Phytochemistry, 1980, Vol. 19, pp. 701–704. © Pergamon Press Ltd. Printed in England. 0031-9422/80/0401-0701$02.00/0

PAVOPHYLLINE, A NEW SAPONIN FROM THE STEM OF *PAVONIA ZEYLANICA*

K. P. Tiwari and P. K. Minocha

Department of Chemistry, University of Allahabad, Allahabad, India

(*Received* 23 *July* 1979)

Key Word Index—*Pavonia zeylanica;* Malvaceae; pavophylline; lup-20(29)-en-28-oic-3-*O*-β)D-glucopyranosyl(4 → 1)-*O*-α-L-rhamnopyranosyl(2 → 1)-*O*-α-L-arabino furanoside.

INTRODUCTION

Pavonia zeylanica (N.O. Malvaceae) is a small tree found in North India, West Rajasthan, West Peninsula, Mauritius, Sind etc. In Zambia the plant is used as vermifuge and a purgative by Mandingo natives [1]. Since no work seems to have been previously done on the chemical analysis of *Pavonia zeylanica*, in spite of its medicinal properties, we have therefore examined the stem of this plant in order to provide a detailed study of its constituents and have isolated a new ketoester [2] and saponin, pavophylline. This paper presents the structure of pavophylline.

RESULTS AND DISCUSSION

Pavophylline, $C_{47}H_{76}O_{16}$, responded to colour reactions characteristic of saponins [3], gave a copious lather when shaken with water and heomolysed red blood cells. It was hydrolysed with 7% H_2SO_4 where upon the sapogenin precipitated out. It was separated from the hydrolysate by filtration and purified by the usual potassium salt method [4].

Structure of the sapogenin

The sapogenin was characterized as 3β-hydroxy-lup-20(29)-en-28-oic acid (betulinic acid) by comparison of IR, ^{1}H NMR and MS data of the sapogenin, its acetyl and methyl derivatives with those reported in the literature [5–9] and by mmp and co-TLC with an authentic sample.

Characterization of the sugar moiety

The presence of D-glucose, L-rhamnose and L-arabinose in the hydrolysate was confirmed by co-PPC with authentic sugar samples (*n*-BuOH–HOAc–H_2O, 4:1:5; spray—aniline hydrogen phthalate, R_f 0.18, 0.20 and 0.37, respectively). Quantitative hydrolysis of pavophylline indicated the genin content was *ca* 48% and the sugar moiety 52%. A quantitative estimation [10] of the sugar present in the hydrolysate revealed that the three sugars were present in equimolecular proportions. Therefore, it was concluded that pavophylline contained 1 mol each of sapogenin, D-glucose, L-rhamnose and L-arabinose.

Position of attachment of sugars to the sapogenin

In betulinic acid, only the OH at C-3 and the COOH at C-17 are available for glycosidic linkage with sugar residues. From the following observations, it was concluded that all three sugars were linked as a trioside unit to the C-3 hydroxyl group of the sapogenin. (i) The saponin was not hydrolysed with 5 N NH_4OH, which is a specific reagent [11] for hydrolysis of sugar esters, without attacking other glycosidic linkages. Thus, sugars were not present in the ester combination with the COOH group of the sapogenin. (ii) On methylation with CH_2N_2, the saponin furnished a methyl ester, $C_{48}H_{78}O_{16}$, which on hydrolysis with 7% H_2SO_4 yielded betulinic acid methyl ester but no betulinic acid and thus indicating that there was a free COOH group in the genin.

Sequence of sugars in pavophylline

The saponin on partial hydrolysis (see Experimental) [12] yielded a mixture of two prosapogenins designated as PS_1 and PS_2 which were separated by a column of Si gel using methanol as solvent. The hydrolysate on paper chromatographic examination revealed the presence of L-arabinose and L-rhamnose (identified by co-PPC with authentic samples).

Study of the prosapogenin PS_1

PS_1 on hydrolysis with 7% H_2SO_4 yielded betulinic acid and D-glucose. The genin content was found to be 72% indicating that PS_1 was a monoglucoside and was formed from 1 mol each of betulinic acid and D-glucose. Permethylation of PS_1 followed by hydrolysis showed the presence of 2,3,4,6-tetra-*O*-methyl-D-glucose in the hydrolysate (identified by co-PPC with an authentic sample) indicating that C-1 of D-glucose was involved in the formation of a glycosidic linkage and also suggesting that D-glucose was present as a pyranoside. Thus PS_1 may be assigned the structure lup-20(29)-en-28-oic-3-*O*-β-D-glucopyranoside.

