Published in final edited form as:

Chem Res Toxicol. 2009 December ; 22(12): 1905–1912. doi:10.1021/tx9000418.

Structural Identification of Diindole Agonists of the Aryl Hydrocarbon Receptor Derived from Degradation of Indole-3-pyruvic Acid

Goutam Chowdhury^{‡,§}, Miroslav Dostalek^{‡,§,||}, Erin L. Hsu^{¶,**}, Linh P. Nguyen^{¶,⊥}, Donald F. Stec[†], Christopher A. Bradfield[¶], and F. Peter Guengerich^{‡,*}

[§]Department of Biochemistry and Center in Molecular Toxicology, Vanderbilt University School of Medicine, Nashville, Tennessee 37232-0146

[⊥]Department of Pathology, Laboratory of Immunology and Vascular Biology, School of Medicine, Stanford University, Stanford, CA 94305

[¶]McArdle Laboratory, Department of Oncology, University of Wisconsin School of Medicine and Public Health, Madison, Wisconsin, 53706-1599

[†]Department of Chemistry, Vanderbilt University, Nashville, Tennessee 37240

Abstract

Aerobic incubation of the tryptophan transamination/oxidation product indole-3-pyruvic acid (I3P) at pH 7.4 and 37 °C yielded products with activity as *Ah* receptor (AHR) agonists. The extracts were fractionated using HPLC and screened for AHR agonist activity. Two compounds were identified as agonists: 1,3-di(1*H*-indol-3-yl)propan-2-one (**1**) and 1-(1*H*-indol-3-yl)-3-(3*H*-indol-3-ylidene)propan-2-one (**2**), with the potency of **2** being 100-fold > **1** (Nguyen *et al.* (2009) *Chem. Res. Toxicol.* 22, xxx-xxx, accompanying paper). Both **1** and **2** showed UV spectra indicative of indole. The molecular formulae were established by high-resolution mass spectrometry (HRMS) and the structures were determined by a combination of NMR methods, including ¹H, natural abundance ¹³C, and two-dimensional methods. An intermediate in the oxidation of I3P to **1** is 3-hydroxy-2,4-di(1*H*-indol-3-yl)butanal (HRMS established the presence of a compound with the formula C₂₀H₁₉N₂O₂). Compound **1** was converted to **2** in air or (faster) with mild oxidants, and **2** could be further oxidized to 1,3-di(3*H*-indol-3-ylidene)propan-2-one. Determination of the structures allowed estimation of the molar *Ah* receptor agonist activity of these natural products, similar in potency to known classical AHR inducers.

*Address correspondence to: Prof. F. Peter Guengerich, Department of Biochemistry and Center in Molecular Toxicology, Vanderbilt University School of Medicine, 638 Robinson Research Building, 2200 Pierce Avenue, Nashville, Tennessee 37232-0146, Telephone: (615) 322-2261, Fax: (615) 322-3141, f.guengerich@vanderbilt.edu.

[‡]These two authors contributed equally to the work and should be considered co-first authors

^{||}Current address: Dept. of Clinical Pharmacology, University Hospital and Faculty of Health Studies, Ostrava University, Ostrava, Czech Republic.

^{**}Current address: Northwestern University, 303 E Chicago Ave., Ward 10-258, Chicago, IL 60611.

Supporting Information Available: Figure S1, NMR spectra of compound **1**. Figure S2, NMR spectra of compound **2**. Figure S3, Extracted ion chromatogram of incubation reaction mixture of I3P showing the formation of compounds **1**, **2**, **3**, and **4**. Figure S4, LC-MS/MS spectra of the oxidation product **3** and the putative precursor (**4**) of compound **1**. Figure S5, Extracted ion chromatogram of an incubation of indole-3-acetaldehyde (**5**), showing the presence of compounds **1** and **2**. Figure S6, LC-HRMS spectra of compounds **1** and **2** from an incubation of indole 3-acetaldehyde (**5**) (from Figure S5). This material is available free of charge via the Internet at <http://pubs.acs.org>.

Introduction

The aryl hydrocarbon receptor (AHR) plays a central role not only in the toxicity of dioxins and other xenobiotics (1-3) but also in normal physiology (4-6). A number of endogenous ligands have been proposed (7-11). Included among these are several indoles. In previous work (12) it was demonstrated that air oxidation of the Trp transamination product indole-3-pyruvic acid (I3P) generated what appeared to be powerful AHR agonists. However, the identity of these products remained unknown, and consequently, their potencies could not be established.

In the accompanying paper (13), an HPLC system was developed for the initial separation of the AHR agonists derived from I3P, with screening using a luciferase reporter system. Two of the putative ligands appeared to have MS $[M+H]^+$ ions of m/z 289 and 287. Compounds corresponding to these MS parameters were fractionated using preparative HPLC and their structures were determined to be 1,3-di(1*H*-indol-3-yl)propan-2-one (**1**) and 1-(1*H*-indol-3-yl)-3-(3*H*-indol-3-ylidene)propan-2-one (**2**) using a combination of high resolution (HR) MS, 1- and 2-dimensional NMR, and UV and fluorescence spectroscopy. As we described, the purified compounds showed strong AHR agonist activity in reporter assays (13). A plausible mechanism for the formation of **1** and **2** from I3P is presented, with some support from HRMS studies.

