



Chapter 4

Biosynthetic pathway for flavonoids in cell suspension cultures of *Helichrysum aureonitens*

4.1 Abstract

Most phenolic compounds are derived from phenylalanine via the core phenylpropanoid pathway leading to formation of the flavonoids and other phenolic compounds. Flavonoids, which represent most of the medicinal properties, are the major phenolic compounds in the genus *Helichrysum*. Flavonols are present in *H. aureonitens*, and are represented by three compounds, kaempferol, quercetin, and galangin. Chemically the difference between galangin and kaempferol is the presence of an OH at the 4' position on the B ring of kaempferol. The only enzyme catalyzing the hydroxylation at the 4' position is cinnamate 4 hydroxylase (C4H) (Boudet, 2003). The chemical analysis of some of the intermediates by GC-MS showed that, except for cinnamic acid and kaempferol, the 4'-OH other intermediates including *p*-coumaric acid and naringenin, were not present in the leaf samples of *H. aureonitens*, while other intermediate compounds without OH at the 4' position on the B ring were detected. It was therefore proposed that for the production of galangin, an alternative pathway (non 4'-OH pathway) might be present in the *H. aureonitens* leaf extracts, which starts with phenylalanine and then moves to cinnamic acid. To produce kaempferol in the leaves of *H. aureonitens* plants, C4H, the only enzyme that can specifically catalyze hydroxylation at the 4' position, is necessary. Although C4H should be active in the leaves of *H. aureonitens*, it is proposed to be able to use galangin also as a substrate.

4.2 Introduction

Flavonoids represent an important class of natural products in plants. This category of secondary metabolites has important functions in the biochemistry and ecology of plants (Forkman and Heller, 1999). Flavonoids consist of several subgroups of compounds which include, anthocyanins, flavonones and flavonols. The phenylpropanoid pathway provides a variety of biologically important secondary metabolites through several important branches (Chen et al., 2007). The major one synthesizes lignins, which play a fundamental role in mechanical support, solute conductance and disease resistance in higher plants (Harakava, 2005; Chen et al., 2007). Another important branch of the pathway synthesizes various flavonoid compounds, which are involved in the attraction of pollinators and for protecting plants from UV irradiation and attacks by fungi and animals (Betz et al., 2001; Chen et al., 2007).

Flavonoids represent the major phenolic compounds in the genus *Helichrysum* (Czinner et al., 2001) and mainly confer the medicinal properties to this genus (Dombrowicz et al., 1994). Flavonoids are an important class of low molecular weight plant secondary metabolites which are widespread throughout the plant kingdom ranging from mosses to angiosperms (Koess et al., 1994). They are most commonly known for their antioxidant activity, as most of the polyphenols are ideal for free radical scavenging. *In vitro* experiments have showed that the majority of flavonoids are more effective in antioxidant activity than vitamin E and C (Rice-Evans et al., 1996; Rice-Evans et al., 1997).

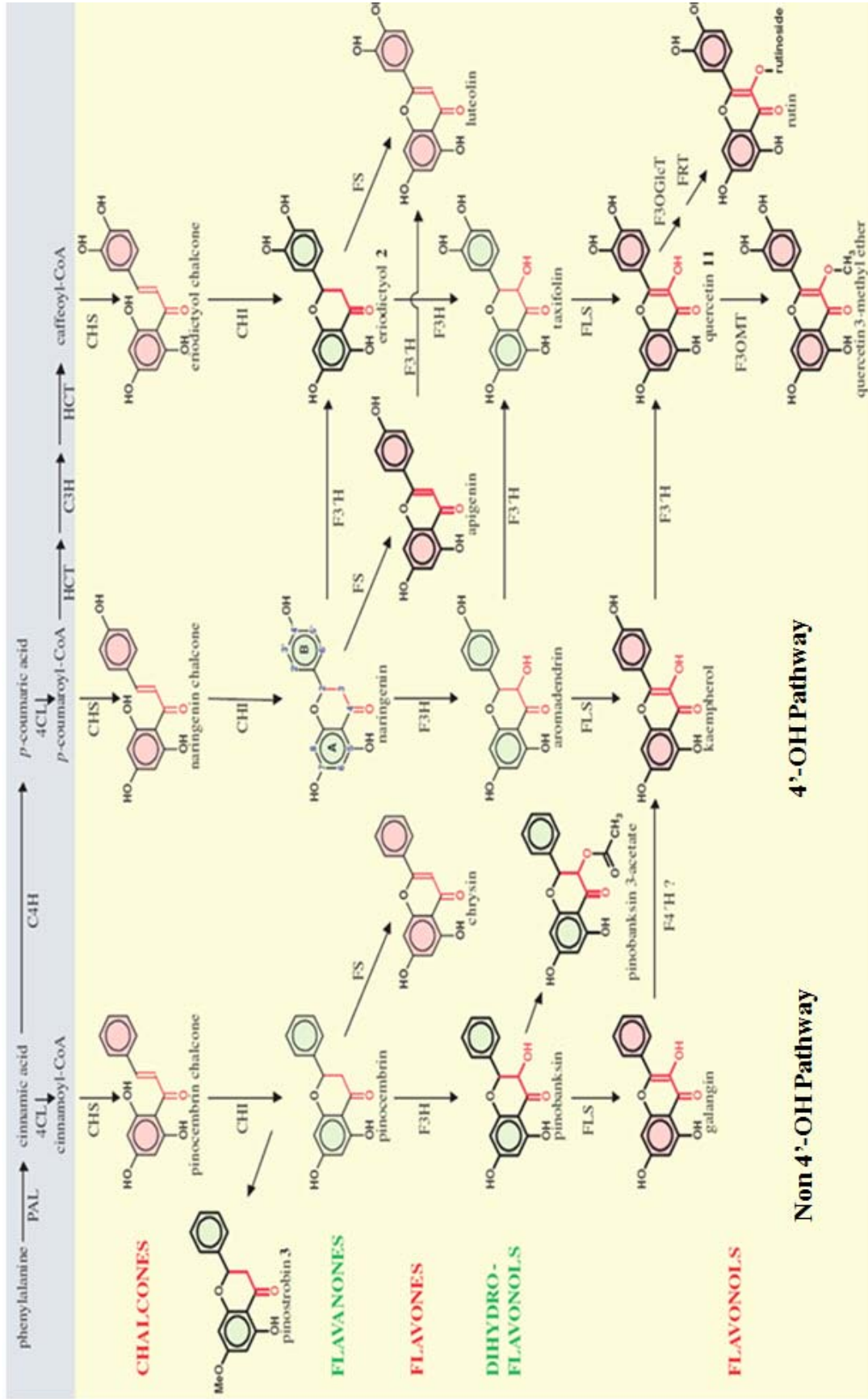
After the discovery of phenylalanine ammonia-lyase by Koukol and Conn (1961), rapid and substantial progress has been made to elucidate the metabolic pathways of flavonoids. Although the main reactions are known, there are still a few gaps.

Flavonoids are synthesized by the phenylpropanoid metabolic pathway in which phenylalanine is used as a substrate to produce 4-coumaroyl-CoA. This then reacts with

malonyl-CoA to synthesise the backbone of flavonoids. They all share a basic skeleton that contains two aromatic rings (A and B) and a heterocyclic ring (C) that contains one oxygen atom (Fig. 4.1). Based on the modification of the C-ring, they can be classified into several groups such as, flavonones, isoflavones, flavones, flavonols, flavanols and anthocyanines (Schijlen et al., 2004). Along this pathway, many products can be formed, including the flavonols, flavan-3-ols, proanthocyanidins (tannins) and a host of other polyphenolics (Scheme 4.1).

The flavonol group contains galangin, kaempferol, quercetin and myricetin (Fig. 4.2). Most phenolic compounds are derived from phenylalanine via the core phenylpropanoid pathway. Chemically the difference among flavonoids is mainly the number and the position of the hydroxyl groups on the B ring. The main difference between galangin and most other members of the flavonols is the absence of the 4'-OH of ring B in galangin whereas most other flavonols have been hydroxylated. Although there is no clear evidence on the whole pathway of galangin biosynthesis in plants, it is logical to propose that there is an alternative pathway (non 4'-OH pathway) for the production of galangin. This pathway has already been tested in an engineered combinatorial biosynthetic pathway in *E. coli* in which phenylalanine was converted to galangin (Miyahisa et al., 2006).