Study of the prosapogenin PS_2

PS_2 on hydrolysis with 7% H_2SO_4 yielded betulinic acid, D-glucose and L-rhamnose. The genin content was 59.6% and the ratio of the sugars in PS_2 was found to be 1:1 by colorimetric estimation [10] suggesting that PS_2 was a bioside of betulinic acid and was formed from 1 mol each of betulinic acid, D-glucose and L-rhamnose. Permethylation of PS_2 followed by hydrolysis and chromatographic examination of the hydrolysate showed the presence of 2,3,4-tri-*O*-methyl-L-rhamnose and 2,3,6-tri-*O*-methyl-D-glucose. Since the structure for PS_1 had already been established, the formation of 2,3,6-tri-*O*-methyl-D-glucose and 2,3,4-tri-*O*-methyl-L-rhamnose can only be explained by assuming that C-4 of the glucose moiety was involved in the formation of the glycosidic linkage

with C-1 of L-rhamnose. The formation of 2,3,4-tri-O-methyl rhamnose further suggested that L-rhamnose was present as a pyranoside. Thus the glycosidic linkages in PS_2 were assigned the following configuration: lup-20(29)-en-28-oic-3-O-β-D-glucopyranosyl(4 → 1)-O-α-L-rhamnopyanoside.

Nature and position of the glycosidic linkages in pavophylline

Pavophylline, on complete permethylation followed by hydrolysis formed 2,3,5-tri-O-methyl-L-arabinose, 3,4-di-O-methyl-L-rhamnose and 2,3,6-tri-O-methyl-D-glucose (identified by co-PPC). The release of 3,4-di-O-methyl-L-rhamnose and 2,3,5-tri-O-methyl-L-arabinose indicated that in the saponin molecule C-1 of L-arabinose was linked to C-2 of L-rhamnose. The easy elimination of L-arabinose and the release of a 2,3,5-tri-O-methyl-L-arabinose unit from the saponin revealed the arabinose was present as a furanoside thus leaving only the anomeric hydroxyl group for glycosidic linkage with the L-rhamnose unit. Pavophylline on enzymatic hydrolysis with diastase yielded L-rhamnose and L-arabinose. Therefore, these two sugars were involved in an α-glycosidic linkage. The exact configuration of the sugar linkages in the saponin was established by consideration of the molecular rotation value in the light of Klyne's rule [13] and the four possible combinations of the sugar linkages are shown in Table 1.

Table 1.

β-D-glucose + α-L-rhamnose + α-L-arabinose	−66	−110	−205 = −381
β-D-glucose + α-L-rhamnose + β-L-arabinose	−66	−110	−77 = −253
β-D-glucose + β-L-rhamnose + α-L-arabinose	−66	+168	−205 = −103
β-D-glucose + β-L-rhamnose + β-L-arabinose	−66	+168	−77 = −25