Materials and Methods

Materials

Unless otherwise mentioned, all reagents were of the highest purity commercially available and were obtained from Sigma-Aldrich (St. Louis, MO). HPLC grade solvents were obtained from Fisher Scientific (Pittsburgh, PA).

Preparation of Agonists

A 4 L solution of I3P (0.5 to 2 mM) was prepared by dissolving I3P in 5 mL of CH₃OH and then adjusting the volume with 1X PBS (phosphate-buffered saline; 15 mM potassium phosphate buffer (pH 7.4) containing 150 mM NaCl), as described previously (12). The solution was incubated at 37 °C for 24 h in a 5 L amber glass bottle (to protect the I3P incubation mixture from light) with constant shaking (250 rpm, Innova gyratory incubator, New Brunswick, NJ). The incubation mixture was divided into eight fractions and each fraction was applied to a C₁₈ SepPak cartridge (Sep-Pak[®] Vac 20 mL, 5 g, pre-washed and equilibrated with 5 mL of CH₃OH and then 5 mL of H₂O). After loading of the incubation mixtures, the cartridges were washed with 5-10 mL of H₂O and eluted with 10 mL of CH₃OH. The eluants (which contained significant amount of water) were pooled together and extracted three times into an equal volume (approximately 100 mL) of CH₂Cl₂. The combined organic layers were dried with anhydrous Na₂SO₄ and concentrated to dryness *in vacuo*. The dried samples were stored under Ar at -70 °C until further processing.

HPLC Purification of Agonists

Compounds **1** and **2** were separated from a crude extract of I3P by semi-preparative HPLC using a Hitachi L-7100 HPLC pumping system linked to a Milton Roy Spectro Monitor 3100 UV detector. A dried crude extract of I3P extract was dissolved in a mixture of CH₃OH:H₂O (2:1, v/v) and centrifuged at $1.6 \times 10^4 \times g$ for 5 min. The supernatant was injected onto a ProntoSIL C18-ace-EPS HPLC Column (8.0 mm $\times$ 250 mm, 5.0 μ m) (Bischoff Chromatography, Leonberg, Germany). The mobile phase consisted of Solvent A (0.1 % HCO₂H in H₂O, v/v) and Solvent B (0.1 % HCO₂H in CH₃OH, v/v) utilizing the following gradient program: 0 min, 100% A; 0-5.0 min, linear gradient to 50% B; 5-25 min, linear gradient to 100% B; 25-35 min, hold at 100% B; 35-36 min, linear gradient to 100% A; 36-45 min, hold

at 100% A. The flow rate was 3.0 mL min⁻¹ and UV absorbance was monitored at 280 nm. The retention times were 14.4 and 24.1 min for compounds **2** and **1**, respectively (this system is similar to that used in the preliminary work in the accompanying paper (13), but should not be compared directly). The eluted samples were extracted into CH₂Cl₂ and evaporated to dryness under a stream of N₂. Due to the instability of the compounds, amber glass vials were used for all steps of the experimental work. The dried samples were stored under Ar at -70 °C until further processing.

The partially purified compounds were purified to homogeneity by HPLC on an analytical Bischoff ProntoSIL C18-ace-EPS HPLC column (4.6 mm × 150 mm, 3.0 μm) using a mobile phase consisting of mixtures of H₂O and CH₃CN, following a gradient program with a flow rate of 1.5 mL min⁻¹: 0-5.0 min, hold at 45% CH₃CN; 5-20 min, linear gradient from 45% to 47% CH₃CN(v/v); 20-25 min, linear gradient to 100% CH₃CN(v/v); 25-26 min, linear gradient to 45% CH₃CN(v/v); 26-30 min, hold at 45% CH₃CN(v/v). The UV absorbance was monitored at 280 nm. The retention times were 11.5 and 18.3 min for compounds **2** and **1**, respectively. Compound **2** was purified by HPLC a third time using a narrow bore analytical ProntoSIL C18-ace-EPS HPLC column (2.0 mm × 150 mm, 3.0 μm) column with a flow rate of 0.35 mL min⁻¹. The solvents and gradient used were as above. Dried samples were stored in a glass desiccator under Ar at -70 °C until further processing.

UV-Visible and Fluorescence Spectroscopy

UV-vis spectra of purified samples were recorded in CH₃OH using a modified Cary14-OLIS spectrophotometer (On-Line Instrument Systems, Bogart, GA). Because of the stray light characteristics of this instrument, the absorbance measurements are accurate to >3.0. Fluorescence spectra were collected using a Spec Fluorolog instrument (Horiba Jobin Yvon Inc, Edison, NJ), with excitation at 280 nm.

