Pergamon

Bioorganic & Medicinal Chemistry Letters 9 (1999) 637–640

BIOORGANIC & MEDICINAL CHEMISTRY LETTERS

ENZYMATIC SYNTHESES OF 6-(4H-SELENOLO[3,2-*b*]PYRROLYL)-L-ALANINE, 4-(6H-SELENOLO[2,3-*b*]PYRROLYL)-L-ALANINE, AND 6-(4H-FURO[3,2-*b*]PYRROLYL)-L-ALANINE

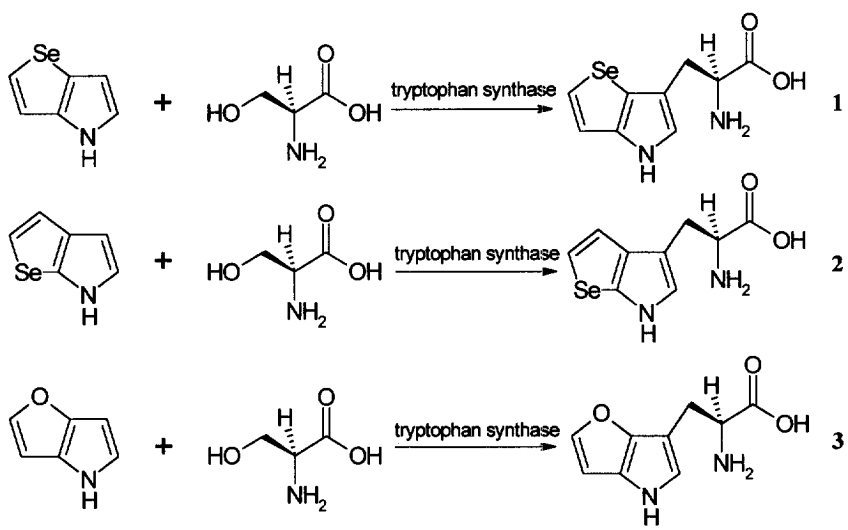
Michael Welch and Robert S. Phillips*

Department of Chemistry, Department of Biochemistry and Molecular Biology, and Center for Metalloenzyme Studies, University of Georgia, Athens, GA 30602-2556, U.S.A.

Received 12 October 1998; accepted 24 December 1998

Abstract: 6-(4H-Selenolo[3,2-*b*]pyrrolyl)-L-alanine **1**, 4-(6H-selenolo[2,3-*b*]pyrrolyl)-L-alanine **2**, and 6-(4H-furo[3,2-*b*]pyrrolyl)-L-alanine **3** have been synthesized via reactions of selenolo[3,2-*b*]pyrrole, selenolo[2,3-*b*]pyrrole, and furo[3,2-*b*]pyrrole, respectively, with L-serine. The reactions are catalyzed by *Salmonella typhimurium* tryptophan synthase.

L-Tryptophan is an essential amino acid required for the biosynthesis of proteins, hormones (serotonin and melatonin), and alkaloids. Thus, novel synthetic analogs of L-tryptophan are of continuing interest. Previously, we have used tryptophan synthase from *Salmonella typhimurium* to synthesize several unnatural analogs of L-tryptophan, namely the aza-,[1] chloro-,[2] and thia-L-tryptophans,[3] the latter of which contain thiophene rings. We reasoned that replacement of the thiophene rings with selenophene rings would afford L-tryptophan analogs that might be useful as heavy atom derivatives in X-ray crystallographic studies of proteins. Selenomethionine has been used extensively for this purpose in recent years,[4] as selenium atoms are easily distinguishable from typical atoms found in proteins (C, H, O, N, S) due to their larger electron densities. In this paper, we describe the first syntheses of 6-(4H-selenolo[3,2-*b*]pyrrolyl)-L-alanine **1** and 4-(6H-selenolo[2,3-*b*]pyrrolyl)-L-alanine **2**. In addition, we report the first synthesis of 6-(4H-furo[3,2-*b*]pyrrolyl)-L-alanine **3**.[5]



