Tetrahedron Letters, Vol.32, No.8, pp 991-994, 1991
Printed in Great Britain

0040-4039/91 $3.00 + .00
Pergamon Press plc

1,1'-ETHYLIDENEBIS(L-TRYPTOPHAN), STRUCTURE DETERMINATION OF CONTAMINANT "97" - IMPLICATED IN THE *EOSINOPHILIA-MYALGIA* SYNDROME (EMS)

M. J. Smith,*† E. P. Mazzola,† T. J. Farrell,# J. A. Sphon,† S. W. Page,†
D. Ashley,§ S. R. Sirimanne,§ R. H. Hill, Jr.,§ and L. L. Needham§
Natural Products and Instrumentation† & Food Formulation# Branches, Center for Food Safety and Applied Nutrition, Food and Drug Administration, Washington, D.C., 20204; Toxicology Branch, Center for Environmental Health and Injury Control, Centers for Disease Control, Atlanta, GA, 30333§

The condensation of tryptophan and acetaldehyde affords a bisindolylaminal of tryptophan which is indistinguishable from contaminant "97." The formation and decomposition of aminals is addressed in light of epidemiological evidence linking them to a recent outbreak of EMS.

Considerable effort worldwide is focused on identifying the toxic agent(s) linked to dietary supplements of the common amino acid, L-tryptophan (trp). The symptoms of this disease, clinically known as the *eosinophilia-myalgia* syndrome, include abnormally high numbers of eosinophils (a type of white blood cell) and pathogenesis of neural, muscular, and connective tissues. More than 1500 people in the U.S. have been afflicted, with 27 fatalities.[1] Comprehensive HPLC analyses[a,2] have uncovered a trace contaminant with an epidemiologically significant correlation to high-priority cases, namely, peak "97"[a] (or "E"),[2] which constitutes less than *ca.* 0.01% of bulk trp. This contaminant has been tentatively identified as *N,N'*-ethylidenebis(L-tryptophan) (**1**, *vide infra*),[a,b] an unusual aminal formed by the condensation of trp and acetaldehyde. Our application of Sakimoto and co-workers' (private communication) one-step, aqueous acid-catalyzed condensation of trp and acetaldehyde forms a compound which is indistinguishable from contaminant "97," according to our high resolution MS, HPLC / MS, and NMR data. The purpose of this communication is to provide data showing that "97" is aminal **2**,[c] which differs from aminal **1** by the position of the ethylidene linkage, despite the inaccessibility of sufficient quantities of analytically pure contaminant "97."

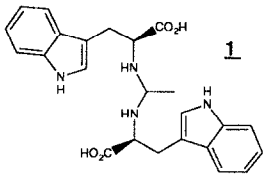
Recent discovery[c] of the ethylidene methine quartet at 7 ppm in the 400-MHz NMR spectrum of synthetic "97" (in the 300-MHz spectrum of contaminant "97" the signal is not resolved from the 5 / 5' or 6 / 6' aromatic protons), raised questions regarding the tentative structure of "97." Selective spin-decoupling of the C_{12} methyl doublet at 2.1 ppm (J=6.5 Hz) collapsed the extraordinarily deshielded C_{11} methine quartet (J=6.5 Hz) of synthetic "97." In regard to the co-identity of synthetic and contaminant "97," as well as their exposure to aqueous acids during preparation, a recent review[3] on aminals notes that most are unstable in aqueous acids and 1° aliphatic amines do not lead to stable (open-chain) varieties. In contrast, indolylaminals such as 1,1'-methylenebis(3-propylindole) are distinguished by their formation in aqueous acids,[4,5] and by highly deshielded aminalic protons (*e.g.*, 6-ppm N-CH_2-N).[4]

To deduce the position of the ethylidene linkage, the *N, N'*-bis(2,4-dinitrophenyl) derivative of synthetic "97" was prepared [**3**; λ_{max} 350 nm (ε $2.01 \cdot 10^4$), 1:1 CH_3CN : H_2O + 0.1% TFA]. By analogy to DNP-

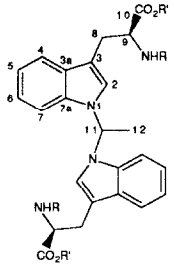
[a] *Morbidity and Mortality Weekly Report* **1990** *39*, 589.
[b] *N. Eng. J. Med.* **1990** *323*, 992.
[c] This structure determination has been reported in the proceedings of the *Tryptophan-EMS conference* (E. Sternberg), American College of Rheumatology, October 29, **1990**, Seattle, WA, and in *Morbidity and Mortality Weekly Report* **1990** *39*, 789. A full article is in preparation.

