# Contribution of circulating host and microbial tryptophan metabolites towards Ah receptor activation


## Authors

Ethan W. Morgan[a], Fangcong Dong[b], Andrew Annalora[c], Iain A. Murray[b], Trenton Wolfe[d], Reece Erickson[d], Krishne Gowda[e], Shantu G. Amin[e], Kristina S. Petersen[f], Penny M. Kris-Etherton[f], Craig Marcus[c], Seth T. Walk[d], Andrew D. Patterson[b], and Gary H. Perdew[b]

## Affiliations

[a] Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA, USA; [b]Department of Veterinary and Biomedical Sciences and the Center for Molecular Toxicology and Carcinogenesis, The Pennsylvania State University, University Park, PA, USA; [c]Department of Environmental and Molecular Toxicology, Oregon State University, Corvallis, OR, 97331, USA; [d] Department of Microbiology and Cell Biology, Montana State University, Bozeman, MT, USA; [e]Department of Pharmacology Penn State College of Medicine, Hershey, PA, USA; [f]Department of Nutritional Sciences, The Pennsylvania State University, University Park, PA, USA;

CONTACT: Gary H. Perdew, ghp2@psu.edu, Department of Veterinary and Biomedical Sciences and the Center for Molecular Toxicology and Carcinogenesis, The Pennsylvania State University, University Park, PA, USA.



## Abstract

The aryl hydrocarbon receptor (AHR) is a ligand activated transcription factor that plays an integral role in homeostatic maintenance by regulating cellular functions such as cellular differentiation, metabolism, barrier function, and immune response. An important but poorly understood class of AHR activators are compounds derived from host and bacterial metabolism of tryptophan. The commensal bacteria of the gut microbiome are major producers of tryptophan metabolites known to activate the AHR, while the host also produces AHR activators through tryptophan metabolism. We used targeted mass spectrometry-based metabolite profiling to determine the presence and metabolic source of these metabolites in the sera of conventional mice, germ-free mice, and humans. Surprisingly, sera concentrations of many tryptophan metabolites are comparable between germ-free and conventional mice. Therefore, many major AHR-activating tryptophan metabolites in mouse sera are produced by the host, despite their presence in feces and mouse cecal contents. AHR activation is rarely studied in the context of a mixture at relevant

concentrations, as we present here. The AHR activation potentials of individual and pooled metabolites were explored using cell-based assays, while ligand binding competition assays and ligand docking simulations were used to assess the detected metabolites as AHR agonists. The physiological and biomedical relevance of the identified metabolites was investigated in the context of cell-based models for cancer and rheumatoid arthritis. We present data here that reframe AHR biology to include the presence of ubiquitous tryptophan metabolites, improving our understanding of homeostatic AHR activity and models of AHR-linked diseases.

**(5-10) Key words:**

Aryl hydrocarbon receptor; tryptophan metabolism; homeostasis; microbiome; metabolomics; indole; kynurenine pathway

## Introduction

Maintaining homeostasis is fundamental to life. Homeostasis is a broad term that includes the many cellular activities that facilitate normal bodily functioning. Evolution has contributed exceptionally complex biochemical and molecular schemes to tightly regulate cellular activity through cell-cell signaling, environmental sensors, transcription factors, post-translational modifications and more. Increased complexity enables greater regulatory control and management of large systems, as seen in mammalian organisms with highly specialized cell types. An important factor in maintaining the specificity of mammalian homeostasis is the ligand activated transcription factor, aryl hydrocarbon receptor (AHR). The AHR, a member of the basic helix-loop-helix/Per-Arnt-Sim family, possesses a promiscuous ligand binding pocket, which allows the recognition of diverse ligands, thus providing flexibility in its function. Upon ligand-mediated activation, the AHR translocates from the cytoplasm to the nucleus where it heterodimerizes with AHR nuclear translocator. This heterodimer binds dioxin response elements in the promoter region of AHR-responsive genes, thereby influencing gene expression.[1] The genes under the AHR's influence span a variety of cellular functions, notably these involved in cellular differentiation, xenobiotic metabolism, and immune function.

The AHR is often studied in terms of facilitating the metabolism of toxic environmental pollutants including polycyclic aromatic hydrocarbons like 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD) or benzo[*a*]pyrene.[1] It is also known to play an important role mediating immune responses by regulating the production of anti- or pro-inflammatory cytokines.[2,3,4] Additionally, the AHR is involved in cell differentiation of multiple cell types including T cells, B cells, and goblet cells.[3,5,6,7] Furthermore, the AHR is widely expressed throughout multiple organ systems and is highly expressed in barrier sites

where it plays a role in maintaining epithelial barrier integrity.[2,8] These expression patterns allow the body to recognize ligands and perform cell-specific roles that contribute to the regulation of homeostasis.