It probably starts with phenylalanine and then goes to cinnamic acid. Since the only enzyme catalyzing the hydroxylation at the 4' position is cinnamate-4-hydroxylase (C4H) and galangin does not have an OH⁻ at this position, it can be concluded that C4H is not involved in this pathway. This means that cinnamic acid reacts directly with malonyl-CoA to synthesise cinnamoyl-CoA and is then converted to pinocembrin chalcone. This pathway also produce pinocembrin which itself will be used as a substrate to synthesise pinobanksin, which is then probably converted to galangin. All the enzymes involved in this pathway would be similar to the 4'-OH pathway except for the C4H. The presence of the alternative



Scheme. 4. 1. Biosynthetic pathway of flavonoids. (Morreel *et al.*, 2006).

biosynthetic pathway is logical, but to date, some of its intermediates have not been tested for the production of the relevant products. It has not been shown by experimentation that pinobanksin is the substrate used to synthesize galangin and that this reaction is catalyzed by flavonol synthase (FLS) (Forkman and Heller, 1999).

In some plants the intermediates and products of both types of pathways have been identified in one plant species. For instance in the extract of Mexican oregano (*Lippia graveolens*), quercetin, naringenin, pinocembrin and galangin was identified by direct comparison with standards (Lin et al., 2007). Identification of some flavonoids such as chrysin, galangin and pinocembrin as major components has been reported in bud coats of *Populus* (Wollenweber, 1975). This information raises some questions. Is there an alternative pathway to synthesize galangin and its intermediates? If yes, are there links between the galangin pathway and the normal biosynthetic pathway for flavonols?

Since galangin has already been isolated and identified from *H. aureonitens* (Meyer et al., 1997), it represented a good opportunity to study the relationship between the galangin production pathway and the normal pathway for flavonols in this plant. In this chapter, a study on the biosynthetic pathway of flavonols in *H. aureonitens*, by chemical analysis of some intermediates based on the 4'-OH flavonoids biosynthetic pathway was done.

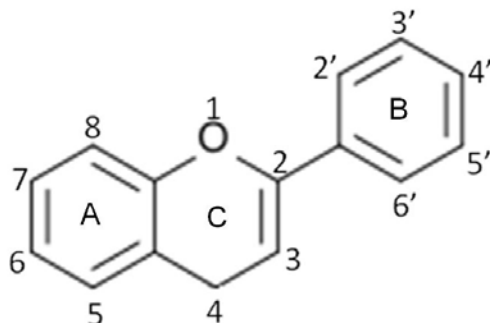
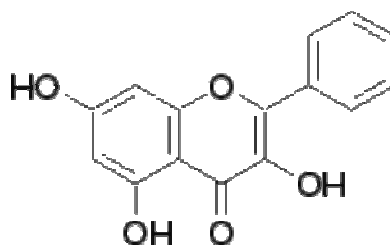
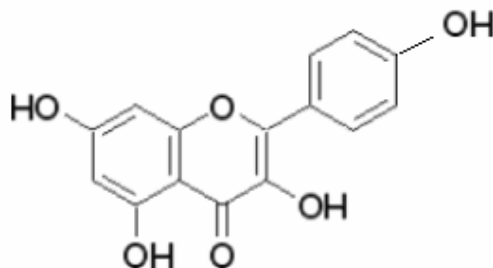


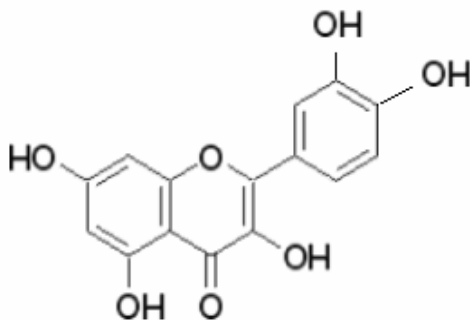
Fig. 4. 1. Molecular structure of the flavonoid backbone.



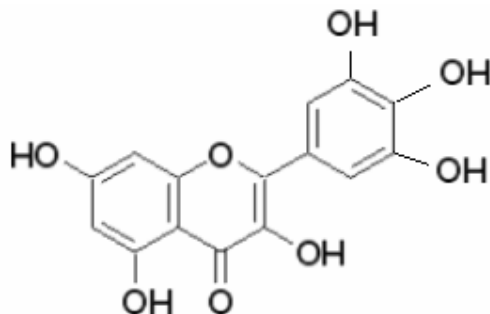
Galangin



Kaempferol



Quercetin



Myricetin

Fig. 4. 2. Chemical structures of some members of the flavonols. The differences between compounds are based on the presence of OH-groups at the 4' position and the number of OH-groups on the B aromatic ring.

4.3 Materials and methods

4.3.1 Plant materials

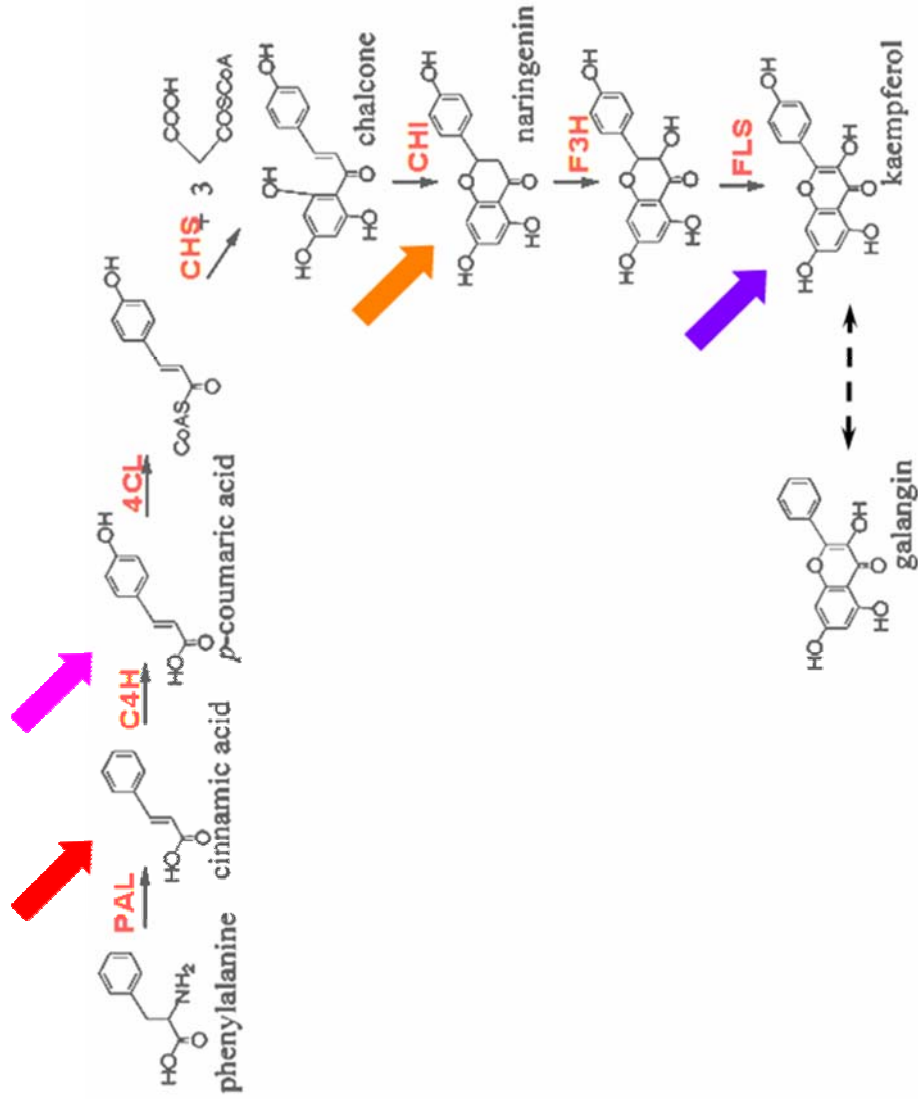
Helichrysum aureonitens plants were collected from the Kwazulu-Natal province of South Africa. For the chemical analysis the air dried leaves were extracted with ethanol. The combined extracts were concentrated under reduced pressure to produce a brown, oily residue. A voucher specimen (Afol. 2001) of the herb was deposited at the South African National Biodiversity Institute, Pretoria.

