Indolyl Carboxylic Acids by Condensation of Indoles with α-Keto Acids

Thomas R. Garbe,*,† Miki Kobayashi,† Naoki Shimizu,‡ Naohide Takesue,‡ Masatomi Ozawa,‡ and Hideaki Yukawa†

Mitsubishi Chemical Corporation, Tsukuba Research Center, 8-3-1, Chuo, Ami, Inashiki, Ibaraki, 300-0332 Japan

Received October 15, 1999

The novel indole derivatives 2,2-bis(3,3′-indolyl)propionic acid (**1**); 1,1,1,-tris(3,3′,3′′-indolyl)ethane (**2**); and 2,2-bis(3,3′-indolyl)isocaproic acid (**3**) were isolated from solvent extracts of indole-supplemented supernatants of *Escherichia coli* and corynebacteria. The compounds were also obtained by chemical synthesis: compounds **1** and **2** from indole and pyruvic acid and compound **3** from indole and α-ketoisocaproic acid, following incubation at 37 °C in aqueous medium. Tryptophan and pyruvic acid gave the novel 2-(2-tryptophanyl)lactic acid (**4**). The condensation reaction between indoles and α-keto acids was of general nature, and the mild reaction conditions suggested it may proceed in vivo. Examples for endogenous occurrence may be the neuro-degenerative diseases phenylketonuria and maple syrup urine disease, both characterized by elevated blood levels of α-keto acids.

Basic and higher life functions depend on indole, which provides the heterocyclic frame for the amino acid tryptophan, numerous alkaloids, plant growth factors, hormones, and neurotransmitters. Because of the broad importance of indole, evidence for bacterially transformed indole was investigated in depth.

Results and Discussion

Thin-layer chromatographic (TLC) analyses of solvent extracts from indole-supplemented cultural supernatants of *Escherichia coli* JM109 and *Brevibacterium flavum* ES consistently gave several fermentation-specific orange-red spots with Ehrlich's reagent.[1] Extracts from *E. coli* supernatants, occasionally also showed a prominent blue spot. The color reactions indicated indole derivatives.[1]

Structural analyses of the blue compound (**1**) from *E. coli* culture extract necessitated prepurification, which was achieved by Si gel column chromatography. On-line HPLC mass spectrometry revealed peaks at *m*/*z* of 188, 259 (M − COOH) (base peak), and 304 ([M]$^+$). HRFABMS gave a mass of 304.1230, consistent with the empirical formula $C_{19}H_{16}O_2N_2$. The results obtained by MS and 1D NMR experiments suggested that the compound consisted of a pair of indole units and a propionic acid as a bridging structure (**1**). In the COSY spectrum, a correlation between H-2 and H-1 was observed. In addition, the HMBC spectrum showed a correlation signal between H-4 and C-3. Neither of these correlations could be observed if the propionic acid were attached to C-2 of the indole unit, demonstrating that the compound was 2,2-bis(3,3′-indolyl)-propionic acid (**1**).

The indole moiety of **1** probably originated in the indole supplement of the growth medium. The origin of its propionic acid moiety was less obvious. To test whether pyruvic acid was the precursor of the propionic acid moiety, the indole-supplemented growth medium was additionally supplemented with 0.2% sodium pyruvate. This resulted in consistently high levels of **1** in the culture supernatants. However, with added pyruvate, **1** occurred to the same extent in the bacteria-free control medium, providing evidence that synthesis of **1** did not rely on an enzyme-catalyzed process, but on a chemical process between indole and pyruvate. If, previously, **1** occurred only in *E. coli* supernatants of indole-supplemented growth medium, and not in sterile control experiments, then this was apparently solely due to bacterial production of pyruvate.

As anticipated, incubation at 37 °C of mixtures of indole and pyruvic acid in aqueous ethanol also gave **1**, according to elemental analysis and TLC. Increasing amounts of water of up to 50% increased the reaction rate, as did excessive pyruvic acid. Both observations are consistent with a formal description of the synthesis of **1**, where the hydrated, geminal diol form of pyruvic acid[2] condenses with indole, resulting in 2-(3-indolyl)lactic acid. Condensation of this hypothetical intermediate with a second indole would yield **1**.

