

CAFFEIC ACID DERIVATIVES AND FLAVONOIDS FROM APIACEAE AND *SALVIA* SPECIES

Ph.D. thesis

Adila Nazli

Semmelweis University Doctoral College

Pharmaceutical Sciences and Health Technologies Division



Supervisors: Szabolcs Béni, Ph.D.
Imre Boldizsár, Ph.D.

Official reviewers: Károly Mazák, Ph.D.
Sándor Gonda, Ph.D.

Head of the Complex Examination Committee: Zoltán Benyó, Ph.D.

Members of the Complex Examination Committee: Farkas Ágnes, Ph.D.
Gergő Tóth, Ph.D.

Budapest
2026

1. INTRODUCTION

Investigating the phytochemical composition of underexplored medicinal plants holds significant promise for novel drug discovery. A wide range of well-known therapeutic agents, such as aspirin, paclitaxel, digitoxin, cocaine, pilocarpine, codeine, quinine, artemisinin, and silymarin, were isolated from medicinal plants. Hence, investigating the phytochemical composition of underexplored medicinal plants holds significant promise for novel drug discovery.

Anthriscus cerefolium (French parsley), *Anthriscus sylvestris* (cow parsley), *Chaerophyllum bulbosum* (tuberous-rooted chervil), and *Myrrhis odorata* (sweet cicely) are edible species from the Apiaceae family. These species are commonly used as flavoring agents, salad dressings, and seasoning herbs. *Anthriscus caucalis* (burr parsley) and *Anthriscus nitida* (glossy cow parsley) are the Apiaceae species whose culinary uses have not been observed. The phytochemical profiles of all these species are still under explored.

Salvia is the largest genus in the family Lamiaceae while *Salvia* species are usually consumed in the form of infusions or decoctions for health benefits. *Salvia officinalis* L. (sage) is a well-studied Nepetoideae species, widely recognized for its high content of rosmarinic acid. *Salvia officinalis* tea possesses strong antioxidant potential, which improves brain function and prevents depression, dementia, diabetes, and obesity. In contrast, *Salvia nemorosa* (woodland sage), *Salvia pratensis* (meadow sage), *Salvia sclarea* (clary sage), *Salvia verticillata* (lilac sage) has been less extensively studied for their specialized metabolites.

2. OBJECTIVES

The primary aim of this study was the phytochemical characterization of selected, poorly investigated Apiaceae and *Salvia* species, followed by the isolation and structural elucidation of novel bioactive compounds, as well as the evaluation of their bioactivities.

To achieve this, the following objectives were defined for this study:

- 1) Identification of compounds in crude extracts prepared from different organs of selected Apiaceae and *Salvia* species through HPLC-UV-Orbitrap-HR-MS/MS analysis.
- 2) Quantifying the compounds in the organs of selected species collected in different phases of the vegetation cycle by an HPLC-UV-based external calibration method.
- 3) Identifying optimal tissues with relatively high levels of identified compounds, enabling their isolation using a one-step preparative HPLC method.
- 4) Investigating the impact of various treatment/extraction conditions on the metabolite composition of *Salvia* species, considering the consumption of these species in the form of teas and infusions.
- 5) Optimizing the treatment/extraction conditions, enabling the isolation of conversion products.
- 6) Detailed structural characterization of all the isolated compounds through NMR spectroscopy, with particular focus on the differentiation of regioisomers and stereoisomers.
- 7) Analyze the MS/MS fragment ion profiles of compounds to identify diagnostic ions suitable for distinguishing isomeric compounds.
- 8) Evaluate the in vitro antioxidant and cytostatic activities of the isolated compounds to demonstrate their potential for drug development.

3. METHODS

Myrrhis odorata, *Salvia sclarea*, and *S. officinalis* were cultivated in home gardens from plants purchased from commercial horticultural nurseries and were collected during the flowering stage and autumn between 2022 and 2025. Other plant samples were collected from various locations in Hungary during different phases of the vegetation cycle between 2022 and 2025. Plant parts (inflorescence, leaf, and root) were separated, washed, lyophilized, and ground to a fine powder.

