Structure elucidation of two tryptophan-derived, high affinity Ah receptor ligands

U Rannug[1], A Rannug[2], U Sjöberg[1], H Li[3], R Westerholm[3] and J Bergman[4]*

[1]Department of Genetic and Cellular Toxicology, Wallenberg Laboratory, Stockholm University, S-106 91 Stockholm, Sweden, [2]Department of Toxicology, National Institute of Work and Health, S-171 84 Solna, Sweden, [3]Department of Analytical Chemistry, Arrhenius Laboratory, Stockholm University, S-106 91 Stockholm, Sweden and [4]Department of Organic Chemistry, Novum-CNT, S-141 57 Huddinge, Sweden

Background: Environmental contaminants, such as 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) and other structurally related 'environmental hormones', exert their harmful biological effects through the Ah receptor signaling pathway. Several naturally occurring substances also bind to this receptor, but its natural role is still obscure. Tryptophan derivatives of the indolo[3,2-*b*]carbazole type, earlier suggested by us to be endogenous ligands for the receptor, should be a powerful tool in understanding receptor function. We therefore set out to determine their identity.

Results: The two tryptophan-derived Ah receptor ligands have been chemically analyzed and characterized by means of mass spectrometry, ^{1}H NMR and ^{13}C NMR. UV, infra-red and fluorescence spectra were also recorded. All data are in accordance with the two compounds being closely related indolo[3,2-*b*]carbazole derivatives. Evidence is presented that compound A (MW = 312) is the symmetrical 6,12-diformylindolo-[3,2-*b*]carbazole, and compound B (MW = 284) is the monosubstituted 6-formylindolo[3,2-*b*]carbazole.

Conclusions: The elucidation of the structures of the two high affinity Ah receptor ligands 6,12-diformylindolo[3,2-*b*]carbazole and 6-formylindolo[3,2-*b*]carbazole provides the necessary basis for further mechanistic studies of this important group of compounds, and will help in determining the natural role of the Ah receptor.

Chemistry & Biology December 1995, **2**:841–845

Key words: Ah receptor, photoproducts, formylindolo[3,2-*b*]carbazoles, tryptophan

Introduction

2,3,7,8-Tetrachlorodibenzo-*p*-dioxin (TCDD) is the prototype for a group of hormonally active xenobiotica that includes coplanar polychlorinated dibenzo-*p*-dioxins and furans with tumorigenic, immunologic, reproductive, neurobehavioral, and acute toxic effects at low doses [1]. Many of the observed effects are consistent with the concept that these compounds act as 'environmental hormones'. From studies, primarily performed in mice (reviewed in [2,3]) it is known that the toxic syndrome in experimental animals proceeds through binding of the inducer compound to the aryl hydrocarbon (Ah) receptor. The ligand-bound receptor activates transcription of several genes that encode proteins involved in the metabolism of xenobiotica and in cell growth and differentiation. Cytochrome P-4501A1 (CYP1A1) is the most extensively studied enzyme induced by TCDD-like substances. It is also induced by naturally occurring substances like flavones [4] and by indole derivatives such as acid condensation products of indole-3-carbinol [5,6] and photooxidized derivatives of tryptophan [7–9]. Additionally, studies in rats and mice have indicated that UV light induces CYP1A1 in the skin [10] and the liver [11].

Of the UV-absorbing amino acids, cysteine, tyrosine and tryptophan, tryptophan absorbs most strongly at near-UV wavelengths [12]. Direct photolysis and dye-sensitized photooxidation generate numerous photoproducts of tryptophan, and some of these have been identified [12,13]. The majority of these products, including the larger molecules, are yet to be identified, however. From our experiments we know that several photoproducts derived from tryptophan are biologically active. Some of these photoproducts are mutagenic, but these are structurally different from the major mutagenic compounds formed after pyrolysis of tryptophan [14]. Products with affinity for the Ah receptor are also formed, two of which are of particular interest. As reported previously [15], these tryptophan-derived compounds (purity >99%) possess extremely high binding affinity for the Ah receptor, with K_d values of 0.07 nM and 0.44 nM, respectively, compared to a K_d of 0.48 nM for TCCD itself under the same conditions. Later experiments have confirmed that the binding affinity of the first compound is 5–8 times higher than that of TCCD and the latter compound has a similar affinity for the receptor as TCCD. Since these compounds showed such high binding affinities — one of them having the highest binding affinity for the Ah receptor reported so far for any compound — and were produced by UV irradiation of the amino acid tryptophan, we suggested [15] that they might be the endogenous ligands for the Ah receptor. The compounds act as both inducers and inhibitors of cytochrome P-4501A1, and show antimutagenic properties [7,16]. It is of importance to identify the chemical structure of the two compounds, which have molecular weights of 284 and 312, respectively. As we suggested previously [15], oxidized amino acids may be responsible for the UV induction of cytochrome P-450. The seasonal

*Corresponding author.