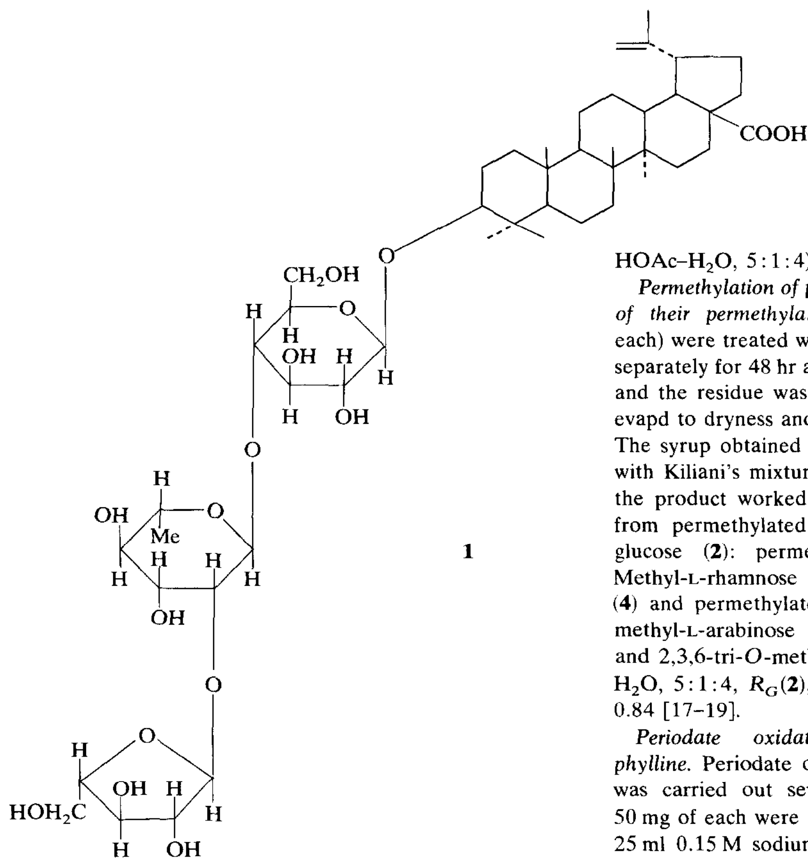
The observed M_D value [13] for the saponin was −319°. The M_D value of the genin is known to be +63.8°. The difference −383° is closed to the first combination of Table 1. Therefore, the exact configuration of the sugar linkages was D-glucose-β, L-rhamnose-α and L-arabinose-α. After elucidating the nature of the glycosidic linkages and position of attachment of all the sugars in the saponin, a complete structure of pavophylline may be assigned by taking into consideration the elimination of L-arabinose during partial hydrolysis of saponin which indicated the arabinose was present as a furanoside [14]. Furthermore, release of L-arabinose as 2,3,5-tri-O-methyl-L-arabinose on hydrolysis of permethylated pavophylline, 2,3,4-tri-O-methyl-L-rhamnose from the hydrolysis of permethylated PS_2 and 2,3,4,6-tetra-O-methyl-D-glucose after hydrolysis of permethylated PS_1 clearly suggested that the glucose and L-rhamnose moieties were present as pyranosides and L-arabinose as a furanoside. The above facts were confirmed by the periodate oxidation [15] of pavophylline, PS_1 and PS_2. Hence pavophylline is lup-20(29)-en-28-oic-3-O-β-D-glycopyranosyl(4 → 1)-O-α-L-rhamnopyranosyl-(2 → 1)-O-α-L-arabinofuranoside (**1**).

EXPERIMENTAL

Extraction. The defatted powdered plant material (6 kg) was exhaustively extracted with EtOH. The EtOH extract (7.5 l.) was concd (4 l.) and kept overnight in a refrigerator when a colourless ppt. was observed. The ppt. was separated by filtration and identified as methyl-19-tetracosanoate [2]. The filtrate was further concd and the residual amount of EtOH was removed *in vacuo*. The residue thus obtained was successively washed with Et_2O, $CHCl_3$ and Me_2CO, finally dissolved in MeOH, filtered and the filtrate poured into excess Et_2O whereby a light brown mass was pptd. The ppt. was separated by filtration and purified by repeating the above process of dissolution in MeOH and precipitation with Et_2O. It was recrystallized from MeOH to yield microcrystals (5.1 g) of pavophylline, mp 230°. The purity of pavophylline was checked by PPC (BuOH–HOAc–H_2O, 4:1:5, spray 25% CCl_3.COOH in Et_2O: Yellow spot, R_f 0.58). (Found: C, 63.12; H, 8.42. $C_{47}H_{76}O_{16}$ requires: C, 62.94; H, 8.48%).

Isolation and study of sapogenin from pavophylline. Pavophylline (2.9 g) was hydrolysed by refluxing with 7% H_2SO_4 (250 ml) for 6 hr on a steam both. The aglycone part was separated from the aq. hydrolysate and was purified by the K salt method [4]. It was crystallized from $CHCl_3$ into colourless crystals, mp 315–316°, $[\alpha]_D^{27}$ + 14° (Py). The purity of the sapogenin was checked by TLC (C_6H_6–Me_2CO, 1:1, spray $SbCl_3$ in $CHCl_3$, R_f 0.55). (Found: C 78.88, H, 10.54; MW (456 (MS). $C_{30}H_{48}O_3$ requires: C, 78.94; H, 10.52%). IR ν_{max}^{KBr} cm^{-1}: 3448, 2933, 2850, 1700, 1641, 1460, 1392, 1381, 1370, 1360, 1295, 1274, 885; MS *m/e*: 456 (M^+), 441, 438, 423, 411, 248, 220, 219, 207 and 189 (base peak). UV λ_{max}^{EtOH} nm: 250. Acetate: $C_{30}H_{47}O_2$.$OCOCH_3$, mp 290–291°. IR ν_{max}^{KBr} cm^{-1}: 2940, 2855, 1739, 1704, 1642, 1240, 885. Methyl ester; $C_{31}H_{50}O_3$, mp 220–222°, IR ν_{max}^{KBr} cm^{-1}: 3540, 1718, 1641, 1460, 1362, 882; ^{1}H NMR (in $CDCl_3$): δ 0.75 (3H), 0.77 (3H), 0.90 (3H), 1.00 (3H), 1.15 (3H), 1.75 (3H), 3.62 (3H), 4.70 (2H); MS *m/e*: 470 (M^+), 455, 452, 437, 411, 262, 233, 220, 207 and 189 (base peak).

