Indole and tryptophan metabolism: endogenous and dietary routes to Ah receptor activation

Troy D. Hubbard, Iain A. Murray, and Gary H. Perdew

Graduate Program in Biochemistry, Microbiology, and Molecular Biology, The Pennsylvania State University, University Park, PA (T.D.H); Center for Molecular Toxicology and Carcinogenesis and the Department of Veterinary and Biomedical Sciences, The Pennsylvania State University, University Park, PA (T.D.H., I.A.M., G.H.P.)
Running Title:
Indole and Tryptophan Derived Endogenous AHR Ligands

Meta Data
Text Pages: 30
Tables: 1
Figures: 5
References: 141
Abstract: 150

Abbreviations
AHR, aryl hydrocarbon receptor
ARNT, aryl hydrocarbon receptor nuclear translocator
Cyp, cytochrome P450 enzyme
TCDD, 2,3,7,8-tetrachlorodibenzo-p-dioxin
DRE, dioxin response element

Address correspondence to: Dr. Gary H. Perdew, CMTC, 309 LSB, Penn State University,
University Park, PA 16802. E-mail: ghp2@psu.edu
ABSTRACT:

The aryl hydrocarbon receptor (AHR) is a ligand-activated transcription factor recognized for its role in xenobiotic metabolism. The physiological function of the AHR has expanded to include roles in immune regulation, organogenesis, mucosal barrier function, and cell cycle. These functions are likely dependent upon ligand-mediated activation of the receptor. High affinity ligands of the AHR have been classically defined as xenobiotics, such as polychlorinated biphenyls and dioxins. Identification of endogenous AHR ligands is key to understanding the physiological functions of this enigmatic receptor. Metabolic pathways targeting the amino acid tryptophan and indole can lead to a myriad of metabolites, some of which are AHR ligands. Many of these ligands exhibit species selective preferential binding to the AHR. The discovery of specific tryptophan metabolites as AHR ligands may provide insight concerning where the AHR is activated in an organism, such as at the site of inflammation and within the intestinal tract.
Introduction

The aryl hydrocarbon receptor (AHR) is a ligand activated transcription factor of the basic region helix-loop-helix-PER/ARNT/SIM homology family (Bersten et al., 2013). AHR was first identified in the early 1970s as the mediator of 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD/dioxin) toxicity (Poland and Knutson, 1982). Overt dioxin exposure in humans leads to a chronic inflammatory condition known as chloracne, which is characterized by persistent painful skin lesions. The receptor contains a promiscuous ligand binding pocket, which can bind multiple classes of exogenous compounds such as PAH’s and polychlorinated biphenyls (Poland and Knutson, 1982).

The prototypic signaling pathway of AHR mediated transcriptional activity is characterized by transcription of a battery of drug metabolizing enzymes, which includes cytochrome P450 (CYP) enzymes 1A1, 1A2, and 1B1 (Denison et al., 1988; Strom et al., 1992; Rowlands and Gustafsson, 1997; Zhang et al., 1998; Nukaya et al., 2009). In the absence of ligand the AHR resides within the cytoplasm in a stable complex with two molecules of heat shock protein 90 (HSP90), one molecule of X-associated protein 2, and one molecule of the HSP90 co-chaperone p23 (Fig. 1) (Perdew, 1988; Meyer et al., 1998; Kazlauskas et al., 1999). The AHR undergoes nucleocytoplasmic shuttling in the absence of ligand, but the significance of this ability remains to be established (Ikuta et al., 2000). AHR ligands enter the ligand binding domain (LBD) of the receptor; facilitating conformational changes that expose the nuclear localization sequence, which mediates active transport into the nucleus (Ikuta et al., 1998). Once localized to the nucleus, the AHR dissociates from its cytoplasmic complex in the presence of the aryl hydrocarbon receptor nuclear translocator (ARNT), thus generating a functional heterodimeric transcription factor complex (Reyes et al., 1992). Formation of the AHR/ARNT heterodimer facilities binding to a consensus DNA sequence (5′-TNGCGTG-3′), referred to as a dioxin response element (DRE) (Patel et al., 2009). AHR/ARNT binding to a DRE
leads to coactivator recruitment to the transactivation domain of the AHR, which facilitates chromatin remodeling and subsequent target gene transcription (Fig. 1). Most of the genes identified as direct targets of AHR activation were identified through exposure to potent AHR ligands, either in cell culture or in vivo without the presence of other stimuli. However, more recently it has been recognized that the AHR can regulate gene transcription in the presence of other mediators, such as those involved in inflammatory signaling. Genes regulated during inflammation, such as \textit{IL6}, \textit{IL22}, \textit{PTGS2}, and in angiogenesis, such as \textit{VEGFA}, are known to be regulated by activation of the AHR in a combinatorial fashion in the presence of inflammatory mediators (DiNatale et al., 2010b; Monteleone et al., 2011; Lahoti et al., 2013; Terashima et al., 2013). The diverse list of genes modulated by receptor activity suggests broad physiological implications for the AHR, beyond that of xenobiotic metabolism.
The Physiological Significance of the AHR