4.3.2 GC-MS analysis

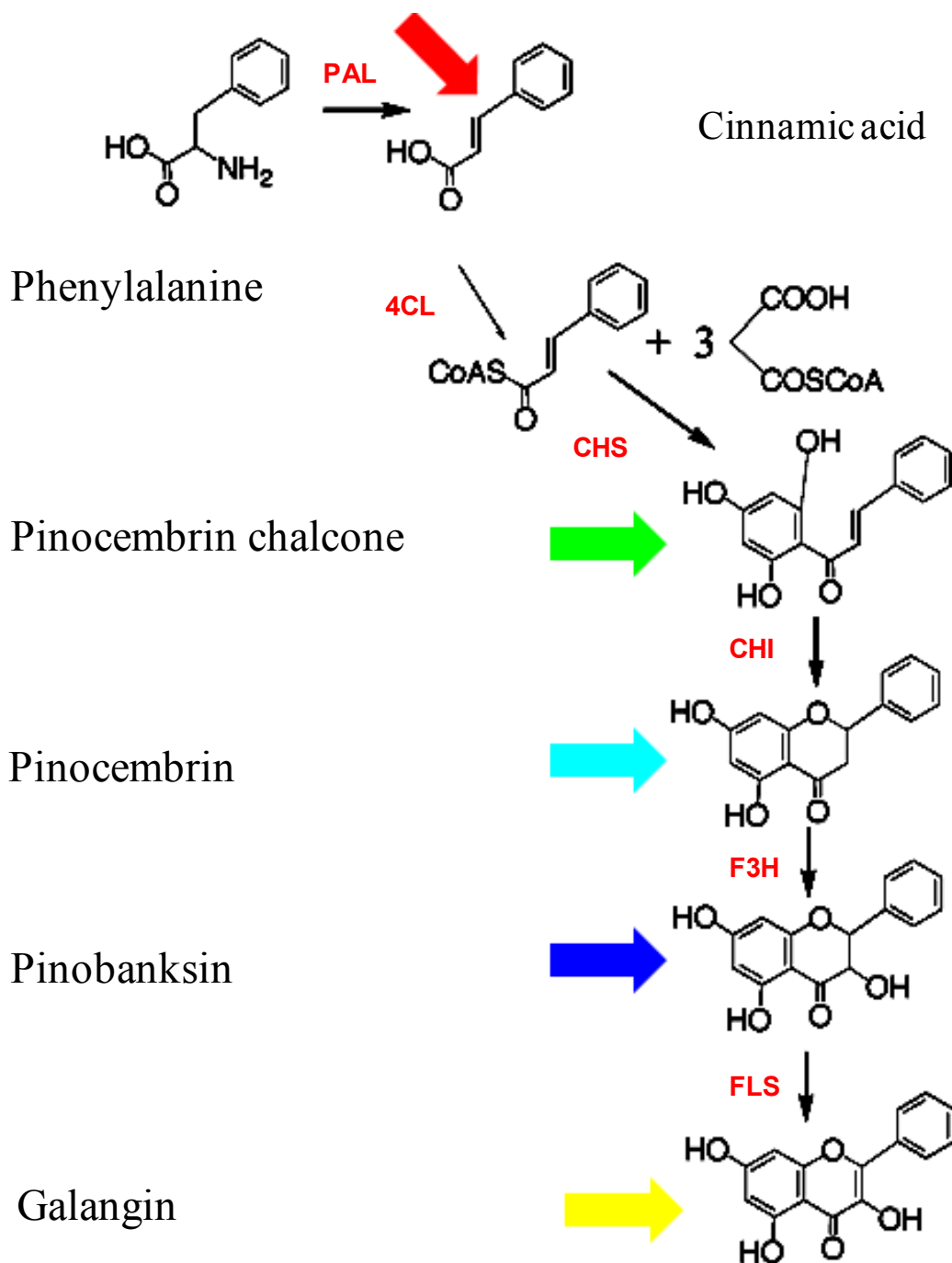
According to the 4'-OH biosynthetic pathway of flavonoids (Scheme 4.1), the presence of some intermediates including cinnamic acid, *p*-coumaric acid, naringenin and kaempferol (as a final product in flavonols biosynthetic pathway) were chemically analyzed by GC-MS. The standard chemicals for each of the intermediates were obtained from Wako, Japan. A part of the residue of the samples were trimethylsilylated with *N*-methyl-*N*-trimethylsilyltrifluoroacetamide (MSTFA, purchased from SIGMA) at 80°C for 30 min followed by GC-MS analysis. GC-MS analyses were carried out under the following conditions. A mass spectrometer (JMS-AM SUN200, JEOL) was connected to a gas chromatograph (6890A, Agilent Technologies), EI (70 eV), source temperature 250°C, HP-5 (30 m × 0.32 mm, 0.25-μM film thickness, J&W), injection temperature 250°C, column temperature program: 80°C for 1 min, then raised to 280°C at a rate of 20°C min⁻¹, and held at this temperature for 8 min; interface temperature 300°C, carrier gas He, flow rate 1.2 mL min⁻¹, splitless injection.

4.4 Results

To assess the pathway for galangin production in *H. aureonitens*, the presence of some intermediates such as cinnamic acid, *p*-coumaric acid, naringenin, galangin and kaempferol were analyzed by GC-MS (Scheme 4.2). The GC-MS analysis of methylated ethanolic leaf extracts of *H. aureonitens* showed the presence of identical peaks with the methylated derivatives of cinnamic acid (identified peaks 220, 205, 161, 131), galangin (471, 399, 327, 105 and 73) and kaempferol (574, 559, 415, 272) (Fig. 4.3 to 4.7). This means that three of the tested intermediates, cinnamic acid, galangin and kaempferol were detected, but *p*-coumaric acid and naringenin, were not present in the ethanolic leaf extracts of *H. aureonitens*. There were however some other peaks which illustrated the production of several compounds in association with galangin in the ethanolic leaf extracts of *H. aureonitens* (Fig. 4.7-A). The GC-MS analysis of the compounds revealed that these compounds are pinocembrin chalcone, pinocembrin and pinobanksin (Fig. 4.8). The difference of these intermediates with ones in the 4'-OH biosynthetic pathway is only in the OH group at the 4' position on the B ring (Scheme 4.1).



Scheme 4. 2. 4'-OH biosynthetic pathway for flavonols originating from *p*-coumaric acid. The arrows indicate the intermediates that were GC-MS analyzed to confirm the presence of this pathway in the leaf samples of *H. aureonitens*. Only kaempferol and galangin were detected.