tryptophan, these values demonstrate that synthetic "97" contains two 1° amines; DNP derivatives of 1° and 2° amines are distinguished by a λ_{max} *ca.* 350 and 380 nm, respectively.[6,7] Further proton, carbon-13, and HETCOR NMR experiments on synthetic "97" have permitted the assignment of critical resonances, *viz.*, those of protons and carbons 2, 2', and 11. Direct evidence that the two tryptophan moieties are linked *via* the 1- and 1'-indole nitrogens was obtained by a long-range carbon-hydrogen chemical shift correlation experiment.[8] Three-bond connectivities were observed between H-11 and carbons 2 and 2' of the tryptophan units and between C-11 and protons 2 and 2' of the same groups. No evidence for the alternative (open-chain) *N, N'*-aminal (**1**) was found; neither coupling between H-11 and C-9/9', nor between C-11 and H-9/9', could be detected. The absence of a plane of symmetry gives rise to the two observed sets of NMR resonances.

***N, N'*-aminal**



1



1,1'-indolylaminals

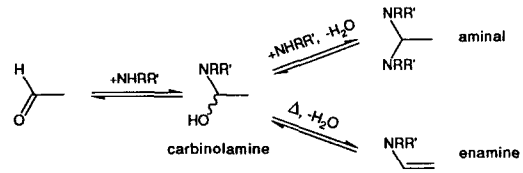
2 R = R' = H
3 R = 2,4-dinitrophenyl, R' = H
4 R = Ac, R' = Me
5 R = Cbz, R' = H

NMR data on Compound 2

Position	C-13 δ	n_H	H-1 δ (J)
2, 2'	124.1, 124.3	1	7.27, 7.34
3, 3'	113.8, 113.9	0	-
3a, 3a'	130.0, 130.1	0	-
4, 4'	120.4	1	7.65, 7.69 d (7.8)
5, 5'	120.7	1	7.01, 7.02 dt (7.8, 1)
6, 6'	123.1, 123.2	1	7.06, 7.08 dt (7.8, 1)
7, 7'	110.7	1	7.30, 7.35 d (7.8)
7a, 7a'	137.4, 137.5	0	-
8, 8'	32.6, 32.7	2	2.83, 2.87 dd (14.2, 10.5) 3.25, 3.28 dd (14.2, 6.5)
9, 9'	57.7, 57.8	1	3.51, 3.52 dd (10.5, 6.5)
10, 10'	182.3, 182.4	0	-
11	63.8	1	6.97 q (6.5)
12	21.0	3	2.13 d (6.5)

Additional confirmation for the fact that contaminant "97" was aminal **2** and not aminal **1** came from the following: (a) condensation of the *N*-acetyl methyl ester of tryptophan with acetaldehyde afforded **4**, which displays the diagnostic 1,1'-ethylidene proton spin system [2.1-ppm methyl doublet (J=6.5 Hz)/6.8-ppm methine quartet (J=6.5 Hz)], with retention of the $NHCOCH_3$ and 2/2'-proton resonances; (b) likewise, condensation of *N*-(carbobenzyloxy)-L-tryptophan with acetaldehyde formed a compound with HPLC/MS features of that for **5**; and (c) an aminal was not formed from the reaction of D-phenylalanine and acetaldehyde, but was formed with 3-methylindole, based on HPLC/MS. Moreover, compounds **2**, **3**, and **4** are missing the intense NH stretching frequency *ca.* 3100-3400 cm^{-1} found in the IR spectra of indoles with an unsubstituted nitrogen,[5] thereby establishing them to be 1,1'-indolylaminals.

Thus, the aqueous acid-catalyzed formation of "97" favors nucleophilic attack by the free indole nitrogen ($pK_{N\text{-}1} << 1$) on the carbonyl carbon; the α-NH_2 group being converted to NH_3^+ is not available for reaction. The general order of reactivity on the indole ring is C-3 > N-1 > C-2, if unsubstituted.[5,9] Indole derivatives *via* N-1 attack on carbonyl groups are not uncommon,[3-5,9-13] *e.g.*, strychnine.[13] Substituent migration from N-1 to C-2 is also known to occur,[4,5,12] which may be relevant to the presence[a,b] of β-carbolines within the same contaminated batches of trp, although the direct insertion of acetaldehyde between the α-NH_2 and C-2 would also effect ring closure. In regard to the reactivity of contaminant "97," the conditions under which indolylaminals are made, namely, diverse (*e.g.*, aqueous acid-catalyzed *vs.* the non-polar, aprotic, anhydrous system for **4**, *vide infra*) and drastic (*i.e.*, strongly acidic), adequately account for its appearance and persistence. In comparison, according to a recent review,[3] the few aminals of 1° amines known are derived from 1° *aromatic* amines and are rather labile, being formed under strictly defined synthetic protocols. While the condensation of **4** occurred at ambient temperatures, at 50°C only starting material was recovered. Aminals, particularly open-chain varieties with an α-hydrogen, readily decompose into an amine and enamine upon heating.[3,14] Aminal-enamine equilibria[3,14] are also common, especially for aminals derived from acetaldehyde and 2° amines.



