Metabolic changes in Arabidopsis thaliana plants overexperssing chalcone synthase

Dao, T.H.H.

Chapter 5

Identification of metabolites in *Arabidopsis thaliana*

T.T.H. Dao[1, 2], H.K. Kim[1], H.J.M. Linthorst [3], Y. H. Choi[1] and R. Verpoorte[1]

[1] Section Metabolomics, Institute of Biology, Leiden University, Leiden, The Netherlands
[2] Traditional Pharmacy Department, Hanoi Pharmacy University, Hanoi, Vietnam
[3] Section Plant Cell Physiology, Institute of Biology, Leiden University, Leiden, The Netherlands

Abstract

Identification of *Arabidopsis thaliana* Col.0 metabolites by use of NMR spectroscopy is described in this chapter. Among the different extraction solvent tested, MeOD was the best solvent to extract phenolic compounds from *Arabidopsis*. By column chromatography using Sephadex LH-20 and prep HPLC, several flavonoids were isolated from the methanol extract of *Arabidopsis*, and their structures were identified by LC-MS and NMR spectroscopy as kaempferol 3-*O*-glucopyranoside-7-*O*-rhamnopyranoside, kaempferol 3-*O*-rhamnosyl (1–2) glucoside-7-*O*-rhamnopyranoside, kaempferol 3,7-*O*- dirhamnopyranoside and quercetine 3-*O*-rhamnopyranoside. Twenty four major metabolites of *Arabidopsis thaliana* Col.0 including amino acids, organic acids, sugars, phenylpropanoids, and flavonoids were identified and their NMR cheracteristics are also summerized in this study.

Keywords: NMR, metabolites, phenolics, flavonoids, extraction method, *Arabidopsis thaliana*

5.1. Introduction

Arabidopsis thaliana has become an extremely popular model system for studying plant biology. The biosynthesis of plant secondary metabolites represents a complex cellular network involving the transcription, translation and post-translational modification of many gene products. Analysis of whole plant metabolomes is a difficult task due to the huge number and great diversity of primary and secondary metabolites present in plant tissues [Dixon and Strack, 2003; Sumner *et al.,* 2003; Stobiecki and Kachlicki, 2005]. A good profiling method should be simple and detect as many of the metabolites as possible in a single extract of the material. In addition, the method should be reproducible to archive the data for future datamining.

The chromatography and spectroscopic technologies as HPLC-UV, GC/MS, LC/MS and NMR employed in plant metabolomics have been extensively reviewed [Fan, 1996; Fiehn *et al.,* 2000; Wagner *et al.,* 2003]. NMR is a tool to analyze the metabolome with a lot of advantages. Sample preparation for NMR measurement is usually simple and rapid, measurement times are short and readily automated and advanced data analysis methods are available. The 1D and 2D-NMR spectra of complex mixtures may provide sufficient information for the structures of unknown components to be elucidated, either from the NMR spectrum of the mixture itself, or after some purification. Another advantage of NMR is the linearity of quantitative responses on increasing metabolite concentrations, irrespective of the chemical compound class. Large signals in NMR can directly be interpreted as high level concentrations, whereas in MS, quantitative responses strongly rely on the ionization potential of each metabolite. Therefore, quantitation in MS is limited to relative abundances of a given metabolite between samples, or requires calibration curves if absolute comparison of different metabolites is needed. NMR either generates a metabolite profile, in which the NMR signals are assigned to specific metabolites, or a metabolite fingerprint, in which the analysis is based on the distribution of intensity in the NMR spectrum rather than the assignment of the signals [Krishnan et *al.,* 2005]. However, NMR has some limitations such as low sensitivity in comparison with MS. But new technology in NMR equipment such as higher-field spectrometers and use of cryogenically cooled probes improved the NMR sensitivity many folds.

Large numbers of metabolites of *Arabidopsis* have been identified. Understanding a significant part of *Arabidopsis* biology requires methods allowing the sensitive detection and quantification as well as the identification of secondary metabolites. Applying such techniques to various genetic backgrounds and to different environmental and developmental conditions then would help elucidate the function of such compounds and of the genes involved in their biosynthesis. Metabolic profiling of Arabidopsis and other plants have been developed in recent years. Most commonly used are gas chromatographies (GC)-mass spectrometry (MS)-based approaches. Several hundred of metabolites can be robustly and reliably detected but most of them are primary metabolites such as sugars, amino acids, organic acids [Roessner *et al.,* 2000; Fiehn *et al.,* 2000; Wagner *et al.,* 2003]. Beisdes that, liquid chromatography (LC)-MS based metabolomic has been used for profiling of metabolites [Roessner *et al.,* 2000]. Every analytical procedure is necessarily limited as to what type of compounds can be separated and detected. GC-MS is predominantly applied to very polar or unpolar substances though requiring derivatization to obtain volatile derivatives, whereas the main application of LC-MS is more related to compounds of medium polarity. About 300 metabolites were detected in *A. thaliana* leaf extracts and about half of them were identified by using GC-MS [Fiehn *et al.,* 2000], LC-MS and NMR [Hendrawati *et al.,* 2006; Le Gall *et al.,* 2005; Von Roepenack *et al.,* 2004].

