

A CONCISE SYNTHESIS OF AN AHR ENDOGENOUS LIGAND WITH THE INDOLECARBONYLTHIAZOLE SKELETON

Pawel Grzywacz, Rafal R. Sicinski, and Hector F. DeLuca*

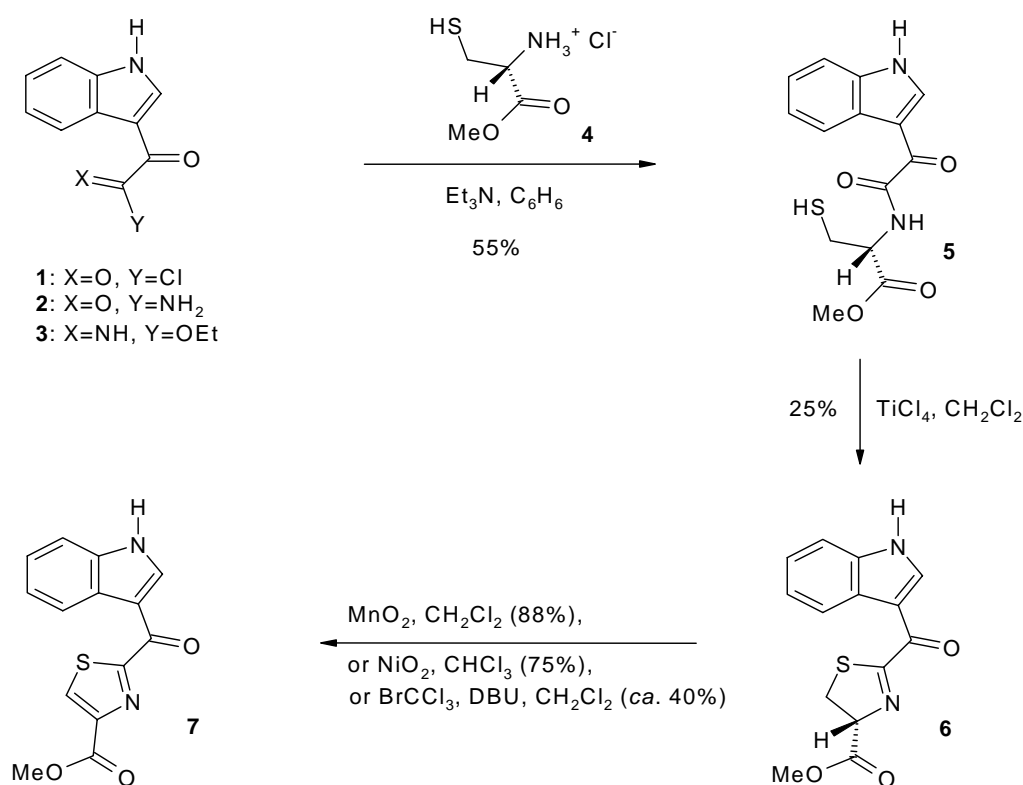
Department of Biochemistry, University of Wisconsin-Madison, 433 Babcock Drive, Madison, WI 53706, USA

Abstract – Short synthesis of an endogenous ligand for the arylhydrocarbon receptor (AHR), 2-(1'*H*-indole-3'-carbonyl)thiazole-4-carboxylic acid methyl ester (**7**), has been described. *N*-Acylation of the L-cysteine methyl ester (**4**) with indoleglyoxylyl chloride (**1**) provided glyoxylamide (**5**) which underwent the TiCl₄-mediated cyclization to thiazoline compound (**6**), followed by dehydrogenation to the final indole thiazole ketone (**7**).

The arylhydrocarbon receptor (AHR)¹ is an important member of the basic helix-loop-helix/Per-Arnt-Sim (bHLH-PAS) family.² It has been established that the AHR is a ligand inducible transcription factor that, upon binding to xenobiotic ligands (3-MC, benzo[α]pyrene or halogenated aromatic hydrocarbons),^{1a,3} regulates the expression of different genes. Thus, receptor activation has been shown to influence cell proliferation,⁴ apoptosis,⁵ and many other physiological functions.⁶ We have recently reported the isolation of an endogenous ligand for the AHR from porcine lung.⁷ The ligand, 2-(1'*H*-indole-3'-carbonyl)thiazole-4-carboxylic acid methyl ester (ITE, **7**, Scheme 1), has been isolated in very small quantities (*ca.* 20 μ g) and identified through extensive spectral studies. Given the biological importance of the ligand, its chemical synthesis was obviously necessary for confirmation of the structural assignment and preparation of larger amounts of compound needed for studies of its physiological activity. Since the molecule of this aromatic ketone consists of two heterocyclic fragments, indole and 4-carbomethoxythiazole attached to the carbonyl group, we sought possible synthetic routes involving intermediate indole glyoxylamides.