Scheme 4. 3. The proposed alternative biosynthetic pathway for galangin production in *Helichrysum aureonitens* originating from cinnamic acid. The arrows indicate the intermediates detected by GC-MS analysis. This pathway was already tested in *E. coli* by construction of a combinatorial cluster gene (Miyahisa et al., 2006), but it has not been reported from plants

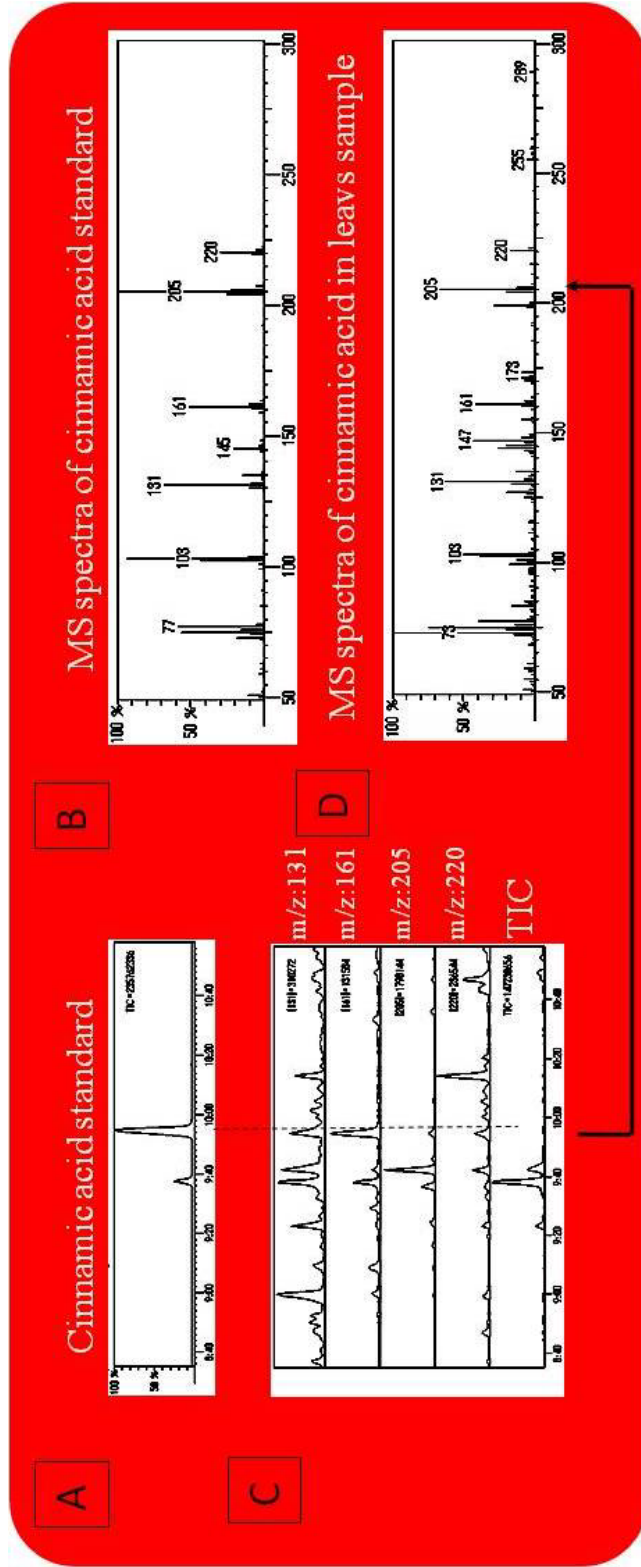


Fig. 4. 3. GC-MS analysis of the presence of cinnamic acid in leaf extracts of *Helichrysum aureonitens*. **A:** Cinnamic acid standard chromatogram; **B:** MS chromatograph of cinnamic acid standard; **C:** Total ion chromatogram. **D:** MS chromatograph of arrowed pick in TIC chromatogram.

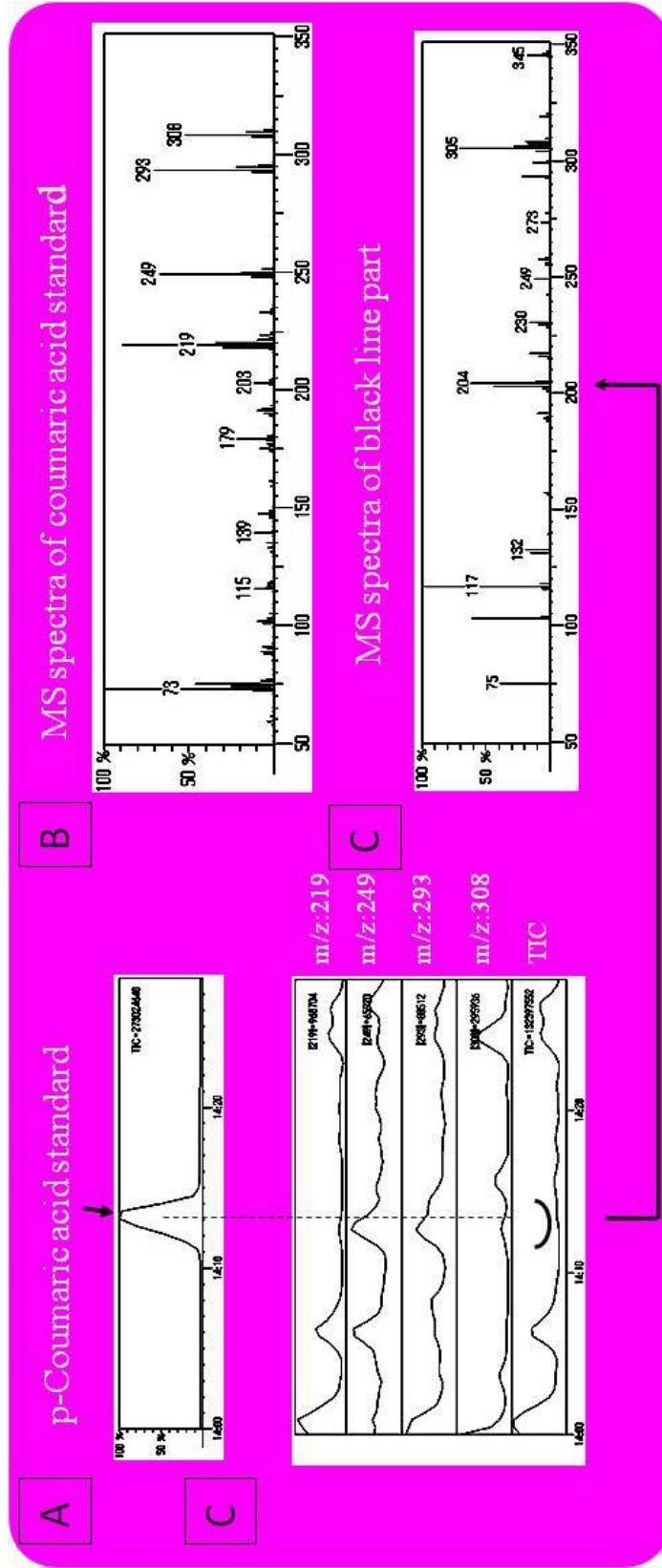


Fig. 4. 4. GC-MS analysis of the absence of *p*-coumaric acid in leaf extracts of *Helichrysum aureonitens*. **A:** *p*-coumaric acid standard chromatogram; **B:** MS chromatogram of *p*-coumaric acid standard; **C:** Total ion chromatogram. **D:** MS chromatogram of arrowed pick in TIC chromatogram.

