Phytochemistry, 1972, Vol. 11, pp. 791 to 797. Pergamon Press. Printed in England.

ABSOLUTE CONFIGURATION OF β-HYDROXY-γ-METHYL-GLUTAMIC ACIDS FROM *GYMNOCLADUS DIOICUS*

G. A. DARDENNE
Faculté des Sciences Agronomiques de l'Etat, Gembloux, Belgium

E. A. BELL and J. R. NULU
Department of Botany, University of Texas at Austin
and
C. CONE
Department of Chemistry, University of Texas at Austin, Austin, Texas 78712, U.S.A.

(*Received* 29 *March* 1971)

Abstract—The new amino acids isolated from the seeds of *Gymnocladus dioicus* (L) Koch (the Kentucky coffee tree) have been identified as 2(S), 3(S), 4(R)-β-hydroxy-γ-methylglutamic acid and 2(S), 3(R), 4(R)-β-hydroxy-γ-methylglutamic acid.

INTRODUCTION

THE ISOLATION of two new amino acids from the seeds of *Gymnocladus dioicus* has been described in the preceding paper.[1] The amino acids were shown to be isomeric forms of β-hydroxy-γ-methylglutamic acid. The present paper is concerned with the characterization of these compounds by physical and biochemical methods and particularly with the determination of their absolute configurations.

RESULTS AND DISCUSSION

Mass Spectrometry

The fragmentation patterns of G_1 and G_2 were very similar although small differences were observed in the relative abundances of some of the fragments. No molecular ion or amine[2] ion was observed because of the very ready (thermal) loss of water to form the corresponding cyclic lactam. The highest mass ion observed in the spectra of both amino acids, *m/e* 159, corresponds to the molecular ion of this lactam. This ion fragments by loss of carboxyl radical to give the amine ion of the lactam at *m/e* 114 and by loss of water (m* 125·0) to give the ion at *m/e* 141. Further loss of water or carboxyl radical respectively then gives rise to the fragment at *m/e* 96. The lactam amine ion also fragmented with loss of HN=C=O to give the ion at *m/e* 71 (m* 44·2), a process analogous to that observed in 2-pyrrolidone.[3]

IR Absorption Spectra

The IR spectra of the two monohydrates (Fig. 1) were quite distinctive though the

[1] G. A. DARDENNE, J. CASIMIR, E. A. BELL and J. R. NULU, *Phytochem.* **11**, 787 (1972).
[2] K. BIEMANN, *Mass Spectrometry: Organic Chemical Applications*, p. 262, McGraw-Hill, New York (1962).
[3] A. M. DUFFIELD, H. BUDZIKIEWICZ and C. DJERASSI, *J. Am. Chem. Soc.* **86**, 5536 (1964).

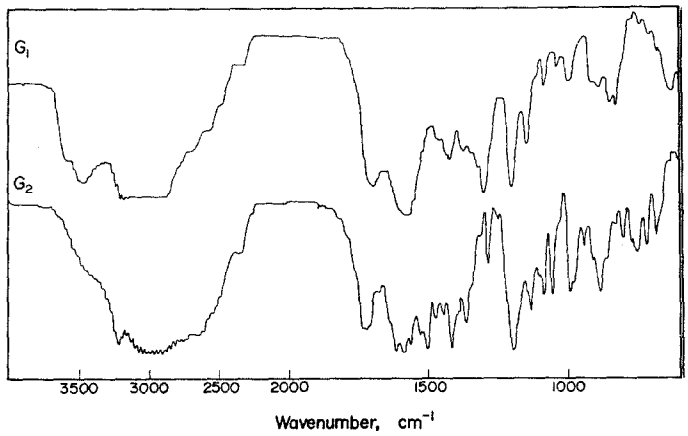


FIG. 1. IR ABSORPTION SPECTRA OF G_1 AND G_2 IN KBr.

