Tryptophan glycoconjugates in food and human urine*

Birgit GUTSCHE†, Christoph GRUN†, Dieter SCHEUTZOW‡ and Markus HERDERICH†¹

†Lehrstuhl für Lebensmittelchemie, Universität Würzburg, Am Hubland, D-97074 Würzburg, Germany, and ‡Institut für Organische Chemie, Universität Würzburg, Am Hubland, D-97074 Würzburg, Germany

Evaluating the formation of tryptophan glycoconjugates other than well-established Amadori rearrangement products, HPLC–tandem MS (MS/MS) analysis of human urine collected from several healthy individuals proved the presence of one distinct tryptophan C-glycosyl compound [Horiuchi, Yonekawa, Iwahara, Kanno, Kurihara and Fujise (1994) *J. Biochem. (Tokyo)* **115**, 362–366]. After isolation, unambiguous identification of this novel tryptophan metabolite as 2-(α -mannopyranosyl)-L-tryptophan was achieved by tandem MS combined with NMR spectroscopy including homonuclear COSY, heteronuclear multiple-bond connectivity and ¹H-detected heteronuclear multiple-quantum coherence experiments. Remarkably, a thorough evaluation of vicinal proton–proton coupling constants in different solvents and nuclear Overhauser effect experiments demonstrate the predominant axial orientation of the hydroxymethyl group of the hexopyranosyl residue. Likewise this spatial arrangement indicates that the respective α -anomeric C-mannosylhexopyranose is preferentially adopting a ¹C₄ con-

formation in acidic methanol. Whereas only one distinct tryptophan mannoconjugate could be observed in human urine, HPLC–MS/MS analysis of food samples for the first time led to the identification of numerous *N*¹-(β -D-hexopyranosyl)-L-tryptophan, 2-(β -D-hexopyranosyl)-L-tryptophan and 1-(1,2,3,4,5-pentahydroxypent-1-yl)-1,2,3,4-tetrahydro- β -carboline-3-carboxylic acid derivatives derived from the condensation of tryptophan with aldohexoses. Taking into consideration the significant differences between profiles and configurations of tryptophan glycoconjugates originating from dietary sources and human urine, C-2 mannosylation of tryptophan residues [de Beer, Vliegthart, Loeffler and Hofsteenge (1995) *Biochemistry* **34**, 11785–11789] represents a novel enzymic pathway in tryptophan metabolism in humans.

Key words: conformational analysis, C-glycosides, tandem mass spectrometry, tetrahydro- β -carboline.

INTRODUCTION

Studying tryptophan degradation in biological systems, our interest in tryptophan glycoconjugates was stimulated by recent reports on proteins that contained post-translationally glycosylated tryptophan residues. Gaede et al. [1] described glycosylation of tryptophan in a neuropeptide isolated from the insect *Carausius morosus*. Furthermore, Hofsteenge and co-workers identified for the first time a modification of human RNase by C-glycosidic attachment of α -D-mannopyranose to the C-2 of the indole ring of Trp⁷ [2–4]. This novel kind of protein glycosylation required the sequence Trp-Xaa-Xaa-Trp and proved to be enzyme-catalysed, employing dolichyl phosphate mannose as a precursor [5,6].

In addition, a novel tryptophan glycoconjugate has been detected in human urine. This compound was proposed to have a tetrahydro- β -carboline structure, but the identity of the sugar part remained to be established [7]. As the occurrence of related tryptophan glycoconjugates in dietary sources has not yet been investigated, an exogenous origin for the tryptophan metabolite could not be excluded.

Here we describe the identification and conformational analysis of the tryptophan glycoconjugate in human urine. In addition, we report on the characterization and structure elucidation of

novel tryptophan glycoconjugates as well as their occurrence in commercial food samples.

EXPERIMENTAL

Reagents

Water, methanol and acetonitrile (all of HPLC gradient grade), trifluoroacetic acid (TFA) (spectroscopic grade), L-tryptophan, D-glucose, pyridine and Lichroprep C₁₈ (43–60 μ m particle size) were supplied by Merck (Darmstadt, Germany). D-Talose was from ICN (Eschwege, Germany), and D-altrose was obtained from Fluka (Buchs, Switzerland). D-Allose, D-idose, D-gulose, acetic anhydride and Dowex 1 $\times$ 8 (100–200 mesh) were purchased from Sigma (Steinheim, Germany). D-Mannose, D-galactose, cellulose, butan-1-ol, methoxyhydroxylamine hydrochloride and 1-dimethylaminopropan-2-ol were obtained from Aldrich (Deisenhofen, Germany). All other chemicals were of analytical purity. Membrane filters with pore size 0.2 μ m were from Ziemer (Mannheim, Germany). Food samples were purchased at local markets.

HPLC with UV detection

The HPLC system consisted of two Knauer (Berlin, Germany) HPLC pumps 64 equipped with micro pump heads and a

Abbreviations used: CID, collision-induced dissociation; ESI, electrospray ionization; HH-COSY, homonuclear COSY; HMQC, ¹H-detected heteronuclear multiple-quantum coherence; HMBC, heteronuclear multiple-bond connectivity; MOA, methoxyoximacetate; MS/MS, tandem MS; NOE, nuclear Overhauser effect; SRM, selected reaction monitoring; TFA, trifluoroacetic acid.

¹ To whom correspondence should be addressed (e-mail mherderi@pzc.uni-wuerzburg.de).

* Dedicated to Professor Dr. Helmut Werner, Würzburg, on the occasion of his 65th birthday.

Figure 1 Tryptophan glycoconjugate in human urine (compound 1)

Shimadzu (Kyoto, Japan) UV detector SPD-10A. Data processing was performed with a Knauer HPLC data system. For sample injection a Rheodyne 7725 equipped with a 5 μ l loop was used. Chromatographic separation was performed on a Euro-spher 100 C₁₈ column (100 mm $\times$ 2.0 mm internal diam., 5 μ m particle size) (Knauer). Solvent A was 0.05 % (v/v) TFA in water; solvent B was acetonitrile. Compounds were separated by using the following gradient programme: 0 min, 5 % (v/v) solvent B; 30 min, 30 % solvent B; 31 min, 100 % solvent B; 40 min, 100 % solvent B. The flow rate was 200 μ l/min and the injection volume was 5 μ l. UV was monitored at 220 nm.