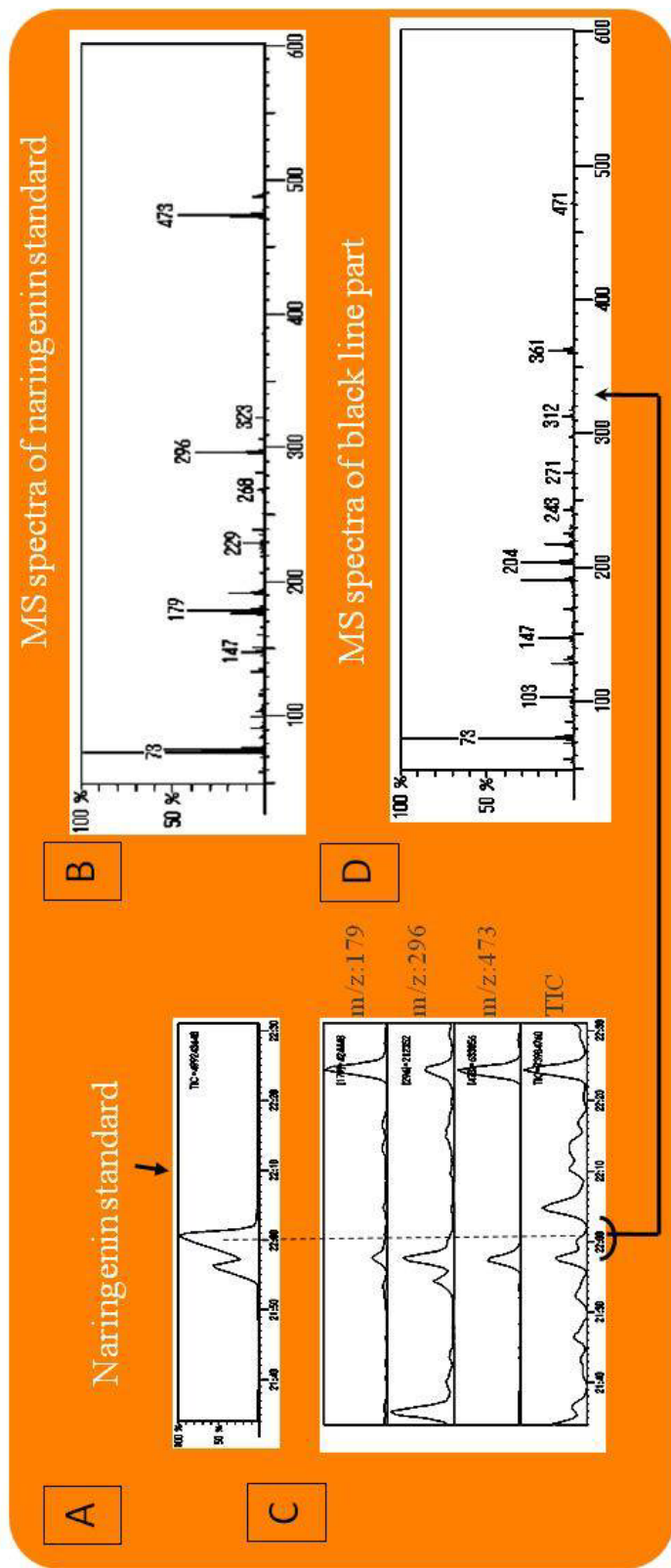


Fig. 4. 5. GC-MS analysis of the absence of naringenin in leaf extracts of *Helichrysum aureonitens*. **A:** Naringenin standard chromatogram; **B:** MS chromatogram of naringenin standard; **C:** Total ion chromatogram. **D:** MS chromatogram of arrowed pick in TIC chromatogram.

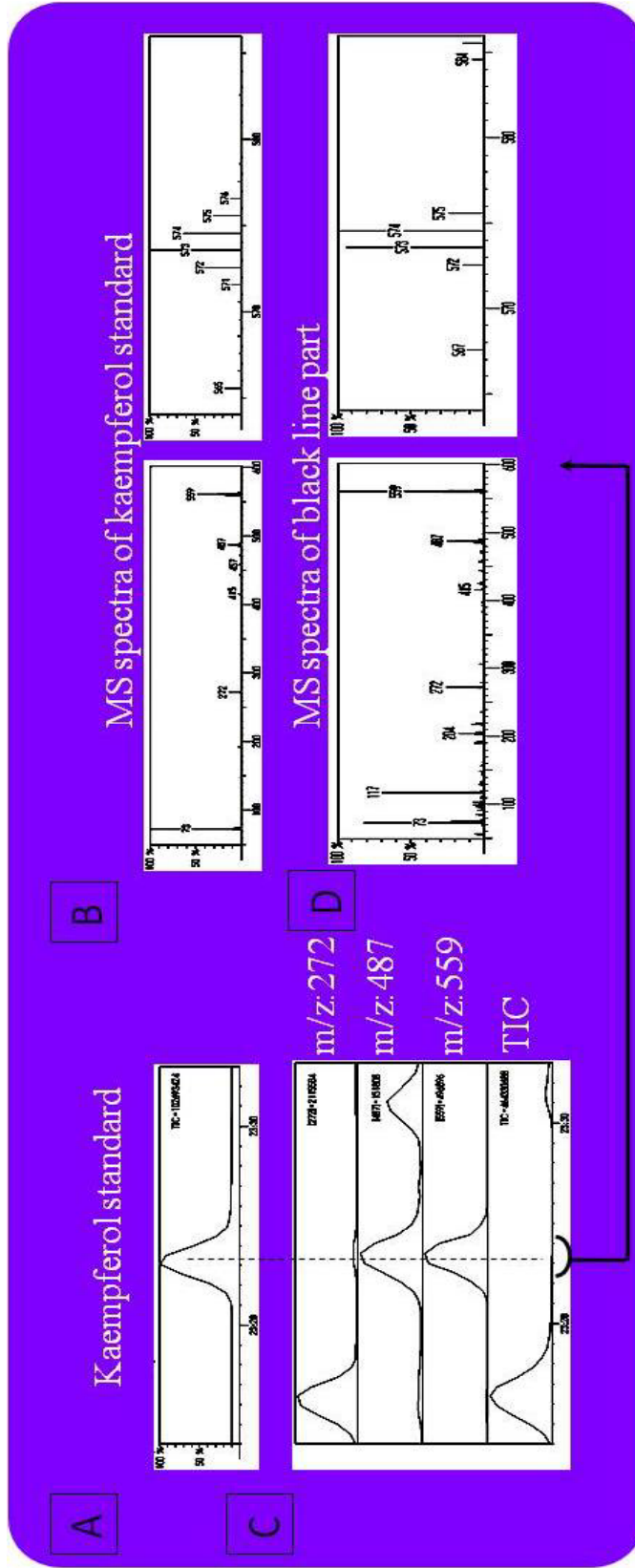
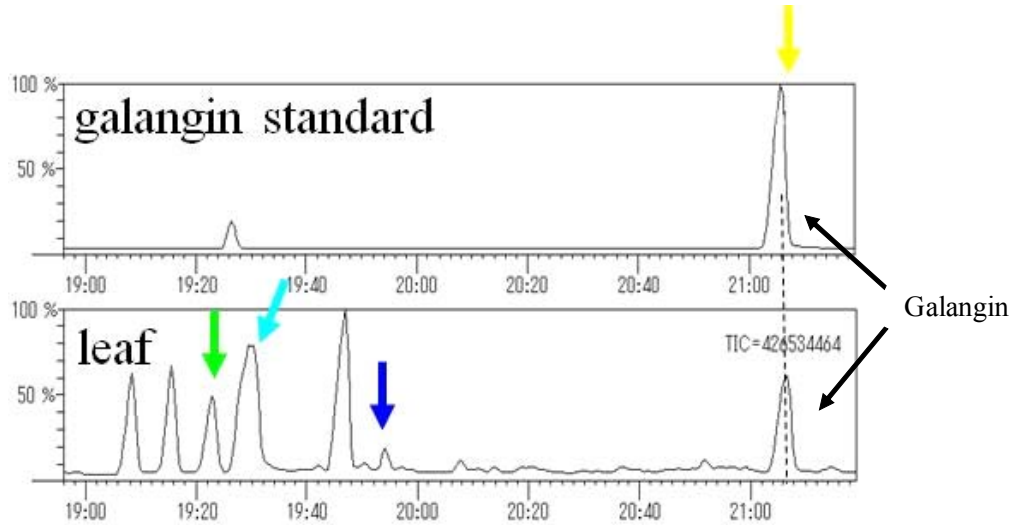


Fig. 4. 6. GC-MS analysis of the presence of kaempferol in leaf extracts of *Helichrysum aureomitens*. **A:** Kaempferol standard chromatogram; **B:** MS chromatogram of kaempferol standard; **C:** Total ion chromatogram. **D:** MS chromatogram of arrowed pick in TIC chromatogram.