The physiological effects of AHR activation have expanded to include multifaceted roles in immune regulation, intestinal homeostasis, and carcinogenesis. The AHR integrates into both innate and adaptive immunity through diverse mechanisms; including repression of acute phase response genes, differentiation of T_{reg} lymphocytes and B-cell differentiation (Funatake et al., 2005; Funatake et al., 2008; Patel et al., 2009; Bhattacharya et al., 2010; Mohinta et al., 2015). The AHR also promotes intestinal homeostasis through regulation and development of intraepithelial lymphocytes and innate lymphoid cells (Kiss et al., 2011; Li et al., 2011; Qiu et al., 2012). These cell types maintain a critical line of defense against infiltrating pathogenic microbes and facilitates gut homeostasis at the basolateral surface of the intestinal epithelium. The involvement of the AHR in inflammatory signaling and cell cycle progression suggests it may play a role in various stages of tumorigenesis (Gasiewicz et al., 2008; Safe et al., 2013; Murray et al., 2014). Evidence for this includes high levels of constitutive receptor activity and nuclear localization observed in aggressive tumors and tumor cell lines (Yang et al., 2008; DiNatale et al., 2012; Richmond et al., 2014). The measurement of AHR activity in the tumor microenvironment has been investigated as a possible diagnostic indicator of cancer aggressiveness (Murray et al., 2014). Results indicate a positive or negative correlation between AHR activity and a poor prognosis that depends on the type of cancer. For example, studies have suggested that increased AHR activity is associated with diminished aggressiveness of hormone-dependent breast tumors, most likely due to AHR transrepression of estrogen receptor signaling (Saito et al., 2014). Conversely, elevated AHR transcriptional activity is observed with non-small cell lung cancer and correlates with increased malignancy and poor prognosis (Su et al., 2013).
These findings indicate the importance of ligand-mediated AHR signaling in vertebrates and suggest the existence of endogenous ligands that modulate inducible receptor activity.
The case for endogenous AHR ligands

The expression of genes, which exhibit a degree of AHR dependency, are observed in the absence of exogenous xenobiotic exposure, suggesting either the existence of endogenously derived AHR ligands or additional mechanisms of regulation. For example, AHR/ARNT heterodimer has been observed to reside constitutively in the nucleus of HeLa cells (Singh et al., 1996). The advent of targeted gene disruption in mammalian systems enabled the creation of the \( Ahr^{-/-} \) mouse model to examine the in vivo physiological functions of the AHR in the absence of xenobiotic exposure (Schmidt et al., 1996). The AHR null mice were found to be resistant to dioxin toxicity and failed to significantly express AHR target genes, such as \( Cyp1a1 \) and \( Cyp1a2 \) (Fernandez-Salguero et al., 1995; Mimura et al., 1997). However, the absence of AHR expression during development led to a multitude of unexpected phenotypic abnormalities (Schmidt et al., 1996; Fernandez-Salguero et al., 1997). For example, the neonatal closure of the portal ductus venosus was found to depend on the presence of AHR in developing neonatal mice (Lahvis et al., 2005). Furthermore, altered organ development, including decreased liver size, impaired hepatic vascular development, cardiac hypertrophy, gastric lesions, and dermal fibrosis were observed in aged \( Ahr^{-/-} \) mice. In addition, these mice are more susceptible to intestinal challenge, supporting a critical role for the AHR in gut immune function and barrier function (Sutter et al., 2011). The decreased fitness of AHR null mice suggests extensive physiological roles for the AHR in developing mammals and supports the existence of endogenous ligands to invoke constitutive receptor activity.
Search for Endogenous Ligands

The search for putative endogenous AHR ligands and their physiological role has been an expanding area of interest in the field of AHR biology (Denison and Nagy, 2003; Nguyen and Bradfield, 2008). The generation of endogenous ligands and subsequent AHR activation may occur by diverse mechanisms, such as dietary consumption, synthesis by host commensal microorganisms, free radical formation, and endogenous enzymatic activity. Interestingly, lack of the AHR leads to accumulation of AHR ligands in the lung of AHR null mice (Chiaro et al., 2007). Numerous compounds have been proposed as putative endogenous AHR ligands, many of which are generated through pathways involved in the metabolism of the amino acid tryptophan and indole (Bittinger et al., 2003; Chung and Gadupudi, 2011) (Table 1). In studies attempting to assess the agonist potential of a compound, it is important to test whether physiologically relevant concentrations can mediate AHR activity. The identification of endogenous metabolites that mediate induction of AHR target genes and nuclear translocation of the receptor should not be used as a proxy for ligand binding assays in order to be able to definitively identify them as actual AHR ligands. A compound must be shown to occupy the ligand-binding domain of the AHR in order to be characterized as an AHR ligand. To be defined as an AHR ligand a given compound should be confirmed through either a binding assessment using a radioactive analog of the compound in question or a competitive ligand binding assay using a known radioactive AHR ligand (Ramadoss and Perdew, 2004). Identification of novel routes of endogenous ligand synthesis is critical to understanding mechanisms of inducible receptor activation and how such activation promotes various states of health and disease.
Dietary exposure to indole-3-carbinol

The absence of a primary endogenous regulator of AHR activation suggests multiple routes of ligand accumulation within an organism, including dietary intake. Consumption of AHR ligands could play a significant role in AHR-mediated effects on intestinal homeostasis and the local metabolism of dietary and endogenous compounds. The physiological significance of AHR signaling in the gastrointestinal tract has been demonstrated using dextran sodium sulfate inducible intestinal injury models in which AHR null mice exhibit increasingly severe symptoms and lethality (Benson and Shepherd, 2010; Arsenescu et al., 2011). Activation of AHR by TCDD in intestinal disease models mediated increased survivability and decreased severity of symptoms (Takamura et al., 2010). However, colitis symptoms were significantly reduced in AHR heterozygous compared to AHR homozygous mice, suggesting overstimulation of the AHR may have mixed consequences (Arsenescu et al., 2011). Dietary intake of AHR agonists would likely provide a diverse array of potential ligands of varying potency and metabolic half-lives that could mediate systemic or localized receptor activation.