 ISSN 1074-5521

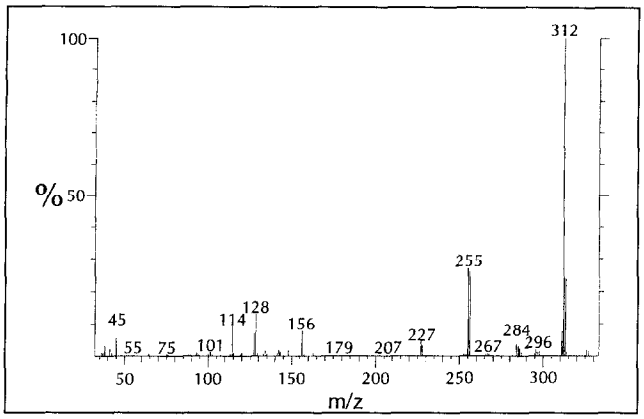

Fig. 1. Mass spectrum (70 eV) of compound A. The molecular ion is at 312.

variation of CYP1A1 activity in human lymphocytes [17] may be UV-light related and explicable in an analogous way. The purpose of the present study was to characterize these two photoproducts and to elucidate their structures.

Results and discussion

UV irradiation of aqueous solutions of L-tryptophan gives rise to a large number of photoproducts. At least three of these components show a high Ah (or TCDD) receptor affinity. Two of these compounds, A and B, have been studied in more detail and have shown interesting biological effects. The mass spectra of the two compounds are shown in Figures 1 and 2. Low voltage (11 eV) mass spectra for the two compounds gave molecular weights of 312 and 284. The 70 eV spectra also showed a strong molecular ion at 312 for compound A and mass fragments of m/z 284, 285, 255, 256 and 128. The corresponding molecular ion for B was at 284, with fragments of m/z 255, 256 and 128.

The UV spectra of the two compounds in chloroform are shown in Figure 3. Compound A showed a double peak around 300 nm and two peaks at 450 and 540 nm. The same absorption peaks were also seen with DMSO as solvent. Compound B showed three peaks at 310, 390 and

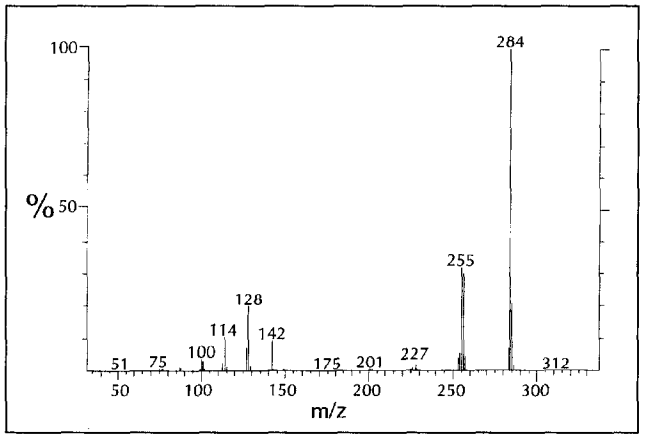

Fig. 2. Mass spectrum (70 eV) of compound B. The molecular ion is at 284.

