Phytochemistry, 1973, Vol. 12, pp. 407 to 411. Pergamon Press. Printed in England.

N-FERULOYLGLYCYL-L-PHENYLALANINE: A SEQUENCE IN BARLEY PROTEINS

C. F. VAN SUMERE, H. DE POOTER, HAIDER ALI and M. DEGRAUW-VAN BUSSEL

Laboratory of Plant Biochemistry, University of Ghent, Belgium

(*Received* 10 *July* 1972. *Accepted* 9 *September* 1972)

Key Word Index—*Hordeum vulgare*; Gramineae; barley; protein sequence; *N*-feruloylglycyl-L-phenylalanine

Abstract—*N*-Feruloylglycyl-L-phenylalanine was obtained from barley globulins by partial hydrolysis with 4 N HCl. It was isolated by means of preparative and 'multiple elimination' TLC (METC) and further identified by TLC comparison with a synthetic sample. Additional proof for its identity was obtained by UV, fluorescence and IR spectroscopy and by the action of carboxypeptidase A. The possible role of *N*-feruloylglycine as a starter in protein biosynthesis in the barley seed is discussed.

INTRODUCTION

FREE and bound phenolic acids and coumarins have been shown to occur in barley husks[1,2] and isolated barley embryos.[3] Very recently the quantitative extraction and determination of the free compounds via 'multiple elimination' TLC,[3–5] combined with fluorimetry or spectrophotometry has been described, and the role of ferulic acid as a possible natural germination regulator in barley seed has been discussed.[3] In addition, the alkali- and β-glucosidase-labile phenolics have been determined in both the isolated embryos and the whole barley seeds. Ferulic acid, the main phenolic, was found to be present combined in alkali-labile linkage(s), probably via ester bonds.[3]

In this connection it is interesting that Fausch *et al.*[6] and Neukom *et al.*[7] have discovered that ferulic acid is a component of a glycoprotein from wheat flour. Moreover, Neukom has suggested that ferulic acid is probably linked to one pentose residue in about 50, in the arabinoxylan moiety of the glycoprotein. Other forms of bound hydroxycinnamic acids involving nitrogen are derivatives of *p*-hydroxyphenyl-β-ethylamine, putrescine, guanidinobutane, and histamine found respectively in *Evodia belahe*,[8] citrus leaves and fruits,[9] barley[10] and in the seeds of *Casimiroa edulis*.[11] From tobacco, chlorogenic acid—protein complexes have been isolated by Wright *et al.*,[12,13] while Meredith and Tkachuck[14]

[1] C. F. VAN SUMERE, H. HILDERSON and L. MASSART, *Naturwissenschaften* **12**, 292 (1958).
[2] C. F. VAN SUMERE, in *Phenolics in Plants in Health and Disease* (edited by J. B. PRIDHAM), p. 25, Pergamon Press, Oxford (1960).
[3] C. F. VAN SUMERE, J. COTTENIE, J. DE GREEF and J. KINT, in *Recent Advances in Phytochemistry* (edited by V. C. RUNECKLES and J. WATKIN), Vol. IV, p. 165, Appleton–Century–Crofts, New York (1972).
[4] C. F. VAN SUMERE, J. KINT and J. COTTENIE, *Arch. Int. Physiol. Biochim.* **76**, 396 (1968).
[5] C. F. VAN SUMERE, *Rev. Ferm. Ind. Alim.* **24**, 91 (1969).
[6] H. FAUSCH, W. KÜNDIG and H. NEUKOM, *Nature, Lond.* **199**, 287 (1963).
[7] H. NEUKOM, L. PROVIDOLI, H. GREMLI and P. A. HUI, *Cereal Chem.* **44**, 238 (1967).
[8] J. RONDEST, B. C. DAS and J. POLONSKY, *Bull. Soc. Chim. Fr.* 2411 (1968).
[9] T. A. WHEATON and I. STEWART, *Nature, Lond.* **206**, 620 (1965).
[10] A. STOESSL, *Phytochem.* **4**, 973 (1965).
[11] C. DJERASSI, C. BANKIEWIECZ, A. L. KAPOOR and B. RINIKER, *Tetrahedron* **2**, 168a (1958).
[12] H. E. WRIGHT, JR., W. W. BURTON and R. C. BERRY, JR., *Arch. Biochem.* **86**, 94 (1960).
[13] H. E. WRIGHT, JR., W. W. BURTON and R. C. BERRY, JR., *Phytochem.* **3**, 525 (1964).
[14] W. O. S. MEREDITH and R. TKACHUK, *J. Inst. Brew.* **70**, 411 (1964).