Compound **1** was the sole reaction product when excessive concentrations of pyruvic acid were used. At 2-fold stoichiometric excess of pyruvic acid, however, a second major compound (**2**) was formed. On the basis of TLC analyses, **2** also appeared in cultural supernatants of *B. flavum* when the medium was supplemented with 0.2% sodium pyruvate and twice the indole concentration as used in *E. coli* cultures. Structural analysis by on-line HPLC mass spectrometry revealed peaks at *m*/*z* 142, 243, 259, 360, and 375 ([M]$^+$). The data suggested a tris(indolyl)-ethane (**2**). Both the ^{1}H and the ^{13}C NMR spectra of chromatographically purified **2** were similar to the spectra of **1**, but simpler, indicating that all three indole moieties were attached at position 3. The lack of aliphatic signals, except for a methyl group, suggested that **2** was 1,1,1-tris-(3,3′,3′′-indolyl)ethane.

Another compound was regularly found in supernatants from indole-supplemented cultures of both *E. coli* and *B. flavum* and gave a rust-red reaction with Ehrlich's reagent. On-line HPLC mass spectrometry revealed peaks at *m*/*z* 118, 130, 142, 174, 188, 230, 289, 301 (M − COOH) (base peak), and 346 ([M]$^+$), compatible with a bis(indolyl)caproic acid (**3**). ^{1}H NMR analysis indicated that attachment of the indolyl moiety to the caproic acid bridge was again at position 3, due to an HMBC correlation between H-10 and C-3. The presence of an isocaproic acid bridge was deduced from the 9-fold signal of a single "aliphatic" proton, which can only be expected from the single proton at the branched carbon in position 11. Furthermore, COSY correlated H-11

* Corresponding author, at the Research Institute of Innovative Technology for the Earth, 9-2, Kizugawadai, Kizu-cho, Soraku-gun, Kyoto, 619-0292 Japan. Tel.: (0774) 75-2308. Fax: (0774) 75-2321. E-mail: garbetr@rite.or.jp.
† Biochemicals Laboratory.
‡ Materials Characterization and Analysis Laboratory.

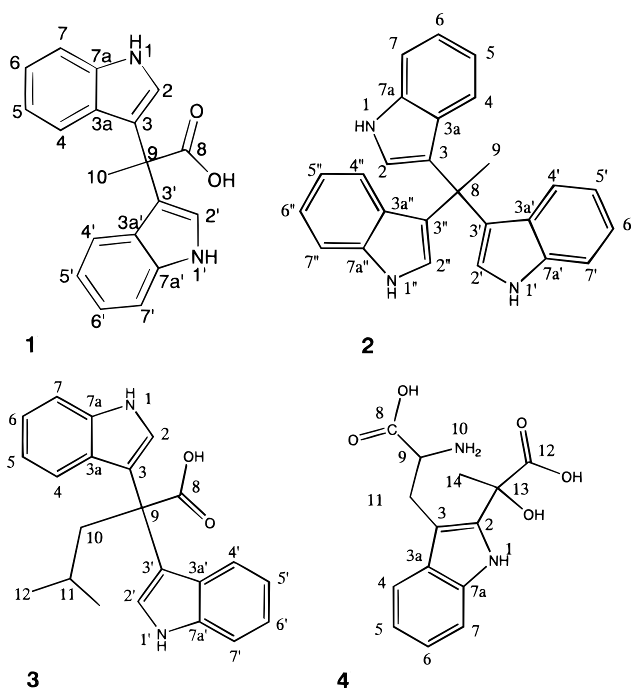
10.1021/np990517s CCC: $19.00 © 2000 American Chemical Society and American Society of Pharmacognosy
Published on Web 04/15/2000

to H-12 and H-10, having integrals of 1, 6, and 2, respectively. Hence, it was concluded that **3** was 2,2-bis(3,3′-indolyl)isocaproic acid.

Subsequently, **3** was recognized in virtually all *E. coli* cultures when grown in LB medium, independent of any indole supplement. Synthesis of **3** probably relied on enzymatically generated indole from tryptophan,[3] which is abundantly present in LB medium. Exclusive formation of **3** may suggest a specific detoxification role of α-keto isocaproic acid for endogenous indole.[4]

Synthesis of **1** and **3** apparently relied on the presence of the α-keto acid function. Other α-keto acids reacted with indole also. All of the α-keto acids tested, including phosphoenolpyruvate, oxalacetic acid, α-ketoglutaric acid, α-ketobutyric acid, and phenylpyruvic acid, readily combined with indole. The reaction with oxalacetic acid coincided with rapid development of gas and yielded mainly compound **1**, according to TLC analyses.