Hence, prolonged exposure of "97" to warm aqueous acids and oxidizing agents warrants attention, *vis-a-vis* ordinary digestion and catabolism, especially since enamines are reactive alkylating agents and quite capable of forming peroxides.[15] In summary, the compatible data on synthetic and contaminant "97" (*e.g.*, relative stability, high resolution MS, virtually identical proton NMR spectra and HPLC / MS) lead to the conclusion that "97" is 1,1'-ethylidenebis(L-tryptophan) (**2**). Studies to further clarify the chemistry and toxicology of "97" are in progress.

The contributions of many colleagues at the FDA, CDC and NIH are gratefully acknowledged, especially those of J.T. Bernert, M. Trucksess, J. Betz, S. Bailey, and M. Gregg for the isolation of contaminant "97" and other HPLC work; J.-Y. T. Chen and M. Mossoba for IR analyses; J. A. G. Roach and W. J. Driskell for MS data. The assistance of Showa Denko K.K. in providing samples of L-tryptophan and in helpful discussions of their technical data, and B. Horenstein of the Albert Einstein School of Medicine for advice on synthesis, are also gratefully acknowledged.

Proton NMR spectra of contaminant "97" were acquired on a Varian XL-300 spectrometer operating at 300 MHz. All NMR experiments on synthetic "97" were acquired on a Varian VXR-400 spectrometer operating at 399.9 MHz (^1H) or 100.6 MHz (^{13}C). Secondary ion mass spectra (SIMS) were recorded on a VG Autospec Q, using Cs^+ ions and magic bullet (MB; *i.e.*, dithioerythritol and dithiothreitol mix) as matrix. IR spectra were recorded on a Nicolet 740 FT-IR spectrometer. A list of representative data follow.

Compound 2 (Synthetic "97"). 1, 1'-ethylidenebis(L-tryptophan): amorphous white powder; 2 eq. L-trp, 2eq. HCl, 1 eq. CHCHO, 0.02 eq. H_2O, 20 min.; neutralize with NH_4OH, redissolve ppt. in H_2O / NH_4OH, neutralize, wash ppt. with H_2O, followed by EtOH (20-30%). ^{1}H-NMR assignments made in CD_3OD : 40% aq. NaOD (15:1) are tabulated (*vide supra*); those which follow are in D_2O [δ(ppm), *n*H, J(Hz)]: 7.7 (dd, 2H), 7.5 (dd, 2H), 7.4 (two singlets, 2H), 7.2 [t, 2H; t, 2H; q, 1H, (6.6), ethylidene CH], 4.0 (dd, 2H), 3.4 (dd, 2H), 3.3 (m, 2H), 2.2 [d, 3H, (6.6), ethylidene CH_3]. $C_{24}H_{26}N_4O_4$=434; SIMS (MB): 435 (M+H); IR (cm^{-1}; KBr): 3048 (M, B) NH_3^+ str, 1627 (VS, B) CO_2^- asym str & NH_3^+ asym ipd. HPLC conditions: Waters Delta-Pak C_{18}, 300 Å, 3.9 x 150 mm; 0-2 min 100% A (0.1% TFA in H_2O), 2-37 min linear gradient to 80% B (0.1% TFA in 8 : 2 CH_3CN : H_2O), elution time (2)=20 min.

Contaminant "97". ^{1}H-NMR (D_2O): 7.7 (dd, 2H), 7.5 (dd, 2H), 7.4 (two singlets, 2H), 7.2 (m, *ca.* 5H), 4.0 (dd, 2H), 3.4 (dd, 2H), 3.3 (m, 2H), 2.2 [d, 3H, (6.6), ethylidene CH_3]. $C_{24}H_{26}N_4O_4$=434; High resolution SIMS (MB): meas. 435.2020 (M+H), calc. 435.2032. HPLC elution time ("97")=20 min.