Quantitative colorimetric estimation of sugars in saponin hydrolysate. The ratio of sugars in the saponin was determined colorimetrically [11] in a Klett–Summerson photoelectric colorimeter using green filter 470 nm with the help of standard curves of authentic sugars. Then solns (5, 10, 15–50 μg in 0.03 ml H_2O) of each of three sugars, D-glucose, L-arabinose and L-rhamnose were applied on Whatman No. 1 filter paper (50 × 55 cm, spot distance 4 cm). The chromatograms were developed by descending technique with BuOH–HOAc–H_2O (4:1:5) for 24 hr, dried in air, sprayed with aniline hydrogen phthalate on both the sides, and dried at 110° for 15 min. The coloured spots were cut out in equal rectangles eluted by immersion in 50% HOAc (10 ml each) and the colour intensity of each eluate measured.



Estimation of sugars in saponin hydrolysate. The saponin (80 mg) was hydrolysed by refluxing with 7% H_2SO_4 (30 ml) for 6 hr on a steam bath. The reaction mixture was extracted with $CHCl_3$ to yield the sapogenin (36.2 mg). The hydrolysate was neutralized with $BaCO_3$, filtered and concd to a syrup (2 ml). 0.2 ml of this syrup was dissolved in 2 ml of H_2O and aliquots applied to Whatman No. 1 filter paper. The chromatograms were developed, sprayed, dried and the coloured spots were cut out in equal rectangles, eluted separately and assayed as described above.

Methyl ester of the pavophylline. Pavophylline (300 mg) in MeOH was treated with an Et_2O soln of CH_2N_2 until a permanent yellow colour was obtained. The reaction mixture was kept overnight at 0° and the product isolated and crystallized from $CHCl_3$–MeOH (1:2) into light yellow crystals, mp 145°. (Found C, 62.78; H, 8.34, $C_{48}H_{78}O_{16}$ requires: C, 63.28; H, 8.57%).

Hydrolysis of methyl ester of pavophylline. Pavophylline methyl ester (100 mg) was refluxed with 7% ethanolic H_2SO_4 (20 ml) on a steam bath for 4 hr. EtOH was removed and the reaction mixture diluted with H_2O. The ppt. was separated by filtration and crystallized from MeOH, mp 221–222°. (Found: C, 78.88; H, 10.54. $C_{31}H_{50}O_3$ requires: C, 79.14; H, 10.64%).

Partial hydrolysis of pavophylline: isolation of PS_1 *and* PS_2. Pavophylline (650 mg) was treated with 0.02 N H_2SO_4 and the reaction mixture was kept at room temp. for 8 days. It was then extracted with BuOH. The BuOH extract, after concn, was chromatographed over a column of Si gel. Elution with MeOH yielded two substances designated as PS_1 (190 mg) and PS_2 (225 mg) which were separately crystallized from MeOH. The hydrolysate contained L-rhamnose and L-arabinose (PPC with their authentic samples, BuOH–HOAc–H_2O, 5:1:4).