Tryptophan and pyruvic acid reacted to form a precipitate, whose mass spectrum revealed peaks at *m*/*z* of 185, 229, and 273 [M − H_2O]. The data indicated a tryptophan derivative with an acrylic acid or lactic acid side chain (**4**). Analysis by ^1H NMR spectroscopy gave a methyl singlet, indicating a methyl group in position 14, presumably from lactic acid. Attachment of lactic acid to the tryptophan moiety at position 1 was ruled out by the existence of an NH proton signal (δ 10.72). Evidence for attachment to position 2 is two-fold: (a) by the absence of an HMQC correlation signal between H and C at position 2, and (b) by the presence of an HMBC correlation signal between the methyl protons of lactic acid (H-14) and C-2. The data suggest that **4** is 2-(2-tryptophanyl)lactic acid. The attachment of the lactic acid substituent at C-2 rather than at the more reactive N-1[5] suggests substituent migration.[5]



The condensation reaction may also proceed in vivo. Examples of α-keto acid toxicity may be the heritable disorders phenylketonuria and branched-chain ketonuria (maple syrup urine disease).[6,7] Both of these neurologic degenerative disorders show increased serum levels of either phenylalanine or leucine, isoleucine, and valine, as well as of their corresponding α-keto acids. An established molecular explanation for the deleterious effects of any of these metabolites does not exist to date.[6−9]. The condensation reaction suggests that the α-keto acids from ketonuria react with indole metabolites such as tryptophan, melatonin, and serotonin, thus interfering with the balance of those metabolites and creating foreign compounds. Furthermore, the condensation reaction may cause posttranslational modification of tryptophan residues, which is anticipated to result in foreign immunogenic determinants on those proteins.

Experimental Section

General Experimental Procedures. NMR spectra were recorded at 297 K on a JEOL JNM-GSX400 spectrometer, operating at 400 and 100 MHz for ^1H and ^{13}C, respectively. Samples were prepared by dissolving 25−60 mg of compound in 0.6 mL of $CDCl_3$ in a 5-mm NMR tube, except **4**, which was dissolved in DMSO-d_6. The chemical shift of $CDCl_3$ (δ 7.26) was used as the internal standard for **1**, other samples were added with TMS as internal standard. ^1H and ^{13}C NMR spectra were assigned using conventional NMR techniques such as $^1H-^1H$ correlation spectroscopy (COSY), distortionless nuclei enhanced by polarization transfer (DEPT), $^1H-^{13}C$ heteronuclear correlation spectroscopy (HECTOR), $^1H^{13}C$ heteronuclear multiple quantum coherence (HMQC), and ^1H-detected heteronuclear multiple bond correlation spectroscopy (HMBC)]. The field gradient method was used on COSY and HMBC experiments. Chemicals were obtained from Sigma, Aldrich, and Wako.

Fermentation and Solvent Extraction. *E. coli* JM109 was supplied by Takara Corp. (Japan). *B. flavum* ES is a derivative of strain MJ-233, which has been cured of its native plasmid.[10] Extracts from bacterial supernatants were prepared as follows: 1 L of LB medium (per liter of water: 10 g of tryptone, 5 g of yeast extract, and 5 g of NaCl), containing 5 mL of 1 M indole solution in EtOH, was inoculated with 100 mL of overnight culture in LB and incubated at constant shaking at 37 °C (*E. coli*) or 30 °C (*B. flavum*). After 5 days, bacterial supernatants were acidified to pH 4 by addition of 6 mL of 6 M HCl and extracted with EtOAc. After drying with anhydrous $MgSO_4$, the extracts were concentrated in vacuo. The oily residues were extracted with boiling hexane for removal of the bulk of indole. The solid residues were dissolved in 1.5 mL of EtOAc. Extract from sterilely incubated indole−LB medium was prepared analogously. It was used as a control in TLC analyses in order to differentiate between biotransformed and degraded or oxidized indole derivatives. A molecular-biologic analysis of the bacterial responses following indole treatments has been published elsewhere.[4]

Chromatography. TLC was performed on Si gel plates (E. Merck, Darmstadt) using a solvent made fresh from MeOAc, 2-propanol, and 25% ammonia (45:35:20) or a solvent made from isopropyl ether and acetic acid (95:5). Ehrlich's spray reagent[1] was modified and consisted of 2% *p*-dimethylaminobenzaldehyde in concentrated formic acid, rather than ethanolic HCl, providing brighter colors and less background.

The EtOAc extracts from cultural supernatants mostly contained nutrients from the LB medium, which obscured the fermented indole derivative **1** for HPLC purification. Separation was accomplished by column chromatography, using a 90-cm long bed of Si gel 60 (E. Merck) in EtOAc. Compound **1** eluted only after acidification of EtOAc by 1% acetic acid. Further purification by HPLC at room temperature relied on a Migtysil $RP_{18}GP$ 5 μm 150 × 3.0 mm i.d. column. The eluent used was: (A) 0.1% acetic acid in $MeOH/H_2O$ (5:95 [v/v]) and (B) MeOH/acetonitrile (50:50 [v/v]) at a flow rate of 0.6 mL/min.