Phenolic compounds are ubiquitous constituents of higher plants found in a wide range of commonly consumed plant foods such as fruits, vegetables, cereals and legumes, and in beverages of plant origin, such as wine, tea and coffee [Cheynier, 2005; Manach *et al.,* 2004]. These compounds are secondary metabolites of plants that are generally involved in defense against ultraviolet radiation or often attack by pathogens. They constitute an important class of plant secondary metabolites and are mostly present as glycosidic conjugates. The major flavonoid compounds in *A. thaliana* are the kaempferol glycosides flavonols [Rohde *et al.,* 2004; Veit and Pauli, 1999], but quercetin glycosides can also accumulate after exposure to UV radiation [Graham, 1998]. Another group of flavonoids present in *A. thaliana* are the anthocyanins, the major red, purple and blue pigments of plants best known from flowers and fruits. The major anthocyanin in *A. thaliana* has a cyanidin core with four attached sugars [Bloor and Abrahams, 2002]. Some flavonoids from green tissues of *A. thaliana* have been

fully structurally characterized with various physicochemical methods. Kaempferol 3-*O*-*β*-[*β*-D-glucosyl(1–6)D-glucoside]-7-*O*-*α*-L-rhamnoside, kaempferol 3-*O*-*β*-D-glucoside-7-*O*-*α*-L-rhamnoside, kaempferol 3-O-*α*-L-rhamnoside]-7-*O*-*α*-L-rhamnoside, kaempferol 3-*O*-*β*-[*α*-L-rhamnosyl (1–2)D-glucoside]-7-*O*-*α*-L-rhamnoside were identified [Veit and Pauli, 1999; Bloor and Abrahams, 2002]. Most studies on flavonoid characterization have been done by analytical procedures using the isolated flavonoids. In further experiments in this thesis a profiling of metabolites in crude extract will be applied so an identification of *Arabidopsis* flavonoids in plant crude extracts needs to be developed. So far the flavonoids of *Arabidopsis* are not available commercially so isolation and identification of flavonoids in *Arabidopsis thaliana* Col. 0 were done in this study in order to have reference compounds for further analysis.

Profiling applications of NMR in plant tissues have usually focused on the identification of particular metabolites, and so the extraction techniques need to be considered for optimal extraction recovery of the compounds of interest. The aim of this study is the application of NMR to identify *Arabidopsis thaliana* Col. 0 metabolites, focusing on phenolic compounds in plant crude extracts because we studied the effect of CHS expression, a key enzyme in flavonoid biosynthesis pathway (**Chapter 2**), on the *Arabidopsis* metabolome. Thus a suitable extraction method for this purpose was developed in this study.

5.2. Methods and Materials

5.2.1. Plant materials and extraction for flavonoid isolation

Arabidopsis thaliana above ground parts were used as a plant material for extraction. 500 ml of CH_3OH was added to 256 mg of dried and ground leaves and ultrasonicated for 30 minutes and then vacuum filtered. The procedure was repeated for 3 times and all the supernatants were pooled and dried using a rotary evaporator. The dried extract was redissolved in 100 ml of deionized water and partitioned with different solvents like *n*-hexane, chloroform, and *n*-butanol. All the fractions were separately dried by rotary evaporator and stored at 4 °C until further use.

5.2.3. Sample Fractionation

The *n*-butanol extract (1.2 g) was selected for fractionation as high flavonoids content was expected in this fraction. Sephadex column LH-20 (145 cm length x 16 mm diameter) was used for sample fractionation with 100% CH_3OH as a mobile phase. Total 84 fractions were collected of 5 ml each. TLC indexing was performed for every fourth fraction and observed under 254 nm and 366 nm. The solvent system for TLC indexing was composed of ethyl acetate, formic acid, acetic acid, and water, in the ratio of 100:11:11:27 (v/v). The fractions that showed a similar pattern under UV were pooled and seven combined fractions (from A to G) were obtained. Fraction A contained fractions from 1-19, B from 20-30, C from 31-34, D from 35-38, E from 39-48, F from 49-71, and G from 72-84. ^{1}H-NMR analyses were performed for all the pooled fractions and on the basis of flavanoids signals, fraction C, D, E, and F were selected for further purification by HPLC. Sixty sub-fractions (C1-4, D1-4, E1-4, and F1-4) were collected and analysed with H-NMR. The results show that flavonoids are mainly in sub-fraction F2 and F4. F4 sub-fraction contained more than one flavonoid so we applied one more HPLC step to fractionate F4 and four fractions (F4.1, 4.2, 4.3, 4.4) were collected each mainly containing a single compound.

5.2.4. HPLC analysis

The selected fractions were further separated using an Agilent 1100 series HPLC (Agilent, Waldbronn, Germany) equipped with a UV detector operating at 254 nm. A semi-preparative reversed phase column (Phenomenex Luna 5µ C18; 250 x 10 mm, 5µ) was used for separations, with a solvent gradient of 0.1 % formic acid with water and 0.1 % formic acid with CH_3OH. The gradient starts from H_2O-CH_3OH in the ration of 60:40 for the first 30 minutes, then shifted towards 20:80 for two minutes. After this the columm was reequilibrated again for the next analysis by ruinning the initited solvent 60:40 for eight minutes. Total time for each run is fourty minutes with the flow rate of 2 ml/min.