For the analysis of naturally occurring metabolites in Apiaceae and *Salvia* species, pulverized plant tissues (20 mg) were extracted with 3 mL of 80% methanol at 60 °C for 30 min. The impact of various extraction conditions was investigated on *Salvia species*. For this purpose, 20 mg of pulverized tissue was suspended in 600 µL distilled water and subjected to three treatments: (1) incubation at room temperature for 5, 15, 30, 45, 60, or 135 min; (2) incubation at room temperature for 15 min, followed by heating at 100 °C for 2 min and further incubation for 120 min; and (3) suspension in hot distilled water, heating at 100 °C for 2 min, and incubation at room temperature for 45 or 60 min.

Reversed-phase (RP) HPLC methods were used for analysis and isolation, utilizing analytical and preparative Dionex Ultimate 3000 HPLC systems. An external standard method was employed for the quantification of compounds using an HPLC-DAD technique. For structure elucidation, the NMR spectra of isolated compounds were recorded using a Bruker Avance III HD 500 (500/150 MHz) spectrometer. Biological activities of isolated compounds were determined by performing the DPPH radical scavenging assay.

4. RESULTS

4.1. Identification of metabolites in methanolic extracts by HPLC-UV-Orbitrap-HR-MS

4.1.1. Identification of metabolites in Apiaceae species

The HPLC-UV analysis of methanolic extracts from selected Apiaceae species showed thirteen peaks corresponding to the compounds of interest. Four compounds were identified as malonyl-glucoside derivatives of quercetin (**1**; C₂₄H₂₂O₁₅), kaempferol or luteolin (**2**, **3**; C₂₄H₂₂O₁₄), and dimalonyl-glucoside of kaempferol or luteolin (**8**; C₂₇H₂₄O₁₇). All these flavonoids were detected in *C. bulbosum*.

Four (**4–7**) compounds possessed the same molecular formula C₂₈H₂₆O₁₅, which aligns with malonyl dicaffeoylquinic acid (MdiCQA) isomers. Compounds **9** and **10** (C₃₁H₂₈O₁₈) represent dimalonyl-dicaffeoylquinic acid (DiMDiCQA) isomers. Compound **4** was present in *A. cerefolium*, *C. bulbosum*, and *A. sylvestris*. Compounds **5** and **6** were detected in *A. sylvestris*. Compounds **7** and **9** were present in *A. cerefolium* and *C. bulbosum*, while compound **10** was detected in *A. caucalis*.

Compound **11** (C₂₅H₂₀O₁₃) was identified as dicaffeoylcaucic acid (DiCDA), while compounds **12** and **13** (C₂₆H₂₂O₁₃) were identified as feruloyl-caffeoylcaucic acid (FCDA) isomers. Compounds **11–13** were present in *M. Odorata*, while compound **11** was also detected in *A. nitida*.

4.1.2. Identification of metabolites in *Salvia* species

The HPLC-UV analysis of methanolic extracts from selected *Salvia* species showed four major peaks, indicating the presence of main metabolites. Compound **16** (C₁₈H₁₆O₈) was identified as rosmarinic acid which was detected in all selected *Salvia* species. Compound **14** (C₂₄H₂₆O₁₃) contained an additional C₆H₁₀O₅ subunit relative to rosmarinic acid, suggesting the

presence of a glucosyl unit. Rosmarinic acid glucoside was detected in *S. verticillata*, *S. pratensis*, *S. nemorosa*, and *S. officinalis*.

Compounds **15** and **17** possess the same molecular formula $C_{27}H_{24}O_{13}$, suggesting the presence of salvianolic acid K and its isomer. These compounds were classified as caffeoyl-dihydroxyphenyllactic acids (CDHPLAs). Compound **15** was present in *S. pratensis*, *S. sclarea*, and *S. nemorosa*, while compound **17** was detected in *S. verticillata*.