Although it is most studied for its recognition of industrial pollutants, the AHR ligand binding pocket can bind many compounds produced by host, bacterial, and plant metabolism of tryptophan (Trp). An estimated 90% of dietary Trp enters the kynurenine pathway, playing an important role in regulating homeostatic Trp concentrations, while also producing compounds capable of activating the AHR, including kynurenine, kynurenic acid, and xanthurenic acid.[9,10,11] A much smaller proportion of dietary Trp (~2%) enters the serotonin pathway, although these compounds are not known to directly interact with AHR and do not represent a current significant area of research.[12] Bacteria in the gut microbiome use an estimated 5% of Trp to produce indole-based derivatives. The trillions of cells comprising the gut microbiome produce numerous Trp metabolites for diverse physiological functions, including energy metabolism, cell signaling, and stress response.[13,14,15] A subset of these compounds are reported to activate the AHR and are major subjects of research, notably, indole, indole-3-propionic acid (IPA), and indole-3-acetic acid (IAA).[10,14,16,17] Additionally, IAA and IPA are known in the plant sciences as auxins, a class of plant hormones, and thus may be present in our diets.[18,19] Trp metabolism provides a robust and consistent pool of AHR ligands that could be contributing to the regulation of systemic AHR activity and, accordingly, are contributors to homeostatic maintenance. We refer to this diverse class of metabolites as pseudo-endogenous due to the difficulty in distinguishing the metabolite progenitor from host, microbiome, or diet. Previously, our group has determined the presence of Trp metabolites in feces and cecal contents of mice that at the concentration present can generate AHR activity.[20] However, the absorption and circulation of these metabolites is still poorly understood in the context of homeostatic AHR activity. Our work provides new information and clarity to this area of research.

In this study, we used liquid chromatography paired with tandem mass spectrometry (LC-MS/MS) to identify and quantify Trp metabolites circulating in human, mouse, and germ-free mouse sera. We identified metabolites frequently attributed to bacterial metabolism that are, in fact, produced independent of bacteria. Further, by profiling circulating metabolites we found that fecal metabolite concentrations do not serve as an accurate proxy measurement for Trp metabolites that are circulating in a host. We demonstrate that circulating Trp metabolites can generate AHR activity, but do not have a synergistic or additive effect and produce varied outcomes in different cell types. Our work, identifying relevant AHR agonists, provides a basis to recontextualize physiologic AHR activation and may help in examining the link between host metabolism and AHR-linked diseases.

## Results

*Tryptophan metabolite profiling of the sera of conventional and germ-free mice*

We initially aimed to build upon our previous work analyzing murine fecal and cecal contents for the presence of AHR activating Trp metabolites by analyzing murine sera to investigate into which Trp metabolites in the host circulation are derived from diet, host metabolism and/or the gut microbiome.[20] Blood was collected from the portal vein of conventional C567BL/6J mice and sera analyzed by targeted LC-MS, with the exception of indole, which was analyzed by GC-MS . Guided by our previous work, a compilation of 16 Trp metabolites known to elicit AHR activity were targeted for quantification (Table 1). Of the 16 metabolites, nine were detected and quantified through interpolation of standard curves and corrected for recovery and matrix effects (Fig. 1, Table S1). Analysis revealed IPA was the most abundant Trp metabolite with a mean concentration of 8.4 ± 5.69 μM and comprising 50.92% of the total composition of examined metabolites by concentration. IPA represents a singularly large proportion of the total examined metabolites with the remaining metabolites composing substantially smaller proportions: 15.68% indole-3-lactic acid (ILA), 11.87% kynurenine (KYN), 7.38% IAA, 5.99% 5-hydroindole-3-acetic acid, 3.12% kynurenic acid (KA), 2.91% indole-3-carboxaldehyde (I3C), 1.10% indole-3-acrylic acid, and 1.02% 2-oxindole. The remaining seven targeted metabolites were undetectable based on their limits of detection (Table 1). These data present a remarkably different metabolic profile compared to the analysis of murine feces or cecal contents.[20] Additionally, these metabolites, especially the indole-containing compounds, have been linked to the gut microbiome and bacterial metabolism, indicating that the microbiome may indeed serve as an important source of AHR activators in mouse sera.[10,20]