HRMS, HRMS/MS, LC-MS and LC-MS/MS

HRMS data for Compounds **1** and **2** were obtained using an LTQ FT-Orbitrap mass spectrometer equipped with a nanospray-ESI ion source (Thermo Fisher, San Jose, CA). Analysis was performed in the nanospray-ESI positive ion mode by direct infusion of the sample at a flow rate of 1 μL min⁻¹ in CH₃OH/H₂O solution (1:1, v/v) containing 0.1% formic acid. A silica-packed capillary was used to infuse the sample. Tune parameters were as follows: spray voltage, 1.8 kV; capillary voltage 50 V; capillary temperature, 160 °C; tube lens voltage 120 V. Centroid full MS scans were acquired and averaged for three minutes over a mass range of *m/z* 130 – 500. Centroid MS/MS data were collected in the positive ion mode over a mass range from *m/z* 75-300 using an isolation width of 2 *m/z* (parent masses), normalized collision energy of 30%, activation time of 30 ms, and an activation *Q* of 0.250.

LC-HRMS data for Compounds **3** and **4** were obtained using a Synapt QTOF mass spectrometer equipped with an Acquity UPLC system using an Acquity C18 BEH shield RP18 column (1.7 μm, 1.0 mm × 100 mm) (Waters, Milford, MA). LC conditions were as follows: Solvent A: 10 mM NH₄CH₃CO₂, 2% CH₃CN, 98% H₂O (v/v); Solvent B: 10 mM NH₄CH₃CO₂, 95% CH₃CN, 5% H₂O (v/v). The following gradient program was used with a flow rate of 150 μL min⁻¹: 0-2 min, 75% A; 2-10 min, linear gradient to 100% B; 10-12 min, hold at 100% B; 12-12.5 min, linear gradient to 75% A; 12.5-15 min, hold at 75% A. Temperature of the column was maintained at 50 °C. Samples (10 μL) were infused with an auto-sampler. MS data were collected in the positive ion W-mode, 6V collision energy, 0.2 s scan time and inter-scan delay of 0.02 s.

LC-MS analyses were performed on ThermoFinnigan LCQ or TSQ Quantum (ThermoFisher, Watham, MA) instruments connected to an Agilent or ThermoFinnigan HPLC system.

Analysis was performed in the ESI positive ion mode using a Bischoff ProntoSIL C18-ace-EPS HPLC column (2.0 mm $\times$ 100 mm, 3.0 μ m). All analyses were performed using a mobile phase consisting of Solvent A (0.1 % HCO₂H in H₂O, v/v) and Solvent B (0.1 % HCO₂H in CH₃OH, v/v) following a gradient program with a flow rate of 1.0 mL min⁻¹ (0.3 mL min⁻¹ to source; 0.7 mL min⁻¹ to waste): 0 min 100% A; 0-5.0 min, linear gradient to 50% B; 5-25 min, linear gradient to 100% B; 25-35 min, hold at 100% B; 35-36 min, linear gradient from to 100% A; 36-45 min, hold at 100% A. Samples were injected in a mixture of CH₃OH:H₂O (2:1, v/v). Retention times were 9.3 and 11.4 min for Compounds **2** and **1**, respectively. ESI conditions were as follows: source voltage, 4 kV; source current, 100 lA; auxiliary gas flow rate setting, 20; sweep gas flow rate setting, 5; sheath gas flow setting, 34; capillary voltage -49 V; capillary temperature, 350 °C; tube lens voltage -90 V. MS/MS conditions were as follow: normalized collision energy, 35%; activation Q, 0.250; activation time, 30 ms. Data were acquired in positive ion mode using the Xcalibur software package (ThermoElectron).

NMR

NMR experiments were acquired using a 14.0 T Bruker magnet equipped with a Bruker AV-III console operating at 600.13 MHz. Samples were dissolved in CDCl₃ and all spectra were acquired in 3 mm NMR tubes using a Bruker 5 mm TCI cryogenically cooled NMR probe. Chemical shifts were referenced internally to CHCl₃ (7.27 ppm), which also served as the ²H lock solvent. For 1-dimensional ¹H NMR, typical experimental conditions included 32K data points, 13 ppm sweep width, a recycle delay of 1.5 s and 32-256 scans depending on sample concentration. For 2-dimensional ¹H-¹H COSY, experimental conditions included 2048 $\times$ 512 data matrix, 13 ppm sweep width, recycle delay of 1.5 s, and 4 scans per increment. The data was processed using squared sinebell window function, symmetrized, and displayed in magnitude mode. Multiplicity-edited HSQC experiments were acquired using a 1024 $\times$ 256 data matrix, a $J_{(C-H)}$ value of 145 Hz (which resulted in a multiplicity selection delay of 34 ms), a recycle delay of 1.5 s, and 16 scans per increment along with GARP decoupling on ¹³C during the acquisition time (150 ms). The data were processed using a $\pi/2$ shifted squared sine window function and displayed with CH/CH₃ signals phased positive and CH₂ signals phased negative. $J_{(C-H)}$ filtered HMBC experiments were acquired using a 2048 $\times$ 256 data matrix, a $J_{(C-H)}$ value of 9 Hz for detection of long range couplings resulting in an evolution delay of 55 ms, $J_{(C-H)}$ filter delay of 145 Hz (34 ms) for the suppression of one-bond couplings, a recycle delay of 1.5 s, and 128 scans per increment. The HMBC data were processed using a $\pi/2$ shifted squared sine window function and displayed in magnitude mode.