0960-894X/99/$ - see front matter
PII: S0960-894X(99)00067-0

Selenolo[3,2-*b*]pyrrole, selenolo[2,3-*b*]pyrrole, and furo[3,2-*b*]pyrrole were prepared by improved versions of the syntheses originally reported by Soth, Farnier, and Paulmier.[6,7] We could not isolate 5-carboxyfuro[2,3-*b*]pyrrole from an attempted saponification of its ethyl ester, confirming the conclusion of Soth and co-workers that furo[2,3-*b*]pyrrole is inaccessible via this series of reactions. The pyrroles were used immediately or stored in the freezer, as they decomposed readily at room temperature to black residues (even at 0 °C, furo[3,2-*b*]pyrrole decomposes within a week, though the selenolopyrroles are noticeably more stable). The pyrroles were incubated in phosphate buffer containing L-serine and tryptophan synthase at 37 °C in the dark and reaction progress monitored by reversed-phase TLC. After filtration of the reaction solution through Celite, the filtrate was concentrated in vacuo, subjected to reversed-phase flash chromatography, and lyophilized to afford **1–3** in 35–50% yield.[8] It is unclear to what extent the low yields are attributable to the marked instability of the starting pyrroles, which assuredly decompose even faster at the reaction temperature than at 0 °C. ^{13}C and ^{1}H NMR spectra of the products are consistent with the assigned structures, **1–3**.[9] Attempts to obtain the M + 1 peaks for these compounds via electrospray ionization impact mass spectrometry (ESI-MS) failed. We wondered if this failure to detect **1–3** was due to their possible contamination with potassium phosphate, the buffer used in the enzyme reaction (see latter part of reference 7). The slightest presence of potassium buffer salts can prevent the detection of analytes when using ESI-MS.[10] To test the viability of this theory, laser desorption mass spectrometry (LD-MS) was tried. While fragmentation of the analyte is uncommon in LD-MS, replacement of its protons by metal cations, such as potassium, is not.[11] In support of our suspicions, LD-MS afforded spectra that contain peaks at *m/z* values matching those expected for **1–3** which have associated with one or more potassium cations.[9,12] The UV spectra of 6-(4H-selenolo[3,2-*b*]pyrrolyl)-L-alanine and 4-(6H-selenolo[2,3-*b*]pyrrolyl)-L-alanine are similar to those exhibited by the thia-L-tryptophans.[9]

The use of 6-(4H-selenolo[3,2-*b*]pyrrolyl)-L-alanine and 4-(6H-selenolo[2,3-*b*]pyrrolyl)-L-alanine as heavy atom derivatives for X-ray crystallography requires that they be bioincorporated into proteins in place of natural L-tryptophan. In preliminary experiments with *E. coli* SVS370, a tryptophan auxotroph [W3110*b*gl R551Δ(TrpEA)$_2$*Tna*A270:*Tn*5], we have found that **1**, at 50 μg/mL, supports cell growth on glucose in a minimal medium. The growth rate, not unexpectedly, is several-fold slower than that with natural L-tryptophan under the same conditions. Nevertheless, this result suggests that 6-(4H-selenolo[3,2-*b*]pyrrolyl)-L-alanine and 4-(6H-selenolo[2,3-*b*]pyrrolyl)-L-alanine can be incorporated into proteins as isomorphous replacements for L-tryptophan.

Acknowledgement: Funding for this research was provided by the National Institute of Health (GM42588). We thank Dr. Edith Wilson Miles for the generous gift of *E. coli* CB149 carrying plasmid pSTB7, which contains the genes which code for *S. typhimurium* tryptophan synthase.