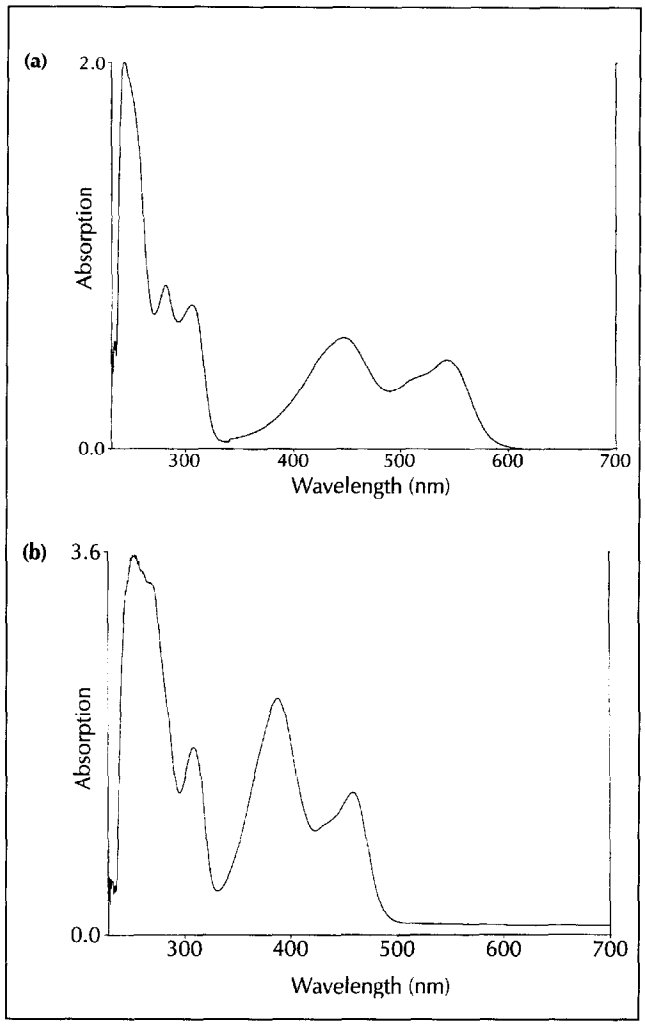

Fig. 3. UV absorption spectra. The UV absorption spectra for **(a)** compound A and **(b)** compound B, both in chloroform, are shown.

460 nm, and similar absorption maxima were observed in methanol and ethanol. Compound B also showed strong fluorescence. The fluorescence spectrum in ethanol is shown in Figure 4. The maximum fluorescence in chloroform was recorded at an emission wavelength of 500 nm and excitation wave lengths of 310, 390 and 460 nm (identical to the peaks in the UV-absorption spectrum).

The solubilities of both compound A (MW = 312, $C_{20}H_{12}N_2O_2$) and compound B (MW = 284, $C_{19}H_{12}N_2O$) are very low. In spite of this it was possible to get quite good NMR spectra of both compounds using DMSO as a solvent (and TMS as standard) provided that the number of scans was high enough (>30.000). The ^{1}H-NMR spectrum of compound A (Fig. 5a) shows two doublets at 8.59 (2H) and 7.86 (2H) ppm and two double doublets at 7.53 (2H) and 7.29 (2H) ppm in the aromatic region. This indicates that we are dealing with a symmetrical structure. The two additional signals at 11.47 (2H) and 12.05 (2H, broad) ppm are consistent with this assumption. Consideration of these data plus information from the

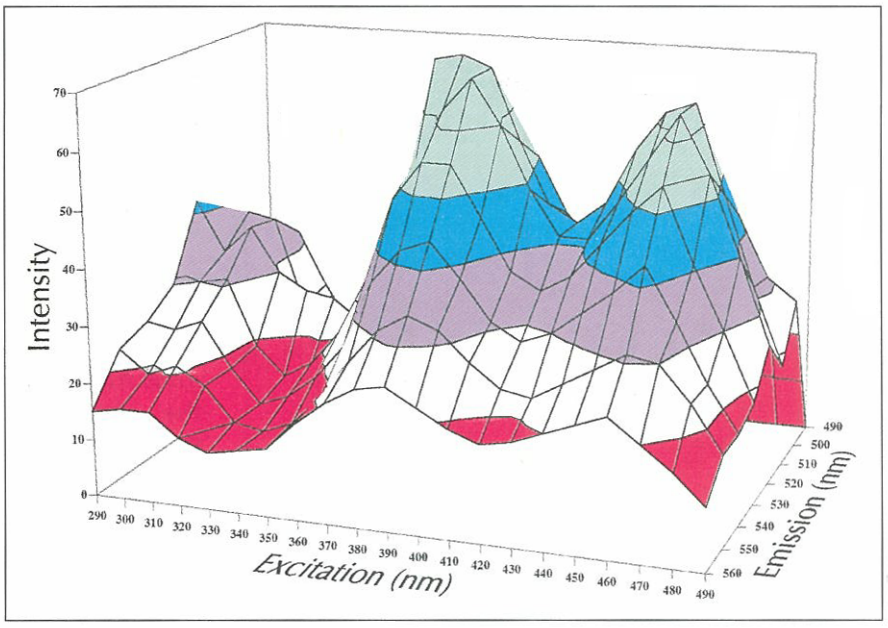

Fig. 4. Fluorescence spectrum of compound B in ethanol.