Oxidation of Compound **1** to Compounds **2** and **3**

A 100 mM solution of **1** in PBS buffer (pH 7.4) was incubated with various oxidizing agents for 20 h in the dark under aerobic conditions at 37° C. The oxidants used were i) 500 mM K₄Fe(CN)₆ ii) 500 mM KMnO₄ iii) 500 mM K₂Cr₂O₇ and iv), 440 mM (0.0015% v/v) H₂O₂. Following incubation, the reaction mixtures were analyzed by LCMS as mentioned in the HPLC Purification of Agonists section (vide supra).

Results and Discussion

Preparation and Purification of Compounds **1** and **2**

Aqueous incubation of I3P has previously been reported to result in the generation of agonists of the AHR (12). HPLC fractionation of the incubation solution of I3P followed by reporter assays revealed that some fractions activated the AHR (13). Initial LC-MS analysis of two such fractions indicted that the major component in those fractions have [M+H]⁺ ions with m/z 289 (**1**) and 287 (**2**). Accordingly, compounds **1** and **2** were prepared by incubating a solution of I3P in PBS buffer at 37 °C for 24 h in the dark with constant shaking under aerobic conditions. Upon incubation, I3P solutions turned dark orange with the formation of a large array of

compounds as judged by HPLC (Figure 1). Consistent with the previous report (12), we found that the yield of compounds **1** and **2** was significantly decreased in the absence of air.

Initial purification and concentration of the incubation mixture were achieved by a solid phase extraction of the crude incubation mixture using octadecylsilane (C-18) columns. The CH₃OH eluent, which contained a significant amount of H₂O, was extracted with CH₂Cl₂ to give a dark red-colored substance. Semi-preparative HPLC of this material using a CH₃OH/H₂O mixture led to the partial purification of **1** and **2**, respectively (Figure 1), which were monitored by LC-MS to identify compounds corresponding to previous HPLC separation work and the [M+H]⁺ ions at *m/z* 289 and 287 (13). Further semi-preparative HPLC using a CH₃CN/H₂O mixture allowed purification of compounds **1** and **2** to homogeneity (Figure 1).

Structural Determination of Compound 1

The elemental composition of compound **1** was determined to be C₁₉H₁₆N₂O based on HRMS (nano-ESI, observed [M+H]⁺ at *m/z* 289.1338; calculated [M+H]⁺ 289.1335). The UV-vis spectrum showed peaks at 275, 282, and 290 nm, clearly indicating the presence of indole, and the fluorescence emission spectrum was also that of an indole (Figure 2). Collision-induced-dissociation (CID) of **1** gave major peaks at *m/z* 172.0754, 158.0598, and 130.0649, consistent with the neutral loss of indole (observed 117.0584; calculated 117.0578), 3-methylindole (observed 131.074; calculated 131.0735), and indole 3-acetaldehyde (observed 159.0684 and calculated 159.0689), respectively (Figure 3A). The ¹H-NMR spectrum revealed the presence of seven peaks including one singlet, two doublets, and two triplets in the region of 7.0-7.5 ppm, consistent with the presence of an indole moiety (Table 1; Figure S2, Supporting Information). The peaks were assigned to the H2, H4, H5, H6, and H7 atoms of the indole ring based on the COSY and HMBC data (Figure 4). The singlet at δ_H 3.9 was indicative of a methylene group and the broad singlet at δ_H 8.06 was assigned to the NH proton of the indole ring. The presence of seven peaks, integrating to eight protons suggested that the molecule is symmetrical. Similarly, the ¹³C NMR spectrum showed the presence of ten peaks; one at δ_C 206.85 that is typical of a carbonyl group, eight in the aromatic region between 108.84 and 136.07 ppm that were assigned to the indole ring, and one at δ_C 38.49 attributed to the methylene carbon (Table 1 and Supporting Information, Figure S2). The fact that there are 19 carbon atoms (HRMS) and only 10 peaks in the ¹³C NMR spectrum suggested that the molecule is symmetric. The ¹³C-¹H HSQC spectrum led to the assignment of each proton to a carbon atom (Supporting Information, Figure S2C). The spectrum also showed that each of the aromatic resonance corresponds to a CH, while the sole aliphatic resonance corresponds to a CH₂ group. Together these results suggested that there are two indole, two methylene, and one carbonyl groups in the molecule. In accord with these observations and the mass spectral data, the structure of compound **1** was concluded to be 1,3-di(1*H*-indol-3-yl)propan-2-one (Scheme 1). A review of the literature revealed that compound **1** has been previously discovered in crude *Malassezia* (yeast) extracts and termed malassezione (14).