Biological analysis of extracts from herbal supplements, fruits, and vegetables has revealed significant activation of DRE-driven cell-based reporter systems, target gene expression, and DNA-binding functions, indicative of the presence of AHR activators (Jeuken et al., 2003). Cruciferous vegetables, such as broccoli and brussels sprouts have been found to be rich sources of potent indole-based glucobrassicin (Fig. 2). Mastication promotes the enzymatic cleavage of glucobrassicin by myrosinases to yield indole-3-carbinol (I3C) and indole-3-acetonitrile (I3ACN) both of which exhibit the capacity to bind and activate the AHR, albeit weakly (Loub et al., 1975; Bjeldanes et al., 1991; Shapiro et al., 2001; Ito et al., 2007). The presence of myrosinase activity within plant cells and resident microflora may contribute to the enzymatic cleavage of glucobrassicin to I3C (Dosz et
al., 2014; Luang-In et al., 2014). Interestingly, the in vivo potency of I3C and I3ACN with regard to AHR activation has been demonstrated to depend on the mode of exposure. Administration through intraperitoneal injection renders these compounds less potent compared to oral gavage (Bradfield and Bjeldanes, 1987). Subsequent studies revealed that in the presence of gastric acid both I3C and I3ACN undergo non-enzymatic acid condensation reactions to generate a number of additional AHR ligands, including 3,3′-diindolylmethane (DIM), 2-(indol-3-ylmethyl)-3,3′-diindolylmethane (Ltr-1), and indolo[3,2-b]carbazole (ICZ) (Bjeldanes et al., 1991). Of these glucobrassicin-derived compounds (I3C, I3ACN, DIM, Ltr-1 and ICZ), ICZ exhibits the highest affinity and activation potential for the AHR and thus probably represents a physiologically relevant glucobrassicin-derived AHR ligand. Through the use of a competitive binding assay, ICZ was found to have a $K_d$ of 190 pM, 27-fold less potent than TCDD (Bjeldanes et al., 1991; Jellinck et al., 1993).

The beneficial effects of I3C and its condensation products are wide ranging. Effects of I3C treatment upon suppression and initiation of carcinogenesis vary and appear to be dependent upon the experimental models used. Inhibition of tumorigenesis by I3C has been observed in various target tissues, including liver, thyroid, skin, lung, and colon (Fares, 2014). Anti-carcinogenic activity within animal models of mammary and uterine tumorigenesis suggest I3C may have therapeutic potential in the treatment of human breast cancer (Bradlow et al., 1991; Kojima et al., 1994; Reed et al., 2005). This therapeutic function may be due to its anti-estrogenic activity and inhibition of cell proliferation by estrogen signaling (Michnovicz and Bradlow, 1990; Marconett et al., 2010). Dietary I3C is reported to suppress in vivo models of prostate carcinogenesis by inducing cell cycle arrest and apoptosis of prostate cancer cells (Chinni et al., 2001; Nachshon-Kedmi et al., 2003; Sarkar and Li, 2004). Dietary I3C or its condensation products have displayed therapeutic activity in mouse models of multiple sclerosis and dextran sodium sulfate induced colitis (Huang et al., 2013; Rouse et
al., 2013). Both murine disease models are characterized by the induction of harmful inflammatory states, which were alleviated through the AHR-dependent promotion of anti-inflammatory T_{reg} cell differentiation. The dietary metabolite I3C has been shown to limit pro-inflammatory responses through additional mechanisms such as suppression of NF-κB signaling pathways (Safa et al., 2015). Additionally, I3C or its metabolites exhibit anti-microbial properties, such as inhibition of E.coli and S.aureus biofilm formation (Monte et al., 2014). These effects may have significant implications upon host intestinal microbiota composition and may be a deterrent to opportunistic infection.

Despite the beneficial effects of I3C, its metabolites or condensation products, long term in vivo treatment studies indicate that, consistent with its activity as an AHR agonist, I3C may be procarcinogenic and aid in the initiation of tumorigenesis. I3C administration within animal models has been shown to enhance the development of colonic lesions in rats and promote aflatoxin B-1 initiated hepatocarcinogenesis in trout models (Oganesian et al., 1999; Exon et al., 2001). Recent investigations by the National Toxicology Program division of NIH indicated that I3C administration to rodents correlated with an increase in hepatocarcinogenesis (NTP, 2014). This effect may be due to systemic overstimulation of the AHR facilitated by excessive amounts of ligand. Currently, I3C is available as a dietary supplement at dosages exceeding levels achievable with normal consumption of vegetables. Dietary levels of I3C most likely would facilitate beneficial localized activation of the AHR in the intestinal tract, while avoiding the potentially detrimental effects of systemic activation, further supporting the notion that “less is more”.
Microbiota Derived AHR Ligands

The gastrointestinal tract is composed of a complex microbial ecosystem that contributes significantly to the first-pass metabolism of dietary constituents. The co-evolutionary commensalism between host and microbes may include the catabolism of tryptophan to AHR ligands. Such interkingdom signaling pathways may have evolved to promote the maintenance of intestinal health through the sensing of bacteria at the epithelial interface. In contrast, the excessive production of AHR ligands could play a role in the etiology of various intestinal disease states. Microflora mediated activation of the AHR has been observed in murine models following oral gavage of heat killed *Lactobacillus bulgaricus* OLL1811, which was found to attenuate dextran sodium sulfate induced colitis (Takamura et al., 2011). The metabolite responsible for the observed therapeutic activity remains to be identified. This concept is of further interest in the field of prebiotics and probiotics, where elucidation of the mechanisms facilitating their beneficial effects remains to be fully explored.