absorption due to an undissociated carboxyl group (a characteristic of dicarboxylic acids) was clearly visible in both. One unusual feature of the IR spectra was the difference between the spectrum of anhydrous G_1 and the spectrum of its monohydrate. The spectrum of the anhydrous form of G_1, like the spectra of both forms of G_2, showed a strong absorption band at 1523 cm^{-1} which is characteristic of NH_3^+ and common to the spectra of most amino acids. This band was missing from the spectrum of the G_1 monohydrate suggesting that the water of crystallization is involved in hydrogen bonding with the amino group. If the amino and β-hydroxy groups in G_1 are *erythro* with respect to each other it is possible to have water forming a hydrogen-bonded link between them and interfering with the usual protonization of the amino group.

ACID DISSOCIATION CONSTANTS

The pK_{a_2} values of G_1 and G_2 were determined by titration and found to be 3·90 and 3·85 respectively. These values were in agreement with ionic mobilities of the two compounds at pH 3·6. The pK_{a_1} values of the amino acids could not be determined accurately by titration but were calculated from the pK_{a_2} values and the values for pI which were determined by electrophoresis. These calculations showed the pK_{a_1} values of G_1 and G_2 to be 2·6 and 2·35 respectively. These values agree with the observed rates of ionic migration of the two compounds at pH 1·9. These pK_{a_1} values suggest that in G_1 the amino and hydroxy groups are in the *erythro* relationship, for in this relationship the close proximity of the hydroxyl group would be expected to reduce the basicity of the amino group and consequently the pK_{a_1} value.[4]

NMR Spectra

The NMR spectra of the two amino acids in acidic solution could not be readily interpreted apart from the methyl doublet at high field which showed clearly that the methyl group was bound to a carbon atom directly bonded to only one proton. In basic solution the spectra of both amino acids were essentially first order, as shown in Fig. 2. The analysis

[4] P. J. PETERSON, *J. Chromatog.* **38**, 301 (1968).

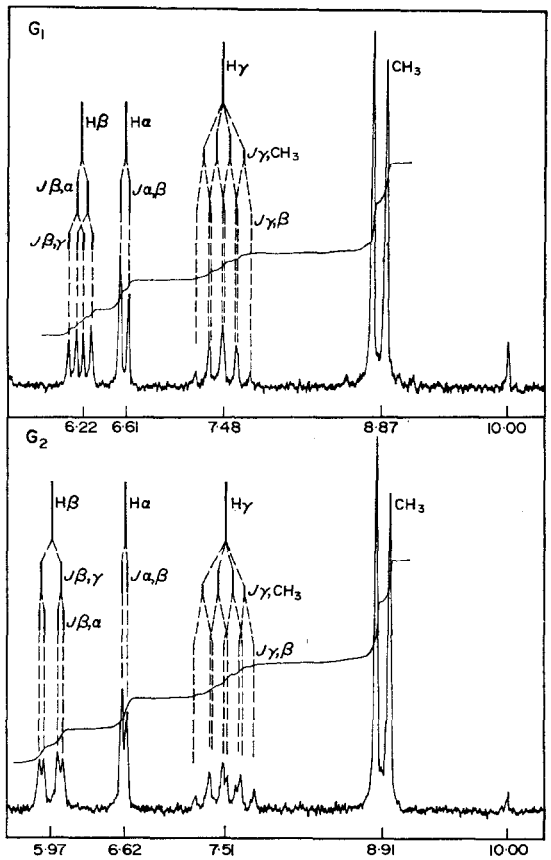


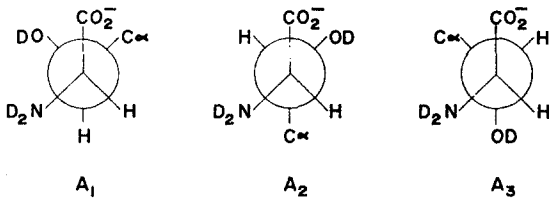
FIG. 2. NMR SPECTRA OF G_1 AND G_2 DETERMINED IN 2 N NaOD IN D_2O.

follows from the splittings and multiplicities observed and both amino acids are clearly shown to have the structure assigned in the previous paper:[1] they are stereoisomers of β-hydroxy-γ-methylglutamic acid.