Permethylation of pavophylline, PS_1 *and* PS_2 *and hydrolysis of their permethylated derivatives.* The glycosides (70 mg each) were treated with MeI (2 ml) and Ag_2O in DMF (4 ml) separately for 48 hr at room temp. The mixtures were filtered and the residue washed with a little DMF. The filtrate was evapd to dryness and the residue taken up in EtOH (25 ml). The syrup obtained after removal of EtOH was hydrolysed with Kiliani's mixture (HOAc–HCl–H_2O, 7:3:10) [16] and the product worked up in the usual way. The hydrolysate from permethylated PS_1 contained 2,3,4,6-tetra-*O*-methyl glucose (**2**): permethylated PS_2 contained 2,3,4-tri-*O*-Methyl-L-rhamnose (**3**) and 2,3,6-tri-*O*-methyl-D-glucose (**4**) and permethylated pavophylline contained 2,3,5-tri-*O*-methyl-L-arabinose (**5**), 3,4,-di-*O*-methyl-L-rhamnose (**6**) and 2,3,6-tri-*O*-methyl-D-glucose (**4**). (PPC, BuOH–EtOH–H_2O, 5:1:4, R_G(**2**), 1.00; (**3**) 1.01; (**4**) 0.83; (**5**) 0.95; (**6**) 0.84 [17–19].

Periodate oxidation of PS_1, PS_2 *and pavophylline.* Periodate oxidation of PS_1, PS_2 and pavophylline was carried out separately by the method of ref. [15]. 50 mg of each were separately dissolved in 25 ml EtOH and 25 ml 0.15 M sodium metaperiodate soln were added. The oxidation was allowed to take place at room temp. for 60 hr. Aliquots (5 ml) were withdrawn in duplicate from the reaction mixture at different time intervals and analysed for periodate and formic acid.

Acknowledgements—The authors are indebted to Drs. D. S. Bhakuni (CDRI, Lucknow), R. S. Grewal and S. Selvavinayakam (CIBA-GEIGY Research Centre, Bombay) for their kind help in the spectral studies of the compounds reported in this paper. One of the authors (P.K.M.) is grateful to the U.G.C. New Delhi for a junior fellowship.

REFERENCES

1. Kirtikar, K. R. and Basu, B. D. (1933) *Indian Med. Plant,* **3,** 324.
2. Tiwari, K. P., Minocha, P. K. and Masood, M. (1978) *Proc. Natl. Acad. Sci. India Sect. A* **48,** 157.
3. Sannie, C. H. (1948) *Analyt. Biochem. Med.* **9,** 175.
4. Dodge, F. D. (1918) *J. Am. Chem. Soc.* **40,** 1917.
5. Shamma, M., Glick, R. E. and Mumma, R. O. (1962). *J. Org. Chem.* **27,** 4512.
6. Bhattacharjee, S. R. and Chatterjee, A. (1962) *J. Ind. Chem. Soc.* **39,** 276.
7. Heilbron, I. and Bunbury, H. M. (1953) *Dictionary of Organic Compounds,* Vol. 1, p. 283. Eyre & Spottiswoode, London.
8. Budzikiewicz, H., Wilson, J. M. and Djerassi, C. (1963) *J. Am. Chem. Soc.* **85,** 3688.
9. Budzikiewicz, H., Djerassi, C. and Williams, D. H. (1964) *Structural Elucidation of Natural Products by Mass Spectrometry,* Vol. 2, p. 121. Holden-Day, San Francisco.
10. Misra, S. B. and Rao, V. K. M. (1960) *J. Sci. Ind. Res.* **19C,** 173.

11. Hariharan, V. and Rangaswami, S. (1970) *Phytochemistry* **9**, 409.
12. Dutta, T. and Basu, U. P. (1968) *Indian J. Chem.* **6**, 471.
13. Klyne, W. (1950) *Biochem. J.* **47**, 4.
14. Jones, J. K. N. and Smith, F. (1946) in *Advances in Carbohydrate Chemistry*, Vol. 2, p. 243. Academic Press, New York.
15. Hirst, E. L. and Jones, J. K. N. (1949) *J. Chem. Soc.* 1659.
16. Kiliani, H. (1930) *Ber. Dtsch. Chem. Ges.* **63B**, 2866.
17. Adams, C. A. (1960) *Can. J. Chem.* **38**, 220.
18. Lederer, E. and Lederer, M. (1957) in *Chromatography*, p. 251. Elsevier, Amsterdam.
19. Williams, S. C. and Jones, J. K. N. (1967) *Can. J. Chem.* **45**, 275.