infrared spectrum (C = 0 at 1655 cm^{-1}) lead us to propose structure **1** (6,12-diformylindolo[3,2-*b*] carbazole; Fig. 6) for compound A.

This structure agrees with the fact that the ^{13}C NMR spectrum exhibited five quaternary signals between 117 and 145 ppm, plus a signal from the formyl groups at 192 ppm and four signals from the carbon atoms bearing hydrogen atoms. The mass spectrum, which showed consecutive losses of m/z 28 and 29, is also consistent with the assigned structure.

The ^{1}H-NMR spectrum of compound B (Fig. 5b) immediately reveals that the compound is asymmetric and closely related to compound A. Structure **2** (6-formylindolo[3,2-*b*]carbazole; Fig. 6) was therefore suggested and corroborated by further analysis of the ^{13}C spectrum and the mass spectrum. All aromatic protons are different in the monoformyl compound and the protons close to the fused pyrrole rings appear as four well-separated 1H doublets. The four remaining protons on the flanking benzene rings appear as four pair-wise, coinciding double doublets. The sharp 1H singlet at 8.63 ppm is due to the isolated aromatic hydrogen atom in the middle ring. The two broad 1H singlets at 10.75 and 11.25 ppm are due to the two different NH groups, and the sharp 1H singlet at 11.70 ppm emanates from the formyl group. Apart from the carbonyl group (at 186.8 ppm) the ^{13}C-NMR spectrum of compound **2** showed, between 145 and 112 ppm, nine signals for both the methine and the quaternary carbon atoms, which partly coincided.

Both compounds are simple derivatives of indolo-[3,2-*b*]carbazole, a nitrogen heterocycle that has been known for a long time. In particular, it has been demonstrated that indole and formaldehyde will, under acidic conditions and in the presence of air, light and a sensitizer, be converted to indolo[3,2-*b*]carbazole [18] through a series of transformations. The two reactants (indole and formaldehyde) can be replaced with 3,3'-diindolylmethane.

Interestingly, the parent compound for compounds **1** and **2**, indolo[3,2-*b*]carbazole, also interacts strongly with the Ah receptor, whereas the corresponding 6,12-dihydroxy indolo[3,2-*b*]carbazole or indolo[3,2-*b*]carbazole-6,12-quinone show much lower affinity, with IC_{50} values >1500 and >150 nM, respectively [19,20].

The formation of indole-3-acetaldehyde (compound **3**) as an intermediate upon irradiation of tryptophan has been reported [21]. The formation of compounds A and B after heat treatment of indole-3-acetaldehyde solutions

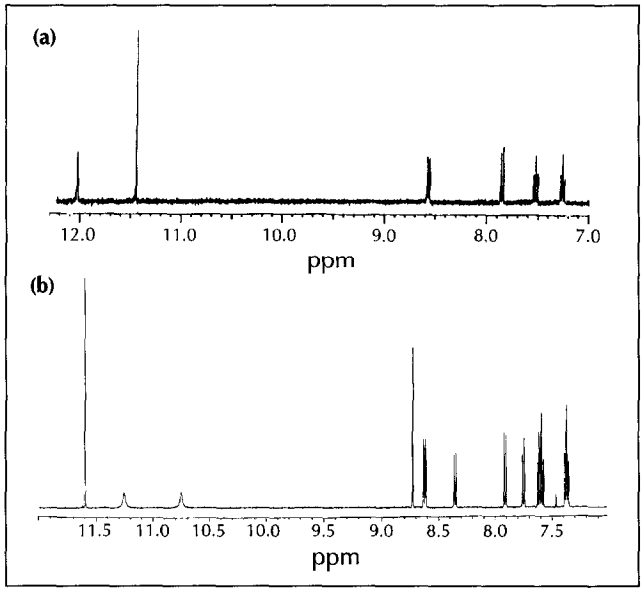

Fig. 5. ^{1}H-NMR spectra. The ^{1}H-NMR spectra for **(a)** compound A (MW = 312) and **(b)** compound B (MW = 284) are shown.

