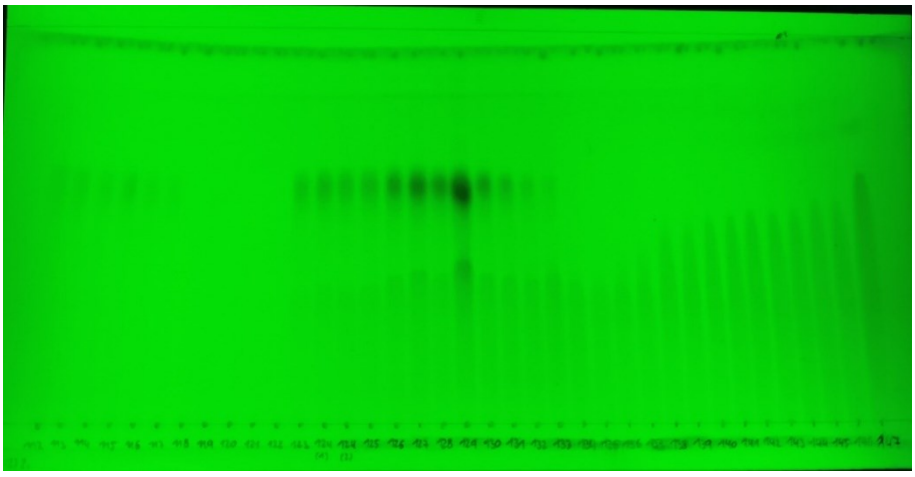


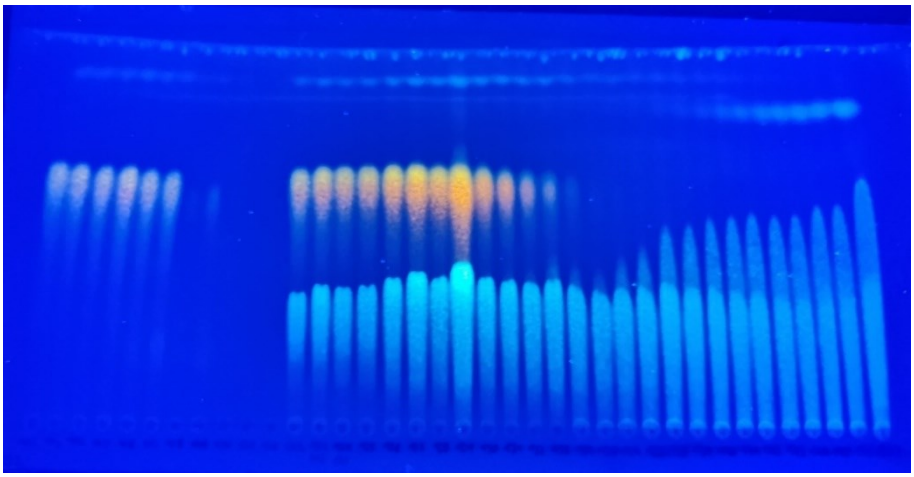
- 2nd run, fractions 75-111:



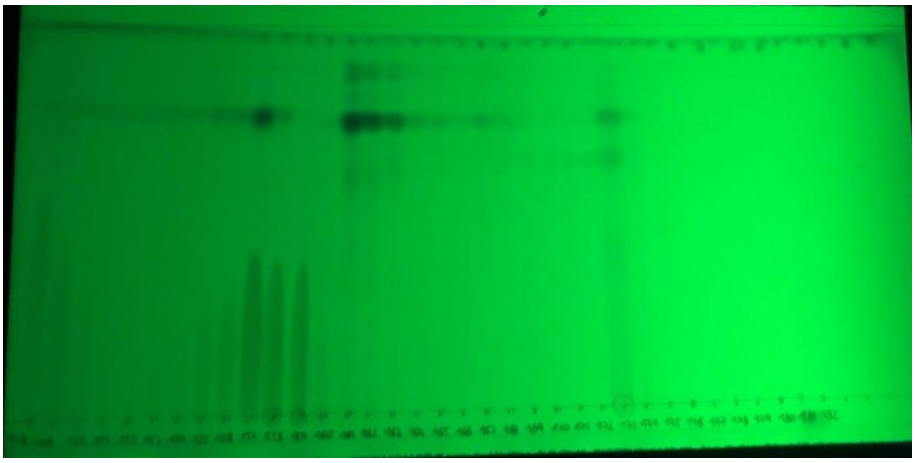


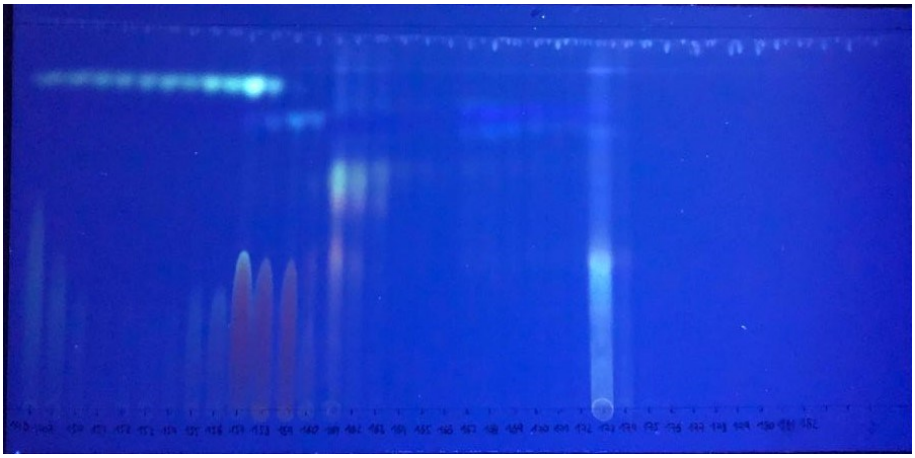
- 2nd run, fractions 112-147:



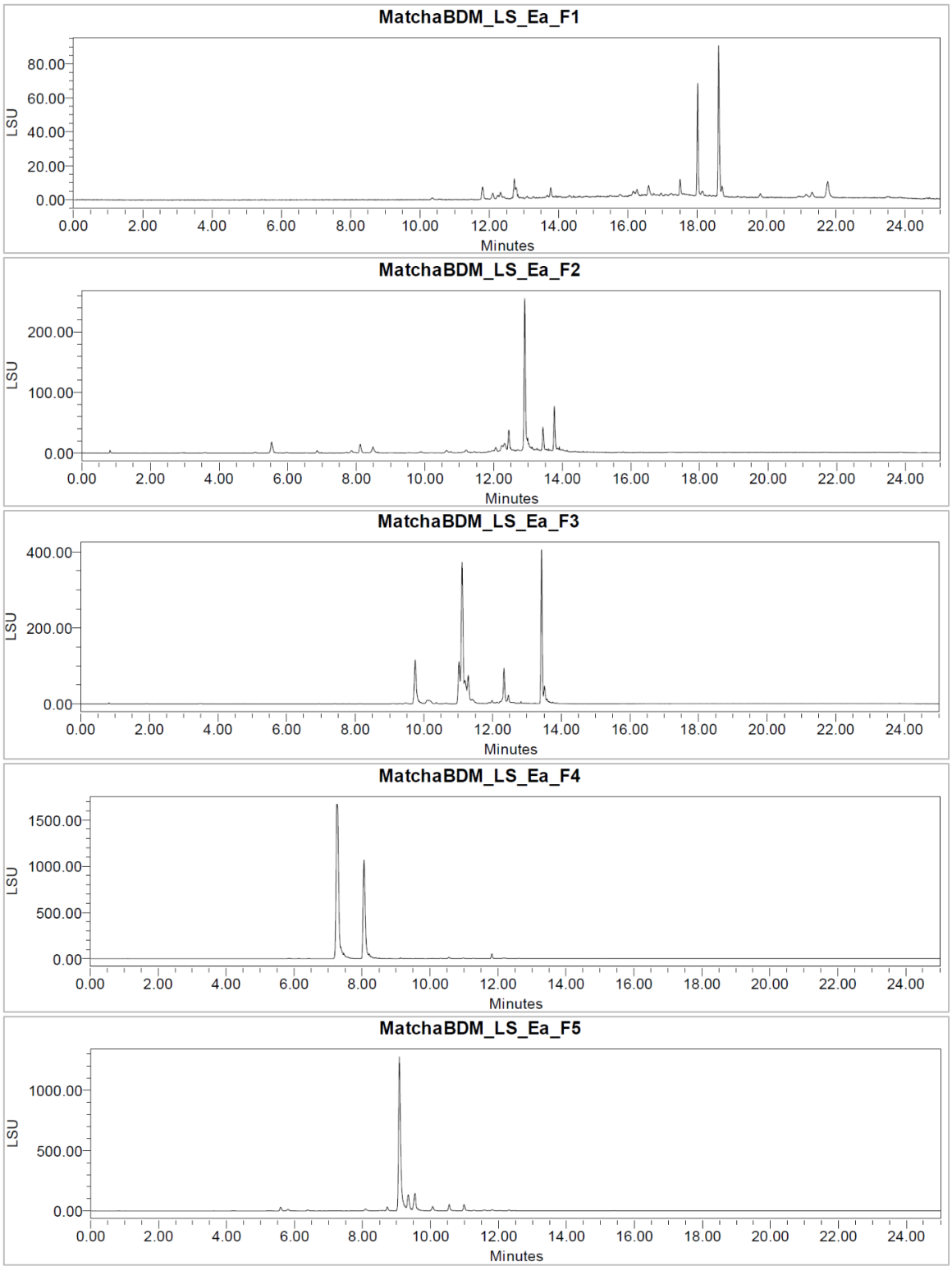


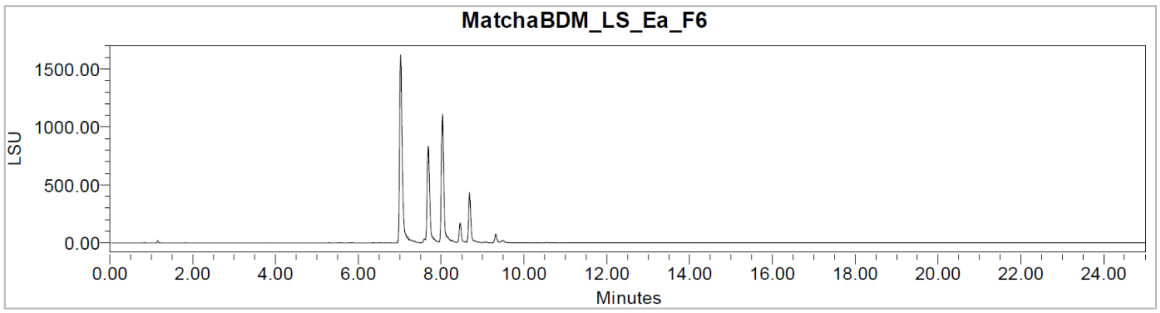
- 2nd run, fractions 148-175:

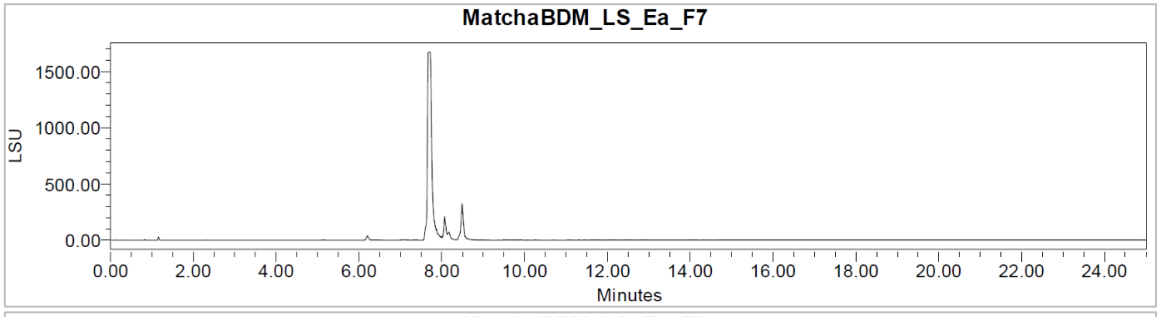


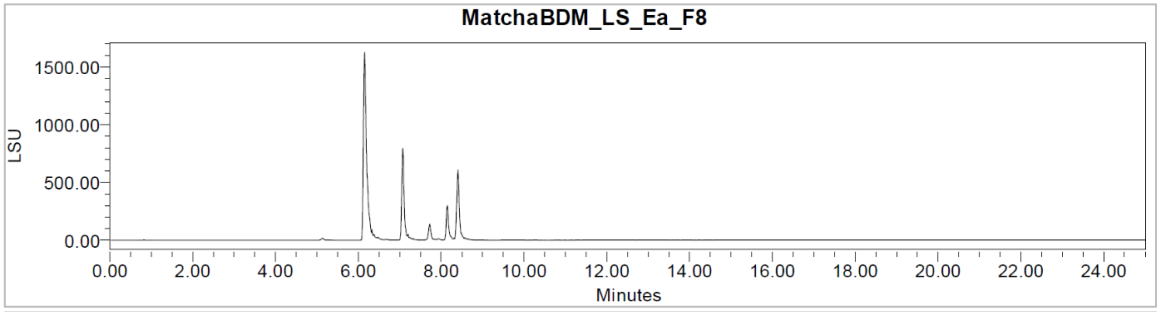


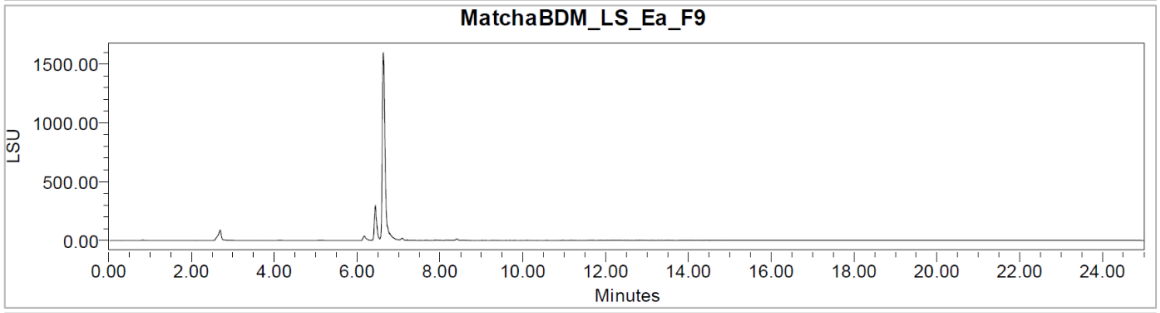
HPCCC fractions (F1-F12) ELSD chromatograms

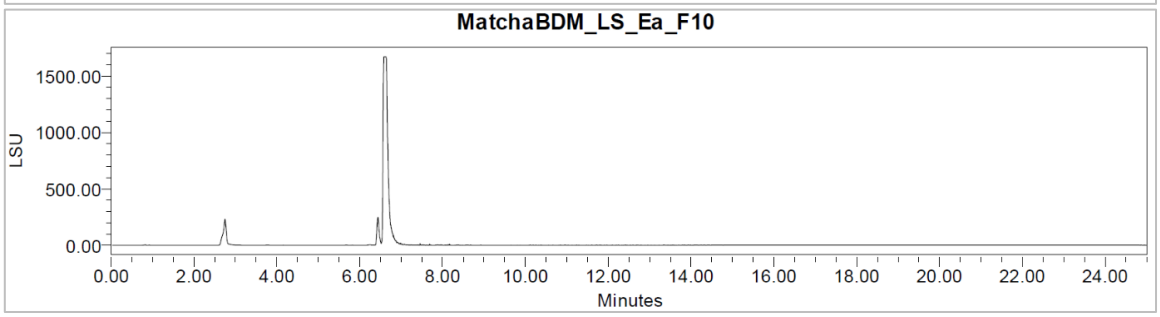


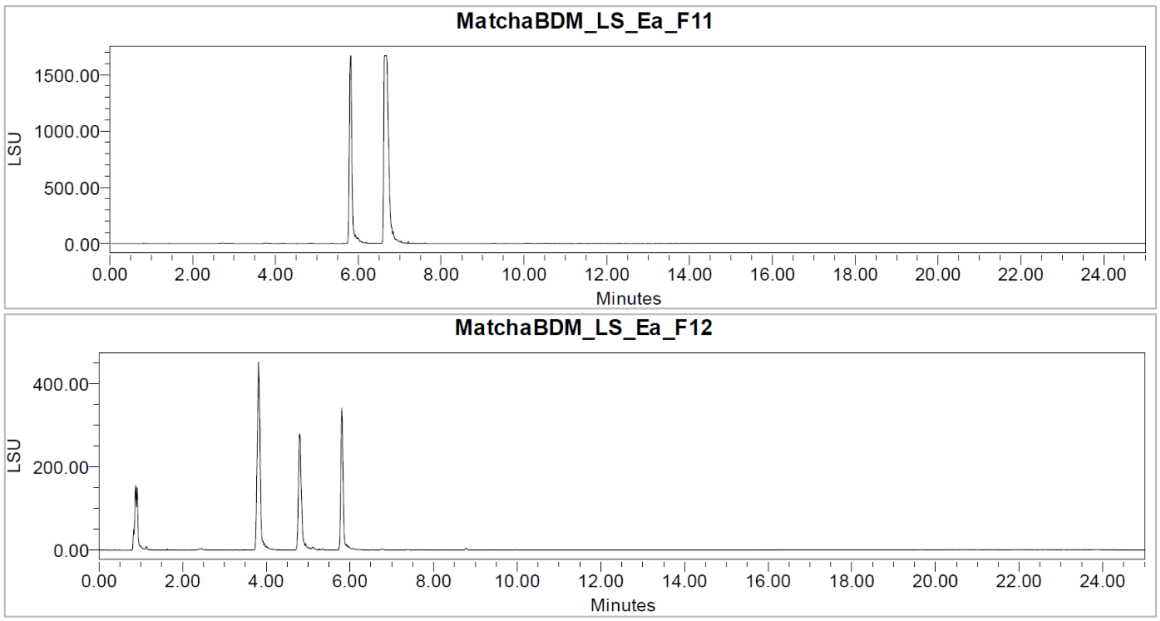

MatchaBDM_LS_Ea_F6
LSU
1500.00
1000.00
500.00
0.00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes


MatchaBDM_LS_Ea_F7
LSU
1500.00
1000.00
500.00
0.00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

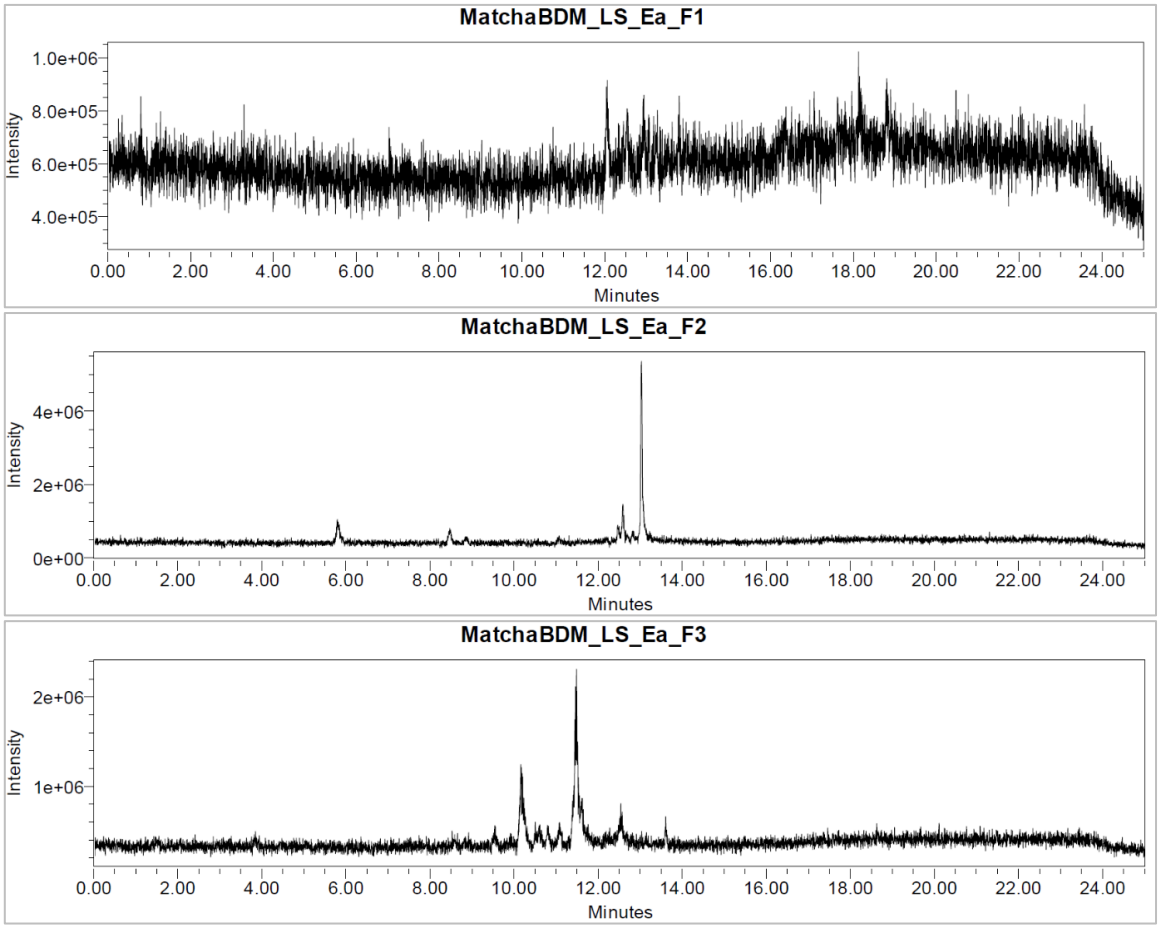

MatchaBDM_LS_Ea_F8
LSU
1500.00
1000.00
500.00
0.00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

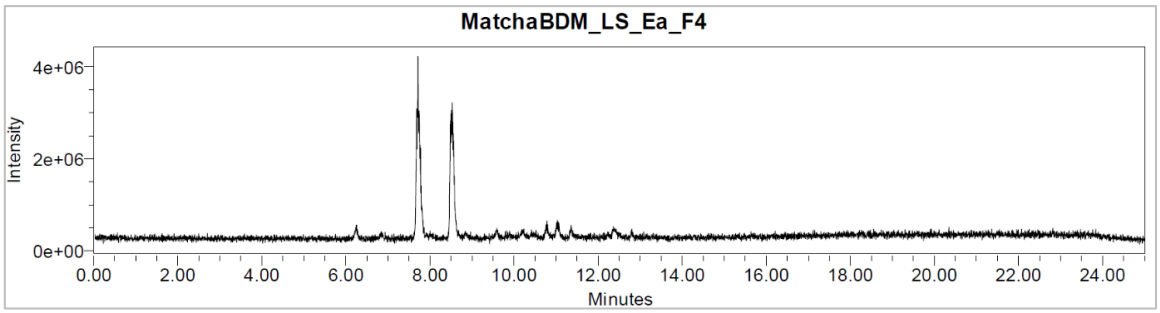

MatchaBDM_LS_Ea_F9
LSU
1500.00
1000.00
500.00
0.00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