4.2. Identification of conversion products in aqueous-treated *Salvia* species extracts by HPLC-UV-Orbitrap-HR-MS

Upon aqueous treatment of pulverized tissues, all identified metabolites were transformed by the theoretical loss of a formic acid moiety through oxidative decarboxylation involving the concomitant removal of C-8' carboxyl group and C-7' hydrogen from parent compounds.

Rosmarinic acid (**16**) was transformed into two isomeric products (**18**, **21**) with molecular formula $C_{17}H_{14}O_6$, supporting their identification as nepetoidins A and B. The transformation of rosmarinic acid glucoside (**14**) may involve loss of the glucose moiety, yielding rosmarinic acid. The conversion products (**19**, **20**), generated from the isomeric compounds **15** and **17**, showed a similar molecular formula, $C_{26}H_{22}O_{11}$, and were named as salvianolic acid P and salvianolic acid Q.

The metabolites remained intact when the powdered tissues were directly incubated in hot water at 100 °C. This observation suggests that the aqueous environment activates endogenous enzymes responsible for metabolite conversion, while heat treatment inactivates these enzymes and prevents the transformation of metabolites.

4.3. Structural characterization of isolated compounds through NMR analysis

The structural elucidation of isolated natural compounds (**1–18**, **19**, **21**) was based on 1D and 2D NMR experiments.

4.3.1. Structural characterization of malonyl dicaffeoylquinic acids (**4–7**, **9**, **10**)

The ^1H NMR spectrum of compounds **4–7** showed six aromatic proton signals, which appeared as two ABX systems. Moreover, four trans-oriented protons formed two vinylene groups. The HMBC correlations confirmed that these subunits form two caffeoyl groups. A pair of sp^3 methylene protons was correlated with two carbonyl carbons, thus forming a malonyl moiety. Four sp^3 methylene protons and three oxymethine protons were considered a part of the core structure, which was identified as quinic acid in compounds **4**, **5**, **7**, and *epi*-quinic acid in compound **6**. A spectroscopic comparison was provided for the differentiation of quinic and *epi*-quinic acid, where H-4 served as a diagnostic proton to decide the relative configuration. A large trans-diaxial coupling constant ($^3J_{4/5} \approx 9$ Hz) in DMSO- d_6 indicated a quinic acid moiety, while a smaller coupling constant ($J < 8$ Hz) pointed to an *epi*-quinic acid moiety. The connecting positions of substituents with the core structure were confirmed based on HMBC correlations, which supported the identification of compounds **4**, **5**, **6**, and **7** as 1,5-dicaffeoyl-3-malonylquinic acid, 3,5-dicaffeoyl-1-malonylquinic acid, 3,5-dicaffeoyl-4-malonyl-*epi*-quinic acid, and 1,5-dicaffeoyl-4-malonylquinic acid, respectively. Compounds **9** and **10**, possessing an additional malonyl subunit, were identified as 1,5-dicaffeoyl-3,4-dimalonylquinic acid and 3,5-dicaffeoyl-1,4-dimalonylquinic acid, respectively. All these compounds (**4–7**, **9**, **10**) were considered novel compounds.

4.3.2. Structural characterization of caffeoyl daucic acids (11–13)

The NMR spectra of compound **11** confirmed the presence of two caffeoyl groups; however, it possessed a different core structure. This structural unit was based on four protons, which were attached to three oxygenated sp^3 methine carbons and a sp^2 methine carbon, supporting the presence of daucic acid moiety. The connecting positions of caffeoyl groups were confirmed based on HMBC correlations, which indicated that compound **11** is 4,5-di-*O*-(*E*)-caffeoyl daucic acid. Compounds **12** and **13**, possessing an additional methyl group, were identified as 4-*O*-(*E*)-caffeoyl-5-*O*-(*E*)-feruloyl daucic acid and 4-*O*-(*E*)-feruloyl-5-*O*-(*E*)-caffeoyl daucic acid. Compound **12** represents a novel natural product.