reported on protein–tannin complexes from barley, malt and beer. More recently Alibert *et al.*[15] have found phenolic acids in a protein fraction from leaves of *Quercus penduculata* and Brieskorn and Mosandl[16] have isolated a caffeic acid containing protein from fruits of Umbelliferae. However, although phenolic acid containing proteins seem to be widespread,[17] there is little knowledge about the type of linkage by which ferulic acid or any other phenolic acid may be bound.

Of several possibilities, it seemed to us, that by analogy with naturally occurring *N*-acylamino acids,[18–26] pseudopeptide bonds, in which the carboxylic group of the phenolic is attached to the amino function of the *N*-terminal amino acid of a polypeptide chain should be seriously considered. No such *N*-acylamino acid derivatives of the substituted benzoyl- or cinnamoyl-type has been found to occur in plants, although small amounts of benzoylaspartic acid were formed when pea epicotyl sections were incubated in a solution of benzoic acid.[27,28] Reports by Booth *et al.*[29] and Armstrong *et al.*[30] indicate furthermore that *N*-feruloylglycine and *N*-*m*-coumaroylglycine may be components of human urine.

In this paper we show that barley globulins contain *N*-feruloylglycyl-phenylalanine as a sequence. The possible importance of this finding for the biosynthesis of proteins in plants will be discussed.

RESULTS AND DISCUSSION

Barley globulins, isolated by standard procedures and freed from free and alkali-labile bound phenolics (see Experimental), proved to contain *N*-feruloylglycyl-L-phenylalanine (I) as a sequence. The phenolic dipeptide was obtained from the above proteins by partial hydrolysis with 4 N HCl, followed by neutralization, salt removal, preparative and 'multiple elimination' TLC (see Experimental). Identification was based on comparison with synthetic material[31] by chromatography (Table 1) by excitation and fluorescence maxima, and by UV and IR spectra. Additional evidence for identity was obtained by the action of carboxypeptidase A (peptidyl-L-amino acid hydrolase; E.C. 3.4.2.1) which gave, with both synthetic and natural samples, *N*-feruloylglycine and phenylalanine.

From these results it seems that, in barley globulins, ferulic acid is partly bound to the *N*-terminal glycine and that *N*-feruloylglycyl-L-phenylalanine is a sequence of barley proteins. Moreover, the fact that the pseudopeptide bond of *N*-feruloylglycine is very stable towards alkali, acid and carboxypeptidase A, shows that more ferulic acid is present in the barley grain than can be calculated from the figures published by Van Sumere *et al.*[3] Indeed,

[15] G. ALIBERT, G. MARIGO and A. BOUDET, *Compt. Rend.* **267D,** 2144 (1968).
[16] C. H. BRIESKORN and A. MOSANDL, *Tetrahedron Letters* **1,** 109 (1970).
[17] C. F. VAN SUMERE, A. DEDONDER, I. PE, M. VAN BUSSEL and H. DE POOTER, unpublished results.
[18] A. MEISTER, in *Biochemistry of the Amino Acids*, Vol. I, p. 558, Academic Press, New York (1965).
[19] N. ROSA and A. C. NEISH, *Can. J. Biochem.* **46,** 797 (1968).
[20] A. I. VIRTANEN and P. LINKO, *Acta Chem. Scand.* **9,** 531 (1955).
[21] C. GILVARG, *Biochim. Biophys. Acta* **24,** 216 (1957).
[22] L. FOWDEN, *Nature, Lond.* **182,** 406 (1958).
[23] I. LISS, *Phytochem.* **1,** 87 (1962).
[24] R. GMELIN, A. KJAER and P. O. LARSEN, *Phytochem.* **1,** 233 (1962).
[25] N. E. GOOD and W. A. ANDREAE, *Plant Physiol.* **32,** 561 (1957).
[26] J. P. AUBERT, J. MILLET, E. PINEAU and G. MILHOUD, *Biochim. Biophys. Acta* **51,** 529 (1961).
[27] W. A. ANDREAE and N. E. GOOD, *Plant Physiol.* **32,** 566 (1957).
[28] M. A. VENIS, *Plant Physiol.* **49,** 24 (1972).
[29] A. N. BOOTH, O. H. EMERSON, F. T. JONES and F. DEEDS, *J. Biol. Chem.* **229,** 51 (1957).
[30] M. D. ARMSTRONG, K. N. F. SHAW and D. E. WALL, *J. Biol. Chem.* **218,** 293 (1956).
[31] H. DE POOTER, HAIDER ALI and C. F. VAN SUMERE, *Bull. Soc. Chim. Belge* in press.