5.2.2. Plant materials and extraction of crude extract for NMR measurements

Plants were ground in liquid nitrogen and pooled before subjected to freeze-drying. Twenty-five milligrams of freeze-dried material were transferred to a micro-centrifuge

tube before adding 600 µl of CH_3OH-d_4. The mixture was vortexed for 2 minutes and sonicated for 20 minutes, followed by centrifugation at 13,000 rpm for 5 minutes at room temperature. Five hundred microliters of the supernatant were then transferred into 2ml micro-centrifuge tubes and were added two hundred fifty microliters of KH_2PO_4 buffer, pH 6.0, containing 0.1% trimethyl silyl propionic acid sodium salt (w/v). The mixture was left for 30 minutes in 4°C and followed by centrifugation at 6000 rpm for 5 minutes at room temperature. Seven hundred microliters of the supernatant were then transferred into 5 mm NMR tubes for analysis.

5.2.5. NMR measurements

The dried sub-fractions were redissolved in 1.0 ml of 50% CH_3OH-d_4 in D_2O (KH_2PO_4 buffer, pH 6.0) containing 0.05% TMSP (trimethyl silyl propionic acid sodium salt, w/v) and then 800 µl of the supernatant was transferred to a 5 mm NMR tube. ^{1}H-NMR and 2D *J*-resolved spectra were recorded at 25 °C on a 500 MHz Bruker DMX-500 spectrometer (Bruker, Karlsruhe, Germany) operating at a proton NMR frequency of 500.13 MHz. CH_3OH-d_4 was used as the internal lock. Each ^{1}H-NMR spectrum consisted of 128 scans requiring 10 min and 26 s acquisition time with the following parameters: 0.16 Hz/point, pulse width (PW) = 30° (11.3 µs), and relaxation delay (RD) = 1.5 s. A pre-saturation sequence was used to suppress the residual H_2O signal with low power selective irradiation at the H_2O frequency during the recycle delay. FIDs were Fourier transformed with LB = 0.3 Hz. The resulting spectra were manually phased and baseline corrected, and calibrated to TSP at 0.0 ppm, using XWIN NMR (version 3.5, Bruker). 2D *J*-resolved NMR spectra were acquired using 8 scans per 128 increments for F1 and 8 k for F2 using spectral widths of 5000 Hz in F2 (chemical shift axis) and 66 Hz in F1 (spin–spin coupling constant axis). A 1.5 s relaxation delay was employed, giving a total acquisition time of 56 min. Datasets were zero-filled to 512 points in F1 and both dimensions were multiplied by sine-bell functions (SSB = 0) prior to double complex FT. *J*-Resolved spectra tilted by 45°, was symmetrized about F1, and then calibrated, using XWIN NMR (version 3.5, Bruker). ^{1}H–^{1}Hcorrelated spectroscopy (COSY) and heteronuclear multiple bonds coherence (HMBC) spectra were recorded on a 600 MHz Bruker DMX-600 spectrometer (Bruker). The COSY spectra were acquired with 1.0 s relaxation delay, 6361 Hz spectral width in both dimensions. Window

function for COSY spectra was sine-bell (SSB = 0). The HSQC spectra were obtained with 1.0 s relaxation delay, 6361 Hz spectral width in F2 and 27,164 Hz in F1. Qsine (SSB = 2.0) was used for the window function of the HSQC. The HMBC spectra were recorded with the same parameters as the HSQC spectrum except for 30,183 Hz of spectral width in F2. The optimized coupling constants for HSQC and HMBC were 145 Hz and 8 Hz, respectively.

5.3. Results and discussions

5.3.1. Optimization of extraction method

The aim of this thesis is to study the effect of overexpression of CHS, a key enzyme of the flavonoid biosynthesis pathway, on the metabolism in *Arabidopsis*. As this enzyme will result in the production of flavonoids and related compounds, the focus is on the phenolic compounds. This includes also the compounds from earlier part of the phenylpropanoid pathway which might be affected because of competitive for the same precusors. Thus an efficient extraction method with good yield and reproducibility which provides reliable metabolic profiling data on phenolic compounds by using NMR spectroscopy was investigated. Due to the large differences of metabolites e.g. in molecular weight and polarity, in general CH_3OH-H_2O is usually used as extraction solvent for metabolic profiling because it is medium polar, and it penetrates cell walls and membranes quite effectively. In order to optimize the extraction for metabolic profiling, different ratios of water were mixed with CH_3OH, from 0 % to 100 % following the gradient 0%, 25%, 50%, 75%, 100%. The solvent CH_3OH/H_2O ratio of 1/1(v/v) give both signals of primary and secondary metabolites in the ^{1}H-NMR spectra, whereas the solvent 100% CH_3OH preferably extracts the secondary metabolites, such as phenolic compounds. Multivariable data analysis of the various extracts (data not shown) only revealed a clear difference between wild type and CHS transgenic plants with the 100% CH_3OH extracts. Therefore CH_3OH-d_4 was chosen as extraction solvent. As CH_3OH-d_4 also extracts chlorophyll, D_2O was added (30% in total volume) to the primary crude CH_3OH-d_4 extract to precipitate chlorophyll before NMR analysis. The final supernatant was analyzed directly by NMR.