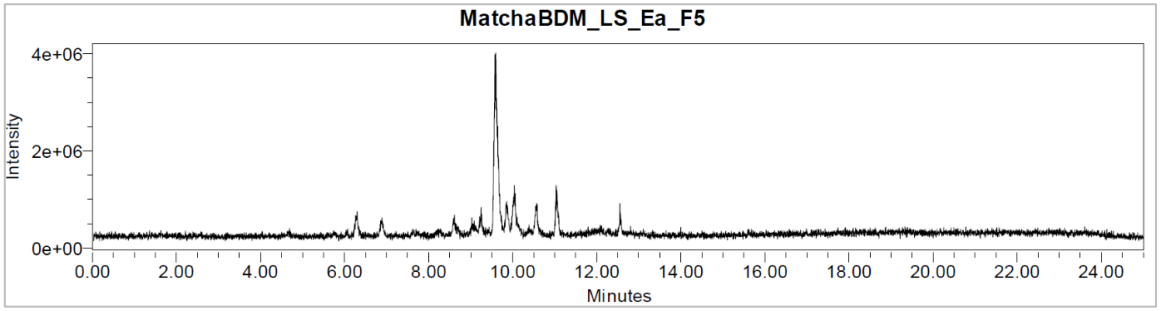

MatchaBDM_LS_Ea_F10
LSU
1500.00
1000.00
500.00
0.00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

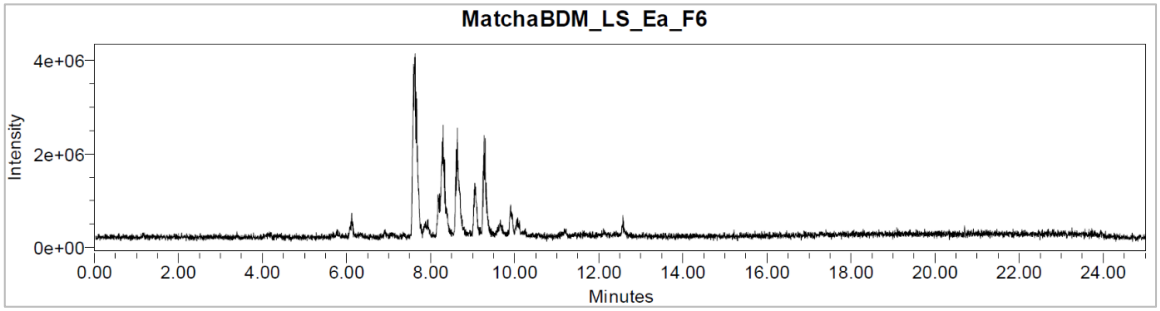


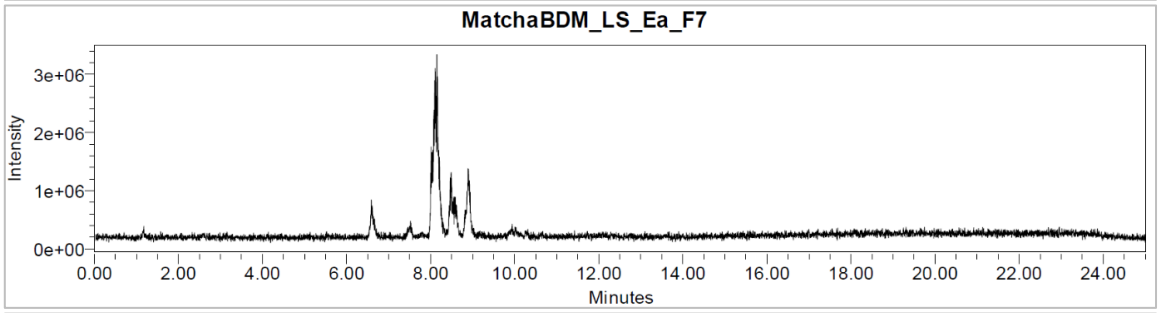
HPCCC fractions (F1-F12) QDa scans (negative mode)