The assignment of the structure was subsequently confirmed based on the 2-dimensional NMR data. The ¹H-¹H COSY spectra clearly showed couplings between H4-H5, H5-H6, H6-H7, and H1-H2, consistent with the presence of an indole ring (Figure 4B). A long range coupling between H2-H8 was also observed. Although surprising, a similar observation has been made in case of 3,3'-diindoleacetic acid (15). In the ¹H-¹³C HMBC spectrum, distinct couplings between H8 and C9, C8, C3, C3a, and C2 and between H2 and C8, C3a, and C7a were also observed (Figure 4C). These couplings were indicative of the presence of a propan-2-one group as a linker that connects two indole rings at C3 (Scheme 1). The ¹H-¹³C HMBC spectrum also showed couplings in the aromatic region that further confirmed the presence of an indole.

Structural Determination of Compound 2

The UV-vis spectrum of Compound **2** was found to be similar to that of Compound **1**, with peaks at 275, 282, and 290 nm, which clearly indicate the presence of an indole moiety, with more near-UV absorbance than **1** (Figure 2). The fluorescence emission spectrum also had characteristics of an indole. The elemental composition of compound **2** was calculated to be C₁₉H₁₄ON₂ based on HRMS (nano-ESI, observed [M+H]⁺ at *m/z* 287.1181; calculated [M+H]⁺ 287.1179). CID of **2** gave major peaks at *m/z* 259.1229 and 170.0599 that are consistent with the neutral loss of CO (observed 27.9952; calculated 27.9949) and indole (observed 117.0582; calculated 117.0578), respectively (Figure 3B). Such a loss of CO in a cleavage/recombination process has precedent in the literature (16). CID of **2** also gave a peak at 130.0650, similar to **1**, which can be attributed to a 3-methylindole carbocation or protonated 3-methylene indole (Figure 3B). Unlike **1**, Compound **2** is not symmetrical.

The ¹H-NMR spectrum revealed ten distinct peaks in the aromatic region between 6.6 and 7.6 ppm, indicating the presence of two indole-like moieties (Table 1 and Figure S2, Supporting Information). The peaks in the aromatic region could be assigned to the hydrogen atoms of two indoles based on the COSY and HMBC data (Figure 5). The singlet at δ_H 5.57, which integrated to a proton, indicated the presence of a methine group. The two doublets at 3.8-3.9 ppm, each of which integrated to one proton, were attributed to a methylene group (*vide infra*). In the ¹³C NMR spectrum, a total of 19 peaks were observed, one at δ_C 208 that clearly indicated the presence of a carbonyl group and sixteen in the aromatic region between 108 and 137 ppm that could be assigned to two indole groups (Table 1 and Supporting Information, Figure S2). The peak at 35 ppm is attributable to the methylene and the peak at 72 ppm to the methine group. Each of the protons was assigned to a carbon atom based on the HSQC data (Figure 5C). The HSQC spectrum also revealed that the resonance at 35 ppm corresponds to a CH₂ while that at 72 ppm to a CH group. The ¹H-¹H COSY spectrum showed couplings between the H4-H5, H5-H6, H6-H7, H4'-H5', H5'-H6', and H6'-H7' protons, clearly suggesting the presence of two indole groups (Figure 5B). Together these results indicate there are probably two indoles, one carbonyl, one methylene, and one methine group present in the molecule. The connectivities of the above mentioned groups were established from the ¹H-¹³C HMBC spectrum (Figure 5D). Couplings between the H8 and C9, and the H8' and C9 atoms clearly suggested that the C8 and C8' atoms are linked to the carbonyl (C9) group (Figure 5A and 5D). In addition, coupling between the H8 and C3, C3a, C2 atoms and between the H8' and C3', C3a' atoms indicated that the C8 and C8' atoms are attached to the C3 position of the two indole moieties (Figure 5A and 5D). Based on the above findings and the observation that Compound **1** could be oxidized to **2** (*vide infra*), we propose the structure of compound **2** to be 1-(1*H*-indol-3-yl)-3-(3*H*-indol-3-ylidene) propan-2-one (Scheme 1, 2). To the best of our knowledge **2** is a previously unknown compound.

In the proposed structure, the two methylene C8 protons of Compound **2** were assigned to the two doublets at 3.8 and 3.9 ppm in the proton NMR spectrum (Scheme 1, Table 1, and Supporting Information, Figure S2). Thus, the two geminal C8 protons are coupled to each other with a coupling constant of 17.6 Hz. Although unusual, similar observations were previously made in case of malassesialactic acid (14). The proton signals of the methylene group at C3 and C2 positions of the indole in malassesialactic acid have been reported to have distinct chemical shifts, with a coupling constant in the range of 14.0-16.0 Hz (14).