5.3.2. Isolation and characterization of Arabidopsis thaliana Col. 0 flavonoids

The flavonoid glycosides of *Arabidopsis thaliana* Col. 0 were isolated and structure elucidated by use of HPLC, NMR and LC/MS.

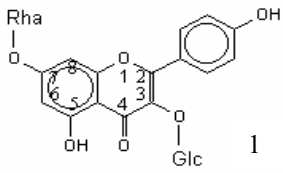
The NMR spectrum of the F2 sub-fraction shows two signals at δ 6.84 (d, J=2.0 Hz) δ 6.82 (d, J=2.0 Hz). Those are H-6, H-8 characteristic signals of a flavonoid glycoside. Another two signals at δ 6.99 (d, J=8.8 Hz) δ 8.11 (d, J=9 Hz) are charactistic signals of H-3'& 5', H-2' & 6' in kaempferol. The signals at δ 5.56 (d, J=1.5 Hz) and δ 1.25 (d, J=6.2 Hz) are characteristic signals of 7-*O*-rhamnose and the signals at δ 5.77 (d, J=7.8 Hz) and δ 0.95 (d, J=6.2 Hz) were identified as signals of a futher 3-*O*- (rhamnosyl (1–2) glucoside) [Kerhoas *et al.,* 2006]. This compound was assigned as kaempferol 3-*O*-[rhamnosyl (1–2) glucoside]-7-*O* -rhamnopyranoside which also fit with its $[M–H]^-$ signal on LC-ESI-MS is *m/z* 739.

The F4.1 fraction also showed the characteristic signals of a keampferol derivative at δ 6.52 (H-6, d, J=2.0 Hz) δ 6.82 (H-8, d, J=2.0 Hz) δ 7.0 (H-3'& 5', d, J=8.8 Hz) δ 8.09 (H-2' & 6', d, J=9 Hz). The signals at δ 5.56 (d, J=1.6 Hz) and δ 1.25 (d, J=6.2 Hz) are characteristic signals of 7-*O*-rhamnose and the signals at δ 5.77 (d, J=7.8 Hz) and δ 5.33 (d, J=7.7 Hz) were identified as signals of a 3-*O*- glucoside [Kerhoas *et al.,* 2006]. $[M–H]^-$ signal in LC-ESI-MS of F4.1 showed a *m/z* 577 which was confirmed that F4.1 is kaempferol (3-*O*-glucopyranoside-7-*O*-rhamnopyranoside) [Kerhoas *et al.,* 2006].

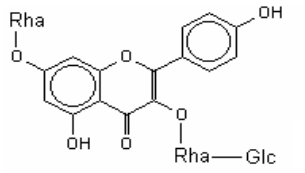
The NMR signals of the F4.3 fraction at δ 6.43 (H-6, d, J=2.0 Hz) δ 6.81 (H-8, d, J=2.0 Hz) δ 7.83 (H-2'& 6', d, J=9.0 Hz) δ 7.04 (H-3'&5', d, J=9.0 Hz) are in accordance with a kaempferol glycoside moiety. The signals at δ 5.56 (d, J=1.6 Hz) and δ 1.25 (d, J=6.2 Hz) are characteristic signals of a 7-*O*-rhamnose and the signals at δ 0.94 (d, J=6.0 Hz) were identified as signals of 3-*O*-rhamnoside [Kerhoas *et al.,* 2006]. Therefore F4.3 were assigned as kaempferol (3,7-*O*-dirhamnopyranoside) and fits with *m/z* 593 [Kerhoas *et al.,* 2006].

The NMR spectrum of the F4.4 fraction shows typical the quercetin derivative NMR signals at δ 7.72 (H-6, d, J=2.0 Hz) δ 6.97 (H-8, d, J=2.0 Hz) δ 6.89 (H-5', d, J=8.0 Hz) δ 7.27 (H-6', dd, J=8.0, 2.0 Hz) δ 7.32 (H-2', d, J=2.1 Hz). The sugar attached to quercetin was identified as rhamnose with the NMR signal at δ 0.94 (d, J=6.0 Hz). This compound also was confirmed as quercetin 3-*O*-rhamnopyranoside (**Fig. 5.1**) with a $[M–H]^-$ signal is *m/z* 477.

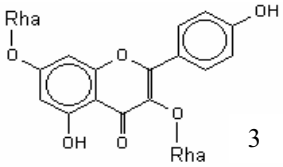
Based on the above mentioned information we could thus identify three kaempferol glycosides and one quercetin glycoside in *Arabidopsis thaliana* Col.0 leaves.