Microbial metabolism of the amino acid tryptophan is recognized as the source for the generation of a number of AHR ligands (Fig. 3). Tryptophan is an essential amino acid, and can only be acquired through the diet due to its inability to be synthesized *de novo*. Resident microbes, such as *E.coli*, may utilize tryptophan as a source of nitrogen and contain up to three distinct permeases responsible for its transport, *Mtr*, *AroP*, and *TnaB* (Yanofsky et al., 1991). Tryptophan can be converted to mono-substituted indole compounds, such as indole acetic acid (IAA) and tryptamine, which have been found to activate the AHR in a yeast DRE-dependent reporter assay and stimulate target gene expression in a human colon-carcinoma (Caco2) cell line (Miller, 1997; Jin et al., 2014). The EC\textsubscript{50} of IAA and tryptamine were determined to be 0.5 mM and 0.2 mM respectively, using a [\textsuperscript{3}H] TCDD competition assay, indicating that they bind to the ligand
binding domain with low affinity (Heath-Pagliuso et al., 1998). The monoamine alkaloid, tryptamine, is derived from the direct decarboxylation of tryptophan. Tryptophan is converted to indole-3-acetic acid (IAA) via the enzymes tryptophan monoxygenase (*IaaM*) and indole-3-acetamide hydrolase (*IaaH*), which constitute the indole-3-acetamide (IAM) pathway (Tsavkelova et al., 2012). The common fecal metabolite 3-methyl indole (skatole), is generated by decarboxylation of IAA and has been shown to facilitate transcription of AHR target genes in human bronchial and colonic epithelial cells (Weems and Yost, 2010; Hubbard et al., In Press). The physiological implications of AHR activation by tryptamine, IAA, or skatole remain to be established.

The catabolism of tryptophan by commensal lactobacilli has been shown to hinder colonization of the intestinal tract by pathogenic *Candida albicans* through AHR dependent expression of IL22 (Zelante et al., 2013). Targeted metabolomic analysis of *L. reuteri* and *L. johnsonii* tryptophan catabolism identified the indole derivative indole-3-aldehyde (Iald), which is generated through the indole pyruvate pathway, catalyzed by the enzyme aromatic amino acid aminotransferase. In vivo treatments with Iald were found to activate the AHR and to promote intestinal homeostasis through induction of *IL-22* (Zelante et al., 2013). The manipulation of pathways that promote Iald production has demonstrated therapeutic promise; although more evidence is needed to address IL-22 mediated activity in human models. It will be important to determine if normal colonization of lactobacillus species can generate adequate amounts of ligand to activate the AHR.

Bacterial metabolism of tryptophan by the tryptophanase pathway has been shown to generate indole and pyruvate (Fig.3). Indole is shuttled in and out of bacteria by passive diffusion or active transport by AcrEF-ToIC and Mtr transporters (Yanofsky et al., 1991; Kawamura-Sato et
al., 1999; Pinero-Fernandez et al., 2011). Interspecies signaling by indole regulates many physiological functions, such as biofilm formation, motility, plasmid stability, virulence and antibiotic resistance in neighboring microflora (Bansal et al., 2007; Lee et al., 2007; Lee and Lee, 2010). Metabolic analysis of human fecal samples has detected indole at concentrations of 250-1100 μM (Karlin et al., 1985; Zuccato et al., 1993). Such high in vivo concentrations of the metabolite suggest it may facilitate a mutually beneficial relationship with the host. In fact, treatment of human intestinal epithelial cells (HCT-8) with indole was shown to increase barrier function and attenuate inflammatory markers (Bansal et al., 2009). In recent competitive ligand binding experiments indole was found to preferentially activate the human AHR compared to the mouse AHR (Miller, 1997; Hubbard et al., In Press). In addition, indole was able to induce DRE-dependent luciferase activity with EC$_{50}$ of ~3 μM, which is well within the range of relevant physiological concentrations. The physiological significance of microbial indole generation upon the intestinal epithelium and its effect upon basal AHR signaling remains to be established in in vivo models.

A lack of intestinal homeostasis has been shown to be a contributing factor in the pathology of many disorders, including inflammatory bowel disease, Crohn’s disease, ulcerative colitis, diabetes, obesity, and cancer. Maintenance of intestinal homeostasis solely by dietary AHR ligands would be transient and vary greatly according to the types of ligands consumed, their respective bioavailability, and half-lives. Thus, synthesis of putative AHR ligands, such as indole, by the intestinal commensal microflora may provide a more consistent source of in vivo AHR stimulation. Such AHR stimulation arising from microbial activity may represent a mechanism by which the AHR acts as a sensor of the microbiota community and, through its established role as a modulator of immune function, maintain host-microbe homeostasis. Clearly,
the implications of microbial ligand synthesis and their effect upon intestinal disease requires further investigation.
Host Metabolism of Indoles

Prevailing evidence suggests that the production of indole is dependent upon microbial metabolic activity (Wikoff et al., 2009). Subsequent absorption of indole by the host occurs via passive diffusion through the colonic epithelia. Hepatic uptake and metabolism of indole is another likely source for the generation of potent AHR ligands (Fig. 3). For example, Phase I metabolism of indole by cytochrome P450 2E1 (Cyp2e1) in rat liver microsomes catalyzes the hydroxylation of indole to form 3-hydroxy-indole (indoxyl) and isatin. In addition, recombinant human CYP450 enzymes within yeast expression systems can also catalyze the same reaction (Gillam et al., 2000; Banoglu et al., 2001). In vivo generation of the indigoids, indigo and indirubin, have been proposed to occur through radical redox mechanisms involving the unstable intermediates, indoxyl and isatin (Gillam et al., 2000). Indigo and indirubin have been identified in human urine and found to be potent activators of the AHR (Adachi et al., 2001; Guengerich et al., 2004; Sugihara et al., 2004). Indirubin was found to be a more potent human AHR agonist when compared to the murine Ahb-1 allele (high affinity for TCDD) of the receptor (Flaveny et al., 2009). Furthermore, in a yeast DRE-driven β-galactosidase reporter expressing the human AHR, indirubin was found to possess an EC50 of ~0.2 nM, 45-fold less than TCDD (Adachi et al., 2001). Whether the liver synthesizes indirubin or indigo in quantities sufficient to activate the AHR has not been established.