these figures account only for the free, alkali- and β-glucosidase-labile ferulic acid present in the barley seed. This new compound may be of considerable interest in connection with the sequential analysis of barley proteins, lignin formation in the barley seedling as well as with the regulation of protein biosynthesis in the barley seed.

TABLE 1. R_fS OF *N*-FERULOYLGLYCYL-L-PHENYLALANINE (I)

TLC: silica gel-cellulose (1:1) Solvents	R_fs (steamed plates) of I* Synthetic	Natural
MeOH–H_2O (3:1)	0·96	0·96
sec–BuOH–H_2O (4:1)	0·66	0·66
Toluene–HCO_2Et–HCO_2H (5:4:1)	0·46	0·46
Benzene–HOAc–H_2O (125:73:2)	0·39	0·39

* Detected by blue fluorescence in UV, changing to blue-green when plate sprayed with 2 N NaOH. Also gives slate blue colour with diazotized *p*-nitroaniline, followed by 5% Na_2CO_3.

In connection with protein biosynthesis in barley it is possible that *N*-feruloylglycine could play, in the seed or the embryo, the role which formylmethionine (HCO–Meth) plays in *E. coli*.[32,33] In the same context Pearlman and Bloch[34] and Krishna and Krishnaswamy,[35] have shown that *N*-acetylamino acids (e.g. MeCO–GlyOH) may be involved in protein synthesis in animal tissue and *E. coli*. Indeed Pearlman and Bloch[34] have shown that certain *N*-acetylamino acids undergo the reactions which are generally believed to initiate protein synthesis from free amino acids. The results obtained by the same authors demonstrate that the NH_2-group of the amino acid need not be free for carboxyl-activation and transfer to *t*RNA. However, the nature of the blocking group seems to be of some importance since *N*-carbobenzoxyamino acids and certain peptides are apparently unreactive in the amino acid activating system.[36,37] The presence of *N*-acetylamino acids as *N*-terminal residues of a number of proteins or polypeptides is now recognized,[38] e.g. *N*-acetylglycine has been identified in horse heart cytochrome *c*.[39] Moreover, in a number of proteins free *N*-terminal amino groups cannot be detected, indicating perhaps that the aminoterminal grouping is masked by substitution.[34] The utilization of *N*-acetylamino acids in early steps of protein synthesis raises, according to the same authors, the possibility that the acetylation reaction leading to acetylproteins occurs with the free amino acid and not at the *t*RNA stage or with the complete protein. Whether and to what extent the acetyl- or feruloyl-groups modify the properties of the proteins must still be investigated.[34]

EXPERIMENTAL

Plant material. Union Barley (*Hordeum Vulgaris* cv. Union), originating from the test fields of the European barley convention at PROVEN (West Flanders, Belgium), harvest 1970, was used throughout this work.

[32] J. M. ADAMS and M. R. CAPECCHI, *Proc. Natl. Acad. Sci. U.S.* **55,** 147 (1966).
[33] R. E. WEBSTER, D. L. ENGELHARDT and N. D. ZUIDER, *Proc. Natl. Acad. Sci. U.S.* **55,** 155 (1966).
[34] R. PEARLMAN and K. BLOCH, *Proc. Natl. Acad. Sci.* **50,** 533 (1963).
[35] R. V. KRISHNA and P. R. KRISHNASWAMY, *Life Sci.* **5,** 2053 (1966).
[36] F. LIPMAN, *Proc. Natl. Acad. Sci.* **44,** 67 (1958).
[37] G. D. NOVELLI, *Proc. Natl. Acad. Sci.* **44,** 86 (1958).
[38] A. MEISTER, in *Biochemistry of the Amino Acids*, Vol. I, p. 558, Academic Press, New York (1965b).
[39] E MARGOLIASH, E. L. SMITH, G. KREIL and H. TUFFY, *Nature, Lond.* **192,** 1125 (1961).