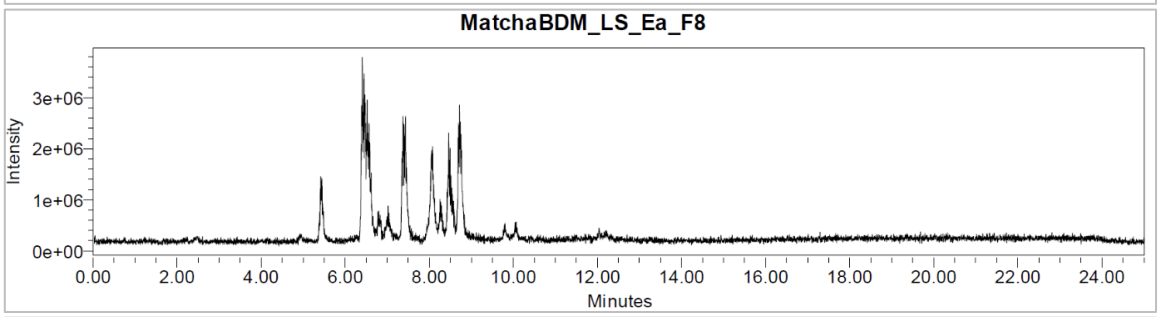


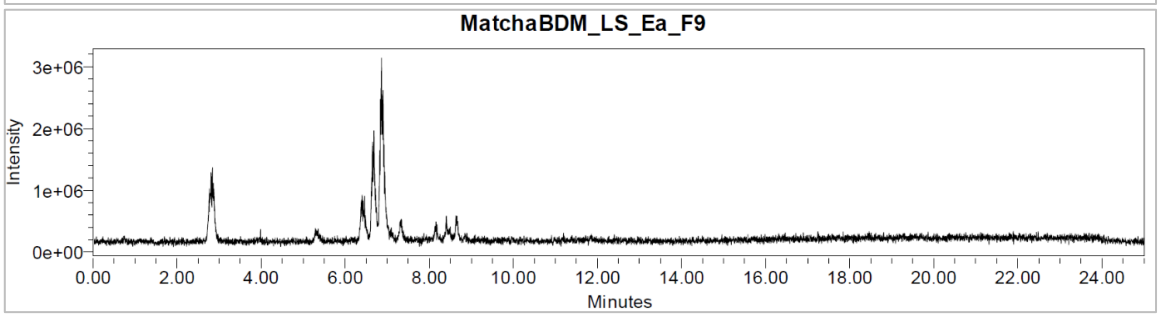

MatchaBDM_LS_Ea_F4
Intensity
4e+06
2e+06
0e+00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes

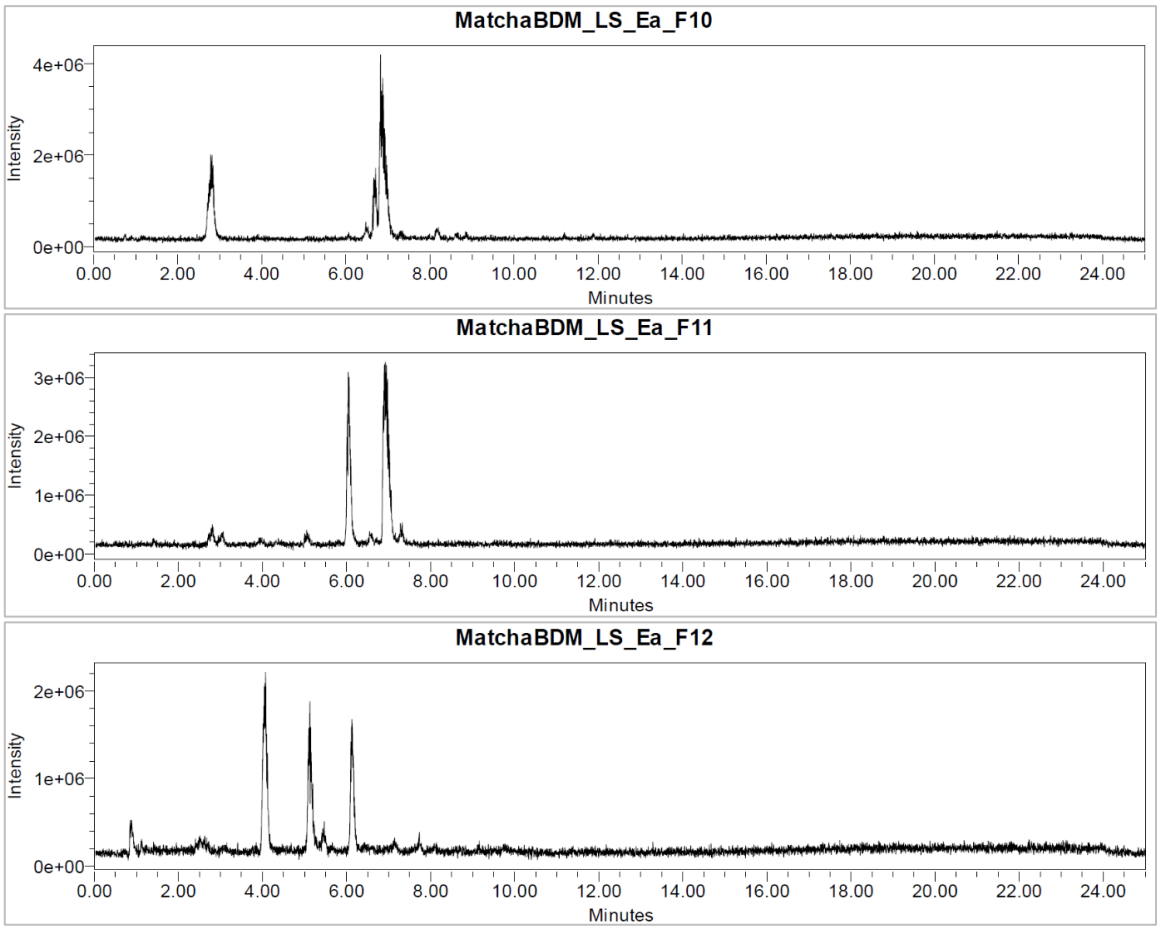

MatchaBDM_LS_Ea_F5
Intensity
4e+06
2e+06
0e+00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes


MatchaBDM_LS_Ea_F6
Intensity
4e+06
2e+06
0e+00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes


MatchaBDM_LS_Ea_F7
Intensity
3e+06
2e+06
1e+06
0e+00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes


MatchaBDM_LS_Ea_F8
Intensity
3e+06
2e+06
1e+06
0e+00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes


MatchaBDM_LS_Ea_F9
Intensity
3e+06
2e+06
1e+06
0e+00
0.00 2.00 4.00 6.00 8.00 10.00 12.00 14.00 16.00 18.00 20.00 22.00 24.00
Minutes



HPCCC fractions (F4-F12) annotation table

F1-F3 not annotated

fraction	RT [min]	m/z negative mode	m/z positive mode	MW [g/mol]	proposed substance(s)
4	7.728	515/353.19		516.4	Dicaffeoylquinic acid
	8.532	515.26/353.17		516.4	Dicaffeoylquinic acid
5	9.599	517.25 [M-H]		~518	Flavonoid glycoside
	9.855	269.18 [M-H]		270.24	Apigenin
	10.045	315.10			Isorhamnetin? 6-Methoxykaempferol?
6	7.643	515.28		516.4	Dicaffeoylquinic acid
	8.205	477.26 [M-H]		~478	Isorhamnetin-7-O-glucoside? Isorhamnetin-3-O-glucoside? Quercetin 3-O-glucuronide?

	8.297	431.27 [M-H]		432.4	Apigenin-7-O-glucoside
	8.664	477.24		~478	Isorhamnetin-7-O-glucoside? Isorhamnetin-3-O-glucoside? Quercetin-3-O-glucuronide?
	9.050	655.34			?
	9.269	655.39			?
7	8.008	477.25 [M-H]		~478	Isorhamnetin-7-O-glucoside? Isorhamnetin-3-O-glucoside? Quercetin 3-O-glucuronide?
	8.127	431.30 [M-H]		432.4	Apigenin-7-O-glucoside
	8.488	477.26 [M-H]		~478	Isorhamnetin-7-O-glucoside? Isorhamnetin-3-O-glucoside? Quercetin-3-O-glucuronide?
	8.896	655.40			?
8	6.498	463.27 [M-H]		464.4	Quercetin-?-O-glycoside
	7.423	641.35			?
	8.090	431.29 [M-H]		432.4	Apigenin-7-O-glucoside
	8.509	461.29			Kaempferol 3-O-glucuronide?
	8.725	507.27 [M-H]		~508	Flavonoid glycoside
9/10	2.824	191.15		354.3	Chlorogenic acid
	6.670	447.29 [M-H]		448.4	Luteolin-7-O-glucoside
	6.884	493.28 [M-H]		494.4	Patuletin-7-O-glucoside
11	6.044-6.072	193.15	379.25 / 177.16 / 195.19		trans-GMCA
	6.985	331.13		332.3	Patuletin
12	0.845	341.25			?
	4.064-4.073	193.17 / 355.18	379.25/ 177.15/ 195.19		cis-GMCA
	5.130	479.28	481.25	480.4	Quercetagetin-?-O-glucoside

	6.134	193.17	379,23/ 177,14/ 195,23		trans-GMCA