Hepatic phase II drug metabolism of indoxyl via sulfation by sulfotransferases (SULT) generates the uremic toxin indoxyl-3-sulfate (I3S). I3S is recognized as a human AHR agonist with 500-fold greater potency in human versus murine hepatoma cell lines (Schroeder et al., 2010). Increased serum concentrations of I3S have been correlated with vascular dysfunction in chronic kidney disease (CKD) patients (Barreto et al., 2009). Assessment of serum from dialysis patients indicates an accumulation of I3S to high μM concentrations. Failed excretion and subsequent
accumulation of I3S may contribute to the overt toxicity seen in dialysis patients and has been shown to contribute to increased risk of tumor growth (Wong et al., 2013). Antagonism of the AHR or targeted inhibition of I3S synthesis may have a beneficial therapeutic effect in CKD patients (Sallee et al., 2014). Oxidative stress from the accumulation of I3S has also been shown to increase the risk of cardiovascular disease in CDK patients through oxidative modification of low density lipoproteins (LDL) (Cao et al., 2014). The in vitro oxidation of LDL by indoxyl radicals has been shown to generate the AHR activator tryptanthrin (Praschberger et al., 2014). Tryptanthrin synthesis has also been characterized previously in Candida lipolytica and Malassezia yeast cultures through the catabolism of tryptophan (Schrenk et al., 1999; Vlachos et al., 2012). Furthermore, the production of AHR agonists within the epidermis by Malassezia yeasts have been shown to contribute to seborrhoeic dermatitis disease progression (Gaitanis et al., 2008). However, AHR activity mediated by tryptanthrin appears to be localized to the skin, limiting its in vivo relevance as an endogenous regulator.

The metabolic profiling of tissue extracts has led to the identification of endogenous AHR ligands that contain indolyl moieties. Organic extractions of porcine lung tissue led to the discovery of the AHR ligand 2-(1’H-indole-3’-carbonyl)-thiazole-4-carboxylic acid methyl ester (ITE), which was shown to be a direct high affinity ligand for human, mouse and fish receptors (Song et al., 2002). However, the presence of ITE in vivo has not been established and thus it is possible that the generation ITE may be an artifact of the extraction process in which the methanol extract was heated to high temperatures, as described in the initial report. Nevertheless, in mouse models of experimental autoimmune uveitis, ITE was found to inhibit the onset of symptoms through an AHR dependent mechanism of T_{reg} cell differentiation (Nugent et al., 2013). ITE has also been shown to possess AHR independent biological activity, such as the inhibition of TGFβ1
mediated myofibroblast differentiation, which prevents fibrosis within human tissue (Lehmann et al., 2011). The AHR ligand ITE has many biological activities; however its classification as an “endogenously” derived compound remains to be established.
Host Metabolism of Tryptophan

Over the past century, biochemical analyses found that the enzymatic reactions that account for greater than 90% of the peripheral metabolism of tryptophan in mammals reside in the kynurenine pathway (Leklem, 1971). The kynurenine pathway is composed of multiple enzymatic steps that facilitate the metabolism of tryptophan to nicotinamide adenine dinucleotide and a number of bioactive intermediate byproducts, some of which have been characterized as activators of the AHR (Fig. 4) (Bohar et al., 2015). In addition, the kynurenine pathway is thought to play a key role in cancer-immunity through signaling cascades that promote immune suppression or tolerance. The depletion of tryptophan and subsequent generation of ligands could activate the AHR, which can alter naive CD4+ T-helper cell differentiation pathways to favor an anti-inflammatory T\textsubscript{reg} rather than a Th\textsubscript{17} phenotype (Quintana et al., 2008; Veldhoen et al., 2008). This alteration in adaptive immunity functions to hinder immune surveillance within the tumor microenvironment. Increased expression of the enzymes in this pathway has been observed in tumors, as evidenced by an increase in the kynurenine:tryptophan ratio, which has been shown to positively correlate with cancer progression (Suzuki et al., 2010; Pilotte et al., 2012).

The kynurenine pathway is initiated by host metabolism of tryptophan via indoleamine-2,3-dioxygenase (IDO1) and tryptophan-2,3-dioxygenase (TDO2). TDO2 and IDO1 are analogous enzymes that perform an identical catabolic conversion of tryptophan to kynurenine, but are structurally divergent and exhibit distinct expression patterns in vivo (Ball et al., 2014). Interferon-γ (IFN\textsubscript{γ}) signaling has been shown to regulate IDO1 expression via activation of a JAK/STAT signaling in cell types of epithelial or monocytic lineage (Yoshida et al., 1981; Nguyen et al., 2010). IDO1 induction was also found to depend on AHR transcriptional
regulation within bone-marrow derived dendritic cells (Nguyen et al., 2010). More recent studies have identified TDO2 enzyme activity as an additional mechanism of tryptophan depletion mainly in hepatocytes and neurons. The upregulated expression of TDO2 in glioblastoma tissues has displayed a positive correlation with the induction of the AHR target gene CYP1B1, which supports the hypothesis of AHR ligand generation within the kynurenine pathway (Opitz et al., 2011). The metabolism of tryptophan by IDO/TDO in glioma cell lines yields significant amounts of kynurenine, which was found to be a low affinity ligand for the AHR (Mezrich et al., 2010; Opitz et al., 2011). Additional metabolism of kynurenine leads to the generation of more potent AHR ligands that may have an increased impact upon immune regulation and tumor malignancy. The metabolite kynurenic acid, generated by the irreversible transamination of kynurenine by kynurenine aminotransferases (KAT’s), was found to be a relatively potent endogenous ligand for the human AHR (Han et al., 2009; DiNatale et al., 2010a). Interestingly, KA can be detected in serum at a concentration of ~5 μM in some CKD patients (Schefold et al., 2009). KA at a physiologically relevant dose of 100 nM mediated activation of the AHR and was found to induce prototypical AHR target genes and synergistically increase the expression of IL6 in the presence of inflammatory signaling (DiNatale et al., 2010a). This investigation also found that the kynurenine metabolite, xanthurenic acid, displayed agonist potential in an AHR dependent luciferase reporter in human hepatoma cells. Xanthurenic acid is generated by transamination of the kynurenine metabolite, 3-hydroxykynurenine. Recently, the kynurenine metabolite cinnabarinic acid (CA) has been identified as an additional endogenous AHR agonist that mediates AHR dependent transcription of IL22 in human and murine CD4+T-cells (Lowe et al., 2014). CA is generated by condensation of 3-hydroxyanthranilic acid, which is produced by kynureninase from 3-hydroxykynurenine (Fig. 4). Recently, CA has also been shown to play a
cytoprotective role in endoplasmic reticulum or oxidative stress induced apoptosis through AHR dependent induction of stanniocalcin 2 (Stc2) (Joshi et al., 2015). Further investigation of the kynurenine pathway may lead to the identification of additional endogenous AHR ligands.