NMR

One-dimensional and two-dimensional NMR spectra were acquired with Bruker (Rheinstetten, Germany) WM 400 and DMX 600 spectrometers. Solvents were C²H₃O²H, (C²H₃)₂SO and C²H₃CN, both acidified with 1 % (v/v) CF₃COO²H, and ²H₂O. Measurements on the WM 400 were performed at 20 °C and those on the DMX 600 at 24 °C. Chemical shifts were calibrated by using the solvent signal (C²H₃O²H, δ_{H} 3.305 p.p.m., δ_{C} 49.0 p.p.m.; (C²H₃)₂SO, δ_{H} 2.580 p.p.m.; C²H₃CN, δ_{H} 1.950 p.p.m.; ²H₂O, δ_{H} 4.700 p.p.m.) as reference. The evaluation of chemical shifts and coupling constants was performed with WINNMR and WINDAISY software (Bruker).

For establishing the conformation of compound 1 (Figure 1), vicinal coupling constants were estimated with an extended Karplus-type equation by the method of Haasnoot et al. [8] with the PC program ALTONA [9], assuming standard dihedral angles of 60 °, 180 ° and 300 °.

HPLC–electrospray ionization–tandem MS (HPLC–ESI–MS/MS)

Chromatographic separation was performed with an Applied Biosystems 140b pump. For sample injection a SunChrom autosampler Triathlon (BAI, Bensheim, Germany) was used. HPLC–ESI–MS/MS analysis was performed with a TSQ 7000 tandem mass spectrometer system equipped with an ESI interface (Finnigan MAT, Bremen, Germany). Data acquisition and evaluation were conducted on a DEC 5000/33 workstation (Digital Equipment, Unterföhring, Germany) with ICIS 8.1 software (Finnigan MAT).

Chromatographic separation for HPLC–ESI–MS/MS was performed on a Symmetry C₁₈ column (150 mm $\times$ 2.1 mm internal diam., 5 μ m particle size; gradient Ia) (Waters, Milford, MA, U.S.A.) or on a Euro-spher 100 C₁₈ column (100 mm $\times$ 2.0 mm internal diam., 5 μ m particle size; gradients Ib and II) (Knauer, Berlin, Germany) with the use of binary gradients. The

Figure 2 Tryptophan glycoconjugates 2–12

flow rate was 200 μ l/min and injection volume was 5 μ l for all analyses.

Gradients Ia and Ib

Solvent A was 0.05 % (v/v) TFA in water and solvent B was acetonitrile. HPLC was programmed as follows. Gradient Ia, pressurizing with 50 % solvent B, equilibration time 15 min at 5 % (v/v) solvent B and linear gradient elution as follows: 0 min, 5 % solvent B; 30 min, 30 % solvent B; 31 min, 100 % solvent B; 40 min, 100 % solvent B. Gradient Ib, pressurizing with 50 % B, equilibration time 15 min at 10 % solvent B and linear gradient elution as follows: 0 min, 10 % solvent B; 20 min, 30 % solvent B; 22 min, 100 % solvent B; 27 min, 100 % solvent B.

Gradient II

Solvent B was acetonitrile and solvent C was water. HPLC was programmed as follows. Pressurizing with 50 % solvent B, equilibration time 15 min at 2 % solvent B and linear gradient elution as follows: 0 min, 2 % solvent B; 20 min, 2 % solvent B; 25 min, 100 % solvent B; 30 min, 100 % solvent B.

Gradient I was used for chromatographic separation when ESI was performed in the positive mode; gradient II was performed with ESI in the negative mode. For pneumatically assisted ESI, the spray capillary voltage was set to 3.5 kV for positive and negative modes. The temperature of the heated inlet capillary was 260 °C. Nitrogen served both as sheath [414 kPa (60 lb/in²)] and auxiliary gas (10 units). Positive or negative ions

were detected with a total scan duration of 1.0 s for a single full spectrum. Product ion scanning (mass range 50–400 atomic mass units) was performed at a collision gas pressure of 267 mPa Ar and collision energy of 19 eV with a total scan duration of 3.0 s for a single spectrum. The most abundant product ions were chosen for the selected reaction monitoring (SRM) experiments; positive ions were detected with a scan duration of 0.5 s each. Selected ion pairs were: SRM I, m/z 367/247, 367/229; SRM II, m/z 367/205, 367/188. Semi-quantitative evaluations were performed with standard solutions (0, 0.1, 0.5, 1.0 and 10.0 $\mu\text{g/ml}$) of reference compounds **7**, **12a/b** (Figure 2) and **13** (Figure 6) in water for external calibration. Values were calculated from signal intensities of the respective product ions as obtained by HPLC–ESI–MS/MS analysis with SRM.

High-resolution gas-chromatography MS

A gas chromatograph GC 8000 series equipped with a split injector (1:10) was combined by direct coupling to a MD 800 spectrometer (Fisons, Valencia, CA, U.S.A.) with a Mass Lab data system. A fused silica capillary column J&W DB-5 (30 m $\times$ 0.25 mm internal diam.; d_r = 0.25 μm) was used for chromatographic separation. The temperature programme was as follows: 60–280 °C at 5 °C/min, followed by 10 min at 280 °C. The flow rate of the He carrier gas was 1.0 ml/min. The temperature of the ion source was 230 °C and that of the connection parts was 250 °C; the electron energy was 70 eV and the cathodic current was 1800 μA . The scan duration was set to 0.45 s at a mass range of 40–480 atomic mass units.

Sample preparation

Seasoning sauces, alcoholic beverages and fruit syrup were filtered through membrane filters of pore size 0.2 μm ; the resulting solutions were directly subjected to HPLC–ESI–MS/MS analysis.

Synthesis of 1-[(1-carboxy-2-indole-3-yl-ethyl)amino]-1-deoxy-D-fructose (**13**)

The Amadori rearrangement product **13** of tryptophan with D-glucose was synthesized by established procedures [10]. Structural assignment was confirmed by ESI–MS, ESI–MS/MS and NMR spectroscopy.

Model reactions

The reaction conditions were optimized for the formation of either C-glycosyl derivatives (model reaction A) or N-glycosides (model reaction B) besides polyol-tetrahydro- β -carboline. Model reaction A: 250 mg of L-tryptophan and 450 mg of aldohexose (D-glucose, D-galactose or D-mannose) were dissolved in 10 ml of water and adjusted to pH 1 with 1 M H_2SO_4 . Reaction mixtures were left at 80 °C for 12 days. Model reaction B: 500 mg of L-tryptophan and 900 mg of aldohexose (D-glucose, D-galactose or D-mannose), dissolved in 10 ml of water and adjusted to pH 2 with 1 M HCl, were left at 80 °C for 12 days.