Synthesis of N-*feruloylglycyl-*L-*phenylalanine.* The ferulic acid containing dipeptide was synthesized via *N*-(*O*-methoxycarbonyl-feruloyl)-glycyl-L-phenylalanine *t*-butyl ester.[31] Its properties were as follows: $[\alpha]_D^{22}$ +31·4 (*c*, 0·484; 95% EtOH); UV spectrum in 0·1 N HCl–MeOH: λ_{max} 296 nm, ϵ 11 500 and 319 nm, ϵ 14 300; UV spectrum in 0·1 N KOH–MeOH: λ_{max} (sh) 296 nm, ϵ 4300; 308 nm, ϵ 5500 and 365 nm, ϵ 18 900. IR spectrum (KBr-disc): broad band between 3700 and 2200 cm^{-1}, with ν'_{max} 3345, 3075, 2940 and 2610 cm^{-1} (NH, OH, alkyl and CO_2H); broadened peaks at 1725 cm^{-1} (*s*, CO_2H); 1655 (*ms*); 1595 (*m*) and 1510 cm^{-1} (*s*) (conj. olefin, benzene and amide) 1250 cm^{-1} (*s*, aromatic–OMe); 977 cm^{-1} (*w*, *trans*-substituted olefin); 845 cm^{-1} (*w*) and 805 cm^{-1} (*m*) (1,2,4-substituted benzene); 702 cm^{-1} (*m*, *mono*-substituted benzene).

The isolation of the barley albumins, globulins and prolamins. The seed proteins have been isolated according to procedures described by Lontie and his group;[40] Lontie and Voets[41] and Préaux *et al.*[42] Starting from 300 g barley flour the following yields (dry wt) were obtained: albumins (1·2 g), globulins (2·5 g) and prolamins (1·3 g)

Presence and estimation of alkali-labile bound ferulic acid in barley proteins. Barley albumins, globulins or prolamins (each 100 mg) were thoroughly extracted with Et_2O (to remove traces of free phenolic acids), prior to hydrolysis with NaOH. The combined Et_2O fractions were then analyzed by means of TLC on silica gel-cellulose layers (1:1) with toluene–HCO_2Et–HCO_2H (5:4:1) (TEF), but no free ferulic acid could be detected. After gentle boiling of the protein fractions for 2 hr with 3 ml 2 N NaOH and neutralization to pH 3 in the cold, the alkali-labile ferulic acid portion was extracted and determined.[3,5] Thus, the barley albumins, globulins and prolamins contain, per g dry wt, respectively: 150, 300 and 90 μg alkali-labile ferulic acid.

Isolation and purification of N-*feruloylglycyl-*L-*phenylalanine from barley globulin hydrolysates.* After alkaline hydrolysis of 1·0 g barley globulins with 20 ml 2 N NaOH, neutralization and removal of the liberated ferulic acid, the remaining solution was treated with enough conc. HCl to reach 4 M. Subsequently, the solution was boiled under reflux for 6 hr. The partial hydrolysate was then concentrated under vacuum, again neutralized with NaOH and treated with a mixture of MeOH–EtOH (1:1). The NaCl precipitate was removed by centrifugation. This treatment was repeated several times resulting in an almost salt free supernatant. This final solution was then concentrated until 1 ml and spotted, on silica gel-cellulose (1:1) plates in TEF. Synthetic *N*-feruloyl-amino acids and *N*-feruloylglycyl-L-phenylalanine were used as test substances. The natural *N*-feruloylglycyl-L-phenylalanine was removed from the plate, eluted with MeOH and respotted next to synthetic *N*-feruloylglycyl-L-phenylalanine on a silica gel-cellulose (1:1) layer, and purified with the following solvents: MeOH–H_2O (3:1), thin-layer steamed,[43] front at 5 cm; *sec*-BuOH–H_2O (4:1), thin-layer steamed, front at 10 cm; MeOH–H_2O (3:1), thin-layer steamed, front at 5 cm; TEF, thin-layer steamed, front at edge of the plate. In certain experiments the internal H_2O phase of the plates was increased by steaming the thin-layers after spotting and before irrigation. The increase of the stationary H_2O phase was obtained by keeping the plates in H_2O vapour escaping from an ordinary tea kettle. However, this treatment requires care, since over-steaming of the plates can destroy the thin-layers. Steaming of the TLCs can be frequently replaced by keeping the plates for 20 min in a tank saturated with H_2O vapour. The *N*-feruloylglycyl-DL-phenylalanine from the hydrolysate, was removed from the plate, eluted with MeOH and analyzed.