Synthesis of 1 and 2. Equal volumes of 1 M indole solution in EtOH and 1 M pyruvic acid in water were incubated at 37 °C overnight, resulting in amber-red reaction mixtures. The reaction products were precipitated by adding H_2O. Compound **1** was separated from neutral compounds by dissolving the

precipitate in 2.5% NH_4OH and extracting the neutral compounds four times with $CHCl_3$. The aqueous solution, containing mostly **1**, was acidified with 0.6 M HCl, extracted with EtOAc, dried over anhydrous $MgSO_4$, concentrated in vacuo, and crystallized from toluene. MS and NMR data for compound **1** were obtained from HPLC-purified material isolated from 1 L of culture supernatant as described. Other data were obtained from synthetic material crystallized from toluene.

Compound **2** was isolated from the crude synthetic precipitate by chromatography on Si gel in EtOAc. The major fraction containing **2** was of ultramarine opalescence and contained a minor contaminant, which was identified by on-line HPLC mass spectroscopy as the ethyl ester of **1**: *m*/*z* 332 $[M]^+$ (18), 259 (59), 216 (100), 142 (18).

Synthesis of 3. Compound **3** was analogously prepared by mixing 100 μL of indole solution with 200 μL of 1 M sodium-α-ketoisocaproate and 30 μL of acetic acid. The major reaction product was chromatographically indistinguishable from the compound isolated from *E. coli* culture supernatant. MS and NMR data for compound **3** were obtained from column-chromatographed material isolated from 1 L of culture supernatant of *E. coli* JM109.

Synthesis of 4. L-Tryptophan (10 mM) and pyruvic acid (100 mM) in 50 mL of water were stirred with a magnetic bar at 37 °C. A white precipitate formed overnight and was collected after cooling on ice. It was washed by suspending in 0.25 L of cold H_2O, crystallized from 70 mL of 50% EtOH, and dried at 50 °C overnight. Total yield of **4** was 610 mg (21%) of microcrystalline needles with a lemon-yellow tinge.

2,2-Bis(3,3′-indolyl)propionic acid (1): yellow-orange powder (toluene), which contained some microscopic diamond- and plate-shaped crystals, but mostly amorphous spheres clumped together; upon storage at room temperature, the color changed to violet; mp 109−114 °C; UV (MeOH) λ_{max} (log ϵ) 222 (4.75), 281 (4.04), 290 (3.99), 388 (2.02) and 479 (1.23) nm; ^{1}H NMR ($CDCl_3$, 400 MHz) δ 8.03 (s, H-1, H-1′), 7.01 (d, H-2, H-2′), 7.51 (d, H-4, H-4′), 6.99 (t, H-5, H-5′), 7.16 (t, H-6, H-6′), 7.35 (d, H-7, H-7′), 2.15 (s, H-10); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 123.0 (C-2, C-2′), 118.4 (C-3, C-3′), 125.8 (C-3a, C-3a′), 121.2 (C-4, C-4′), 119.4 (C-5, C-5′), 121.9 (C-6, C-6′), 111.2 (C-7, C-7′), 136.7 (C-7a, C-7a′), 178.9 (C-8), 46.3 (C-9), 25.7 (C-10); EIMS *m*/*z* 304 $[M]^+$ (26), 259 (57), 188 (100); HREIMS *m*/*z* 304.1230 (calcd for $C_{19}H_{16}O_2N_2$, 304.3482); *anal.* C 75.6%, H 5.6%, calcd for $C_{19}H_{16}O_2N_2$, C 74.98%, H 5.29%.

1,1,1-Tris(3,3′,3″-indolyl)ethane (2): ^{1}H NMR ($CDCl_3$, 400 MHz) δ 7.66 (s, H-1, H-1′, H-1″), 6.66 (d, H-2, H-2′, H-2″), 7.43 (d, H-4, H-4′, H-4″), 6.87 (t, H-5, H-5′, H-5″), 7.07 (t, H-6, H-6′, H-6″), 7.24, (d, H-7, H-7′, H-7″), 2.44 (s, H-9); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 123.0 (C-2, C-2′, C-2″), 123.8 (C-3, C-3′, C-3″), 126.4 (C-3a, C-3a′, C-3a″), 121.9 (C-4, C-4′, C-4″), 118.5 (C-5, C-5′, C-5″), 121.1 (C-6, C-6′, C-6″), 111.0 (C-7, C-7′, C-7″), 136.9 (C-7a, C-7a′, C-7a″), 39.3 (C-8), 28.1 (C-9); EIMS *m*/*z* 375 $[M]^+$ (12), 360 (37), 259 (100), 243 (10), 142 (22).