The identity of the metabolite(s) primarily responsible for AHR activation in tumor tissue has not yet been confirmed. Increased levels of AHR activity in tumors, characterized by increased CYP1B1 mRNA expression, have been shown to positively correlate with poor survival in glioma patients (Opitz et al., 2011). In addition, studies examining metabolite concentrations in pancreatic adenocarcinomas identified a positive correlation between kynurenic acid levels and tumor size (Botwinick et al., 2014). Such evidence supports the notion that AHR stimulation may increase tumor aggressiveness through metabolites of the kynurenine pathway. Targeted antagonism of the AHR or the kynurenine pathway enzymes responsible for synthesis of higher affinity ligands may be of therapeutic interest as a supplement to traditional chemotherapy and is worthy of exploration. Indeed, pharmaceutical inhibition of IDO function is currently in the initial stages of clinical trials.
Photo-Oxidation of Tryptophan

Studies have found that exposure of culture media to ultraviolet (UV) light leads to increased AHR activation and induction of cytochrome P450 target genes within cultured cells (Oberg et al., 2005). The photo-oxidation of tryptophan was found to generate 6-formylindolo[3,2-b]carbazole (FICZ), which has been established as a potent AHR ligand with an affinity similar to that of TCDD (Fig. 5) (Rannug et al., 1987; Rannug et al., 1995). FICZ treatment of a murine autoimmune encephalitis model displayed an increase in the severity of symptoms, contradictory to that observed with TCDD (Quintana et al., 2008; Veldhoen et al., 2008). Further investigation found that FICZ induced differentiation of Th17 cells and expression of proinflammatory $\text{Il17}$ rather than anti-inflammatory Treg cells. This unique physiological activity of FICZ suggests that it may exhibit selective AHR activity relative to other AHR ligands.

The relevance of FICZ in in vivo systems remains to be firmly established. Epidermal UV-irradiation of both human skin or keratinocyte cultures was found to increase AHR target monooxygenases, but this apparently did not depend on FICZ formation (Katiyar et al., 2000; Gonzalez et al., 2001). The metabolic analysis of human urine indicated the presence of multiple FICZ conjugates, suggesting it may be an endogenous high affinity regulator of the AHR (Wincent et al., 2009). However, confirmation of sub-dermal synthesis of FICZ has not yet occurred. Studies in mouse hepatoma cells indicate a negative feedback loop in which FICZ induces $\text{Cyp1a1}$ expression that in turn leads to its rapid degradation (Wei et al., 2000). Thus, limited in vivo synthesis and rapid degradation make FICZ unlikely to be a significant contributor to AHR-mediated activity within the context of an entire organism. Nevertheless,
FICZ may have significant implications within the context of AHR activation in skin, where constant UV-exposure may generate sufficient quantities of the ligand to activate the AHR.
**Human vs. Mouse AHR Ligand Selectivity**

The identity of the primary mediators of AHR activity remains elusive due to the ability of the receptor to bind a number of different classes of compounds. This issue is further complicated when trying to determine which endogenous ligands may have important implications in human health. While the regulation of monooxygenase expression by the AHR appears to be conserved, there exist a number of structural distinctions between murine and human isoforms of the receptor. Amino acid sequence comparisons between the human and murine receptor (Ahb-1 allele) reveal 85% sequence homology in the amino terminal half of the receptor that contains the ligand-binding domain. These variations lead to significant differences in ligand affinity and selectivity between species. For example, the mouse AHR (Ahb-1) displays a ten-fold increased affinity for TCDD relative to the human AHR. This difference in affinity for dioxin has been shown to be associated with a single amino acid substitution in the ligand binding pocket (Ema et al., 1994; Ramadoss and Perdew, 2004). The human receptor has been shown to have increased affinity for indole based derivatives such as indigoids, indole, indoxyl-sulfate, and kynurenic acid (Flaveny et al., 2009; DiNatale et al., 2010a; Schroeder et al., 2010; Hubbard et al., In Press). Enhanced affinity of the human receptor for particular compounds may indicate an evolutionary adaptation to respond to specific endogenous and exogenous regulators that mediate certain physiological responses (e.g. during inflammation). Considering the differential ligand affinity frequently observed between species, the primary use of in vitro or in vivo mouse systems to catalog potential ligands of the AHR will not always identify ligands that bind to the AHR in other species. For this reason, both human and murine systems should be utilized when trying to assess putative AHR ligand potential of a given compound. Although functionally conserved, the mouse and human AHR exhibit significant differences in their ability to modulate
gene expression. The expression of the human AHR in a transgenic mouse allowed the assessment of whether the human AHR exhibits similar transcriptional activity relative to the mouse AHR in mouse hepatocytes (Flaveny et al., 2010). While many of the AHR-dependent targets are conserved, species-specific targets do exist that appear to be due to inherent differences in AHR structure. Clearly when comparing two species, other differences in transcriptional regulation are also likely, such as where functional DRE are found in various promoter regions. Species-specific differences in AHR activity indicate the need for a suitable transgenic mouse model that possesses tissue-specific or global knock-in of the human AHR at the murine locus. Such ‘humanized AHR models may facilitate the in vivo analysis of putative endogenous AHR ligands and allow for a greater understanding of the impact of endogenous signaling in human health and disease.
Possible mechanism of AHR activation by low molecular weight compounds