A



B

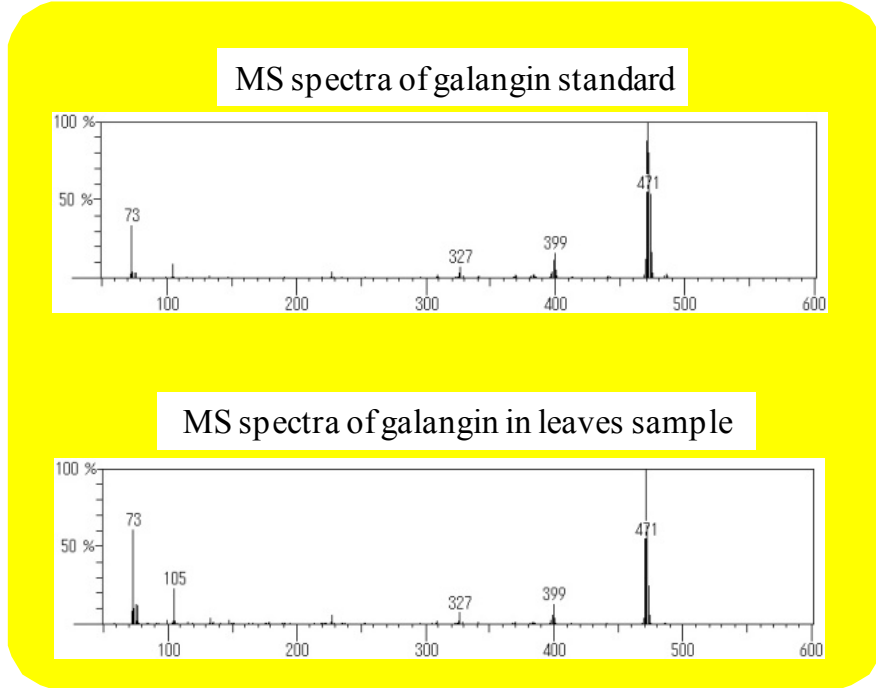


Fig. 4. 7. GC-MS analysis of the presence of galangin in leaf extracts of *Helichrysum aureonitens*. Arrows show the detection of galangin and the precursors (see Fig. 4.8) of the galangin production pathway in leaf extracts by GC-MS. **A:** GC results of standard galangin and detected galangin in the leaf sample; **B:** MS spectra of standard galangin and galangin content in leaf extracts.

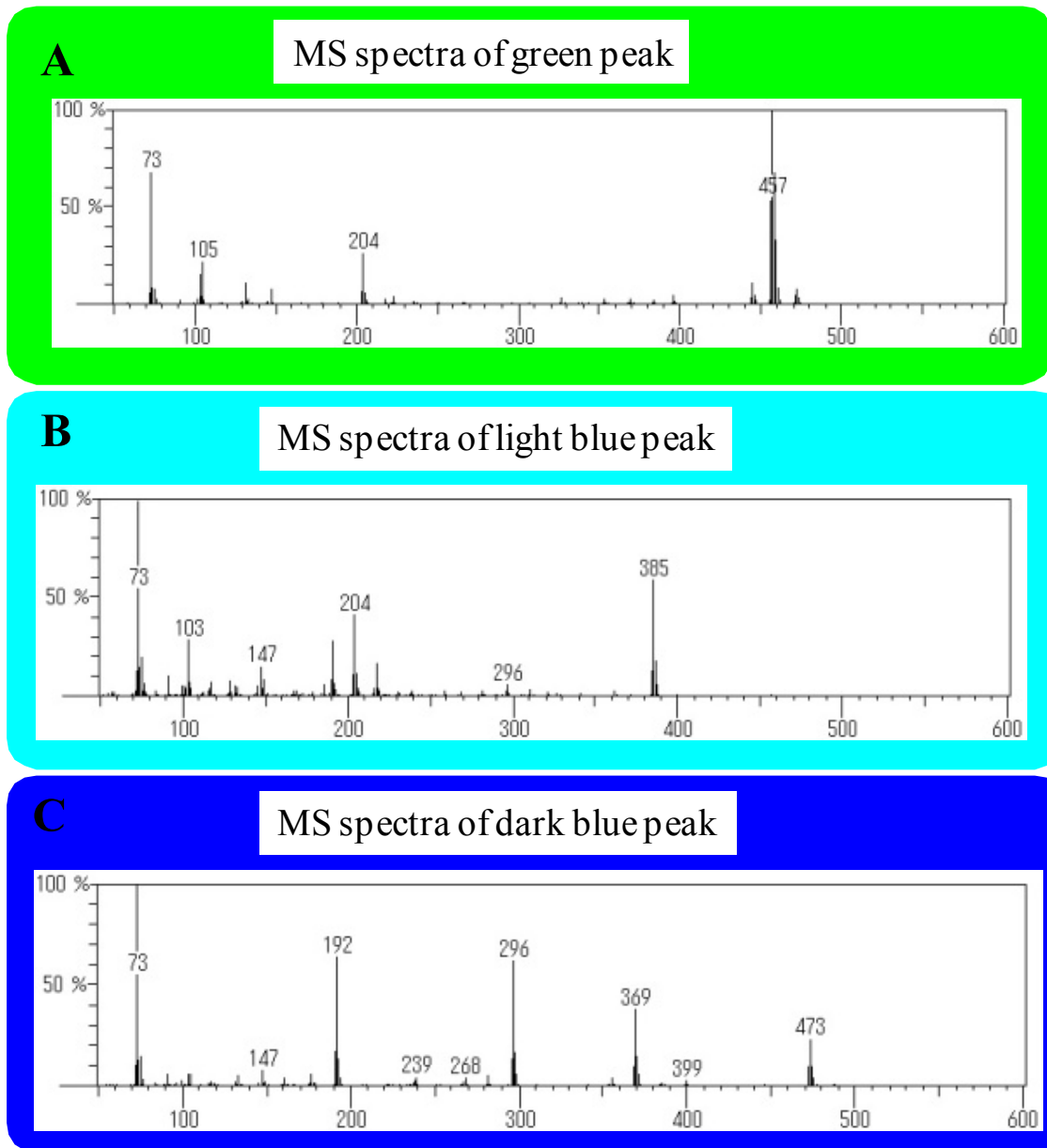


Fig. 4. 8. MS spectra of intermediates (Fig. 4.7) of the non 4'-OH biosynthetic pathway for flavonol production in *Helichrysum aureonitens*. **A:** Pinocembrin chalcone; **B:** Pinocembrin and **C:** Pinobanksin.