Isolation of tryptophan glycoconjugates from model reactions

Model reactions were filtered and applied directly to a cellulose column (17 cm $\times$ 2.5 cm). The column was packed with cellulose powder suspended in water-saturated butan-1-ol and eluted with the same solvent. Fractions (10 ml) were collected and checked by HPLC with UV detection. Fractions containing tryptophan glycoconjugates **2–12** were pooled, freeze-dried and

further separated by liquid chromatography on a Lichroprep C_{18} column (33 cm $\times$ 2.5 cm, 43–60 μm particle size) with water/acetonitrile/TFA (90:10:0.05, by vol.) as solvent. Fractions (5–10 ml each) were analysed by HPLC with UV detection; fractions containing only one tryptophan glycoconjugate were pooled. Mixtures were concentrated and reappplied to a Lichroprep C_{18} column. Because tryptophan glycoconjugates of D-mannose could not be separated by this procedure, semipreparative HPLC was performed on a Eurospher 100 C_{18} column (250 mm $\times$ 16 mm internal diam., 5 μm particle size) with water/acetonitrile/TFA (87:13:0.05, by vol.) as solvent at a flow rate of 7.5 ml/min. Structure elucidation was performed by HPLC–MS, HPLC–MS/MS and also by ^1H -NMR, ^{13}C -NMR and heteronuclear multiple-bond connectivity (HMBC) experiments. Signal assignments were confirmed by distortionless enhancement by polarization transfer, homonuclear COSY (HH-COSY) and ^1H -detected heteronuclear multiple-quantum coherence (HMQC) experiments.

Isolation of tryptophan glycoconjugate (compound **1**) from human urine

Human urine (4 litres) was collected from healthy individuals ($n = 7$). After being freeze-dried, the concentrated samples were adjusted to pH 9–10, filtered and subjected to anion-exchange chromatography on Dowex 1 $\times$ 8 (100–200 mesh, exchange capacity 1.2 meq/ml wet resin, 100 g of dry material). Before sample application, anion-exchange material was washed with methanol, 2 M NaOH and water, then loaded with 2 M ammonium acetate solution (500 ml) and again washed with water (1.5 litres). After loading with sample, the column was washed with 500 ml of water/acetonitrile (95:5, v/v), collecting four fractions of about 125 ml, and then eluted with 500 ml of water/acetonitrile/acetic acid (90:5:5, by vol.). Tryptophan glycoconjugate **1** was eluted during the wash with water/acetonitrile (95:5, v/v). The collected fractions were analysed by HPLC with UV detection; fractions 3 and 4 containing compound **1** were pooled and freeze-dried. Further separation was performed by liquid chromatography on a Lichroprep C_{18} column (33 cm $\times$ 2.5 cm, 43–60 μm particle size) with 1.5 litres of water containing 0.05% TFA (v/v) and 500 ml of water/acetonitrile/TFA (90:10:0.05, by vol.) as solvents. Fractions (5–10 ml each) were analysed by HPLC with UV detection; fractions 145–172 containing compound **1** were collected. After being freeze-dried, compound **1** was further purified by semipreparative HPLC on a Eurospher 100 C_{18} column (250 mm $\times$ 16 mm internal diam., 5 μm particle size) with water/acetonitrile/TFA (95:5:0.05, by vol.) as solvent. The flow rate was 7.5 ml/min; the retention time of compound **1** was 41 min. The separation procedure yielded 3 mg of compound **1**. Structure elucidation of compound **1** was performed by HPLC with UV detection, HPLC–MS, HPLC–MS/MS and also by one-dimensional and two-dimensional NMR experiments (^1H -NMR and ^{13}C -NMR, distortionless enhancement by polarization transfer, HH-COSY, HMQC, HMBC and NOESY experiments), as follows. UV/visible [water/acetonitrile/TFA (90:10:0.05, by vol.)], $\lambda_{\text{max}} = 217, 277 \text{ nm}$; ESI–MS, m/z 367 [$M + \text{H}$] $^+$, m/z 365 [$M - \text{H}$] $^-$; ESI–MS/MS (positive mode, 19 eV, 267 mPa Ar, precursor ion m/z 367), m/z 350 [$M + \text{H} - \text{NH}_3$] $^+$, 332 [$350 - \text{H}_2\text{O}$] $^+$, 314 [$332 - \text{H}_2\text{O}$] $^+$, 296 [$314 - \text{H}_2\text{O}$] $^+$, 276, 271 [$M + \text{H} - \text{C}_2\text{H}_4\text{O}_2 - 2\text{H}_2\text{O}$] $^+$, 258 [$276 - \text{H}_2\text{O}$] $^+$, 254 [$314 - \text{C}_2\text{H}_4\text{O}_2$] $^+$, 247 [$M + \text{H} - \text{C}_4\text{H}_8\text{O}_4$] $^+$, 229 [$247 - \text{H}_2\text{O}$] $^+$, 205 [tryptophan + H] $^+$, 202, 188 [$205 - \text{NH}_3$] $^+$, 144 [$188 - \text{CO}_2$] $^+$; ESI–MS/MS (negative mode, 19 eV, 267 mPa Ar, precursor ion m/z 365), m/z 245 [$M - \text{H} - \text{C}_4\text{H}_8\text{O}_4$] $^-$.

Acid hydrolysis of compound **1** and preparation of methoxyoximacetate (MOA) derivatives

The isolated tryptophan glycoconjugate **1**, and also reference compounds **2**, **6** and **10**, were treated with 2 M TFA in a reaction tube at 100 °C for 6 h [11,12]. After being freeze-dried, the liberated monosaccharides were converted into the methoxyoxime acetates as described [13]. Remaining water was removed with anhydrous Na₂SO₄ and the samples were concentrated under vacuum to 50 µl. To obtain the reference compounds, D-glucose, D-mannose, D-galactose, D-talose, D-altrose, D-allose, D-gulose and D-idose (1 mg of each) were treated in the same manner and diluted with dichloromethane to 0.1 mg/ml before GLC-MS analysis.

RESULTS AND DISCUSSION

By applying HPLC-ESI-MS/MS to the detection and identification of polar and labile metabolites [14], an analysis of human urine reveals the occurrence of a single tryptophan glycoconjugate characterized by a molecular ion m/z 367 [$M+H$]⁺. As no sample preparation and preconcentration step is required before these initial analyses, the formation of compound **1** as an artifact can be excluded. The product ion spectrum of compound **1** as obtained by low-energy collision-induced dissociation (CID) of the protonated molecule (ESI positive mode) shows a fragmentation pattern dominated by the loss of 120 atomic mass units, corresponding to the cleavage of a C₄H₈O₄ moiety characteristic of C-glycosyl compounds [2,15,16]. In addition, sequential retro-aldol cleavage occurs, combined with a loss of water. Low-energy CID of the deprotonated molecular ion (ESI negative mode) yields a product ion m/z 245, again resulting from the characteristic loss of a C₄H₈O₄ moiety. Thus the low-energy CID of compound **1** proves the C-glycosyl attachment of an aldohexose. However, no neutral loss of 73 atomic mass units is observed, which is indicative of tetrahydro-β-carboline-3-carboxylic acids [17]. Consequently, further mass spectrometric differentiation between C-glycosyl- and tetrahydro-β-carboline structures required authentic reference compounds.