4.3.3. Structural characterization of caffeoyl-dihydroxyphenyllactic acids (14–17) and conversion products (18, 19, 21)

The ^1H NMR spectrum of compound **17** showed nine proton resonances in the form of three ABX systems. One ABX system was correlated with $CH_2-CH(OR)-COOH$, corresponding to a 3,4-dihydroxyphenyllactic acid (DHPLA) subunit. The next ABX system was coupled with $CH=CH-COOR$, forming a caffeic acid moiety. The third ABX system was correlated with $CH(OH)-CH(OR)-COOH$, yielding a (3,4-dihydroxyphenyl)glyceric acid residue. The α -oxymethine proton of DHPLA was correlated with the carbonyl carbon of caffeic acid, thus confirming the esterification between these two subunits. In the $^1H-^1H$ NOESY spectrum, α -glyceric acid proton was coupled with H-2 of caffeic acid, indicating its substitution with C-3 hydroxyl of caffeic acid. Compound **17** represents a novel natural product, which was named as salvianolic acid O. In the case of isomeric compound **15**, C-4 hydroxyl of caffeic acid is connected with (3,4-dihydroxyphenyl)glyceric acid moiety, supporting its identification as

salvianolic acid K. Furthermore, compounds **14** and **16** were verified as rosmarinic acid-3-*O*- β -D-glucoside, and rosmarinic acid.

The NMR data of conversion product **19** is in close agreement with its parent salvianolic acid K (**15**), apart from positions 7' and 8', which contained vinylic protons. Conversion product **19** is a novel derivative and was named as salvianolic acid P. The isomeric conversion products of rosmarinic acid (**18**, **21**) also possessed olefinic protons at positions 7' and 8', supporting their identification as nepetoidin A and nepetoidin B, respectively.

4.3.4. Structural characterization of flavonoid-malonyl-glucosides (1–3, 8)

Flavonoid malonyl glucosides (FMGs) **1**, **2**, and **3** were identified as quercetin-3-*O*-(6"-*O*-malonyl)- β -D-glucoside, luteolin-7-*O*-(6"-*O*-malonyl)- β -D-glucoside, and kaempferol-3-*O*-(6"-*O*-malonyl)- β -D-glucoside, based on the comparison of their NMR data reported in earlier studies. Compound **8** is esterified with two malonyl units, which were substituted with a glucoside moiety, supporting its identification as luteolin-7-*O*-(2",6"-di-*O*-malonyl)- β -D-glucoside, which represents a novel flavonoid.

4.4. Quantification of metabolites in selected Apiaceae species

Compounds **1** and **3** were mainly accumulated in inflorescences of *C. bulbosum*, i.e., 25.8 mg/g and 9.6 mg/g, while compounds **2** and **8** were mainly present in leaves of *C. bulbosum*, i.e., 10.6 mg/g and 2.4 mg/g, respectively. The inflorescences of *A. cerefolium* and *C. bulbosum* were found to be the most concentrated source of compounds **4** and **7**, with 28.6 mg/g and 4.6 mg/g in *A. cerefolium*, while 20.2 mg/g and 4.5 mg/g in *C. bulbosum*. The inflorescences of *C. bulbosum* also possessed the highest concentration of compound **9**, i.e., 30.9 mg/g. The highest levels of compounds **5** and **6** were observed in leaves of *A. sylvestris*, i.e., 2.1 and 1.9

mg/g, respectively. DiMDiCQA (**10**) was accumulated in the inflorescences of *A. caucalis* up to 3.24 mg/g. The highest amounts of DiCDA (**11**) were detected in the roots of *M. odorata* collected at the PVP, attaining up to 9.42 mg/g, while the same sample was the only source of FCDAs (**12**, **13**) with 0.75 mg/g and 1.36 mg/g concentrations, respectively.