Exogenous high affinity AHR ligands, such as halogenated dioxins and dibenzofurans are products of the combustion of organic matter and industrial activity. Such ligands generally are not metabolically labile, but are highly lipophilic, thus allowing bioaccumulation. The accumulation of these ligands, combined with their high affinity for the AHR and subsequent persistent modulation of AHR-dependent signaling, contribute significantly to AHR associated toxicity. However, lower endogenous levels of activation of the AHR have been shown be beneficial in the maintenance of immune health and intestinal homeostasis. The adaptation of AHR to bind endogenous compounds with moderate to low affinity may have increased its physiological relevance. Such ligands may not structurally resemble dioxin and are significantly smaller in size than polycyclic molecules. AHR displays decreased affinity for ligands such as indole, kynurenic acid, or indoxyl sulfate, relative to TCDD. The accommodation and interaction of small molecules within the ligand binding domain of AHR is likely to differ from that observed with high affinity ligands such as dioxin (Backlund and Ingelman-Sundberg, 2004). Indeed, recent studies provide evidence that the human AHR has the capacity to bind two indole molecules per ligand binding pocket to facilitate activation (Hubbard et al., In Press). The concept of small molecule activation or combinatorial activation of the AHR by two molecules represents a paradigm shift in AHR ligand binding dynamics and expands the pool of potential ligands. The structural diversity of endogenous ligands may induce differential conformations of the ligand bound receptor that ultimately dictate varying biological activities (Lees and Whitelaw, 1999; Kronenberg et al., 2000; Patel et al., 2009; Murray et al., 2010a; Murray et al., 2010b; Jin et al., 2012). Synthesis or accumulation of distinct endogenous ligands within certain cell types or tissues could mediate tissue/cell selective activity of the AHR. The lack of an x-ray crystal structure for the
AHR hinders efforts to better understand ligand-receptor molecular interactions that mediate agonist, antagonist, and selective modulator activities (Safe and McDougal, 2002).
There are many routes of endogenous AHR ligand generation (Fig. 5) and all likely contribute in some way to modulating in vivo AHR function, with perhaps additional pathways and ligands yet to be identified. Considering the growing list of endogenous ligands that have been identified, it is improbable that a single bona fide endogenous ligand exists that is primarily responsible for mediating the in vivo physiological functions of the AHR. The basal activity of AHR within an organism may also be influenced by antagonists or selective modulators generated endogenously that have not yet been characterized. A more comprehensive and mechanistic understanding of all factors that influence AHR signaling pathways within an organism are critical to elucidating the physiological function(s) of this fascinating receptor.
Acknowledgements. We thank Marcia H. Perdew and Kayla Smith for excellent editorial assistance.
Author Contributions

Participated in research design: Hubbard, Murray, Perdew.

Performed data analysis: Hubbard, Murray, Perdew.

Wrote or contributed to the writing of the manuscript: Hubbard, Murray, Perdew.
References


Funatake CJ, Marshall NB and Kerkvliet NI (2008) 2,3,7,8-Tetrachlorodibenzo-p-dioxin alters the differentiation of alloreactive CD8+ T cells toward a regulatory T cell phenotype by a mechanism that is dependent on aryl hydrocarbon receptor in CD4+ T cells. J Immunotoxicol 5:81-91.


NTP (2014) Toxicology studies of indole-3-carbinol in F344/N rats and B6C3F1/N mice and toxicology and carcinogenesis studies of indole-3-carbinol in Harlan Sprague Dawley rats and B6C3F1/N mice (Gavage Studies). National Institute of Health.


Footnote:

This research was supported by the National Institutes of Health; National Institute of Environmental Health Sciences [Grants ES004869, ES019964].
**Figure Legends**

**Fig. 1.** Canonical pathway of AHR-mediated transcription following ligand binding.

**Fig. 2.** Gastric acid condensation of indole-3-carbinol lead to the production of high-affinity AHR ligands.

**Fig. 3.** Microbial synthesis of AHR ligands and host metabolism of indole.

**Fig. 4.** Endogenous tryptophan metabolism through the kynurenine pathway generates AHR ligands.

**Fig. 5.** Summary of Endogenous Pathways of AHR Ligand Synthesis. (1) Microbial Production of indole, (2) Microbial/Host indole metabolism to generate AHR ligands, (3) Tryptophan UV-photo-oxidation to form 6-formylindolo[3,2-b]carbazole (FICZ), (4) Endogenous tryptophan metabolism via the kynurenine pathway, and (5) Diet derived ligand synthesis, are all pathways of AHR ligand synthesis that mediate (6) AHR activation and transcription of target genes interleukin 6 (*IL6*), cytochrome P450 1A1 (*CYP1A1*), vascular endothelial growth factor A (*VEGFA*), prostaglandin G/H synthase 2 (*PTGS2*), interleukin 22 (*IL22*), and cytochrome P450 1B1 (*CYP1B1*).
Table 1. List of endogenous AHR activators derived from indole or tryptophan metabolism.