Excitation and fluorescence spectra of N-*feruloylglycyl-*L-*phenylalanine and the unknown.* The excitation and fluorescence wavelengths of both compounds in 0·1 N NaOH were recorded by means of a Baird Atomic Fluorescence spectrophotometer SF 1, combined with an Omnigraph recorder. The excitation and fluorescence maxima were identical (excit. λ_{max} 379 nm; fluor. λ_{max} 480 nm) for both compounds.

IR spectroscopy. Transfer of N-*feruloylglycyl-*L-*phenylalanine* (I) *to* KBr *micro disks after METC.* After METC of the partial protein hydrolysate, the spots from the synthetic and natural material were eluted with MeOH in the usual way. Subsequently, the silica gel residues were removed by centrifugation and each supernatant was concentrated in N_2 to a small vol. After respotting of these concentrates, next to synthetic I, on gradient nylon-silica gel layers (the nylon-silica gel layers consisted of 3 g of each component. Both materials were suspended in 30 ml 60% MeOH. For the preparation of the layers the GM (Gradient mixer) applicator of Desaga has been employed), the samples were chromatographed with MeOH. Thereafter the spots were separately removed from the thin layer and extracted with the same solvent. Subsequently, the concentrated solutions were brought in small cuvets containing a KBr triangle (modification of the Garner–Packer technique[44]) and the samples were separately chromatographed on the top of the triangles with dry

[40] R. LONTIE, J. RONDELET and J. DALCINO, *European Brewery Convention: Proc. of the Congress Nice*, p. 33, Elsevier, Amsterdam (1953).

[41] R. LONTIE and TH. VOETS, *European Brewery Convention. Proc. of the Congress Rome*, p. 27, Elsevier, New York (1959).

[42] G. PRÉAUX, A. VAN GIJSEL, C. ROBIN, F. A. WOUTERS and R. LONTIE, *European Brewery Convention: Proc. Congress Brussels*, p. 59, Elsevier, Amsterdam (1964).

[43] C. F. VAN SUMERE, G. WOLF, H. TEUCHY and J. KINT, *J. Chromatog.* **20**, 48 (1965).

[44] H. R. GARNER and H. PACKER, *Appl. Spectros.* **22**, 122 (1968).

MeOH as a solvent. After drying, the KBr tips, containing the blue fluorescing spots, were removed and pressed into a 2 × 11 mm micro disk in the usual way (when necessary some extra dry KBr was added) and the IR spectra were recorded. The KBr triangles 3 × 3 × 3 cm were prepared with a stainless steel mould by application of a pressure of about 500 kg.

The carboxypeptidase A reaction. After METC of the partial protein hydrolysate, the samples were separately eluted with MeOH. The MeOH was then removed in N_2 and 1 ml maleic acid buffer, pH 6·8 (maleic acid 0·1 M–NaOH 0·1 M, 1:2) and 6 μg carboxypeptidase A were added. After 2 hr at 40°, the reaction mixtures were spotted on a silica gel-cellulose layer, together with *N*-feruloylglycyl-L-phenylalanine, *N*-feruloylglycine and phenylalanine, and the chromatogram was developed with *sec*-BuOH–H_2O (4:1). The blue fluorescing *N*-feruloylglycylphenylalanine spots were only weakly visible, indicating that the isolated compound was mainly present in the L-configuration, while in both cases, *N*-feruloylglycine and phenylalanine were formed in relative large quantities. Recently, both *N*-*p*-coumaroylglycyl-L-phenylalanine and *N*-feruloylglycyl-L-phenylalanine have been found to be very useful substrates for the submicrodetermination of Carboxypeptidase A activities.[45]

Acknowledgements—Thanks are due to Mrs. C. Vindevogel-Debacquer, Mrs. Houpeline-Decock, Mrs. Geirnaert-Pilaert and Mr. W. Hutsebaut. The support of our work by the Belgium 'Fonds voor Kollektief Fundamenteel Onderzoek', is gratefully acknowledged.

[45] C. F. VAN SUMERE, H. DE POOTER and HAIDER ALI, in preparation.