4.5 Discussion

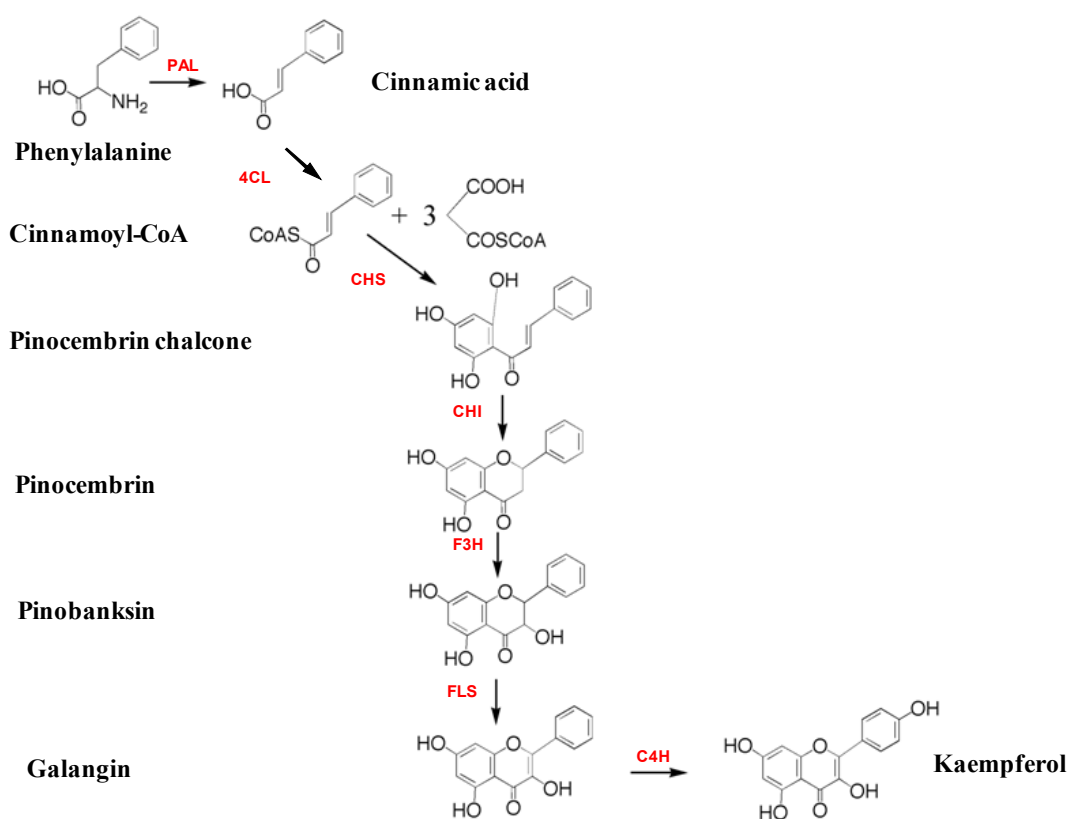
According to scheme 4.1, the production of flavonols are derived from two pathways, one originates from *p*-coumaric acid, with OH at C-4' on the B ring (Fig. 4.2) and the other one originates from cinnamic acid, with a H at the C-4' (Fig. 4.3). Naringenin and pinocembrin from the two separate pathways have been isolated from the bark of *Choerospondias axillaries* (Li et al., 2005), leaves of *Glycyrrhiza glabra* (Biondi et al., 2005), the resinous exudates of *Heliotropium sinuatum* (Torres et al., 1996), the stem extracts of *Viscum liquidambaricolum* (Liu et al., 1993), *Eriodictyon californicum* (Liu et al., 1992), *Artemisia campestris* (Hurabielle et al., 1982) *Prunus cerasus* and *P. domestica* (Nagarajan and Pamar, 1977a; Nagarajan and Pamar, 1977b). In the non 4'-OH pathway cinnamic acid should be converted to cinnamoyl-CoA by the combination with acetyl-CoA. The enzyme that catalyzes this reaction would be *p*-coumarate: CoA-ligase (4CL). One of the questions that can be asked is: Would 4CL accept cinnamic acid instead of *p*-coumaric acid? It has already been reported that 4CL can be activated by cinnamate in some plants (Forkman and Heller, 1999). In cactus elicited cell cultures (*Cephalocereus senilis*), 4CL was activated with cinnamic and *p*-coumaric acids (Liu et al., 1995). Kaneko et al. (2003) also reported 4CL from a Gram-positive filamentous bacterium (*Streptomyces coelicolor* A3) which can convert cinnamic acid to cinnamoyl-CoA, as well as *p*-coumaric acid to *p*-coumaroyl-CoA. It has also been proved that the chalcone synthase (CHS) enzyme is active not only with *p*-coumaroyl-CoA (leading to naringenin), but also with cinnamoyl-CoA (leading to pinocembrin) (Fliegmann et al., 1992; Paré et al., 1992). Chalcone isomerase is the next enzyme involved in the biosynthesis of flavonones and has been reported in the actinomycete, *Streptomyces coelicolor* where it was shown that CHI can also accept pinocembrin chalcone and naringenin chalcone as substrate to produce pinocembrin and naringenin respectively (Liu et al., 1995). In other research on the production of flavonones in *Escherichia coli* an artificial gene cluster

which contained three genes of heterologous origins including PAL, 4CL, CHS and CHI were constructed. *E. coli* cells containing this gene cluster produced two flavonones, pinocembrin from phenylalanine (with 4'-H) and naringenin from tyrosine (with 4'-OH). The same was true for chalcone isomerase (CHI) (Miyahisa et al., 2005). This demonstrates the possibility of the biosynthesis of flavonols from two separate biosynthetic pathways originating from different substrates, cinnamic acid and *p*-coumaric acid.

Based on this information it could be suggested that in some species both pathways are active. According to the results obtained in this study, it is proposed that in the leaves of *H. aureonitens*, the non 4'-OH biosynthetic pathway for the production of galangin and other flavonol compounds is active or that some parts of the 4'-OH pathway is inactive. In the non-4'-OH pathway, cinnamic acid would be the substrate to produce cinnamoyl-CoA directly in a reaction catalyzed by the 4CL enzyme. This means that in this pathway, C4H and *p*-coumaric acid are not involved, but other enzymes, including phenylalanine ammonia lyase (PAL), *p*-coumarate: CoA-ligase (4CL), chalcone synthase (CHS), chalcone isomerase (CHI), flavonone 3-hydroxylase (F3H) and flavonol synthase (FLS) which have similar functions in the 4'-OH pathway, are involved (Scheme 4.3).

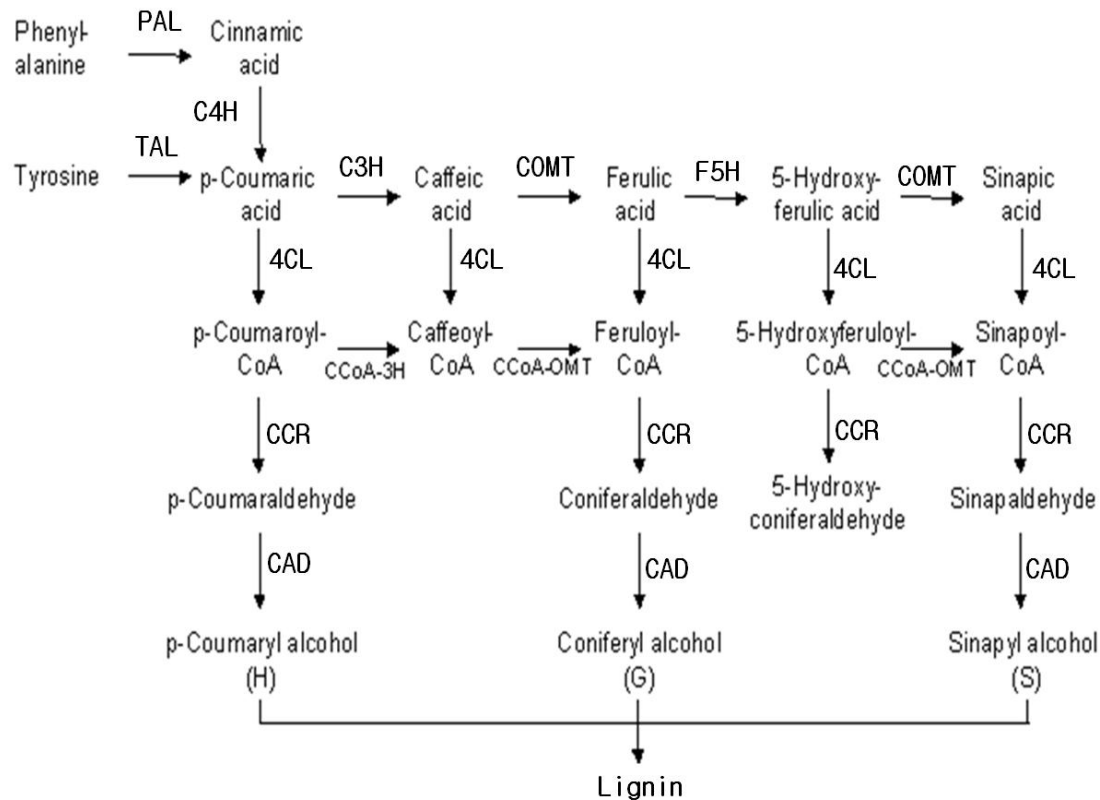
Since we could detect kaempferol in the *H. aureonitens* leaf extracts, the question of whether the proposed pathway for the production of galangin can also produce kaempferol, arose. Although the detection of galangin and kaempferol have been reported in several plants such as *Milletia recemosa* (Ganapaty et al., 1999), *Populus deltoids* and *P. balsamifera* (Kurkin et al., 1990; Kurkin et al., 1990), to date no evidence has been reported to show the conversion of intermediates from the non 4'-OH pathway to the relevant hydroxylated compounds in the 4'-OH pathway. According to the literature, the only enzyme that catalyzes the hydroxylation at the 4' position is cinnamate 4-hydroxylase (C4H) (Boudet, 2003). Based on the results obtained in this part of the study, it is therefore hypothesized that the C4H with the same

function (hydroxylation at 4' position on the B ring) might link the non 4'-OH to the 4'-OH biosynthetic pathways by hydroxylation of galangin to the corresponding intermediate, kaempferol. It is also possible that C4H can make different links between the non 4'-OH and the 4'-OH biosynthetic pathways by hydroxylation of the several intermediates from non 4'-OH pathway to the corresponding intermediates in the 4'-OH biosynthetic pathway (Scheme 4.4).