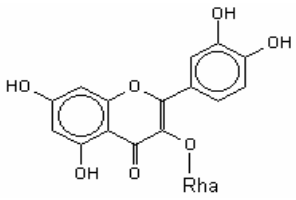
1



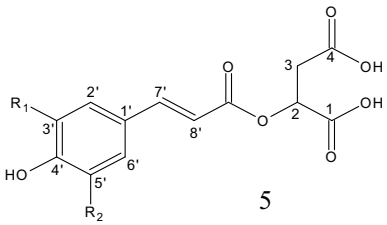
2



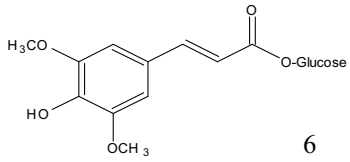
3



4



5



6

1) Kaempferol 3-*O*-glucopyranoside-7-*O*-rhamnopyranoside.
(2) Kaempferol 3-*O*-rhamnosyl (1–2) glucoside-7-*O*- rhamnopyranoside.
(3) Kaempferol 3,7-*O*- dirhamnopyranoside.
(4) Quercetine 3-*O*-rhamnopyranoside.
(5) Phenylpropanoids: $R_1 = OCH_3$, $R_2 = OH$, hydroxyferuloyl malate
$R_1 = OH$, $R_2 = H$, caffeoyl malate
$R_1 = H$, $R_2 = H$, coumaroyl malate
$R_1 = OCH_3$, $R_2 = OCH_3$, sinapoyl malate
(6) Synapoyl glucose

Figure 5.1. Chemical structures of flavonoids and phenylpropanoids in *A. thaliana* Col.0

5.3.3. NMR analysis of Arabidopsis in methanol crude extract

Metabolic profiling of *Arabidopsis* CH_3OH-d_4 crude extracts by using NMR 600MHz will be applied in the next experiments (**Chapter 6, 7, 8**). Identification of compounds is based on NMR spectra as described in this chapter.

Figure 5.2 shows the ^{1}H-NMR spectrum of the *Arabidopsis* Col.0. The combined information gathered from ^{1}H-NMR, COSY, J-resolved and HMBC spectra and the use of a library of ^{1}H-NMR spectra of reference compounds allowed an almost complete assignment. Sugars, organic acids and amino acids signals are present in the high field region of the NMR spectra, between 0.5 to 6.0 ppm (**Figure 5.2 b**). In the amino acid region (δ 0.8–δ 4.0) the main identified signals were alanine δ 1.48 (H-3, d, J=7.0 Hz), glutamic acid δ 2.07 (H-2, m) δ 2.41 (H-3, m), glutamine δ 2.12 (H-2, m) δ 2.48 (H-3, m) , leucine δ 0.96 (H-5, d, J=8.0 Hz), threonine δ 1.32 (H-5, d, J=6.6 Hz), valine δ 1.03 (H-5, d, J=7.8 Hz), aspartic acid δ 2.67 (m) and asparagine δ 2.8 (m). The organic acid regions of the NMR spectrum only show signals of formic acid δ 8.5 (s) and malic acid δ 4.32 (H-2, dd, J=4.0 Hz, 11 Hz) because the other organic acids have very poor solubility in CH_3OH. Hence only formic acid and malic acid can be detected in the CH_3OH extract. The signals of the terminal CH_3 of choline was identified at δ 3.23 (s). For sugars, the anomeric proton of β-glucose at δ 4.57 (H-1, d, J=8.0 Hz), α-glucose at δ 5.18 (H-1, d, J=3.7 Hz), sucrose at δ 5.4 (H-1, d, J=4.0 Hz), rhamnose at δ 5.62 (H-1, d, J=8.0 Hz), and fructose at δ 4.17 (H-1, d, J=9.0 Hz) were assigned.

Signals of four flavonoids present in the low field region (6.8-8.2 ppm) have been analyzed (see above), Quercetine derivatives are present as minor compounds in the crude extract but difficult to detect in the NMR spectrum. Moreover, in the aromatic region, the presence of five major doublets with the same coupling constants (d, J=16.0 Hz) in the range of δ 6.31–δ 6.50 indicate the presence of the *trans* olefinic protons H-8' of phenylpropanoids (**Figure 5.3**) [Liang *et al.,* 2006]. This also was confirmed by the correlation of H-8′ of the phenylpropanoids with the H-7′ (d, J=16.0 Hz) protons at δ 7.54–δ 7.59 in the COSY spectrum (**Figure 5.4**). Five *trans*-phenylpropanoids were elucidated by two dimensional NMR. Those are *trans*-caffeoyl malate (H-8', δ 6.32; H-7', δ 7.66), *trans*-5-hydroxyferuloyl malate (H-8', δ 6.34; H-7', δ 7.66), *trans*-coumaroyl malate (H-8', δ 6.37, H-7', δ 7.66), sinapoyl malate (H-2 & 6, δ 6.99 s; H-8, δ 6.48 d, J=16 Hz; H-7 δ 7.66 d, J=16 Hz), and sinapoyl glucose (H-2 & 6, δ 6.97 s; H-8 δ 6.49 d, J=16 Hz; H-7 δ 7.77 d, J=16 Hz) [Liang *et al.,* 2006] (**Figure 5.3**). The *cis*-form of those phenylpropanoids are present only at very low concentration in crude extract so we could not identify them.