4.5. Quantification of metabolites and their conversion products in variously prepared extracts of *Salvia* species

4.5.1. Enzymatic conversion of rosmarinic acid into nepetoidins A and B

The untreated *S. officinalis* leaf contained rosmarinic acid (**16**) up to 27.4 mg/g (sample 1). Rosmarinic acid was rapidly decomposed during room-temperature aqueous incubation of pulverized tissues. Samples incubated in water at room temperature for 5, 15, 30, 45, and 60 min showed a significant decrease in rosmarinic acid content as incubation time increased: 1.97 mg/g and 0.15 mg/g rosmarinic acid content were detected after 15 and 60 min of aqueous incubation, respectively. On the other hand, thermal pretreatment at 100 °C for 2 min, followed by a 60-min aqueous incubation at room temperature, resulted in a rosmarinic acid content comparable to that observed in untreated samples (29.3 mg/g). While untreated leaves contained negligible amounts of nepetoidin A (**18**) and B (**21**) (sample 2; 0.11 mg/g and 0.08 mg/g, respectively), a room temperature water incubation elevated these values to 0.90 and 3.93 mg/g (30 min incubation). Longer incubation periods caused a steady decline in the levels of both compounds.

4.5.2. Enzymatic Conversion of salvianolic acid K into salvianolic acid P

Salvianolic acid K (**15**) was determined in the root samples of *S. pratensis*, *S. nemorosa*, and *S. sclarea*, ranging from 40.0–43.2, 14.3–17.5, and 5.19–6.37 mg/g, respectively. Upon room-temperature aqueous treatment of root tissues, salvianolic acid K converted into salvianolic acid P (**19**). The highest

amount of salvianolic acid P was determined after 5 min incubation, with concentrations of 0.55 mg/g in *S. sclarea* (sample 2), 0.45 mg/g in *S. nemorosa* (sample 3), and 0.42 mg/g in *S. pratensis* (sample 1). Prolonged aqueous incubations resulted in a decrease in the levels of salvianolic acid P. In the case of heat-treated extracts (100°C, DW45), salvianolic acid K remained intact with levels comparable to those noticed in untreated (UT) samples.

4.5.3. Enzymatic Conversion of salvianolic acid O into salvianolic acid Q

Salvianolic acid O (**17**) accumulated in the roots of *S. verticillata* at concentrations ranging from 20.7–78.6 mg/g. During aqueous incubations of root samples at room temperature, salvianolic acid O converted to salvianolic acid Q (**20**). The maximum concentration of salvianolic acid Q was observed after a 5 min incubation, reaching up to 0.42 mg/g (sample 8). Prolonged aqueous treatments resulted in a decrease in salvianolic acid Q content. Similar to rosmarinic acid and salvianolic acid K, salvianolic acid O also remained intact during room-temperature aqueous incubations when the samples were subjected to a prior thermal treatment at 100 °C for 2 min.

4.6. Evaluation of biological activities of isolated compounds

4.6.1. Determination of antioxidant potential

The DPPH inhibition potential of CDHPLAs (**14–21**) with IC_{50} 8.9–15.1 μ M was quite similar to that of MdiCQAs (**4–7**, **9**, **10**) with IC_{50} 10.1–13.6 μ M. However, CDAs (**11–13**) showed relatively lower antioxidant potential with IC_{50} values of 19.7–26.9 μ M. Among FMGIs, compounds **1**, **2**, and **8** showed noticeable antioxidant potential with IC_{50} 7.8–22.5 μ M. However, FMGI **3** exhibited the least activity with IC_{50} 233.9 μ M.

4.6.2. Determination of cytostatic potential

Among MDiCQAs (**4–7**, **9**) and FMGls (**1–3**, **8**), compounds **3**, **4**, and **9** showed moderate non-selective cytostatic activity towards normal Vero E6 cells and cancerous HT-29, U87, HepG2, HL-60, and A2058 cells. The antiproliferative effects of MDiCQAs **10** and CDAs **11–13** remain to be elucidated in future studies.

Among CDPLAs, the conversion products, namely salvianolic acid P (**19**), nepetoidin A (**18**), and nepetoidin B (**21**), which lack a carboxyl group at the C8' position, showed better cytostatic activity towards the cancer cell lines (HT-29, U87, U251, H838, MM6) than the parent compounds rosmarinic acid (**16**), rosmarinic acid glucoside (**14**), salvianolic acid K (**15**), and salvianolic acid O (**17**) containing a carboxyl group at C8'.