Scheme 4. 4. Proposed biosynthetic pathway of kaempferol in the leaves of *Helichrysum aureonitens*.

If *p*-coumaric acid is not present in *H. aureonitens* the question arises as to how lignin is synthesized in this species. Briefly, high-resolution microautoradiography of developing hardwood xylem has revealed that three monolignol units are incorporated at different stages of cell wall formation. H (*p*-hydroxyphenyl) units are incorporated in the earliest stages mainly in the cell corner and middle lamella during the period of formation of the S1 layer. The deposition of G lignin continues through the early to late stages (Dixon et al., 2001) (Scheme 4.5). All these branches originate from *p*-coumaric acid. The absence of *p*-coumaric acid in *H. aureonitens* brings up this question as to how lignin is synthesized in this plant?



Scheme 4. 5. The lignin biosynthetic pathway. The lignin biosynthetic enzymes are: PAL phenylalanine ammonia-lyase; TAL tyrosine ammonia-lyase; C4H cinnamate 4-hydroxylase; C3H 4-hydroxycinnamate 3-hydroxylase; COMT caffeic acid 3-O-methyltransferase; F5H ferulate 5-hydroxylase; 4CL 4-coumarate: CoA ligase; CCoA-3H coumaroyl-coenzyme A 3-hydroxylase; CCoA-OMT caffeoyl-coenzyme A O-methyltransferase; CCR cinnamoyl-CoA reductase; and CAD cinnamyl alcohol dehydrogenase (Modified from Boudet et al., 1996; Campbell and Sederoff, 1996).

Although researchers have studied lignin for more than a century, many aspects of its biosynthesis remain unresolved. It is also becoming increasingly clear that lignins are derived from several more monomers than just the three monolignols, as shown in scheme 4.5. Many “normal” plants contain lignins substantially derived from other monomers, and all lignins contain traces of units from apparently incomplete monolignol biosynthesis and other (side-) reactions that occur during that biosynthesis (Ralph et al., 2001). Many of these units have recently been identified in transgenic and mutant plants. For example, researchers have long thought that the hydroxylation and methylation reactions occur at the level of the cinnamic acids and that *p*-coumaric, ferulic, and sinapic acid are subsequently converted to the corresponding monolignols by 4-coumarate:CoA ligase (4CL), CCR, and CAD respectively. There is, however, a number of *in vitro* enzymatic assays with heterologously produced enzymes in which the identification of novel genes implicated in the pathway, and analyses of mutant and transgenic plants modified in monolignol biosynthesis have cast doubt on this route, and the pathway had to be redrawn (Humphreys and Chapple, 2002).

Another level of complexity that may affect the synthesis pathway is about down-regulation of PAL and C4H in tobacco which largely reduces lignin content (Elkind et al. 1990; Bate et al. 1994; Sewalt, et al. 1997). PAL down-regulation reduces mainly G units, whereas C4H down-regulation reduces mainly S units in lignin, an observation that cannot easily be explained by the pathway presented in scheme 4.5.

Dixon et al. (2001) believe that three possible explanations can be proposed for this apparent contradiction: (a) the pathway to G lignin may bypass C4H (Fig. 4.9), (b) C4H still catalyzes reactions other than the 4-hydroxylation of cinnamic acid, or (c) C4H may be part of a metabolic channel committed to S lignin biosynthesis. It is indicated that *p*-coumaric acid might not be the only substrate for the lignin pathway. In fact, the results of this study on *H.*

aureonitens confirms their proposed pathway about lignin production, as *p*-coumaric acid was not detected.

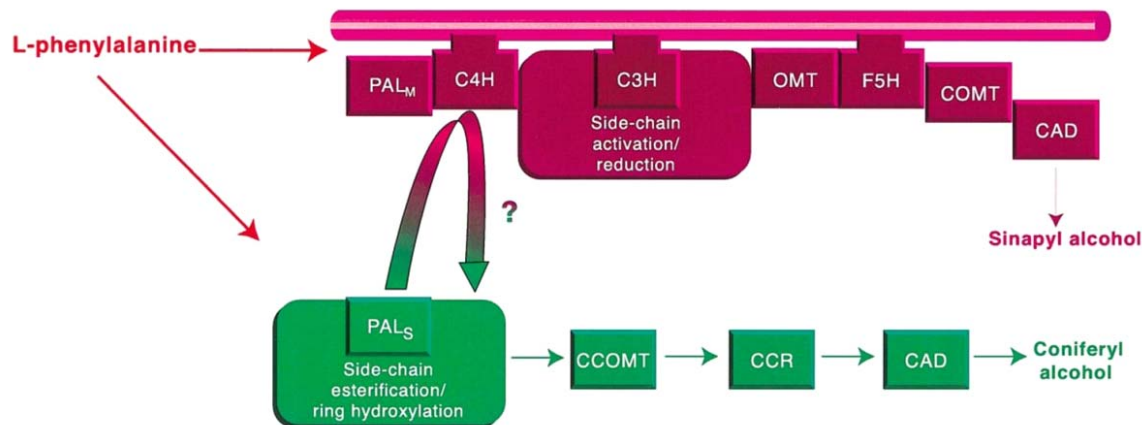


Fig. 4. 9. A metabolic channel model for independent pathways to G and S monolignols (Dixon et al., 2001). The hypothetical metabolic channel leading to S lignin (in red) is proposed to involve a microsomal-associated form of PAL (PAL_M) that directly channels its product, cinnamic acid, to the ER-associated cinnamate 4-hydroxylase (C4H) cytochrome P450. 3- Hydroxylation (possibly via a second P450 enzyme) and CoA esterification and reduction of the side chain then occur, although it is not clear whether hydroxylation occurs before or after side chain modification. These reactions are followed by the first O-methylation reaction (by COMT?) that produces coniferaldehyde, the substrate for the ER-associated F5H. The second methylation reaction is then catalyzed by COMT (perhaps associated with F5H) to yield sinapyl aldehyde, which is reduced by CAD to yield sinapyl alcohol. The hypothetical pathway to G units (in green) is proposed to involve a soluble form of PAL (PAL_S), that either provides cinnamate for C4H in an unchanneled reaction, or bypasses C4H by hydroxylation of the aromatic ring at the level of a side chain esterified intermediate. CCOMT then catalyzes the single O-methylation reaction, and coniferyl alcohol is formed via soluble forms of CCR and CAD. In an alternative model, the pathway to G units might also be membrane associated.

4.6 Conclusions

The results obtained indicated that *p*-coumaric acid and naringenin from the 4'-OH biosynthetic pathway of flavonols are not present, while cinnamic acid, pinocembrin-chalcone, pinocembrin, pinobanksin, galangin and kaempferol were detected in the leaf extracts of *H. aureonitens*. According to these results, we hypothesize that the non 4'-OH biosynthetic pathway for the production of galangin and kaempferol is present in *H. aureonitens*. Since the only enzyme known that can catalyze the hydroxylation at this specific position is C4H, we hypothesize that C4H is present in the leaf extracts of *H. aureonitens* and it is responsible for catalyzing hydroxylation reaction (Scheme 4.4). It is possible that C4H can possibly link the non 4'-OH pathway with the 4'-OH pathway by the hydroxylation of different intermediates. To confirm this hypothesis firstly the presence of C4H in *H. aureonitens* must be proven by the isolation and characterization of the C4H gene, and secondly the functionality of it should be determined.