The ^{1}H chemical shifts (δ) and coupling constants (Hz) of the indentified *Arabidopsis thaliana* Col.0 metabolites are presented in **Table 5.1**.

Table 5.1. ^{1}H chemical shifts (δ) and coupling constants (Hz) of *Arabidopsis thaliana* Col. 0 metabolites identified by references and using 1D and 2D NMR spectra (CH_3OH-d_4–KH_2PO_4 in D_2O, pH 6.0)

Compounds	Chemical shifts (ppm) and coupling constants (Hz)
Amino/organic acids	
Threonine	δ 1.32 (H-5, d, J=6.6 Hz)
Alanine	δ 1.48 (H-3, d, J= 7.0 Hz)
Glutamine	δ 2.12 (H-2, m) δ 2.48 (H-3, m)
Glutamic acid	δ 2.07 (H-2, m) δ 2.41 (H-3, m)
Valine	δ 1.03 (H-5, d, J=7.8 Hz)
Leucine	δ 0.96 (H-5, d, J=8.0 Hz)
Asparagine	δ 2.8 (m), 2.97(m)
Aspartic acid	δ 2.67 (m)
Malic acid	δ 4.32 (H2, dd, J=4.0 Hz, 11 Hz) δ 2.80 (H3, dd, J=8.8 Hz, 16.0 Hz) δ 2.96 (H2, dd, J=3.6 Hz, 16.0 Hz)
Formic acid	δ 8.5 (s)
Sugars	
β-glucose	δ 4.57 (H-1, d, J=8.0 Hz)
α-glucose	δ 5.18 (H-1, d, J=3.7 Hz)
Rhamnose	δ 5.62 (H-1, d, J=8.0 Hz)
Fuctose	δ 4.17 (H-1, d, J=9.0 Hz)
Sucrose	δ 5.40 (H-1, d, J=4.0 Hz)
Phenylpropanoids/Flavonoids	
Kaempferol 3-*O*-glucopyranoside-7-rhamnopyranoside	δ 6.52 (H-6, d, J=2.0 Hz) δ 6.82 (H-8, d, J=2.0 Hz) δ 7.0 (H-3'& 5', d, J=8.8 Hz) δ 8.09 (H-2' & 6', d, J=9 Hz)
Kaempferol 3,7-*O*-dirhamnopyranoside	δ 6.43 (H-6, d, J=2.0 Hz) δ 6.81 (H-8, d, J=2.0 Hz) δ 7.83(H-2'& 6', d, J=9.0 Hz) δ 7.04 (H-5', d, J=9.0 Hz)
Kaempferol 3-*O* -rhamnosyl (1–2) glucoside-7-*O*-rhamnopyranoside	δ 6.84 (H-6, d, J=2.0 Hz) δ 6.82 (H-8, d, J=2.0 Hz) δ 6.99 (H-3'& 5', d, J=8.8 Hz) δ 8.11 (H-2' & 6', d, J=9 Hz)
Quercetine 3-*O*-rhamnoside	δ 7.72 (H-6, d, J=2.0 Hz) δ 6.97 (H-8, d, J=2.0 Hz) δ 6.89 (H-5', d, J=8.0 Hz) δ 7.27 (H-6', dd, J=8.0, 2.0 Hz) δ 7.32 (H-2', d, J=2.1 Hz)
trans-5-hydroxyferuoyl malate	δ 6.34 (H-8', d 16 Hz) δ 7.54 (H-7', d, J=16 Hz)
trans-feruoyl malate	δ 6.42 (H-8', d 16 Hz) δ 7.66 (H-7', d, J=16 Hz)
trans-caffeoyl malate	δ 6.32 (H-8', d 16 Hz) δ 7.55 (H-7', d, J=16 Hz)
trans-coumaroyl malate	δ 6.37 (H-8', d 16 Hz) δ 7.59 (H-7', d, J=16 Hz)
trans-sinapoyl malate	δ 6.99 (H-2 & 6, s) δ 6.48 (H-8, d, J=16 Hz) δ 7.66 (H-7, d, J=16 Hz)
trans-sinapoyl glucoside	δ 6.97(H-2 & 6, s), δ 6.49 (H-8,d, J=16 Hz), δ 7.77(H-7, d, J=16 Hz)
Other compounds	
Choline	δ 3.23 (s)
Inositol	δ 4.1 (H-2, dd, J=2.0 Hz, 13Hz) δ 3.62(H-4 and 6, dd, J=8.8 Hz, 16.2 Hz) δ 3.46 (H-1 and 3, dd, J=6.5 Hz, 13.9 Hz)