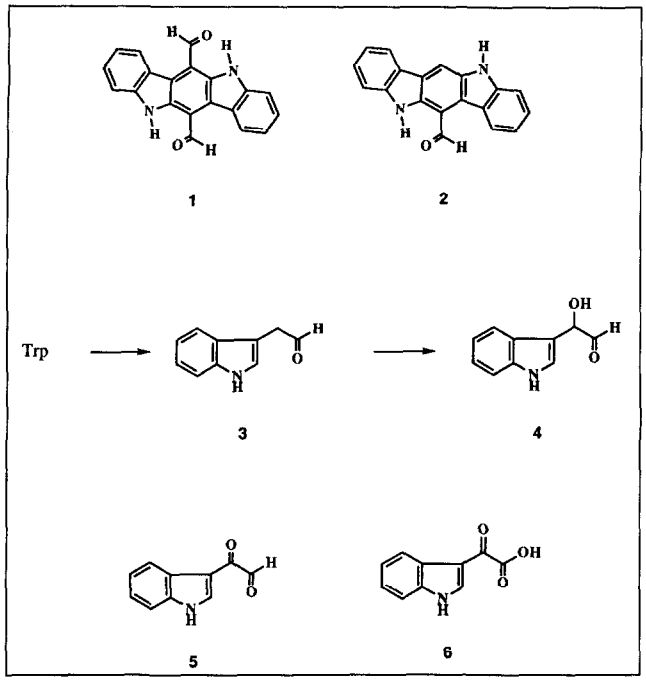

Fig. 6. Structures of the compounds discussed in the text. Structures **1** and **2** are the proposed structures for compounds A and B, respectively.

support this as the precursor of these compounds in the experiments in which tryptophan is irradiated. Oxidation of compound **3** would lead to the α-hydroxy-aldehyde (compound **4**), which should readily undergo acid-catalyzed condensation to yield compound **2**.

Compound **4** has never been described, but has been invoked as an intermediate in a metabolic pathway of tryptophan. This pathway could be initiated by tryptophan side-chain oxidase, leading to indole-3-glyoxal (compound **5**) and consecutive products such as indole-3-glyoxylic acid (compound **6**) [22].

Compounds **1** and **2** are currently available in small amounts only by the photolysis of tryptophan followed by extensive chromatography. Therefore synthetic approaches, involving modification of related indolo[2,3-*b*]carbazoles as well as ring syntheses, are now being studied in our laboratories. This work will also provide rigorous proof that compounds A and B are identical to the assigned structures.

Significance

The importance of the signal pathways involving the Ah receptor in mammals, including man, can hardly be overestimated. Recent data obtained from studies in mice lacking the Ah receptor appear to favor a physiological role for the Ah receptor in addition to its function in xenobiotic metabolism [23]. Ah receptor deficient mice have a low survival rate and show impaired development of the liver and the immune system [23]. Agonists for the Ah receptor show considerable toxicity both in animals and man, as exemplified by TCDD and polycyclic aromatic hydrocarbons (PAH), which are amongst the most toxic and carcinogenic compounds known. There is also a growing concern regarding immunologic and reproductive effects on humans from perinatal low dose exposures to Ah receptor agonists.

Photoproducts of tryptophan have been studied by us with regard to Ah receptor binding and enzyme induction and were proposed to be the endogenous Ah receptor ligands [15]. Compound B, with the suggested structure 6-formylindolo[3,2-*b*]-carbazole, has been tested for Ah receptor binding and was found to be an extremely good ligand for the Ah receptor, showing affinity at picomolar concentrations ($K_d = 7 \times 10^{-11}$ M). This affinity is almost ten times higher than that of the prototype ligand TCDD. The related compound A, 6,12-diformylindolo[3,2-*b*]carbazole, also showed a high binding affinity, in our experiments approximately the same as TCDD. It remains to be shown that these compounds are present in biologically relevant extracts. The fact that indolo[3,2-*b*]carbazole derivatives can be formed from the dietary constituent, indole-3-carbinol, further emphasizes the importance of characterizing this group of compounds in more detail. Additional biological studies using these tryptophan-derived compounds may give valuable information regarding the functional role of the Ah receptor in development and cell differentiation. The toxicity exerted by TCDD and polycyclic aromatic hydrocarbons may be explained by their great structural and physical resemblance to these products of the UV irradiation of tryptophan.