Oxidation of 1 to Compounds 2 and 3

Based on the proposed structures of **1** and **2**, and the fact that oxygen is a necessary reagent for its formation, (12), we hypothesized that **2** may be an oxidation product of **1**. Accordingly, purified **1** was subjected to oxidation by various oxidizing agents including H₂O₂ (440 μM), KMnO₄ (500 μM), K₄Fe(CN)₆ (500 μM), and K₂Cr₂O₇ (500 μM). The results (Figure 6) clearly

show that **1** can be oxidized to **2** by air and all of these oxidants, with varying degrees of efficiency. Oxidation of **1** also resulted in the formation of another product (**3**) with mass 284 (Figure 6 and Scheme 2). Based on HRMS analysis (ESI, observed $[M+H]^+$ at m/z 285.1042; calculated $[M+H]^+$ 285.1028), the elemental composition of the other product was calculated to be $C_{19}H_{12}ON_2$. Similar to compound **2**, CID of the m/z 285 ion gave a major peak at m/z 257, consistent with the release of CO (Supporting Information, Figure S4). Thus, it seems that compound **3** may be an over-oxidation product of **1** (Scheme 2). In accord with these observations the structure of **3** is proposed to be 1,3-di(3*H*-indol-3-ylidene)propan-2-one (Scheme 2).

Proposed Mechanism of Formation of Compounds 1 and 2

Based on the proposed structures of compounds **1** and **2**, a possible mechanism of formation from I3P has been proposed (Scheme 2). In our proposed mechanism, I3P undergoes decarboxylation to indole-3-acetaldehyde (**5**) followed by an aldol reaction to form compound **4**. Subsequent oxidation and decarboxylation of **4** results in the formation of **1**. Finally, **1** undergoes further oxidation to give **2**. Consistent with our proposed mechanism we found that incubation of indole-3-acetaldehyde (**5**) under similar conditions as used with I3P resulted in the formation of products with high resolution mass spectra which we conclude are the same as seen with compounds **1** and **2** (Supporting Information Figures S5 and S6). Moreover LC-HRMS analysis of the crude incubation mixture also led to the detection of intermediate **4** in incubation mixtures of I3P (ESI, observed $[M+H]^+$ at m/z 319.1422; calculated $[M+H]^+$ 319.1441). CID of intermediate **4** from incubation mixtures of I3P is consistent with our proposed structure (Supporting Information Figure S4).

Conclusions

Oxidation of tryptophan by transamination or the action of an amino oxidase yields I3P, which is unstable and forms a number of AHR ligands (12). Two of these were purified and characterized as novel agonists 1,3-di(1*H*-indol-3-yl)propan-2-one (**1**) and 1-(1*H*-indol-3-yl)-3-(3*H*-indol-3-ylidene)propan-2-one (**2**), using a combination of NMR, HRMS, UV, and fluorescence methods. A plausible mechanism for formation of these involves decarboxylation of the unstable α -keto acid I3P to indole 3-acetaldehyde (**5**), which undergoes an aldol reaction to form **4** (Scheme 2). The β -keto aldehyde **4** undergoes oxidation and decarboxylation to form **1**. Compound **1** in turn undergoes a 2-electron oxidation to **2** and also further oxidation to **3**, the AHR activity of which was not measured due to instability issues.

The exact reason for the inherently greater AHR agonist activity of **2** compared to **1** (~100-fold) (13) is unknown. Compound **2** has less flexibility than **1** and may fit the pharmacophore for the AHR better. Nevertheless, these compounds seem slightly larger and more flexible than the classic AHR model (17,18).

Several searches for **1** and **2** were made in extracts of mouse liver, heart, kidney, and lung using LC-MS of the compounds and also dansyl hydrazones, for which the sensitivity was much greater. To date, the results of the searches have been negative (limit of detection ~ 3 pmol/g tissue), but these results may not preclude an *in vivo* role in that I3P is readily formed in normal physiological reactions (13) and, as shown here and in the accompanying paper, undergoes facile air oxidation and condensation reactions to form **1** and **2**. A variety of indole compounds have been isolated previously and are proposed to be AHR ligands (7-11,19-25) and some of the diindole compounds have chemical similarity to our **1** and **2**, including 3,3'-diindolymethane, a hydrolytic product of plant indoyl methyl glucosinate and long recognized as an AHR agonist (20). However, direct comparison of AHR activities of these compounds have not been made, and exactly which is most likely to be the most important ligand *in vivo*, if any are, requires further study. Although potent AHR agonists are readily formed from

common physiologically relevant molecules, which, if any, of those compounds are relevant *in vivo* requires further study.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported in part by NIH Grants R37 CA090426 and P30 ES000267 (F.P.G) and R37 ES005703, T32 CA009135, P30 CA014520, and T32 ES007015 (C.A.B.) and a fellowship from Merck Research Laboratories (G.C.). We acknowledge the NIH grant S10 RR019022 for funding to purchase the Bruker 600 MHz NMR instrument and accessories. We thank D. Hachey, M. W. Calcutt, and S. Hill for assistance with mass spectrometry and M. Voehler for assistance with NMR spectroscopy. Within the past three years, C.A.B. has served on a scientific advisory board related to the toxic action of halogenated dioxins for Dow Chemical Company.