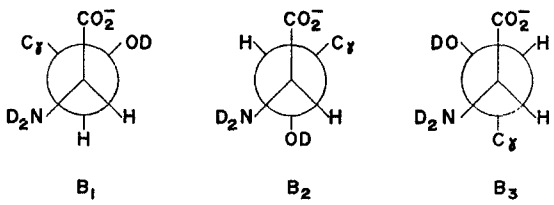
The assignment of the configurations at the β and γ carbon atoms of these two amino acids relative to that at the α carbon atom was based on the observed values of the first order coupling constants (Table 1). The configuration at the α carbon atom was assumed to be *L* (S). This assumption was later proven correct by the biosynthetic studies described in the following section. The relative configurations can be deduced from the Karplus relationship[5] employing arguments essentially similar to those used by Alderweireldt *et al.*[6] to assign configurations to the two diasteroisomeric γ-hydroxy-γ-methylglutamic acids. A large coupling constant is expected between protons having the anti-relationship and a small coupling constant between protons having a gauche relationship.

[5] M. KARPLUS, *J. Chem. Phys.* **30,** 11 (1959).
[6] F. ALDERWEIRELDT, J. JADOT, J. CASIMIR and A. LOFFET, *Biochim. Biophys. Acta* **136,** 89 (1967).

The Newman projection formulae along the C_α—C_β bond for the two possible configurations 2(S), 3(S) or *erythro*, and 2(S), 3(R) or *threo* are:



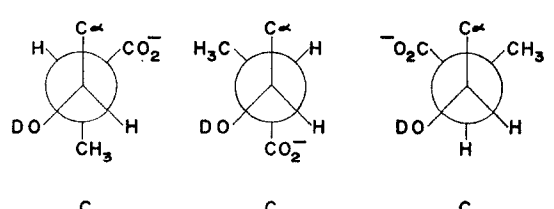
erythro



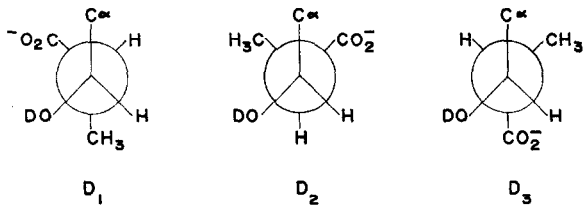
threo

Conformations A_3, B_1 and B_2 are unfavourable because of the gauche relationship of the bulky groups. In the *erythro* configuration the conformations A_1 and A_2 will both be populated although A_2 will be favoured because of the anti relationship of the bulky groups. Hydrogen bonding between the hydroxyl group and the ionized carboxyl group will decrease the steric interactions in A_2 and in A_1. A fairly large coupling constant would be expected for this configuration. In the *threo* configuration only conformation B_3 would be favoured. Hydrogen bonding in conformer B_3 would decrease the coupling constant but increase the steric interaction between the other groups. A relatively small coupling constant would be expected for the *threo* configuration. We can thus assign the *erythro* configurational relationship at these carbon atoms to G_1 and the *threo* configuration to G_2.

For amino acid G_1 the possible configurations at the β and γ carbon atoms are then the 3(S), 4(S) or *erythro* and the 3(S), 4(R) or *threo* configurations. These are shown, as Newman projection formulae, for the three possible conformers in each case, viewed along the C_β—C_γ bond.