Materials and methods

UV irradiation of tryptophan solutions

Water solutions of tryptophan (1 $g l^{-1}$) were irradiated using a high pressure Hg lamp (Black-Ray long-wave UV lamp without filter) for 24 h essentially as described [15].

Sample preparation

The irradiated aqueous solutions, now brown-reddish in color, were extracted either with chloroform (1:9, chloroform:water) or with ethyl acetate. Alternatively, the photoproducts were concentrated on RP-18 columns. After loading the columns, the material of interest was eluted with ethyl acetate. Since this procedure was rather time-consuming, most of the concentration experiments were done by extraction of the water phase. The solvent was then evaporated under reduced pressure.

Further chemical separations and analyses

The dry eluate was dissolved in chloroform and loaded on a gravity-fed glass column packed with silica gel (Merck 60; 0.040–0.063 mm particle diameter). The mobile phase was chloroform. For purification this system was better than the RP-18 system used previously [15] since the nonpolar

compounds of interest eluted first. The known photoproduct A eluted first, followed by compound B and a third, unknown compound with a molecular weight of 248. As compound A is not soluble in methanol, the final purification included methanol washes. Compound B, which is relatively soluble in methanol, was instead washed with cyclohexane. The products thus obtained have a purity of >99% as indicated by the absence of foreign signals in the NMR spectra.

The mass spectrometry was performed using a Jeol D300 mass spectrometer (Jeol Inc., Japan). The mass spectrometer was operated in the electron impact ionization mode using the probe inlet. The probe was heated to a final temperature of 350 °C. The ion source temperature was set to 200 °C, ionization voltage was set to 70 eV, scan range to 35–350 atomic mass units and scan time to 1.0 s. The molecular weights 284 and 312 were obtained using low voltage ionization (11 eV). The mass spectrometer was interfaced to the Finnigan computer system INCOS 2000 (Finnigan Inc., USA).

UV-spectra were recorded on a Hitachi U-320 spectrophotometer, and the fluorescence spectra on a Shimadzu fluorescence spectromonitor RF-510LC. ^{1}H-NMR spectra were recorded on a Varian A60 instrument and ^{13}C-NMR spectra were recorded on a Bruker WP400 instrument.

Acknowledgements: This work was supported by the Swedish Natural Science Research Council.