Phytochemistry, 1980, Vol. 19, pp. 704–705. © Pergamon Press Ltd. Printed in England. 0031-9422/80/0401-0704$02.00/0

N-FERULOYLGLYCYL-L-PHENYLALANINE ISOLATED BY PARTIAL HYDROLYSIS OF BULK LEAF PROTEIN OF LUCERNE, *MEDICAGO SATIVA*, CV EUROPE

CHRISTIAAN F. VAN SUMERE, HILDA J. HOUPELINE-DE, COCK, YOLANDE C. VINDEVAGHEL-DE BACQUER and GODELIEVE E. FOCKENIER

Laboratory of Plant Biochemistry, University of Ghent, K. L. Ledeganckstr., 35, B-9000—Ghent, Belgium

(*Revised received* 30 *July* 1979)

Key Word Index—*Medicago sativa*; Leguminosae; bulk leaf protein; *p*-coumaric acid; ferulic acid; *N*-feruloylglycine; *N*-feruloylglycyl-L-phenylalanine.

Phenolic acids most frequently occur bound as esters, glycosides or amides [1–5]. In addition, a pseudopeptide-type of bonding, in which the carboxylic group of ferulic acid is attached to the amino function of the *N*-terminal amino acid of a polypeptide chain, has been described. This is *N*-feruloylglycyl-L-phenylalanine which was found as a terminal sequence of a barley globulin fraction [1]. The same sequence has now also been recognized in lucerne bulk leaf proteins and this paper describes briefly its identification from this new source.

EXPERIMENTAL

Lucerne bulk leaf proteins were obtained from R. L. M. Synge (Norwich, U.K.). They were extracted with PhOH–HOAc–H_2O (1:1:1) as described [6]. The extract was directly concd, the phenol removed and the concentrate slurried in water and freeze-dried. Before analysis this undialysed protein fraction was first thoroughly extracted with Et_2O in order to remove any remaining free phenolic acids. Thereafter the residue was treated, under N_2, with boiling 2 N NaOH (2 hr). This treatment mainly releases the ester-bound phenolic acids, which, after acidification and subsequent Et_2O extraction, were analysed as described [7]. These analyses showed the presence of 5.40 mg alkali-labile bound ferulic acid and 1.53 mg alkali-labile bound *p*-coumaric acid per g protein material. However, as was shown by pilot experiments synthetic *N*-feruloylglycyl-L-phenylalanine (FGP) [8, 9] (2.5 μmol), proved under the same conditions, to be completely hydrolysed (products formed: *N*-feruloylglycine (FG) [8, 10] (1.25 μmol), ferulic acid (1.25 μmol), glycine (1.25 μmol) and phenylalanine (2.5μmol)). Thus the foregoing alkaline treatment may also result in a partial release of peptide-bound ferulic acid from protein. After alkaline hydrolysis, the remaining protein material was further repeatedly treated (×16) for 15 min periods with boiling 4 N HCl in N_2. Successive partial hydrolyses are required because the ferulic acid-containing peptides, which are formed during hydrolysis, are continuously and rapidly split by boiling acids. Furthermore, liberated ferulic acid is also rapidly converted into several decomposition products [11]. After each hydrolysis, the hydrolysate was cooled and the released peptides and remaining free phenolic acids were thoroughly extracted with EtOAc. These extracts were then pooled from 4 consecutive hydrolyses, washed with H_2O and concd *in vacuo* to a small vol. The resulting concentrates were diluted with 2 ml EtOH and chromatographed by PLC with toluene–HCO_2Et–HCO_2H (5:4:1) on Si gel-cellulose MN (1:1) mixtures [7, 12, 13]. The spot corresponding to the synthetic FGP standard was eluted with EtOH (×6) and then further purified by means of METC (Multiple elimination TLC) [1, 7, 14, 15]. 0.25% NH_4OH, MeOH–H_2O (3:1) (with steaming), *sec*-BuOH–H_2O (4:1) and toluene–HCO_2Et–HCO_2H (5:4:1) (with steaming) were used in sequence. Finally the purified material, showing the same R_f and blue fluorescence (greenish after spraying with 2 N NaOH) as the standard, was removed from the plate and eluted with 0.1 N NaOH. Comparison of the UV spectra, the excitation and fluorescence maxima and the R_f values in four solvents proved the identity of the unknown with synthetic FGP. The combined extract corresponding with the 3rd-hour hydrolysis of lucerne protein contained the highest amount of *N*-feruloylglycyl-L-phenylalanine. Further support for the presence of amino acid-bound ferulic acid in lucerne proteins was obtained by the subsequent isolation of *N*-feruloylglycine