To obtain a deeper insight into the reactivity of tryptophan towards aldohexoses and to obtain spectroscopic data of tryptophan glycoconjugates we performed model reactions, heating tryptophan with D-glucose, D-galactose and D-mannose to 80 °C at pH 1 or 2 for 12 days. HPLC-ESI-MS of each model reaction yielded up to four tryptophan glycoconjugates from each of the

Figure 3 HPLC-MS analysis of model reaction of tryptophan with D-glucose, gradient I-B, full scan 200–500 atomic mass units

Reconstructed ion chromatogram for molecular ion m/z 367 [$M+H$]⁺.

Figure 4 Product ion spectra of tryptophan glycoconjugates, precursor ions m/z 367 [$M+H$]⁺ (positive mode, 19 eV; 267 mPa Ar)

(A) Tryptophan-N-glycosides; (B) tryptophan-C-glycosyl compounds and polyol-tetrahydro-β-carbolines.

Figure 5 Product ion spectra of tryptophan glycoconjugates, precursor ions m/z 365 [$M-H$]⁻ (negative mode, 19 eV; 267 mPa Ar)

(A) Tryptophan-C-glycosyl derivatives; (B) polyol-tetrahydro-β-carbolines.

aldohexoses with molecular ions m/z 367 [$M+H$]⁺ other than the well-known Amadori rearrangement products (Figure 3) [10,18].

Product ion spectra of the tryptophan glycoconjugates **2–5** as obtained by low-energy CID of the protonated and deprotonated molecular ions are outlined in Figures 4 and 5. Low-energy CID of molecular ions m/z 367 [$M+H$]⁺ (Figure 4A) and m/z 365 [$M-H$]⁻ (results not shown) of **3** are dominated by the loss of an

Figure 6 Postulated reaction products of tryptophan with aldohexoses

intact anhydro-sugar moiety $C_6H_{10}O_5$ (loss of 162 atomic mass units). This strongly suggests that glucose is N-glucosidically linked to the indole moiety [19]. Remarkably, product ion spectra of protonated molecular ions m/z 367 $[M + H]^+$ of compounds **2**, **4** and **5** in Figure 4(B) are almost identical with the product ion

spectrum of compound **1**. Subsequently, we succeeded in differentiating between compounds **2**, **4** and **5** by low-energy CID of the deprotonated molecular ions m/z 365 $[M - H]^-$. The product ion spectrum of compound **2** (Figure 5A) is dominated by a typical loss of 120 atomic mass units (m/z 245), whereas the major product ion m/z 215 of compounds **4** and **5** results from the loss of a $C_5H_{10}O_5$ moiety (Figure 5B). Additional fragment ions of **4** and **5** arise from the precursor ion by the subsequent loss of H_2O , CO_2 , CH_2O , $C_2H_4O_2$ and $C_4H_8O_4$. These results suggest that **2**, **4** and **5** are C-glycosidically linked to the indole moiety. In addition, product ion spectra demonstrate that tryptophan glycoconjugate **1** from human urine is structurally related to compound **2**. Whereas MS/MS experiments unequivocally allow identification of the linkage position, it should be noted that epimeric manno-, gluco- and galacto-conjugates cannot be differentiated by low-energy CID. With regard to the reactivity of the indole moiety towards electrophilic attack, the potential reaction products of tryptophan with aldohexoses are summarized in Figure 6.

To confirm the structures of the tryptophan glycoconjugates outlined in Figure 2, the conjugates were isolated from model reactions by RP-18 column chromatography. 1H -NMR and ^{13}C -NMR data are compiled in Tables 1 and 2. Signal assignment was confirmed by HH-COSY, HMQC and HMBC experiments. As an example we discuss the structure elucidation of tryptophan

Table 1 1H NMR (400 or 600 MHz) data for tryptophan glycoconjugates 2–12

The solvent was $C^2H_5O_2H/CF_3COO^2H$, $[(C^2H_5)_2SO/CF_3COO^2H$ for H-1] except where indicated otherwise.

Signal	Chemical shift (p.p.m.)										
	2*	3	4	5	6	7	8	9	10	11	12a (12b)
Tryptophan											
H-1	11.20	—	10.96	11.07	11.14	—	11.03	11.10	11.13	10.87	10.53
H-2	—	7.361	—	—	—	7.404	—	—	—	—	—
H-4	7.597	7.617	7.505	7.493	7.598	7.615	7.506	7.507	7.540	7.501	7.50 (7.49)†
H-5	7.127	7.137	7.056	7.068	7.051	7.131	7.065	7.073	7.066	7.063	7.06†
H-6	7.223	7.231	7.141	7.161	7.145	7.219	7.155	7.167	7.134	7.155	7.15†
H-7	7.466	7.573	7.378	7.382	7.366	7.630	7.357	7.384	7.334	7.364	7.44†
H-8a	3.603	3.521	3.481	3.442	3.555	3.518	3.41†	3.463	3.607	3.451	3.46 (3.45)†
H-8b	3.462	3.305	3.138	3.228	3.431	3.326	3.41†	3.231	3.539	3.235	3.21†
H-9	4.304	4.280	4.402	4.440	4.299	4.274	4.797	4.442	3.95†	4.940	4.42 (4.50)†
Sugar moiety											
H-1'	4.653	5.502	5.015	5.096	4.505	5.432	5.307	5.009	5.068	5.397	5.02 (5.05)†
H-2'	3.770	3.896	4.254	4.624	4.026	4.237	4.618	4.750	4.293	4.290	4.23 (4.24)†
H-3'	3.643	3.637	4.092	4.281	3.644	3.76†	3.62†	4.019	3.373	4.224	4.28 (4.28)†
H-4'	3.461	3.502	4.051	3.935	3.972	4.020	3.812	3.811	3.836	3.764	3.89 (3.88)†
H-5'	3.64†	3.603	3.745	3.80†	3.77†	3.844	3.79†	3.968	3.495	3.746	3.80†
H-6'a	4.00†	3.877	3.804	3.83†	3.77†	3.76†	3.62†	3.72†	3.957	3.834	3.86†
H-6'b	3.70†	3.718	3.681	3.732	3.77†	3.75†	3.62†	3.70†	3.940	3.708	3.73†
Signal	Coupling constant (Hz)										
	2*	3	4	5	6	7	8	9	10	11	12a (12b)
$J_{H-1',H-2'}$	9.4	9.2	8.4	1.6	9.6	9.2	5.3	1.5	0.9	2.2	6.2 (6.1)
$J_{H-2',H-3'}$	9.3	8.9	3.4	4.0	9.2	9.3	1.9	1.9	2.9	5.4	8.4 (8.4)
$J_{H-3',H-4'}$	9.2	9.0	1.4	1.7	2.9	3.2	8.2	8.9	9.5	0	1.0 (—)†
$J_{H-4',H-5'}$	9.2	9.7	7.4	7.4	0.9	0.8	1.6	1.8	9.6	8.4	8.0 (—)†
$J_{H-5',H-6'a}$	—†	2.3	4.1	4.1	—†	—†	—†	—†	2.6	3.9	3.6
$J_{H-5',H-6'b}$	—†	5.4	5.3	4.6	—†	—†	—†	—†	3.4	5.6	5.5
$J_{H-6'a,H-6'b}$	—†	—12.0	—11.0	—10.4	—†	—†	—†	—†	—11.0	—11.4	—11.2

* The solvent was C^2H_5CN/CF_3COO^2H .