References

1. Poland A, Palen D, Glover E. Tumour promotion by TCDD in skin of HRS/J hairless mice. *Nature* 1982;300:271–273. [PubMed: 7144882]
2. Nebert DW. The *Ah* locus: genetic differences in toxicity, cancer, mutation, and birth defects. *CRC Crit Rev Toxicol* 1989;20:153–174.
3. Hankinson O. The aryl hydrocarbon receptor complex. *Annu Rev Pharmacol Toxicol* 1995;35:307–340. [PubMed: 7598497]
4. Schmidt JV, Su GH, Reddy JK, Simon MC, Bradfield CA. Characterization of a murine Ahr null allele: involvement of the Ah receptor in hepatic growth and development. *Proc Natl Acad Sci U S A* 1996;93:6731–6736. [PubMed: 8692887]
5. Fernandez-Salguero P, Pineau T, Hilbert DM, McPhail T, Lee SST, Kimura S, Nebert DW, Rudikoff S, Ward JM, Gonzalez FJ. Immune system impairment and hepatic fibrosis in mice lacking the dioxin-binding Ah receptor. *Science* 1995;268:722–726. [PubMed: 7732381]
6. Lahvis GP, Lindell SL, Thomas RS, McCuskey RS, Murphy C, Glover E, Bentz M, Southard J, Bradfield CA. Portosystemic shunting and persistent fetal vascular structures in aryl hydrocarbon receptor-deficient mice. *Proc Natl Acad Sci U S A* 2000;97:10442–10447. [PubMed: 10973493]
7. Adachi J, Mori Y, Matsui S, Takigami H, Fujino J, Kitagawa H, Miller CA III, Kato T, Saeki K, Matsuda T. Indirubin and indigo are potent aryl hydrocarbon receptor ligands present in human urine. *J Biol Chem* 2001;276:31475–31478. [PubMed: 11425848]
8. Denison MS, Nagy SR. Activation of the aryl hydrocarbon receptor by structurally diverse exogenous and endogenous chemicals. *Annu Rev Pharmacol Toxicol* 2003;43:309–334. [PubMed: 12540743]
9. Guengerich FP, Martin MV, McCormick WA, Nguyen LP, Glover E, Bradfield CA. Aryl hydrocarbon receptor response to indigoids *in vitro* and *in vivo*. *Arch Biochem Biophys* 2004;423:309–316. [PubMed: 15001395]
10. Sugihara K, Kitamura S, Yamada T, Okayama T, Ohta S, Yamashita K, Yasuda M, Fujii-Kuriyama Y, Saeki K, Matsui S, Matsuda T. Aryl hydrocarbon receptor-mediated induction of microsomal drug-metabolizing enzyme activity by indirubin and indigo. *Biochem Biophys Res Commun* 2004;318:571–578. [PubMed: 15120638]
11. Williams, SN.; Dunham, E.; Bradfield, CA. Induction of P450 enzymes. In: Ortiz de Montellano, PR., editor. *Cytochrome P450: Structure, Mechanism, and Biochemistry*. Kluwer Academic/Plenum Publishers; New York: 2005. p. 323–346.
12. Bittinger MA, Nguyen LP, Bradfield CA. Aspartate aminotransferase generates proagonists of the aryl hydrocarbon receptor. *Mol Pharmacol* 2003;64:550–556. [PubMed: 12920190]
13. Nguyen LP, Hsu EL, Chowdhury G, Dostalek M, Guengerich FP, Bradfield CA. D-Amino acid oxidase generates agonists of the aryl hydrocarbon receptor from D-tryptophan. *Chem Res Toxicol* 2009;22:xxx–xxx. accompanying article.
14. Irlinger B, Bartsch A, Krämer H-J, Mayser P, Steglich W. New tryptophan metabolites from cultures of the lipophilic yeast *Malassezia furfur*. *Helv Chim Acta* 2005;88:1472–1485.