To test this hypothesis and evaluate the necessity of microorganisms to produce these nine examined metabolites, sera were collected from germ-free (GF) mice and subjected to the same targeted metabolomic analysis. Surprisingly, only IPA, 2-oxindole and indole-3-acrylic acid were undetected in the serum of GF mice. The remaining six metabolites were present and their concentrations were comparable between GF and conventional mice maintained on the same diet (Fig. 1, Table 2). In addition to routine surveillance of GF mice (see Methods), the absence of bacteria e was supported by undetectable serum concentrations of hippuric acid (Fig. S1), a known microbial metabolite.[21] Taken together, these data indicate that the profile of circulating metabolites contains fewer Trp metabolites than those based on fecal analysis, and indole-containing compounds that are often ascribed to microbial metabolism (e.g. IAA, I3C, ILA) appear to actually originate from host metabolism.

*Diet alters circulating tryptophan metabolites*

126 Having generated Trp metabolite profiles from sera, we asked if altering the diet would
127 significantly change circulating concentrations. One of two groups of conventional C57BL/6J mice with
128 *ad libitum* access to a standard laboratory chow were switched onto a semi-purified diet (AIN-93G) while
129 the second remained on standard chow. After one week, portal vein blood was collected and analyzed
130 using targeted LC-MS (Fig. 2). The abundance of IPA dropped significantly (39-fold) in mice eating the
131 semi-purified diet with (mean concentration 4.71 ± 1.48 μM and 0.12 ± 0.10 μM for standard and semi-
132 purified diets, respectively). Additionally, the abundance of indole-3-acrylic acid also dropped by 3-fold
133 in the same group of mice (mean concentration 0.31 ± 0.10 μM and 0.10 ± 0.02 μM, respectively). Given
134 that bacteria are necessary for the circulation of IPA and indole-3-acrylic acid (IAcA) in serum (Fig. 1),
135 this result indicates that diet can be used to substantially alter microbial Trp metabolism. Furthermore,
136 except for IAA, the dietary shift had no statistically significant influence Trp metabolite abundances. The
137 difference observed in IAA, while significant, was minor (1.6-fold) compared to that observed in IAcA
138 and IPA. The remaining six targeted metabolites did not significantly differ in abundance. While it is not
139 clear which dietary components influence microbial or host Trp metabolism, the data demonstrate that
140 diet can alter the abundance and distribution of certain circulating Trp metabolites.

141 *Activation of murine AHR by circulating tryptophan metabolites*

142 As we and others have observed, the detected metabolites possess the potential to activate the
143 AHR, so we next assessed the ability for these nine metabolites to produce AHR activity at their relevant
144 serum concentrations.[20,22,23,24] We assessed the potential of each metabolite to induce AHR activity at the
145 mean concentrations measured in mouse sera (Table 3). We treated a murine AHR-dependent reporter
146 assay system (Hepa 1.1) to each of the nine metabolites for 4 h. We also generated a synthetic Trp
147 metabolite pool using the mean concentrations of all nine quantified Trp metabolites to model the
148 potential for the complete Trp metabolite profile to generate AHR activity (Fig. 3). The complete Pool
149 generated a ~7-fold induction ($p < 0.0001$) relative to the vehicle treatment. KYN was the only individual
150 metabolite to produce an increase in luciferase activity (~6-fold induction, $p < 0.0001$) relative to the
151 vehicle treatment. Despite KYN comprising just 12% of the total concentration of the Pool, it generated
152 85% the activity produced by the whole Pool.

153 *Tryptophan metabolite profiling of human sera*

154 Trp metabolites circulating in serum could provide insight into homeostatic activity of AHR, and
155 we wished to know if the observations in mice could be replicated in humans. Previous work by our
156 group identified AHR activators in human feces, yet it is not clear if these activators are also present in
157 circulation. To address this question, we quantified Trp metabolites in human serum. To mitigate the