† Owing to strong overlap, exact values of chemical shifts and coupling constants cannot be given.

Table 2 ^{13}C -NMR (150 MHz) data for tryptophan glycoconjugates 2–12The solvent was $\text{C}^2\text{H}_5\text{O}^2\text{H}/\text{CF}_3\text{COO}^2\text{H}$.

Signal	Chemical shift (p.p.m.)										
	2	3	4	5	6	7	8	9	10	11	12a (12b)
Tryptophan											
CO_2H	171.6	171.8	172.6	170.9	171.6	171.6	171.0	170.7	171.8	171.4	171.2 169.7
C-2	136.0	126.2	129.5	128.8	136.1	126.3	127.2/127.1*	128.3/ 127.2*	134.3	128.2/127.4*	128.2/127.0*
C-3	106.8	109.8	107.2	108.7	107.0	109.4	107.0	108.4	107.9	107.8	108.3 (108.0)
C-3a	128.8	129.4	127.1	127.5	128.7	129.2	127.2/127.1*	128.3/ 127.2*	129.1	128.2/127.4*	128.2/127.0* (128.0)
C-4	119.2	119.8	119.1	119.0	119.2	119.5	119.1	118.8	118.2	118.9	118.9
C-5	120.4	121.7	120.3	120.7	120.4	121.4	120.7	120.6	120.4	120.6	120.7
C-6	123.3	124.0	123.5	123.6	123.4	123.6	123.7	123.5	123.2	123.6	123.7
C-7	112.5	112.0	112.4	112.5	112.4	112.0	112.5	112.3	112.2	112.4	112.8
C-7a	137.7	139.0	138.5	138.7	137.6	138.7	138.6	138.6	137.5	138.5	138.4
C-8	26.4	27.6	23.7	23.0	26.2	27.4	23.0	22.8	26.9	22.7	23.4 (23.2)
C-9	54.4	54.4	55.3†	57.1†	54.5	54.2	54.8†	56.8†	55.7	55.2†	57.4 (57.3)†
Sugar moiety											
C-1'	75.4	86.4	56.0	57.6	76.4/75.9*	87.0	57.7	60.4	77.6	55.8	58.7 (59.0)
C-2'	74.8	74.2	72.2	71.4	71.9	71.4	69.9	68.6	73.5	72.0	71.6
C-3'	79.6	79.1	72.9	72.1	76.4/75.9*	75.7	72.6	74.7	75.8	73.3	73.9 (73.7)
C-4'	71.6	71.7	70.9	72.2	71.1	70.6	71.6/71.2*	71.6/71.3*	67.5	72.0	70.8
C-5'	81.6	80.8	73.1	72.9	80.9	79.3	71.6/71.2*	71.6/71.3*	81.3	72.9	73.0
C-6'	62.7	62.8	64.5	64.6	63.0	62.6	64.7	64.6	61.5	64.7	64.7

* Assignments are exchangeable.

† C-1' and C-9 of polyol-tetrahydro- β -carboline correspond to C-1 and C-3 of tetrahydro- β -carboline. For greater clarity, all tryptophan glycoconjugates are numbered in the same way; thus the signal designation of polyol-tetrahydro- β -carboline is not consistent with IUPAC rules.

glucoconjugates **2–5** on the basis of NMR data. The signal at approx. δ_{H} 11 p.p.m. in the ^1H -NMR spectrum of tryptophan in DMSO is assigned to indole H-1. The absence of this signal from the ^1H -NMR spectrum of **3** recorded in DMSO implies that the sugar moiety is N-glucosidically linked to the indole H-1. This is further substantiated by the ^{13}C chemical shift of C-1' (δ_{C} 86.4 p.p.m.), which is comparable to the ^{13}C -NMR data of the anomeric carbon of N^{α} -acetyl-1-(β -D-glucopyranosyl)-D,L-tryptophan amide (δ_{C} 85.6 p.p.m.) [20]. The ^1H -NMR spectrum of **3** reveals that the ^1H chemical shifts and spin-coupling topology are similar to those observed for β -D-glucopyranose; only shifts for protons H-1' and H-2' are higher [21]. Both the chemical shift and the vicinal ^1H – ^1H coupling constant of H-1' (δ_{H} 5.502 p.p.m., $^3J_{\text{H-1',H-2'}}$ 9.2 Hz) confirm the N- β -glucosidic linkage of the sugar moiety [22]. In addition, deshielding of C-5' results in a shift of C-5' to δ_{C} 80.8 p.p.m., which proved to be characteristic of glycopyranose derivatives with replacement of the anomeric hydroxy group by an aromatic residue [23,24]. Thus the identity of **3** was unambiguously established as N^1 -(β -D-glucopyranosyl)-L-tryptophan.

The ^1H -NMR spectrum of **2** dissolved in DMSO reveals the presence of indole H-1 (δ_{H} 11.20 p.p.m.) and the absence of indole H-2. This implies a substitution at C-2 of tryptophan. In addition, ^1H chemical shifts and vicinal ^1H – ^1H coupling constants are similar to those observed for β -D-glucopyranose, indicating a ring-closed $^4\text{C}_1$ conformation of the sugar moiety [21]. This is substantiated by a higher shift of C-5' (δ_{C} 81.6 p.p.m.) [23,24] and HMBC cross-peaks between C-1'/H-5' and C-5'/H-1'. The vicinal ^1H – ^1H coupling constant of H-1' ($^3J_{\text{H-1',H-2'}}$ 9.4 Hz) again indicates the β configuration; remarkably, there is no higher shift of H-1' (δ_{H} 4.653 p.p.m.) as observed for **3**. The chemical shift of C-1' (δ_{C} 75.4 p.p.m.) confirms the C-glucosidic linkage of **2** [25], as does the higher shift of C-2 of the indole moiety (δ_{C} 136.0 p.p.m.), which has also been observed for 2-[2,3-dihydroxy-1-(3-indolyl)propyl]-L-tryptophan and 2-(3-indolylmethyl)-L-

tryptophan (δ_{C} 139.8 and δ_{C} 137.4 p.p.m.) [26]. Accordingly, **2** was identified as 2-(β -D-glucopyranosyl)-L-tryptophan. Note that it is not possible to make a definite assignment of the anomeric configuration of 2-(D-mannopyranosyl)-L-tryptophan **10** on the basis of $^3J_{\text{H-1',H-2'}}$ (α -mannose 1.8 Hz, β -mannose 1.5 Hz [21]). Finally, compound **10** was established as the β anomer according to the characteristic $^1J_{\text{H-1',C-1'}}$ coupling constant (148 Hz) [27].