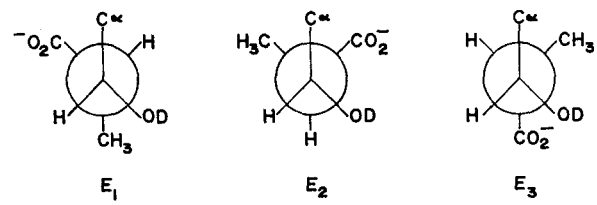
erythro



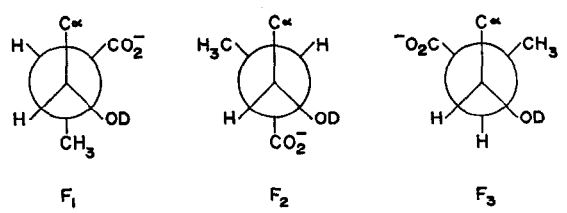
threo

For the *erythro* configuration conformations C_2 and C_3 would be essentially not populated due to the unfavourable steric interactions. A very large coupling constant would be expected for the anti relationship between the protons in conformer C_1. In the *threo* configuration, conformer D_2 would not be populated. Conformers D_1 and D_3 would both be populated with D_3 favoured by the anti relationship of the large groups. Hydrogen bonding would result in an increased value of the coupling constant and a decreased steric interaction for conformer D_1 and in an increased steric interaction but a decreased coupling constant for D_3. A medium averaged coupling constant would consequently be expected for the *threo* configuration, enabling us to assign this configuration to G_1.

In G_2, the two possible configurations would each have the three conformers shown in the following Newman projection formulae:



erythro



threo

For the *erythro* configuration, only conformer E_1 is favourable and a large coupling constant is to be expected. In the *threo* configuration conformer F_3 is highly unfavourable. Conformers F_1 and F_2 would both be populated. Conformer F_2 is favoured by the anti relationship of the large groups. Hydrogen bonding would result in a decreased steric interaction for conformer F_1 and an increased coupling constant. The opposite would be true for conformer F_2. The resulting coupling constant for the *threo* configuration should thus have an intermediate value. The *erythro* configuration can therefore be assigned to G_2.

In summary then, amino acid G_1 has the *erythro–threo* configuration and G_2 has the

threo–erythro configuration as shown below in the classical Fischer formulae, and in the Newman projections viewed along the C_α—C_β and C_β—C_γ bonds.

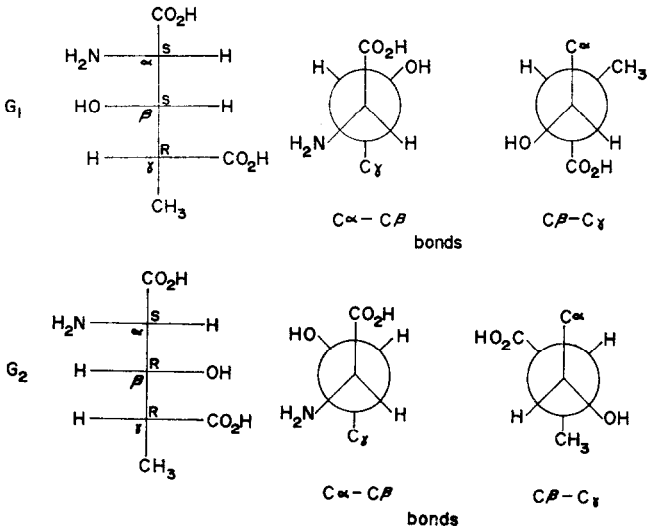


TABLE 1. COUPLING CONSTANTS OF G_1 AND G_2

	$J_{\alpha,\beta}$	$J_{\beta,\gamma}$	J_{γ,CH_3}
G_1	4·2	7·8	7·2
G_2	2·2	9·8	7·2

Determination of Absolute Configuration

The establishment by NMR of the relative configurations of G_1 and G_2 as *erythro–threo* and *threo-erythro* respectively, still left the absolute configurations of the two compounds unresolved. The conditions used to reduce each of the two amino acids to the corresponding γ-methylglutamic acid produced a mixture of racemic derivatives and neither G_1 nor G_2 underwent change when incubated with L or D amino acid oxidase or with L-glutamic acid decarboxylase.