158 variable effect of diet on metabolite profiles, we analyzed sera samples collected from study participants
159 fed a defined diet.[25] As with the previously discussed mouse sera, targeted metabolomics was performed
160 on serum samples using LC-MS and GC-MS (Table S2). Six Trp metabolites, IPA, IAA, ILA, I3C, KYN,
161 and KA, were detected in human serum at considerably varied concentrations (Fig. 4). Additionally, many
162 metabolites were highly variable across the cohort of trial participants (Table 4). IPA was the second most
163 variable metabolite (CoV: 0.571) and had the greatest range of the metabolites with concentrations
164 detected between 0.443-8.417 μM. KA was the most variable (CoV: 0.763), but least abundant
165 metabolite. The abundances of KA and KYN are positively correlated (Fig. S2, Table S4) which may be
166 explained by a necessary coupling due to the kynurenine pathway, while positive correlations between
167 I3C and all other detected metabolites, except IPA, is unclear. We calculated the sum of all six Trp
168 metabolites across study participants to provide broader insights into individual variability. The wide
169 range of Trp metabolite concentrations (5.52-13.32 μM) added together further illustrates the vast
170 individual variation in Trp metabolite profiles, which emphasizes the need for a substantial sample size to
171 establish a pattern.

172 *Circulating tryptophan metabolites generate human AHR activity*

173 Given species differences in the AHR ligand binding pocket, metabolism, and the composition of
174 the gut microbiome, it was unclear if the previously discussed mouse AHR-based activity assays would
175 produce similar results in a human model. To address this question, we used human cell culture models to
176 examine AHR activation potential of the Trp metabolites detected in human sera. We evaluated the
177 potential of the six detected metabolites to activate the AHR using the mean concentrations of each
178 metabolite (IPA: 2.95 μM; IAA: 2.22 μM; ILA: 1.92 μM; I3C: 0.11 μM; KYN: 1.98 μM; KA: 0.08 μM)
179 to assess potential physiologic activation. We also generated a mixture comprised of the mean
180 concentrations of all six detected metabolites, referred to as the "Pool" (Fig. 5a), to provide a more
181 biologically relevant model for assessing systemic AHR activity. AHR activity was measured using a
182 human AHR-dependent luciferase reporter assay system (HepG2 40/6) following 4 h exposure to
183 metabolite treatments (Fig. 5b). The Pool, KA, KYN, and IAA generated significant activity relative to
184 the vehicle treatment, with a 2 or 3-fold induction. Interestingly, the sum of individual metabolites
185 generates far more activity than is seen when cells are exposed to the metabolite Pool, suggesting an
186 antagonistic effect. Furthermore, while the removal of a single metabolite from the Pool does not
187 significantly alter activity (Fig. 5c) relative to the complete Pool; the removal of two dominant
188 metabolites, KA and KYN, does significantly diminish activity. Strikingly, the presence of IAA in this
189 treatment does not protect against a loss of activity. This further supports an antagonistic relationship

190 between the six Trp metabolites and AHR activity. However, it should be noted that the six Trp
191 metabolites were determined to not have cytotoxic effects (Fig. S3).

192 Guided by the potential for an agonist/antagonist dynamic observed in the Trp mixture, we
193 predicted that the total concentration of Trp metabolites would affect the degree of AHR activity. To test
194 this hypothesis, we replicated the Trp metabolite profile of individual trial participants with the three
195 greatest (**1.** 13.32 μM, **2.** 12.9 μM, **3.** 12.53 μM) and lowest total concentrations (**4**. 6.34 μM, **5.** 6.01 μM,
196 **6.** 5.52 μM) of examined metabolites to assess induction of the prototypical AHR target gene, *CYP1A1*
197 (Fig. 5d, e). The three profiles with the highest and lowest total metabolite abundances were grouped
198 accordingly and induced a statistically significant difference in AHR activity. Considering the substantial
199 variation in total metabolite abundance, the surprisingly small difference in AHR activation may suggest
200 that multiple factors may affect AHR activation, but metabolite affinity for the AHR and potential
201 competitive binding dynamics may be a leading factor. We sought to further investigate this question by
202 examining the affinities of AHR for each of the six Trp metabolites.

203 *Tryptophan metabolites directly interact with the AHR*

204 We assessed the affinity of the six Trp metabolites for the AHR by exposing HepG2 40/6 cells to
205 a range of concentrations for each metabolite (Fig. 6). While the luciferase reporter system is not
206 necessarily reflective of a realistic physiological response, it does allow for relative comparisons of
207 activation potential. $EC_{25}$ values demonstrate that KA is the most potent of the six metabolites with an
208 $EC_{25}$ value of 40.8 nM. The $EC_{25}$ values of KYN (1.76 μM) and IAA (2.92 μM) rank these metabolites
209 the second and third most potent metabolites, respectively. Further, the $EC_{25}$ values for KA, KYN and
210 IAA fit within the range of concentrations that were detected in human sera, suggesting that these values
211 are physiologically possible. Meanwhile, the $EC_{25}$ values for IPA (19.2 μM), ILA (25.9 μM), and I3C
212 (109.8 μM) are far above the range of concentrations that were detected in human sera. These data imply
213 that these metabolites are unlikely to directly generate homeostatic AHR activity at typical physiological
214 concentrations.