2,2-Bis(3,3′-indolyl)-isocaproic acid (3): ^{1}H NMR ($CDCl_3$, 400 MHz) δ 8.09 (s, H-1, H-1′), 7.37 (d, H-2, H-2′), 7.32 (d, H-4, H-4′), 6.84 (t, H-5, H-5′), 7.06 (t, H-6, H-6′), 7.29 (d, H-7, H-7′), 2.65 (s, H-10), 1.47 (9-fold, H-11), 0.67 (d, H-12); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 123.5 (C-2, C-2′), 116.7 (C-3, C-3′), 126.3 (C-3a, C-3a′), 121.2 (C-4, C-4′), 119.0 (C-5, C-5′), 121.5 (C-6, C-6′), 110.9 (C-7, C-7′), 136.3 (C-7a, C-7a′), 179.0 (C-8), 50.3 (C-9), 44.7 (C-10), 25.5 (C-11), 23.8 (C-12); EIMS *m*/*z* 346 $[M]^+$ (30), 301 (91), 289 (24), 230 (100), 188 (16), 174 (19), 142 (53), 130 (19), 118 (11).

2-(2-Tryptophanyl)lactic acid (4): microcrystalline needles with a lemon-yellow tinge (50% EtOH); mp 204−208 °C; UV (MeOH) λ_{max} (log ϵ) 222 (4.70), 281 (4.08), 290 (3.97) and 353 (2.24) nm, UV (NaOH, 50 mM) λ_{max} (log ϵ) 223 (4.41) and 281 (3.80) nm; ^{1}H NMR ($CDCl_3$, 400 MHz) δ 10.72 (s, H-1), 7.29 (d, H-4), 6.96 (t, H-5), 7.04 (t, H-6), 7.41 (d, H-7), 4.17 (dd, H-9), 3.18 (dd, H-11), 2.93 (dd, H-11), 1.78 (s, H-14); ^{13}C NMR ($CDCl_3$, 100 MHz) δ 132.7 (C-2), 104.3 (C-3), 125.6 (C-3a), 117.6 (C-4), 118.5 (C-5), 121.2 (C-6), 111.7 (C-7), 136.1 (C-7a), 171.3 (C-8), 52.0 (C-9), 23.3 (C-11), 170.4 (C-12), 60.5 (C-13), 24.0 (C-14); EIMS *m*/*z* 273 $[M - H_2O]$ (100), 229 (19), 185 (75); *anal.* C 59.5%, H 5.4%, calcd for $C_{14}H_{16}O_5N_2$, C 57.53%, H 5.52%.

References and Notes

(1) Dawson, R. M. C.; Elliott, D. C.; Elliott, W. H. *Data for Biochemical Research*, 3rd ed.; Clarendon Press: Oxford, U.K., 1986; p 460.

(2) Dawson, R. M. C.; Elliott, D. C.; Elliott, W. H. *Data for Biochemical Research*, 3rd ed.; Clarendon Press: Oxford, U.K., 1986; p 50.

(3) McFall, E.; Newman, E. B. In *Escherichia coli and Salmonella*; Neidhardt, F. C., Ed.; ASM Press: Washington, DC, 1996; Vol. 1, Chapter 22, pp 358−379.

(4) Garbe, T. R.; Kobayashi, M.; Yukawa, H. *Arch. Microbiol.* **2000**, *173*, 78−82.

(5) Smith, M. J.; Mazzola, E. P.; Farrell, T. J.; Sphon, J. A.; Page, S. W.; Ashley, D.; Sirimanne, S. R.; Hill, R. H.; Needham, L. L. *Tetrahedron Lett.* **1991**, *32*, 991−994.

(6) Levy, H. L. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 1811−1813.

(7) Chuang, D. T. *J. Pediatr.* **1998**, *132(3 Pt. 2),* S17−23.

(8) Levy, H. L.; Ghavami, M. *Teratology* **1996**, *53*, 176−184.

(9) McNaught, K. St. P.; Carrupt, P.-A.; Altomare, C.; Cellamare, S.; Carotti, A.; Testa, B.; Jenner, P.; Marsden, C. D. *Biochem. Pharmacol.* **1998**, *56*, 921−933.

(10) Kurusu, Y.; Kainuma, M.; Inui, M.; Satoh, J.; Yukawa, H. *Agric. Biol. Chem.* **1990**, *54*, 443−447.

NP990517S