First, we examined a direct method of thiazoline ring formation from imino ether derivatives described by North and Pattenden.⁸ A convenient, commercially available starting material, 3-indoleglyoxylyl chloride (**1**), was easily converted into the known amide (**2**) and then alkylated using triethyloxonium hexafluorophosphate. Although the ¹H NMR spectrum of the product seemed to confirm formation of the

desired imino ether (**3**), all attempts of its condensation with L-cysteine methyl ester hydrochloride (**4**) failed to provide the corresponding thiazoline derivative (**6**).



Scheme 1

Thus, we decided to prepare the desired compound (**6**) from glyoxylamide (**5**). The latter compound was easily obtained by acylation of the L-cysteine methyl ester (**4**) with indoleglyoxylyl chloride (**1**) carried out in the benzene solution containing triethylamine.⁹ ¹H NMR spectrum of **5** showed three signals of protons interchangeable with deuterium oxide at δ 8.84 (m, NH), 8.25 (CONH), and 1.50 (CH₂SH), whereas in ¹³C NMR spectrum, the following resonances confirmed a presence of three carbonyl groups: δ 181.85 (COCONH), 171.69 (COOMe) and 165.27 (COCONH). Next, we performed cyclization of the glyoxylamide (**5**) by employing reaction conditions used by Mann *et al.*¹⁰ in their preparation of analogs of curacin A, i.e. treatment with TiCl₄ in dichloromethane. This methodology allowed us to isolate the desired thiazoline ester (**6**) in 25% yield. ¹H NMR spectrum of **6** showed the expected ABX pattern of three protons from thiazoline ring: δ 5.59 (~ t, J = ca. 9.5 Hz, 4-H), 3.67 (dd, J = 11.3, 10.1 Hz, one of 5-H₂), and 3.58 (dd, J = 11.3, 8.9 Hz, one of 5-H₂), whereas in the ¹³C NMR spectrum two carbonyl resonances appeared at δ 179.71 (indole-C=O) and 171.14 (COOMe) as well as a signal of strongly

deshielded thiazoline carbon at δ 173.88 (S-C=N). Despite many attempts, we were not able to increase the yield of **6**. Other acidic conditions (e.g. H₂SO₄ or *p*-TsOH in CHCl₃) were even less efficient. It is very likely that the presence of two neighboring carbonyl groups in the glyoxyamide (**5**) is responsible for the encountered difficulties in its cyclization process. Also (diethylamino)sulfur trifluoride (DAST), a promising reagent used successfully in the synthesis of oxazoline derivatives,¹¹ failed to provide the desired thiazoline compound (**6**).

Finally, three different methods of oxidation of thiazoline (**6**) were examined. Thus, treatment of **6** with MnO₂ or NiO₂ in dichloromethane provided the indolecarbonylthiazole (**7**) in satisfactory yields (88 and 75%, respectively). A mild method of dehydrogenation described by Williams *et al.*,¹² i.e. the use of BrCCl₃ and DBU in dichloromethane, was less efficient (*ca.* 40%) and yielded a mixture of the thiazole (**7**) and a bromine-containing side product whose structure has not been established.

We have found that the HPLC retention time⁷ and spectroscopic properties of the synthesized compound (**7**) are identical in all respects to those of the endogenous AHR ligand isolated in our laboratory from pig lung. Its successful synthesis, therefore, unequivocally confirms the structure and allows for further biological testing aimed at establishing its physiological role in living organisms.

EXPERIMENTAL

Indoleglyoxylyl chloride (**1**) and L-cysteine methyl ester hydrochloride (**4**) were purchased from Aldrich. Melting points (uncorrected) were determined on a Thomas-Hoover capillary melting-point apparatus. Optical rotations were measured in chloroform using a Perkin-Elmer Model 343 polarimeter at 22 °C. Ultraviolet (UV) absorption spectra were recorded with a Perkin-Elmer Lambda 3B UV-VIS spectrophotometer in ethanol. Infrared (IR) spectra were taken with a Nicolet Magna 550 spectrophotometer in KBr pellets. ¹H and ¹³C nuclear magnetic resonance (NMR) spectra were recorded at 500 and 125 MHz with a Bruker Instruments DMX-500 Avance console spectrometer in the solvent noted. Chemical shifts (δ) are reported downfield from internal Me₄Si (δ 0.00). Electrospray ionization (ESI) and electron impact (EI) MS were obtained with a Micromass LCT (Beverly, MA) and a Micromass AutoSpec (Beverly, MA) instruments, respectively. High-performance liquid chromatography (HPLC) was performed on a Waters Associates liquid chromatograph equipped with a Model 6000A solvent delivery system, a Model U6K Universal injector, and a Model 486 tunable absorbance detector.