References and Notes

1. Sloan, M.; Phillips, R. S. *Bioorg. Med. Chem. Lett.* **1992**, *2*, 1053.
2. Lee, M.; Phillips, R.S. *Bioorg. Med. Chem. Lett.* **1992**, *2*, 1563.
3. Phillips, R. S.; Cohen, L. A.; Annby, U.; Wensbo, D.; Gronowitz, S. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 1133.
4. For a review, see Besse, D.; Busida, N.; Karnbrock, W.; Minks, C.; Musiol, H. J.; Pegoraro, S.; Siedler, F.; Weyher, E.; Moroder, L. *Biolog. Chem.* **1997**, *378*, 211.
5. To call attention to the similiarity of compounds **1–3** to L-tryptophan, we suggest the following trivial names: **1** - 4,5-selena-L-tryptophan; **2** - 6,7-selena-L-tryptophan; **3** - 4,5-oxa-L-tryptophan. The numerical locants which begin all three of the trivial names describe the numbers of the two carbon atoms in L-tryptophan which are replaced by either selenium, for compounds **1** and **2**, or oxygen, for compound **3**. These trivial names were generated with reference to the general guidelines for replacement nomenclature for heterocyclic compounds.
6. Original syntheses of the selenolo-, thieno-, and furopyrroles: Soth, S.; Farnier, M.; and Paulmier, C. *Can. J. Chem.* **1978**, *56*, 1429.
7. Improved syntheses of the selenolo-, thieno-, and furopyrroles: Welch, M.; Phillips, R. S. (in preparation).
8. In a typical reaction, 100 mg of either selenolo[3,2-*b*]pyrrole, selenolo[2,3-*b*]pyrrole, or furo[3,2-*b*]pyrrole was combined with 1.1 equiv of L-serine, 0.48 mg pyridoxal-5'-phosphate, and 100 μL (2.2 mg) of *S. typhimurium* tryptophan synthase in 50 mL of 0.1 M solution of potassium phosphate buffer, pH 7.8. The flask was stoppered, wrapped in tin foil, and gently shaken for the duration of the reaction time in a water bath at 37 °C. Reversed-phase TLC, using water to develop the plate(s), was used to monitor the reaction progress. The pyrrole (R_f 0.16) and the product L-tryptophan analog (R_f 0.55) cleanly separated from each other and L-serine, which eluted at the solvent front. The pyrrole and the L-tryptophan analog were visualized using UV light. They were also visualized with Ehrlich's reagent (*p*-dimethylaminobenzaldehyde in 10% HCl). L-serine was not detected in the ultraviolet and did not react with Ehrlich's reagent, but was easily detected by staining with a standard ninhydrin solution. After the reaction was complete (i.e., when no pyrrole was apparent by TLC analysis), the reaction mixture was filtered through Celite to remove tryptophan synthase, concentrated, and subjected to reversed-phase flash column chromatography (C18, Analtech, 2 cm × 15 cm), with water as eluent, to separate the product L-tryptophan analog from excess L-serine and potassium phosphate. Note, however, that this much silica gel does not appear to completely separate the product L-tryptophan analog from potassium phosphate, as evidenced by the mass spectra listed below, though it was easily separated from L-serine. Use of a longer column might be judicious. TLC was again used, as above, to identify the chromatographic fractions containing the L-tryptophan analog, which were pooled and lyophilized. The resultant L-tryptophan analog was stored over $CaCl_2$ at 0 °C.
9. **1**: UV(H_2O), λ_{max} 267 nm (log ε = 4.11); 218 (3.92); ^{1}H NMR (D_2O, ppm, 300 MHz) 7.03(d, 1H, *J* = 6.6 Hz), 6.95(s, 1H), 6.91(d, 1H, *J* = 6.6 Hz), 3.77(t, 1H, *J* = 5.2 Hz, α-H), 3.09(m, 2H, β-Hs); ^{13}C NMR (D_2O, 75 MHz, ppm) 30.5, 57.6, 110.7, 119.0, 121.2, 126.4, 132.2, 136.4, 178.0. LD-MS(*m/z*, no matrix): 331, 332, 333, 335, 336, 337(M + 2K – 1H); 367, 369, 370, 372, 374(M + 3K – 2H).

   2: UV(H_2O), λ_{max}248 nm (log ε = 3.79); 218 (4.20); ^{1}H NMR (D_2O, ppm, 300 MHz) 7.49(d, 1H, *J* = 5.7 Hz), 7.29(d, 1H, *J* = 5.7 Hz), 6.99(s, 1H), 3.79(broad s, 1H, α-H), 3.21(dd, 1H, *J* = 14.9 Hz, 5.0 Hz, β-H), 3.10(dd, 1H, *J* = 14.9 Hz, 7.3 Hz, β-H); ^{13}C NMR (D_2O, 75 MHz, ppm) 31.4, 58.0, 108.5, 113.0, 122.2, 124.6, 126.1, 134.0, 179.7. LD-MS(*m/z*, no matrix): 330, 331, 332, 334, 336(M + 2K – 1H); 370, 372, 374(M + 3K – 2H).

3. UV(H_2O), λ_{max}296 (log ε = 3.21); 253 (3.80); ^{1}H NMR (D_2O, ppm, 250 MHz) 6.81(d, 1H, J = 2.2 Hz), 6.35(s, 1H), 6.12(d, 1H, J = 2.2 Hz), 3.66(wide s, 1H, α-H), 3.08(m, 2H, β-H); ^{13}C NMR (D_2O, 62.5 MHz, ppm) 35.8, 57.5, 93.2, 100.2, 122.2, 127.1, 149.7, 156.2, 182.5. LD-MS(*m/z*, no matrix): 271(M + 2K – 1H); 233(M + 1K); unassigned peaks exist @ 255 and 213. Each is less than 40% the intensity of the two assigned peaks.

10. Fairwell, T.; Fales, H. M.; Dutky, R. C.; Sheeley, D. M. *Proceedings of the 42nd ASMS Conference on Mass Spectrometry and Allied Topics, May 29-June 3,1994*, Chicago, Illinois, **1994**, 640.

11. Silverstein, R. M.; Bassler, G. C.; Morrill, T. C. In *Spectrometric Identification of Organic Compounds*; John Wiley & Sons: New York, 1991; Vol. 5, pp 11–12.

12. The *m/z* peaks conform to the expected formula, M + nK – (n-1)H, where n is an integer from 2–3, M is the molecular weight of the L-tryptophan analog analyzed, and H and K stand for the atomic weights of hydrogen and potassium atoms.[9] In addition, the expected isotopic distribution for selenium-containing compounds is seen in the spectra of 6-(4H-selenolo[3,2-*b*]pyrrolyl)-L-alanine and 4-(6H-selenolo[2,3-*b*]pyrrolyl)-L-alanine, as M has multiple values owing to the six possible atomic weights of selenium (74, 76, 77, 78, 80, and 82).[9]