15. Jones A, Galston AW. Formation of the auxins 3-indolylglycolate and 3, 3'-diindolylacetic acid by reactions between indole and glyoxylate. *J Exp Botany* 1969;20:257–275.
16. Tai Y, Pei S, Wan J, Cao X, Pan Y. Fragmentation study of protonated chalcones by atmospheric pressure chemical ionization and tandem mass spectrometry. *Rapid Commun Mass Spectrom* 2006;20:994–1000. [PubMed: 16479556]
17. Bjeldanes LF, Kim J-Y, Grose KR, Bartholomew JC, Bradfield CA. Aromatic hydrocarbon responsiveness-receptor agonists generated from indole-3-carbinol *in vitro* and *in vivo*: Comparisons with 2,3,7,8-tetrachlorodibenzo-*p*-dioxin. *Proc Natl Acad Sci U S A* 1991;88:9543–9547. [PubMed: 1658785]
18. Henry EC, Kende AS, Rucci G, Totleben MJ, Willey JJ, Dertinger SD, Pollenz RS, Jones JP, Gasiewicz TA. Flavone antagonists bind completely with 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) to the aryl uptake and transformation. *Mol Pharmacol* 1999;55:716–725. [PubMed: 10101030]
19. Nguyen LP, Bradfield CA. The search for endogenous activators for the aryl hydrocarbon receptor. *Chem Res Toxicol* 2008;21:102–116. [PubMed: 18076143]
20. Loub WD, Wattenberg LW, Davis DW. Aryl hydrocarbon hydroxylase induction in rat tissues by naturally occurring indoles of cruciferous plants. *J Natl Cancer Inst* 1975;54:985–988. [PubMed: 1127728]
21. Wille G, Mayser P, Thoma W, Monsees, Baumgart A, Schmitz H-J, Schrenk D, Polborn K, Steglich W. Malassezin—a novel agonist of the arylhydrocarbon receptor from the yeast *Malassezia furfur*. *Bioorg Med Chem* 2001;9:955–960. [PubMed: 11354679]
22. Song J, Clagett-Dame M, Peterson RE, Hahn ME, Westler WM, Sicinski RR, DeLuca HF. A ligand for the aryl hydrocarbon receptor isolated from lung. *Proc Natl Acad Sci U S A* 2002;99:14694–14699. [PubMed: 12409613]
23. Henry EC, Bemis JC, Henry O, Kende AS, Gasiewicz TA. A potential endogenous ligand for the aryl hydrocarbon receptor has potent agonist activity *in vitro* and *in vivo*. *Arch Biochem Biophys* 2006;450:67–77. [PubMed: 16545771]
24. Bergander L, Wincent E, Rannug A, Foroozesh M, Alworth W, Rannug U. Metabolic fate of the Ah receptor ligand 6-formylindolo[3,2-*b*]carbazole. *Chem-Biol Interact* 2004;149:151–164. [PubMed: 15501436]
25. Wincent E, Amini N, Luecke S, Glatt H, Bergman J, Crescenzi C, Rannug A, Rannug U. The suggested physiologic aryl hydrocarbon receptor activator and cytochrome P4501 substrate 6-formylindolo [3,2-*b*]carbazole is present in humans. *J Biol Chem* 2009;284:2690–2696. [PubMed: 19054769]

Abbreviations

CID	collision-induced dissociation
COSY	correlated (NMR) spectroscopy
ESI	electrospray ionization
HMBC	heteronuclear multiple bond correlation
HR	high resolution
HSQC	heteronuclear single quantum coherence
I3P	indole-3-pyruvic acid
AHR	aryl hydrocarbon receptor
PBS	phosphate-buffered saline (15 mM potassium phosphate buffer, pH 7.4, containing 150 mM NaCl)

Figure 1. HPLC separation of Compounds **1** and **2** from an aerobic incubation of I3P. A, HPLC chromatogram of the crude extract in CH₃OH/H₂O; B, HPLC chromatogram of partially purified **1** (t_R 24.1 min peak in Part A) in CH₃CN/H₂O; C, HPLC chromatogram of partially purified **2** (t_R 14.4 min peak in Part A) in CH₃CN/H₂O.

Figure 2. UV-visible and fluorescence spectra of compounds **1** and **2** in CH_3OH . A, UV-vis spectra; B, Fluorescence spectra (excitation at 280 nm). The concentration of **1** was 16.5 μM (ϵ_{282} 148,000 $\text{M}^{-1} \text{cm}^{-1}$) and **2** was 100 μM (ϵ_{282} 36,600 $\text{M}^{-1} \text{cm}^{-1}$).

Figure 3.
HR MS/MS spectra of compounds **1** and **2**. A, **1**; B, **2**.

Figure 4. NMR spectra of compound **1**. A, Observed ^1H - ^1H COSY and HMBC couplings; B, ^1H - ^1H COSY spectrum; C, HMBC spectrum.

Figure 5. NMR spectra of compound 2. A, Observed ^1H - ^1H COSY and HMBC coupling; B, ^1H - ^1H COSY spectrum; C, HSQC spectrum; D, HMBC spectrum.

Figure 6.

Conversion of compound 1 to 2 and 3 by various oxidizing agents. The incubations were done aerobically at 37 °C for 20 h with either no treatment or in the presence of 500 μ M of KMnO₄, K₄Fe(CN)₆, or K₂Cr₂O₇ or 440 μ M H₂O₂. Fold-increase was calculated by comparison with the yields of 2 (□) and 3 (■) under aerobic condition without other oxidants.

Scheme 1.
Assigned Structures **1** and **2** for AHR Agonists

Scheme 2. Proposed Mechanism of Conversion of I3P to **1**, **2**, and **3**

Table 1¹H and ¹³C chemical shifts of compounds **1** and **2**

C	1		2	
	¹ H	¹³ C	¹ H	¹³ C
2	6.99 s	123	6.97 s	123
3		108		108
3a		127		127
4	7.46 d (7.9)	118.8	7.29 d (8.0)	118.8
5	7.11 t (7.6)	119.8	7.06 t (7.3)	119.8
6	7.2 t (7.6)	122.3	7.18 t (7.6)	122.3
7	7.37 d (8.1)	111.1	7.34 d (8.3)	111.2
7a		136		136
8	3.9 s	38.5	3.8 d (16.6) 3.9 d (16.6)	35
9		206.8		208
2'			7.23 d (2.5)	124
3'				113
3a'				126
4'			7.56 d (7.9)	119.2
5'			7.15 t (7.4)	120.5
6'			7.25 t (8.3)	122.8
7'			7.41 d (8.2)	111.5
7a'				137
8'			5.57 s	72