(2R)-2-[2'-(1''H-Indol-3''-yl)-2'-oxo-acetyloamino]-3-mercaptopropionic acid methyl ester (5)

To a stirred suspension of indoleglyoxylyl chloride (**1**) (2.07 g, 10 mmol) and L-cysteine methyl ester hydrochloride (**4**) (2.57 g, 15 mmol) in a dry benzene (150 mL) was added dropwise triethylamine (4.2

mL, 3.03 g, 30 mmol) at 0 °C. Cooling bath was removed, and the mixture was stirred at rt for 20 h, and then refluxed for 2.5 h. The warm solution was filtered, and a precipitate washed with benzene. Filtrate was washed with saturated NaHCO₃ and water, dried (anhydrous Na₂SO₄), and evaporated. The crystalline residue was purified by flash chromatography. Elution with chloroform/methanol (99:1) gave the product that was crystallized from benzene. Yield of pure compound (**5**) = 1.68 g (55%), mp 145-146 °C; [α]_D²² +193° (c 0.8, CHCl₃); UV (EtOH) λ_{\max} 255.5 nm (ϵ 11,000), 266.0 nm (ϵ 9,600), 273.5 nm (ϵ 8,700), 330.5 nm (ϵ 9,700); IR (KBr) 3358, 3262, 2952, 2941, 2538, 1750, 1744, 1662, 1605, 1500, 1493, 1434, 1352, 1314, 1235, 1227, 1157, 1135, 744 cm⁻¹; ¹H NMR (500 MHz, CDCl₃) δ 9.02 (1H, d, J = 3.3 Hz, 2''-H), 8.84 (1H, m, NH_{indole}), 8.44 (1H, m, 4''-H), 8.25 (1H, br d, J = ca. 8 Hz, CONH), 7.45 (1H, m, 7''-H), 7.35 (2H, br m, 5''- and 6''-H), 4.92 (1H, ~ dt, J = 8.2, 4.7 Hz, CHCO₂Me), 3.83 (3H, s, CO₂Me), 3.06 and 3.12 (each 1H, each ~ ddd, J = 14, 9, 4.7 Hz, CH₂SH), 1.50 (1H, t, J = 9.0 Hz, CH₂SH); ¹³C NMR (125 MHz, CD₃OD) δ 181.85 (s, COCONH), 171.69 (s, COOMe), 165.27 (s, COCONH), 139.76 (d, C-2''), 137.98 (s, C-7a''), 127.86 (s, C-3a''), 124.92, 123.93 and 122.99 (each d; C-4'', -5'' or -6''), 113.96 (s, C-3''), 113.14 (d, C-7''), 55.85 (d, CHCO₂Me), 53.16 (q, CO₂Me), 26.48 (t, CH₂SH); MS (EI) m/z (rel. int.): no M⁺, 273 (M⁺ - SH, 0.5), 256 (2), 233 (2), 185 (2), 183 (2), 153 (19), 144 (indole-C(O)-, 22), 136 (8), 115 (8), 107 (25), 91 (100), 77 (63), 59 (51); MS (ESI) m/z 329.0568 (M⁺ + Na), C₁₄H₁₄N₂O₄NaS requires 329.0572; Anal. Calcd for C₁₄H₁₄N₂O₄S: C, 54.89; H, 4.61; N, 9.15; S, 10.47. Found: C, 54.87; H, 4.65; N, 9.17; S, 10.40. The compound gave a single peak on HPLC (20% 2-propanol in hexane, 10 mm x 25 cm Zorbax-Sil column, 4 mL/min) at R_v 27 mL.