^1H -NMR and ^{13}C -NMR spectra of **4** and **5** are quite similar. Differences in ^1H -NMR are observed for chemical shifts of H-1', H-2' and H-3' as well as in the coupling constants of H-1' (**4**, $^3J_{\text{H-1',H-2'}}$ 8.4 Hz; **5**, $^3J_{\text{H-1',H-2'}}$ 1.6 Hz) probably caused by stereochemical effects. In addition, **5** shows higher shifts of C-1' and C-9 than **4** (see Figure 2 for signal notation). This is characteristic of diastereomeric tetrahydro- β -carboline: carbon signals C-1 and C-3 assigned to the *cis* isomers have been shown to shift higher than those of *trans* isomers [28,29]. Again, the presence of indole H-1 and the absence of indole H-2 implies a substitution at C-2 of tryptophan. In particular the ^1H and ^{13}C chemical shifts and also the spin coupling topology of the sugar moiety are different from those observed for **2** and **3**. Vicinal ^1H – ^1H coupling constants except $^3J_{\text{H-1',H-2'}}$ and $^3J_{\text{H-2',H-3'}}$ are similar to the coupling constants of glucitol, the corresponding alditol of glucose [30]. This indicates an open structure of the sugar moiety combined with a tetrahydro- β -carboline residue. Tetrahydro- β -carboline structure is further substantiated by the ^{13}C chemical shifts of C-1' (**4**, δ_{C} 56.0 p.p.m.; **5**, δ_{C} 57.6 p.p.m.) and C-2 (**4**, δ_{C} 129.5 p.p.m.; **5**, δ_{C} 128.8 p.p.m.) which are comparable to the chemical shifts of well-known tetrahydro- β -carboline [26,28]. In addition, HMBC cross-peaks between C-1'/H-9 and C-9/H-1' confirm the closure of the tetrahydropyrido ring system. Thus **4** was established as *trans*- and **5** as *cis*-1-[(1*R*,2*R*,3*S*,4*S*)-1,2,3,4,5-pentahydroxypent-1-yl]-1,2,3,4-tetrahydro- β -carboline-3-carboxylic acid.

To the best of our knowledge, this is the first time that tryptophan C-glycosyl compounds or polyol-tetrahydro- β -carboline have been identified in non-enzymic condensation re-

Figure 7 HPLC–MS/MS analysis of soy sauce and fruit syrup, gradient I–A, SRM experiment SRM 1: m/z 367/247, 367/229 (19 eV; 267 mPa Ar)

(A) Reconstructed ion chromatogram of soy sauce; (B) reconstructed ion chromatogram of fruit syrup. Abbreviation: ARP, Amadori rearrangement product.

actions of tryptophan and aldohexoses. It should be mentioned that product ion spectra of the protonated and deprotonated tryptophan glycoconjugates reveal characteristic fragmentation

patterns (Figures 4 and 5), thus, enabling us to differentiate tryptophan glycoconjugates in complex samples by HPLC–ESI–MS/MS according to their linkage type.

In order to clarify the relevance of the newly identified tryptophan glycoconjugates, several food samples including seasoning sauces, alcoholic beverages and fruit products were analysed by means of HPLC–ESI–MS/MS. Detection limits of approx. 100 ng/ml were achieved with SRM. Most food samples under study contained tryptophan glycoconjugates in the range of 100 ng/ml to 10 μ g/ml. As outlined in Figure 7, the profile pattern of tryptophan glycoconjugates depends strongly on the raw materials and the processing conditions.

As mentioned above, mass spectrometric analysis of human urine demonstrated that tryptophan metabolite **1** is structurally related to compound **2**, which was identified unambiguously as 2-(β -D-glucopyranosyl)-L-tryptophan. In addition, CID of compound **1** was significantly different from the product ion spectra of isolated polyol-tetrahydro- β -carbolines compounds **4** and **5**. Thus the structure of compound **1** remained to be established. Therefore we isolated **1** and verified the structure by ^1H -NMR and ^{13}C -NMR spectroscopy as well as by HH-COSY, HMQC and HMBC experiments (Table 3).

Table 3 ^1H -NMR (600 MHz), ^{13}C -NMR (150 MHz) and HMBC-NMR data for compound **1** (Figure 1)

Abbreviation: n.c., not calculated.

^{13}C -NMR		^1H -NMR			HMBC cross-peaks	
C-atom	$\delta_{\text{C}}^{\dagger}$ (p.p.m.)	H-atom	δ_{H}^* (p.p.m.)	$\delta_{\text{H}}^{\ddagger}$ (p.p.m.)	$^2J_{\text{CH}}^{\ddagger}$ (Hz)	$^3J_{\text{CH}}^{\ddagger}$ (Hz)
Tryptophan		Tryptophan				
CO ₂ H	171.8					
C-2	136.3					
C-3	107.8					
C-3a	128.5					
C-4	119.0	H-4	7.669	7.576		C-3, C-6, C-7a
C-5	120.5	H-5	7.154	7.061		C-7, C-3a
C-6	123.5	H-6	7.252	7.150		C-4, C-7a
C-7	112.5	H-7	7.473	7.382		C-5, C-3a
C-7a	137.8	H-8b	3.519	3.635	C-9, C-3	C-2, C-3a
C-8	27.0	H-8a	3.344	3.360	C-9, C-3	C-2, C-3a
C-9	54.6	H-9	4.063	4.159		C-3
Sugar moiety		Sugar moiety				
C-1'	66.1	H-1'	5.113	5.052	C-2', C-2	C-3', C-5', C-3
C-2'	69.1	H-2'	4.373	4.289	C-1'	C-2
C-3'	72.4	H-3'	4.062	4.058		
C-4'	71.1	H-4'	3.891	3.887	C-3'	C-2'
C-5'	81.5	H-5'	3.835	3.959	C-4', C-6'	C-1', C-3'
C-6'	60.5	H-6'b	4.191	4.329	C-5'	
		H-6'a	3.684	3.682		
Coupling constants (Hz)						
A, B		$J_{\text{A,B}}^*$	$J_{\text{A,B}}^{\dagger}$	$J_{\text{A,B}}^{\ddagger}$		
H-1', H-2'		8.2	9.4	9.3		
H-2', H-3'		3.3	3.2	2.3		
H-3', H-4'		5.0	4.2	4.0		
H-4', H-5'		3.3	1.8	2.4		
H-5', H-6'b		9.0	9.2	n.c.		
H-5', H-6'a		3.1	3.7	n.c.		
H-6'a, H-6'b		–12.6	–12.2	n.c.		