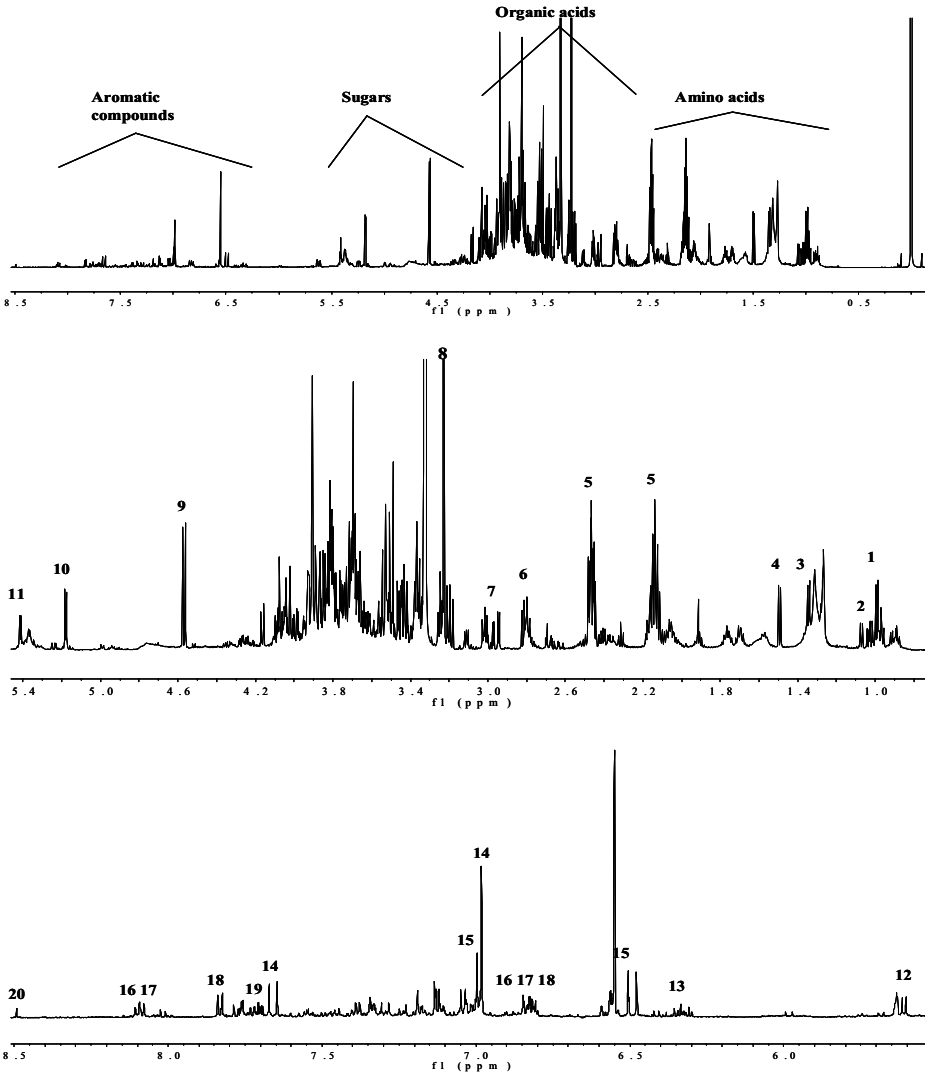


Figure 5.2. ^{1}H-NMR spectra of *A. thaliana* Col. 0 (a), extended high field region 0.0-5.5 ppm (b), extended low field region 5.6-8.2 ppm (c). 1. Leucine., 2. Valine, 3. Threonine, 4. Alanine, 5. Glutamine, 6. Asparagine, 7. Malic acid, 8. Choline, 9. β-glucose, 10. α-glucose, 11. Sucrose, 12. Rhamnose, 13. Phenylpropanoids (*trans*-feruoyl malate, *trans*-caffeoyl malate, *trans*-coumaroyl malate, *trans*-5-hydroxyferuoyl malate) 14. *trans*-sinapoyl malate, 15. *trans*-sinapoyl glucose, 16. Kaempferol 3-*O*-glucopyranoside-7-rhamnopyranoside, 17. Kaempferol 3-*O*-rhamnosyl (1–2) glucoside-7-*O*- rhamnopyranoside, 18. Kaempferol 3,7-*O*- dirhamnopyranoside, 19. Quercetin derivatives, 20. Formic acid.