References

1. Lindström, G., Hooper, K., Petreas, M., Stephens, R. & Gilman, A. (1995). Workshop on perinatal exposure to dioxin-like compounds. I. Summary. *Environ. Health Perspect.* **103**, 135–142.
2. Okey, A.B. (1990). Enzyme induction in the cytochrome P-450 system. *Pharmacol. Ther.* **45**, 241–298.
3. Okey, A.B., Riddick, D.S. & Harper, P.A. (1994). Molecular biology of the aromatic hydrocarbon (dioxin) receptor. *Trends Pharmacol. Sci.* **15**, 226–232.
4. Lu, Y-F., Santostefano, M., Cunningham, B.D.M., Threadgill, M.D. & Safe, S. (1995). Identification of 3'-methoxy-4'-nitroflavone as a pure aryl hydrocarbon (Ah) receptor antagonist and evidence for more than one form of the nuclear Ah receptor in MCF-7 human breast cancer cells. *Arch. Biochem. Biophys.* **316**, 470–477.
5. Bradfield, C.A. & Bjeldanes, L.F. (1984). Effect of dietary indole-3-carbinol on intestinal and hepatic monooxygenase, glutathione S-transferase and epoxide hydrolase activities in the rat. *J. Toxicol. Environ. Health* **21**, 311–323.
6. Bjeldanes L.F., Kim, J-Y., Grose, K.R., Bartholomew, J.C. & Bradfield, C.A. (1991). 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. USA* **88**, 9543–9547.
7. Rannug, A. &. Rannug, U. (1989). UV-irradiation of tryptophan gives rise to AHH-inducing compounds with very high affinity for the Ah receptor. *Chemosphere* **18**, 1101–1106.
8. Helferich, W.G. & Denison, M.C. (1991). Ultraviolet photoproducts of tryptophan can act as dioxin agonists. *Mol. Pharmacol.* **40**, 674–678.
9. Kocarek, T.A., Schuetz, E.G. & Guzelian, P.S. (1993). Transient induction of cytochrome P450 1A1 mRNA by culture medium component in primary cultures of adult rat hepatocytes. *In Vitro Cell Dev. Biol.* **29A**, 62–66.
10. Mukhtar, H., DelTito Jr, B.J., Matgouranis, P.M., Das, M., Asokan, P. & Bickers, D.R. (1986). Additive effects of ultraviolet B and crude tar on cutaneous carcinogen metabolism: possible relevance to the tumorigenicity of the Goeckerman regimen. *J. Invest. Dermatol.* **87**, 348–353.
11. Goertz, G., Merk, H., Bolsen, K., Tsambaos, D. & Berger, H. (1983). Influence of chronic UV-light exposure on hepatic and cutaneous monooxygenases. *Experientia* **39**, 385–386.
12. Creed, D. (1984). The photophysics and photochemistry of the near-UV absorbing amino acids — I. Tryptophan and its simple derivatives. *Photochem. Photobiol.* **39**, 537–562.
13. Borkman, R.F., Hibbard, L.B. & Dillon, J. (1986). The photolysis of tryptophan with 337.1 nm laser radiation. *Photochem. Photobiol.* **43**, 13–19.
14. Sjögren, M., Sjöberg, U. & Rannug, U. (1994). Mutagenicity of tryptophan photoproducts in the Ames *Salmonella* assay. *Mutat. Res.* **321**, 229–239.
15. Rannug, A., *et al.*, & Grafström, A-K. (1987). Certain photooxidized derivatives of tryptophan bind with very high affinity to the Ah receptor and are likely to be endogenous signal substances. *J. Biol. Chem.* **262**, 15422–15427.
16. Rannug, U., Agurell, E., Rannug, A. & Cederberg, H. (1992). Certain tryptophan photoproducts are inhibitors of cytochrome P450-dependent mutagenicity. *Environ. Mol. Mutag.* **20**, 289–296.
17. Paigen, B., *et al.*, & Sartori, P. (1981). Seasonal variation of aryl hydrocarbon hydroxylase activity in human lymphocytes. *Cancer Res.* **41**, 2757–2761.
18. Bergman, J. (1970). Condensation of indole and formaldehyde in the presence of air and sensitizers. *Tetrahedron* **26**, 3353–3355.
19. Gillner, M., Bergman, J., Cambillau, C., Fernstrom, B. & Gustafsson, J-A. (1985). Interactions of indoles with specific binding sites for 2,3,7,8,-tetrachlorobenzo-*p*-dioxin in rat liver. *Mol. Pharmacol.* **28**, 357–363.
20. Gillner, M., Bergman, J., Cambillau, C., Alexandersson, M., Fernstrom, B. & Gustafsson, J-A. (1993). Interactions of indolo[3,2-*b*]carbazoles and related polycyclic aromatic hydrocarbons with specific binding sites for 2,3,7,8-tetrachlorobenzo-*p*-dioxin in rat liver. *Mol. Pharmacol.* **44**, 336–345.
21. Saito, I., Sugiyama, H., Matsuura, T., Fukuyama, K. & Kotsube Y. (1984). Photochemical reaction between tryptophan and thymine. *Tetrahedron Lett.* **25**, 3243–3247.
22. Narumiya, S., Takai, K., Tokuyama, T., Nodo, Y., Ushiro H. & Osamu, H. (1979). A new metabolic pathway of tryptophan initiated by tryptophan side chain oxidase. *J. Biol. Chem.* **254**, 7007–7015.
23. Fernandez-Salguero, P., *et al.*, & Gonzalez F.J. (1995). Immune system impairment and hepatic fibrosis in mice lacking the dioxin-binding Ah receptor. *Science* **268**, 722–726.

Received: **18 Sep 1995**; revisions requested: **10 Oct 1995**; revisions received: **25 Nov 1995.** Accepted: **29 Nov 1995.**