5. CONCLUSIONS

- 1) The present dissertation explored the phytochemical profile of various tissues from six Apiaceae species: *A. cerefolium*, *C. bulbosum*, *A. sylvestris*, *A. caucalis*, *A. nitida*, and *M. odorata*, and from five *Salvia* species: *S. officinalis*, *S. nemorosa*, *S. pratensis*, *S. sclarea*, and *S. verticillata*.
- 2) Metabolites were initially identified in different organs of the analyzed plant species through HPLC-UV-Orbitrap-HR-MS and quantified through an external calibration method. Based on this analysis, optimal tissue and treatment conditions were determined, enabling the isolation of selected compounds via a single-step preparative HPLC method.
- 3) Thirteen compounds were isolated from Apiaceae species, which were classified as malonyl dicaffeoylquinic acids (MdiCQAs) (**4–7**, **9**, **10**), caffeoyl daucic acids (CDAs) (**11–13**), and flavonoid-malonyl-glucosides (FMGs) (**1–3**, **8**).
- 4) Seven compounds (**14–17**) were isolated from *Salvia* species, belonging to the caffeoyl-dihydroxyphenyllactic acids (CDHPLAs) and related derivatives (**18**, **19**, **21**).
- 5) The rapid degradation of rosmarinic acid, accompanied by the transient formation of nepetoidins A and B, followed by their subsequent decomposition, was demonstrated to be a common enzymatic process in *Salvia* species. Furthermore, salvianolic acid K, detected for the first time in the roots of *S. nemorosa*, *S. pratensis*, and *S. sclarea*, and salvianolic acid O, identified as a previously undescribed natural product in the roots of *S. verticillata*, underwent analogous enzymatic transformations in pulverized root tissues. These reactions resulted in the formation of two newly identified compounds, salvianolic acids P and Q, respectively, which were subsequently degraded.

- 6) The structures of all the isolated compounds were confirmed through 1D and 2D NMR experiments, which enabled complete ^1H and ^{13}C assignments for these compounds.
- 7) MS/MS fragment ion profiles of isomeric compounds were evaluated at different CID energies, and energy-dependent differences in diagnostic fragment ion intensities enabled isomer differentiation.
- 8) Among the twenty isolated compounds, ten compounds were considered as novel derivatives, including six MdiCQAs (**4–7**, **9**, **10**), a feruloyl CDA (**12**), a FMGI (**8**), a CDHPLA (**17**), and a related derivative (**19**).
- 9) Both the structure-specific radical scavenging capacity and the cytostatic activities of these compounds were confirmed.

6. BIBLIOGRAPHY OF THE CANDIDATE'S PUBLICATIONS

6.1. Publications related to the thesis

Nazli A, Gáborová M, Ausbuttel T, Stipsicz B, Toth G, Bősze S, Béni S, Boldizsár I. Malonyl-caffeoylquinic acids and malonyl-flavonoid-glucosides from three edible Apiaceae plants *Anthriscus cerefolium*, *Anthriscus sylvestris*, and *Chaerophyllum bulbosum*. **ACS Omega**. 2025;10(38):43656–68.

Nazli A, Szögi-Tatár B, Bősze S, Tóth G, Solymosi K, Béni S, Boldizsár I. Aqueous incubation reveals transformation of rosmarinic acid in medicinally important Nepetoideae species and leads to novel bioactive compounds. **Pharmaceutics**. 2026; 18(8):1043–75.

6.2. Publications not related to the thesis

Nazli A, Khan MZI, Rácz Á, Béni S. Acid-sensitive prodrugs; a promising approach for site-specific and targeted drug release. **Eur J Med Chem**. 2024;276:116699–731.

Nazli A, Malanga M, Sohajda T, Béni S. Cationic cyclodextrin-based carriers for drug and nucleic acid delivery. **Pharmaceutics**. 2025;17(1):81–123.

Naseem A, Liu Y, **Nazli A**, Kuang HX, Yang B.Y. An insight into indigenous ethnobotanical and pharmacological potential of Solanaceae family in Pakistan: a review. **J Herb Med**. 2023;42: 100763–73.