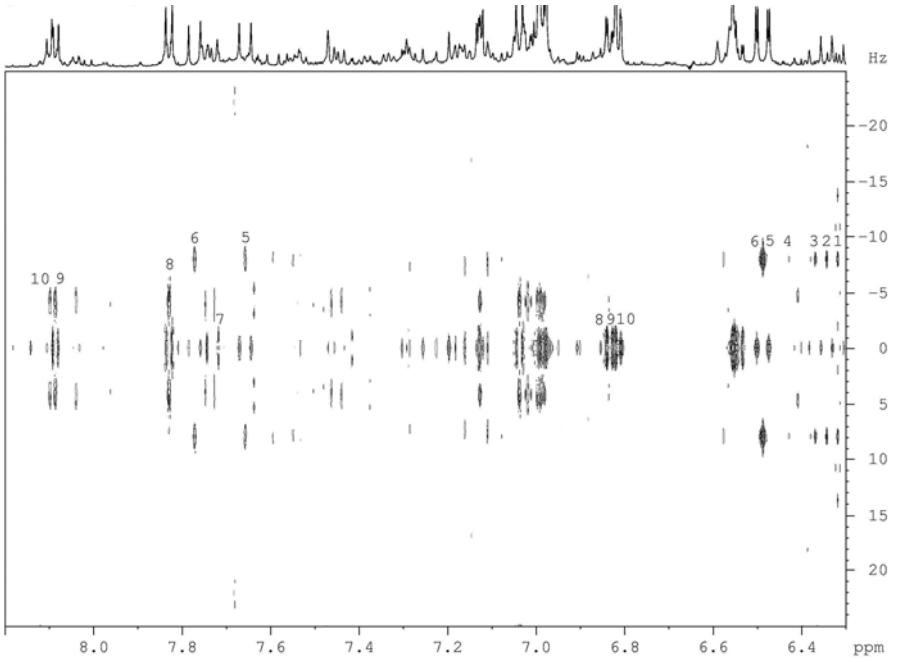


Figure 5.3. 2D NMR J-resolved spectra of *A. thaliana* Col. 0 in aromatic region from 6.2 – 8.2 ppm. 1. *trans*-caffeoyl malate, 2 *trans*-5-hydroxyferuloyl malate., 3. *trans*-coumaroyl malate, 4. *trans*-feruloyl malate, 5. *trans*-sinapoyl malate,,6. *trans*-sinapoyl glucose, 7. Quercetin derivatives., 8. Kaempferol 3,7-*O*- dirhamnopyranoside, 9. Kaempferol 3-*O*-rhamnosyl (1–2) glucoside-7-*O*- rhamnopyranoside, 10. Kaempferol 3-*O*-glucopyranoside-7-rhamnopyranoside.

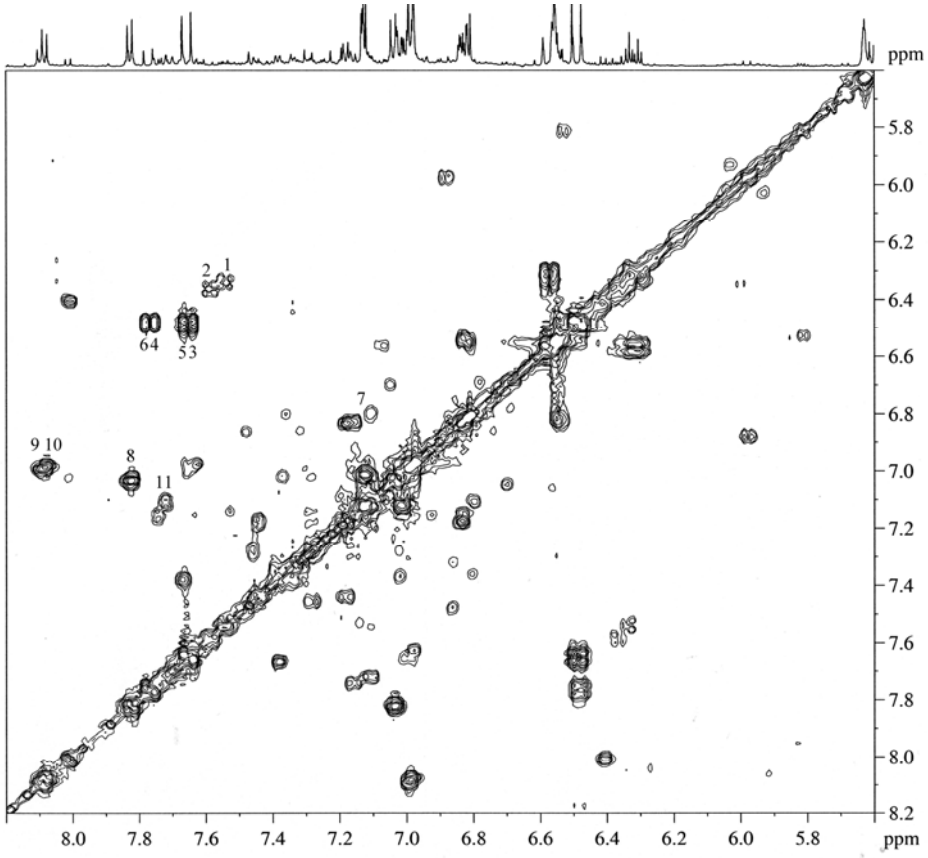


Figure 5.4. 2D NMR Cosy spectra of *A. thaliana* Col. 0 in aromatic region from 5.7 – 8.2 ppm. 1. *trans*-caffeoyl malate, 2 *trans*-5-hydroxyferuloyl malate., 3. *trans*-coumaroyl malate, 4. *trans*-feruloyl malate, 5. *trans*-sinapoyl malate,,6. *trans*-sinapoyl glucose, 7. Quercetin derivatives., 8. Kaempferol 3,7-*O*- dirhamnopyranoside, 9. Kaempferol 3-*O*-rhamnosyl (1–2) glucoside-7-*O*-rhamnopyranoside, 10. Kaempferol 3-*O*-glucopyranoside-7-rhamnopyranoside.

5.4. Conclusion

Three keampferol glycosides and one quercetin glycoside were isolated and identified in this study. Twenty six metabolites of *A. thaliana* Col.0 in methanol crude extract were identified and listed **Table 5.1.** These results are now used as reference for next studies.