(4R)-2-(1'H-Indole-3'-carbonyl)-4,5-dihydrothiazole-4-carboxylic acid methyl ester (6)

To a stirred solution of indoleglyoxamide (**5**) (2.53 g, 8.3 mmol) in anhydrous methylene chloride (300 mL) was added TiCl₄ (1 M sol. in CH₂Cl₂, 8.4 mL, 8.4 mmol) dropwise at rt. The mixture was then refluxed for 5 h, cooled to room temperature, stirred overnight (16 h) and quenched by an addition of saturated NaHCO₃. The organic layer was washed with water, dried (anhydrous MgSO₄) and evaporated. The residue was purified by flash chromatography. Elution with chloroform/methanol (99:1) gave the product that was crystallized from methanol/benzene. Yield of pure compound (**6**) = 0.6 g (25%), mp 190-191 °C; [α]_D²² +64° (c 0.5, CHCl₃); UV (EtOH) λ_{\max} 261.0 nm (ϵ 8,900), 268.5 nm (ϵ 9,200), 275.5 nm (ϵ 8,800), 333.0 nm (ϵ 8,200); IR (KBr) 3420, 3222, 2956, 1748, 1604, 1597, 1580, 1514, 1488, 1458, 1432, 1314, 1233, 1212, 1187, 1133, 1071, 1055, 805, 776, 752 cm⁻¹; ¹H NMR (500 MHz, CD₃COCD₃) δ 11.27 (1H, m, NH), 8.82 (1H, d, J = 3.2 Hz, 2'-H), 8.34 (1H, m, 4'-H), 7.56 (1H, m, 7'-H), 7.27 (2H, m, 5'- and 6'-H), 5.59 (1H, part X of ABX system, ~ t, J = ca. 9.5 Hz, 4-H), 3.80 (3H, s, CO₂Me), 3.67 (1H, part A of ABX system, dd, J = 11.3, 10.1 Hz, one of 5-H₂), 3.58 (1H, part B of ABX system, dd, J = 11.3, 8.9 Hz, one of 5-H₂); ¹³C NMR (125 MHz, CD₃COCD₃) δ 179.71 (s, C=O), 173.88 (s, C-2), 171.14 (s,

COOMe), 138.73 (d, C-2'), 137.39 (s, C-7a'), 127.22 (s, C-3a'), 124.31, 123.32 and 122.45 (each d, C-4', -5' or -6'), 114.06 (s, C-3'), 112.92 (d, C-7'), 80.49 (d, C-4), 52.65 (q, CO₂Me), 33.62 (t, C-5); MS (EI) m/z (rel. int.) 288 (M⁺, 29), 256 (M⁺ - MeOH, 4), 236 (7), 229 (M⁺ - CO₂Me, 6), 202 (4), 144 (100), 137 (15), 116 (M⁺ - C₆H₆O₃SN, 15), 95 (16), 81 (41), 69 (87); MS (ESI) m/z 311.0454 (M⁺ + Na), C₁₄H₁₂N₂O₃NaS requires 311.0466; Anal. Calcd for C₁₄H₁₂N₂O₃S: C, 58.32; H, 4.20; N, 9.72; S, 11.12. Found: C, 58.34; H, 4.09; N, 9.77; S, 10.82. The compound gave a single peak on HPLC (20% 2-propanol in hexane, 10 mm x 25 cm Zorbax-Sil column, 4 mL/min) at R_V 39 mL.

2-(1'*H*-Indole-3'-carbonyl)thiazole-4-carboxylic acid methyl ester (7)