Compound 3. 1, 1'-ethylidenebis(*N*-2,4-dinitrophenyl-L-tryptophan): yellow glass; 1 eq. compound 2, 4 eq. 2,4-dinitrofluorobenzene in H_2O : CH_3CN (1:1), excess $NaHCO_3$, 3 h.[16] Extract aqueous phase with ether *ca.* pH 5, acidify to pH 1 with HCl, extract into EtOAc (50%). ^{1}H-NMR (CD_3CN): 8.8 [m, NH; d, 1H (2.7); d, 1H (2.7), DNP H-3 / 3'], 8.1 [dd, 1H (9.5, 2.7), DNP H-5], 8.0 [dd, 1H (9.5, 2.7), DNP H-5'], 7.4 [dd, 1H. (7.7, 1)], 7.3 [dd, 3H (7.7, 1); s, 1H], 7.2 [s, 1H], 7.1 [dt, 2H (7.7, 1), 7.0 [dt, 2H (7.7, 1)], 6.9 [d, 1H (9.5); d, 1H (9.5) DNP H-6 / 6'], 4.8 [m, 2H α-H's], 3.4 [m, 4H, β-H's], 2.0 [d (6.6), ethylidene CH_3]. $C_{36}H_{30}N_8O_{12}$=766; SIMS (MB): 767 (M+H); IR (KBr): 1727 (WMB) doublet C0 str in acid dimer, 1589 (M) NO_2 asym str, 1337 (VS) NO_2 sym str.

Compound 4. 1, 1'-ethylidenebis(*N*-acetyl-L-tryptophan methyl ester): clear oil, 2 eq. *N*-Ac L-trp Me ester, 1 eq. CH_3CHO in $CHCl_3$, 48 h (45%). ^{1}H-NMR ($CDCl_3$): 7.5 [d, 1H (8.2); d, 2H (7.6)], 7.4 [d, 1H (8.2)], 7.1 [m, 4H], 6.9 [s, 1H, H-2], 6.8 [q, 1H (6.5), ethylidene CH], 6.7 [d, 1H, (7.5), NH], 6.5 [s, 1H, H-2'], 6.4 [d, 1H, (7.4), NH], 5.0 [ddd, 1H, α-H], 4.9 [ddd, 1H, α-H], 3.6 [s, 3H, $COCH_3$], 3.4 [dd, 1H, β-H], 3.3 [s, 3H, CO_2CH_3], 3.3 [m, 3H, β-H's], 2.1 [d, 3H (6.5) ethylidene CH_3], 1.9 [s, 3H, $NCOCH_3$], 1.8 [s, 3H, $NCOCH_3$]. $C_{30}H_{34}N_4O_6$=546; SIMS (MB): 547 (M+H); IR (KBr): 3381 (WM, B) bonded NH (trans) str, 3289 (WM, B), bonded NH (cis) str, 1742 (VS) CO str in ester, 1658 (VS) CO str in amide, 1535 (M, B) NH ipd in amide, 1213 (SM, B) C-O-C str in ester.

1. Swygert, L.A. *et al.* **1990** *J. Am. Med. Assoc. 264*, 1698.
2. Belongia, E.A. *et al.* **1990** *N. Engl. J. Med. 323*, 357.
3. Duhamel, L. in *Chemistry of Amino, Nitroso, and Nitro Compounds and Their Derivatives*, (S. Patai, ed.), Wiley Interscience, vol. 2, pp 849-907, **1982**.
4. Wolinsky, J; Sundeen, J.E. **1970** *Tetrahedron 26*, 5427.
5. Remers, W.A. in *Indoles*, (W.J. Houlihan, ed.), Wiley Interscience, vol. 1, pp 1-226, **1972**.
6. Dubin, D.T. **1960** *J. Biol. Chem. 235*, 783; see also, McIntire, F.C; Clements, L.M.; Sproull, M. **1953** *Anal. Chem. 25*, 1757.
7. Rao, K.R.; Sober, H.A. **1954** *J. Amer. Chem. Soc. 76*, 1328.
8. Reynolds, W.F.; McLean, S.; Perpick-Dumont, M.; Enriquez, R.G. **1989** *Magn. Reson. Chem. 27*, 162.
9. Rees, C.W.; Sabet, C.R. **1965** *J. Chem. Soc.* 680.
10. Robinson, R.; Saxton, J.E. **1950** *J. Chem. Soc.* 3136; ibid. **1952**, 976; ibid. **1953**, 2596.
11. Teuber, H.-J., Glosauer, O.; Hochmuth, U. **1964** *Chem. Ber. 97*, 557.
12. Noland, W.E.; Robinson, D.N. **1958** *Tetrahedron 3*, 68.
13. Woodward, R.B.; *et al.* **1954** *J. Amer. Chem. Soc. 76*, 4749.
14. Haynes, L.W. in *Enamines*, (A.G. Cook, ed.), Marcel Dekker, pp 55-100, **1969**.
15. Pitacco, G; Valentin, E. in *Chemistry of Amino, Nitroso, and Nitro Compounds and Their Derivatives*, (S. Patai, ed.), Wiley Interscience, vol. 1, pp 623-714, **1982**.
16. Sanger, F. *Biochem. J.* **1945** *39*, 33.

(Received in USA 23 November 1990)