Preliminary biosynthetic studies established however that L-leucine but not D-leucine can act as a good precursor of both G_1 and G_2 in the leaves of the plant. While the biosynthetic relationship between leucine and the two isomeric amino acids is still under investigation this finding clearly indicates that the G_1 is 2(S), 3(S), 4(R) β-hydroxy-γ-methylglutamic acid and G_2 is 2(S), 3(R) 4(R) β-hydroxy-γ-methylglutamic acid.

EXPERIMENTAL

Mass spectra. The high resolution mass spectra of G_1 and G_2 were recorded photographically using a Du Pont (CEC) 21-110 high resolution mass spectrometer. The recorded mass spectra were measured with a Grant Instruments plate reader- PDP 8-I system.[7] The determined masses were within two milli-amu of those

[7] C. CONE, presented at *18th Annual Conference on Mass Spectrometry and Allied Topics*, San Francisco Calif. (June 1970). To be published.

calculated for the elemental compositions reported. The strong ions above *m/e* 60 were found to have the following compositions: 159 ($C_6H_9NO_4$), 141($C_6H_7NO_3$), 114 ($C_5H_8NO_2$), 97 (C_5H_7NO), 97 ($C_5H_5O_2$), 96 (C_5H_6NO), 74 ($C_2H_4NO_2$), 71 (C_4H_7O).

NMR spectra. NMR spectra were determined with a Varian Associates A-60 spectrometer at ambient temperature (DSS as internal standard). All chemical shifts are reported as τ-values and coupling constants (first order) as Hz. The symbols s, d, q, and m represent singlet, doublet, quartet and multiplet respectively. Sample concentrations were 10% w/v. For G_1 in 2 N NaOD in D_2O: 8·87,d,H–CH_3; 7·48,m, 1H γ-CH; 6·22,q. 1H β-CH; 6·61,d,1H α-CH. In 2 N CF_3CO_2D in D_2O: 8·68,d,3H, γ–CH_3; 6·85, m, 1H γ-CH; *ca.* 5·65, m,2H, α-CH and β-CH. For G_2 in 2N NaOD in D_2O: 8·91,d,3H,–CH_3; 7·51,m,1H, γ-CH; 5·97,q,1H, β-CH; 6·62,d,1H α-CH. In 2N CF_3CO_2D in D_2O: 8·70,d,3H,–CH_3; 7·10,m 1H, γ-CH; 5·55,m, 2H, α-CH and β-CH. (Coupling constants are given in Table 1.)

IR absorption spectra. The spectra were recorded using the KBr disc method and a Perkin–Elmer (247) instrument.

pK_a values. pK_a values were determined by titration using a Titrator Radiometer (TTTIC′) and a Titrigraph Radiometer (SBB_2C').

Biosynthesis. Young leaves were allowed to take up through their cut stems solutions (0·5 ml) containing variously 10 μc (0·03 mg) of L-leucine-1-^{14}C, 10 μc (0·05 mg) of D-leucine-1-^{14}C and 10 μc (0·1 mg) DL-leucine-1-^{14}C. After the solutions had been absorbed (24 hr) the leaves were kept with their stems in water for a further 48 hr. The leaflets were then removed and ethanolic (70%) extracts prepared from both leaflets and leaf stems. These extracts were analysed by 2D chromatography and high-voltage electrophoresis on paper and the incorporation of radioactivity into the free amino acids determined by autoradiography using Kodirex X-ray film (Kodak Ltd).

Acknowledgements—We thank the National Institutes of Health (P.H.S. Grant No. R01 ES00442, E.A.B.), the Robert A. Welch Foundation (Grant No. F375, E.A.B.) and the North Atlantic Treaty Organization (Grant No. 448) for financial support. C.C. thanks the National Science foundation both for support as a Faculty Associate (U.S.D.P. Grant No. GU-1958) and for funds to purchase the high resolution mass spectrometer (GP-8509). G.A.D. thanks the Institut de Recherches Scientifiques pour l'Industrie et l'Agronomie (IRSIA), Belgium for financial assistance to complete this work in the University of Texas at Austin.

Key Word Index—Gymnocladus dioicus; Leguminosae; β-hydroxymethylglutamic acid; stereoisomers.