To a stirred solution of thiazoline ester (6) (38 mg, 0.13 mmol) in anhydrous methylene chloride (30 mL) was added freshly activated manganese(IV) oxide (115 mg, 1.3 mmol). The mixture was stirred for 3 h at rt, filtered through Celite, and evaporated. The residue was purified by flash chromatography. Elution with chloroform/methanol (99:1) gave the product, that was crystallized from methanol. Yield of pure compound (7) = 33 mg (88%), mp 234-235 °C; UV (EtOH) λ_{max} 271.0 nm (ε 10,500), 278.0 nm (ε 11,400), 356.5 nm (ε 11,700); IR (KBr) 3452, 3241, 3125, 2957, 2927, 1737, 1593, 1577, 1507, 1482, 1466, 1432, 1338, 1234, 1206, 1202, 1128, 1113, 1102, 1063, 816, 782, 776, 755, 642 cm⁻¹; ¹H NMR (500 MHz, CD₃COCD₃) δ 11.33 (1H, m, NH), 9.30 (1H, s, 2'-H), 8.70 (1H, s, 5-H), 8.44 (1H, m, 4'-H), 7.61 (1H, m, 7'-H), 7.30 (2H, m, 5'- and 6'-H), 3.94 (3H, s, CO₂Me); ¹H NMR (500 MHz, CD₃OD) δ 9.25 (1H, s, 2'-H), 8.66 (1H, s, 5-H), 8.36 (1H, m, 4'-H), 7.51 (1H, m, 7'-H), 7.28 (2H, m, 5'- and 6'-H), 3.99 (3H, s, CO₂Me); ¹H NMR (500 MHz, DMSO-d₆) δ 9.08 (1H, s, 2'-H), 8.86 (1H, m, 5-H), 8.30 (1H, m, 4'-H), 7.59 (1H, m, 7'-H), 7.28 (2H, m, 5'- and 6'-H), 3.91 (3H, s, CO₂Me); ¹³C NMR (125 MHz, DMSO-d₆) δ 176.45 (s, C=O), 169.86 (s, C-2), 161.01 (s, CO₂Me), 146.89 (s, C-4), 137.98 (d, C-2'), 136.34 (s, C-7a'), 133.90 (d, C-5), 126.35 (s, C-3a'), 123.63, 122.70 and 121.40 (d, C-4', -5' or -6'), 112.72 (d, C-7'), 112.04 (s, C-3'), 52.35 (q, CO₂Me); MS (EI) m/z (relative intensity) 286 (M⁺, 70), 144 (M⁺ - C₅H₄NO₂S, 100), 116 (M⁺ - C₆H₄NO₃S, 17), 89 (14), 69 (18); MS (ESI) m/z 309.0297 (M⁺ + Na), C₁₄H₁₀N₂O₃NaS requires 309.0310; Anal. Calcd for C₁₄H₁₀N₂O₃S: C, 58.73; H, 3.52; N, 9.78; S, 11.20. Found: C, 58.83; H, 3.46; N, 9.94; S, 11.12. The compound gave a single peak on HPLC (10% 2-propanol in hexane, 10 mm x 25 cm Zorbax-Sil column, 4 mL/min) at R_V 42 mL.

ACKNOWLEDGEMENTS

This work was supported in part by a fund from the Wisconsin Alumni Research Foundation. We thank Dr. Mark Anderson for his assistance in recording NMR spectra.

REFERENCES AND NOTES

1. (a) D. W. Nebert and J. E. Gielen, *Fed. Proc.*, 1972, **31**, 1315; (b) A. Poland, E. Glover, and A. S. Kende, *J. Biol. Chem.*, 1976, **251**, 4936; (c) K. M. Burbach, A. Poland, and C. A. Bradfield, *Proc. Nat. Acad. Sci. USA*, 1992, **89**, 8185.
2. S. I. Karchner, W. H. Powell, and M. E. Hahn, *J. Biol. Chem.*, 1999, **274**, 33814.
3. A. Poland and E. Glover, *Mol. Pharmacol.*, 1973, **9**, 736.
4. G. Elizondo, P. Fernandez-Salguero, M. S. Sheikh, G.-Y. Kim, A. J. Fornace, K. S. Lee, and F. J. Gonzales, *Mol. Pharmacol.*, 2000, **57**, 1056.
5. C. Bonnesen, I. M. Eggleston, and J. D. Hayes, *Cancer Res.*, 2001, **61**, 6120.
6. (a) Y. Gu, J. B. Hogenesch, and C. A. Bradfield, *Ann. Rev. Pharmacol. Toxicol.*, 2000, **40**, 519; (b) J. C. Benedict, T. M. Lin, I. K. Loeffler, R. E. Peterson, and J. A. Flaws, *Toxicol. Sci.*, 2000, **56**, 382.
7. J. Song, M. Clagett-Dame, R. E. Peterson, M. E. Hahn, W. M. Westler, R. R. Sicinski, and H. F. DeLuca, *Proc. Nat. Acad. Sci. USA*, 2002, **99**, 14694.
8. M. North and G. Pattenden, *Tetrahedron*, 1990, **46**, 8267.
9. A. Da Settimo, G. Primofiore, F. Da Settimo, A. M. Marini, E. Novellino, G. Greco, C. Martini, G. Giannaccini, and A. Lucacchini, *J. Med. Chem.*, 1996, **39**, 5083.
10. B. K. D. Martin, J. Mann, and O. A. Sageot, *J. Chem. Soc., Perkin Trans. 1*, 1999, 2455.
11. A. J. Phillips, Y. Uto, P. Wipf, M. J. Reno, and D. R. Williams, *Org. Lett.*, 2000, **2**, 1165.
12. D. R. Williams, P. D. Lowder, Y-G. Gu, and D. A. Brooks, *Tetrahedron Lett.*, 1997, **38**, 331.