* The solvent was $^2\text{H}_2\text{O}$.

$\dagger$ The solvent was $\text{C}_2\text{H}_3\text{O}_2\text{H}/\text{CF}_3\text{COO}^-\text{H}$.

$\ddagger$ Coupling constants were estimated by a Karplus-type equation [8] with the PC program ALTONA [9].

Figure 8 Section of the one-dimensional NMR spectrum of 2-(α -mannopyranosyl)-L-tryptophan (**1**)

The notation is as depicted in Figure 1. The corresponding NOESY spectrum, recorded with a mixing time of 800 ms, is also shown. Selected NOESY cross-peaks are labelled.

The chemical shifts and coupling constants of compound **1** correspond to reported results [7], but the topological assignments of the ^{13}C -NMR signals are different. In the ^1H -NMR spectrum of compound **1**, the signal of the indole H-2 is absent, indicating that tryptophan is substituted at C-2 by the hexose moiety (indole H-1 was not observable owing to proton exchange in the applied solvents $\text{C}^2\text{H}_3\text{O}^2\text{H}$ and $^2\text{H}_2\text{O}$). However, compared with NMR data of the tryptophan-C-glycosyl compounds **2**, **6** and **10** and the polyol-tetrahydro- β -carboline **4**, **5**, **8**, **9**, **11** and **12** (see Figure 2), higher shifts of C-5' (δ_{C} 81.5 p.p.m.) and C-2 (δ_{C} 136.3 p.p.m.) as well as HMBC cross-peaks between C-1'/H-5' and C-5'/H-1' prove the C-glycosidic structure of compound **1**. Consequently, compound **1** was characterized unambiguously as 2-(hexopyranosyl)-L-tryptophan. The coupling constant $^3J_{\text{H-1}',\text{H-2'}}$ of 9.4 Hz in acidic methanol indicates an axial orientation of protons H-1' and H-2', but the configuration of the hexose moiety cannot be determined simply by a comparison of chemical shifts and coupling constants with NMR data of most monosaccharides, disaccharides or reference compounds **2**, **6** and **10** [21,25]. Interestingly, the signal of C-1' (δ_{C} 66.1 p.p.m.) was different from those of other C-glycosyl compounds (δ_{C} 70–80 p.p.m.) [25]. In addition, the shift of the hydroxymethyl proton H-6'b (δ_{H} 4.329 p.p.m. in $\text{C}^2\text{H}_3\text{O}^2\text{H}/\text{CF}_3\text{COO}^2\text{H}$; δ_{H} 4.190 p.p.m. in $^2\text{H}_2\text{O}$) was significantly higher [21]. Configurational and conformational analysis of compound **1** in acidic methanol was performed by two-dimensional NOESY. A section of the ^1H -NMR spectrum of compound **1** and the corresponding NOESY spectrum are shown in Figure 8. In the nuclear Overhauser effect (NOE) spectrum of compound **1**, an intense NOE between H-1' and the hydroxymethyl proton H-6'b is observed. The proximity of the protons H-1' and H-6'b indicates a predominant axial orientation of the hydroxymethyl group. NOE between H-1' and H-6'b also shows that these protons are located on the same side

of the ring; thus the hexopyranosyl moiety of compound **1** should be in the α configuration. According to de Beer et al. [3], the large vicinal coupling constant $^3J_{\text{H-1}',\text{H-2'}}$ of 9.4 Hz in acidic methanol in combination with a weak NOE between H-1' and H-2' excludes the possibility of a *gluco* configuration. In addition, a clear vicinal NOE between H-2' and H-3' and the small coupling constant $^3J_{\text{H-2}',\text{H-3'}}$ indicate a *manno* or *talo* configuration for the hexose moiety. The intense NOE between H-4' and H-6'a shows that H-4' is close to H-6'a, proving an equatorial orientation for H-4' as given in a *manno* configuration. In summary, the carbohydrate moiety of compound **1** was identified as an α -mannosyl derivative. A comparison of the ^1H -NMR and ^{13}C -NMR spectra of compound **1** with NMR data of peptide-bound 2-(α -mannopyranosyl)-L-tryptophan in RNase [2,3] confirmed the structure elucidation.

Interestingly, a change of NMR solvent from acidic methanol to water resulted in a variation in the chemical shifts and coupling constants of compound **1** (Table 3). De Beer et al. [3] claimed that the mannosyl residue of peptide-bound 2-(α -mannopyranosyl)-L-tryptophan in water exists as dynamic equilibrium mixture of several ring conformations. However, with regard to the conformational equilibrium of compound **1**, the equatorial orientation of the tryptophan group at C-1', together with the axial orientation of the hydroxymethyl group, is predominant in acidified methanol. Together with the missing NOE between H-3' and H-5' (Figure 8) [3], this spatial arrangement indicates that the respective α -anomeric C-mannosyl hexopyranose is preferentially adopting a $^1\text{C}_4$ conformation. This was substantiated by a comparison of the experimentally determined coupling constants of compound **1** with coupling constants estimated by applying a Karplus-type equation (see Table 3) [8,9]. Remarkably, model reactions of glucose, galactose and mannose yielded exclusively tryptophan- β -C-glycosyl derivatives with a $^4\text{C}_1$ conformation.

For confirmation of the structure elucidation, the sugar was liberated by treating compound **1** with acid. Subsequent GLC-MS analysis of MOA derivatives revealed the presence of mannose and glucose (the latter probably formed by acid-catalysed epimerization), thereby excluding the presence of talose. In addition, the same pattern of MOA derivatives could be observed after acid treatment, MOA derivatization and GLC-MS analysis of 2-(β -D-mannopyranosyl)-L-tryptophan (compound **10**). Our data do not permit differentiation between D- and L-mannopyranose. However, because compound **1** is a human metabolite, a D-mannose derivative seems much more probable. Consequently, the identity of compound **1** was unambiguously established as 2-(α -mannopyranosyl)-L-tryptophan preferentially adopting a $^1\text{C}_4$ conformation (Figure 1).