<table>
<thead>
<tr>
<th>Origin</th>
<th>Compound</th>
<th>Structure</th>
<th>M.W.</th>
<th>Human AHR Selective Agonist</th>
<th>Mouse AHR Selective Agonist</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exogenous</td>
<td>2,3,7,8-Tetrachlorodibenzo-p-dioxin (TCDD)</td>
<td>![Structure Image]</td>
<td>701.67</td>
<td>✔</td>
<td>✔</td>
<td>Poland and Knudson, 1992</td>
</tr>
<tr>
<td>Dietary</td>
<td>Indole-3-carbinol (IC)</td>
<td>![Structure Image]</td>
<td>144.18</td>
<td>✔</td>
<td></td>
<td>Bjeldanes et al., 1991</td>
</tr>
<tr>
<td>Dietary</td>
<td>Indole-3-acetonitrile (ICAN)</td>
<td>![Structure Image]</td>
<td>156.18</td>
<td>✔</td>
<td></td>
<td>Bjeldanes et al., 1991</td>
</tr>
<tr>
<td>Dietary</td>
<td>3,3'-Diindolylmethane (DIM)</td>
<td>![Structure Image]</td>
<td>246.31</td>
<td>✔</td>
<td></td>
<td>Bjeldanes et al., 1991</td>
</tr>
<tr>
<td>Dietary</td>
<td>2-(indol-3-ylmethyl)-3,3'-diindolylmethane (LPS-1)</td>
<td>![Structure Image]</td>
<td>389.47</td>
<td>✔</td>
<td></td>
<td>Bjeldanes et al., 1991</td>
</tr>
<tr>
<td>Dietary</td>
<td>indole[3,2-b]carbazole (IC)</td>
<td>![Structure Image]</td>
<td>254.28</td>
<td>✔</td>
<td></td>
<td>Bjeldanes et al., 1991</td>
</tr>
<tr>
<td>Dietary</td>
<td>2-(1H-indole-3-carbonyl)-thiazole-4-carboxylic acid methyl ester (ITE)</td>
<td>![Structure Image]</td>
<td>280.31</td>
<td>✔</td>
<td></td>
<td>Song et al., 2002</td>
</tr>
<tr>
<td>Microbial</td>
<td>Indole</td>
<td>![Structure Image]</td>
<td>117.15</td>
<td>✔</td>
<td>✔</td>
<td>Hubbard TD, 2015</td>
</tr>
<tr>
<td>Microbial</td>
<td>Indole-3-acetic acid (IAA)</td>
<td>![Structure Image]</td>
<td>175.10</td>
<td>✔</td>
<td></td>
<td>Miller, 1997</td>
</tr>
<tr>
<td>Microbial</td>
<td>Indole-3-alddehyde (IAA)</td>
<td>![Structure Image]</td>
<td>145.16</td>
<td>✔</td>
<td></td>
<td>Jin et al., 2014</td>
</tr>
<tr>
<td>Microbial</td>
<td>Tryptamine</td>
<td>![Structure Image]</td>
<td>180.22</td>
<td>✔</td>
<td></td>
<td>Miller, 1997</td>
</tr>
<tr>
<td>Microbial</td>
<td>3-methyl-indole (5,6-monoxygen)</td>
<td>![Structure Image]</td>
<td>131.17</td>
<td>✔</td>
<td>✔</td>
<td>Weema and Yost, 2010</td>
</tr>
<tr>
<td>Yeast</td>
<td>Tryptanthrin</td>
<td>![Structure Image]</td>
<td>248.24</td>
<td>✔</td>
<td></td>
<td>Gohrich et al., 1996</td>
</tr>
<tr>
<td>Microbial/Host Metabolism</td>
<td>Indigo</td>
<td>![Structure Image]</td>
<td>248.24</td>
<td>✔</td>
<td>✔</td>
<td>Litchfield et al., 2001</td>
</tr>
<tr>
<td>Microbial/Host Metabolism</td>
<td>Indirubin</td>
<td>![Structure Image]</td>
<td>282.27</td>
<td>✔</td>
<td>✔</td>
<td>Litchfield et al., 2001</td>
</tr>
<tr>
<td>Microbial/Host Metabolism</td>
<td>3-indole-3-sulfate (5,6-disulfate)</td>
<td>![Structure Image]</td>
<td>211.54</td>
<td>✔</td>
<td>✔</td>
<td>Gohrecoor et al., 2019</td>
</tr>
<tr>
<td>Host Metabolism</td>
<td>Kynurenine (Kyn)</td>
<td>![Structure Image]</td>
<td>208.21</td>
<td>✔</td>
<td>✔</td>
<td>Mezrich et al., 2013</td>
</tr>
<tr>
<td>Host Metabolism</td>
<td>Kynurenic acid (KAA)</td>
<td>![Structure Image]</td>
<td>198.17</td>
<td>✔</td>
<td>✔</td>
<td>Djilali et al., 2010</td>
</tr>
<tr>
<td>Host Metabolism</td>
<td>Xanthurenic acid</td>
<td>![Structure Image]</td>
<td>205.17</td>
<td>✔</td>
<td></td>
<td>Djilali et al., 2010</td>
</tr>
<tr>
<td>Host Metabolism</td>
<td>Cinchonanic acid (CA)</td>
<td>![Structure Image]</td>
<td>300.22</td>
<td>✔</td>
<td></td>
<td>Lowo et al., 2011</td>
</tr>
<tr>
<td>UV-Light Oxidation</td>
<td>6-formylindole[3,2-b]benzoxazole (FIC2)</td>
<td>![Structure Image]</td>
<td>284.31</td>
<td>✔</td>
<td></td>
<td>Ranug et al., 1987</td>
</tr>
</tbody>
</table>

* Denotes molecule is an activator of the AHR, but further ligand binding assays are required.
Figure 1

AHR agonists

Transformation

HSP90, XAP2, HSP90

Translocation

p23

Ligand binding

Cytosolic/core complex dissociation

AHR-ARNT dimerization

AHR-ARNT

DRE-specific DNA binding

GCGTGC

AHR-mediated transcription
Figure 2

Vegetable rich diet

Stomach

Myrosinase

Indole-3-carbinol

Glucobrassicin

3, 3'-diindoylmethane (DIM)

Indole-3-carbinol

Glucobrassicin

Indole[3,2b]carbazole

LTR-1

Principal condensation product with AHR agonist activity

Acid condensation reactions

Local & systemic AHR activation

Myrosinase

Indole-3-acetonitrile

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+

H^+