HPLC-ESI-MS/MS analysis of food samples demonstrated the presence of numerous tryptophan glycoconjugates in dietary sources. However, significant differences were observed between the profiles and conformations of tryptophan glycoconjugates originating from dietary sources and human urine. Hence the C-2 mannosylation of tryptophan residues [1] represents a novel enzymic pathway in tryptophan metabolism in humans. This view is supported by the observation that the post-translational modification of RNase relies on a specific enzymic reaction [5,6]. With regard to the formation of compound **1**, one has to consider the enzymic glycosylation of free tryptophan as well as the degradation of peptides containing mannosylated tryptophan residues. Although the clarification of the mechanism of the biosynthesis of compound **1** and its role in human metabolism require further study, the identification of compound **1** will trigger the elucidation of the biosynthetic pathway leading to this unusual human tryptophan glycoconjugate.

Conclusion

In summary, *N*¹-(β -D-hexopyranosyl)-L-tryptophan, 2-(β -D-hexopyranosyl)-L-tryptophan and 1-(1,2,3,4,5-pentahydroxypent-1-yl)-1,2,3,4-tetrahydro- β -carboline-3-carboxylic acids have been identified for the first time in several food samples as well as in non-enzymic condensation reactions of tryptophan with aldohexoses. Configurations were unambiguously established by HPLC–MS/MS and NMR spectroscopy. In addition, the tryptophan glycoconjugate in human urine was identified as 2-(hexopyranosyl)-L-tryptophan by the use of MS/MS as well as one-dimensional and two-dimensional NMR spectroscopy. NOE experiments and acid hydrolysis followed by the GLC–MS analysis of MOA derivatives enabled us to determine the configuration and conformation of the hexopyranosyl moiety. The metabolite was identified as 2-(α -mannopyranosyl)-L-tryptophan with predominant axial orientation of the hydroxymethyl group, indicating that an enzymic pathway is involved in the glycosylation of tryptophan in humans, as has recently been demonstrated for the post-translational modification of RNase [5,6].

We thank E. Ruckdeschel, M. Roth and B. Pink for assistance with NMR spectroscopy, and Professor P. Schreier for providing access to the excellent instrumental facilities at the chair of food chemistry. This study was supported by the Deutsche Forschungsgemeinschaft (DFG), Bonn (He 2599/1-3).

REFERENCES

- Gaede, G., Kellner, R., Rinehart, K. L. and Proefke, M. L. (1992) *Biochem. Biophys. Res. Commun.* **189**, 1303–1309
- Hofsteenge, J., Mueller, D. R., de Beer, T., Loeffler, A., Richter, W. J. and Vliegthart, J. F. G. (1994) *Biochemistry* **33**, 13524–13530
- de Beer, T., Vliegthart, J. F. G., Loeffler, A. and Hofsteenge, J. (1995) *Biochemistry* **34**, 11785–11789
- Loeffler, A., Doucey, M.-A., Jansson, A. M., Mueller, D. R., de Beer, T., Hess, D., Meldal, M., Richter, W. J., Vliegthart, J. F. G. and Hofsteenge, J. (1996) *Biochemistry* **35**, 12005–12014
- Doucey, M.-A., Hess, D., Cacan, R. and Hofsteenge, J. (1998) *Mol. Biol. Cell* **9**, 291–300
- Krieg, J., Hartmann, S., Vicentini, A., Glaesner, W., Hess, D. and Hofsteenge, J. (1998) *Mol. Biol. Cell* **9**, 301–309
- Horiuchi, K., Yonekawa, O., Iwahara, K., Kanno, T., Kurihara, T. and Fujise, Y. (1994) *J. Biochem. (Tokyo)* **115**, 362–366
- Haasnoot, C. A. G., de Leeuw, F. A. A. M. and Altona, C. (1980) *Tetrahedron* **36**, 2783–2792
- Cerda-Garcia-Rojas, C. M., Zepeda, L. G. and Joseph-Nathan, P. (1990) *Tetrahedron Comput. Methodol.* **3**, 113–118
- Yaylayan, V. and Forage, N. G. (1991) *J. Agric. Food Chem.* **39**, 364–369
- Wang, L.-X., Tang, M., Suzuki, T., Kitajima, K., Inoue, Y., Inoue, S., Fan, J.-Q. and Lee, Y.-C. (1997) *J. Am. Chem. Soc.* **119**, 11137–11146
- Hardy, M., Townsend, R. R. and Lee, Y. C. (1988) *Anal. Biochem.* **170**, 54–62
- Herderich, M. and Schreier, P. (1993) *Phytochem. Anal.* **4**, 256–260
- Lehmann, W. D. (1996) *Massenspektrometrie in der Biochemie, Spektrum Akademischer Verlag, Heidelberg*
- Rath, G., Touré, A., Nianga, M., Wolfender, J. L. and Hostettmann, K. (1995) *Chromatographia* **41**, 332–342
- Li, Q. M., Van den Heuvel, H., Dillen, L. and Claeys, M. (1992) *Biol. Mass Spectrom.* **21**, 213–221
- Gutsche, B. and Herderich, M. (1997) *J. Agric. Food Chem.* **45**, 2458–2462
- Yaylayan, V. and Huyghues-Despointes, A. (1994) *Crit. Rev. Food Sci. Nutr.* **34**, 321–369
- Frear, D. S., Swanson, H. R. and Mansager, E. R. (1989) *J. Agric. Food Chem.* **37**, 1408–1412
- Nyhammar, T. and Pernemalm, P.-A. (1985) *Food Chem.* **17**, 289–296
- Collins, P. and Ferrier, R. (1996) *Monosaccharides, Their Chemistry and Their Roles in Natural Products*, pp. 534–535, Wiley, Chichester
- Avalos, M., Babiano, R., Durán, C. J., Jiménez, J. L. and Palacios, J. C. (1992) *J. Chem. Soc. Perkin Trans. II*, 2205–2215
- Voelter, W. and Breitmaier, E. (1973) *Org. Magn. Reson.* **5**, 311–319
- Kato, T. and Morita, Y. (1993) *Heterocycles* **35**, 965–973
- Agrawal, P. K. (1992) *Phytochemistry* **31**, 3307–3330
- Mueller, A., Brusker, E., Guenther, K. and Hoppe, B. (1991) *Bioforum* **14**, 350–354
- Tvaroska, I. and Taravel, F. R. (1995) *Adv. Carbohydr. Chem. Biochem.* **51**, 15–61
- Cox, E. D. and Cook, J. M. (1995) *Chem. Rev.* **95**, 1797–1842
- Bailey, P. D., Hollinshead, S. P., McLay, N. R., Morgan, K., Palmer, S. J., Prince, S. N., Reynolds, C. D. and Wood, S. D. (1993) *J. Chem. Soc. Perkin Trans. I*, 431–439
- Hawkes, G. E. and Lewis, D. (1984) *J. Chem. Soc. Perkin Trans. II*, 2073–2078

Received 5 October 1998/14 June 1999; accepted 15 July 1999