215 Although we observed AHR activity in response to treatments with the examined metabolites, it
216 was unclear if these compounds are truly AHR ligands which directly interact with the ligand binding
217 pocket of the receptor. We sought to address this question with a cell-based AHR ligand competition
218 binding assay.[26] This is a modification of the gold standard of AHR ligand classification that relies on
219 displacement of a high affinity radioactive ligand from the ligand binding site, but incorporates the
220 complexity of cellular biology, including metabolite uptake, efflux, and protein binding. This assay,
221 however, is limited by the potential of further metabolism of the Trp metabolites prior to binding with the

222 AHR. An assay length of 30 min was chosen to mitigate further metabolism. The result of the assay
223 showed that all six metabolites can compete with the photoaffinity ligand (PAL) at 10 μM concentrations
224 (Fig. 7). Importantly, the Pool also demonstrated that it competed with the PAL for binding to the AHR
225 ligand binding site. Furthermore, ligand docking simulation using Autodock Vina provides additional
226 evidence that the Trp metabolites are indeed interacting with the ligand binding site of AHR (see
227 Methods, and Fig. S4). Computational docking analysis of the six Trp metabolites (Fig. 8) revealed that
228 each compound can bind the PAS B domain of the AHR with low micromolar to high nanomolar affinity
229 in the same primary ligand binding pocket, defined by indirubin.[27] ILA (0.81 nM) and KA (0.95 nM)
230 demonstrated the highest computational affinity for the AHR (Table S5), but this did not correlate
231 completely with AHR ligand binding assay data obtained in HepG2 40/6 cells (Fig. 5). However, all six
232 Trp metabolites bound the AHR in the same hydrophobic groove as indirubin (see Fig. S5), with several
233 forming unique electrostatic interactions with Fα helical residue Ser-336, which was not a well-defined
234 element of the primary indirubin binding pocket.[27] The formation of this unique contact with helix Fα
235 stabilized the recognition of the smaller, Trp-related ligands, even in the absence of hydrophobic contacts
236 with Leu-353, the important Gβ-sheet residue that appears to define the opening of the primary ligand
237 binding pocket. Although data from our cell-based assays would imply differences in binding affinity,
238 simulated binding energies and dissociation constants are similar across the six metabolites. This is not
239 surprising given the structural similarities between these compounds. Nevertheless, the combination of
240 modeling and biological activity data clearly supports the concept that the Trp metabolites are direct AHR
241 ligands.

242 *Trp metabolites activate AHR in disease models*

243 We wanted to assess the influence of detected Trp metabolites on models for diseases that have
244 been linked to AHR to provide insight into the physiological relevance of the Trp Pool metabolites. First,
245 we used Caco2 cells to model the impact of Trp metabolites on a cancer model. Previous studies have
246 primarily investigated the role of individual metabolites in isolation, yet the data presented here indicate
247 that many AHR activators likely circulate throughout the body together. Caco2 cells were exposed to
248 individual metabolites at mean serum concentrations, the complete Trp metabolite Pool, and variations of
249 the Pool (excluding individual metabolites) for 4 h prior to RNA isolation. Quantification of AHR target
250 gene, *CYP1A1,* transcripts by qPCR established that only KA, KYN and the Pool generate AHR activity
251 significantly different from the vehicle treatment. Moreover, removing a single metabolite from the Pool
252 does not alter AHR activity, yet removing the two most potent AHR activators, KA and KYN, from the
253 Pool substantially decreases AHR activity.

254 Second, we investigated the impact of the Trp metabolite Pool by exposing a primary cell line of
255 fibroblast-like synoviocytes (FLS) collected from healthy patients to the Trp metabolite Pool. FLS are
256 used as a model in the study of rheumatoid arthritis, an AHR-linked disease, and we have previously
257 shown a synergistic effect on the expression of inflammatory mediators when cells are exposed to IL1β
258 and AHR activators.[4] FLS were co-treated with 0.5 ng/ml IL1β and a saturating dose of TCDD (10 nM)
259 or the Trp Pool for 24 h prior to RNA isolation, to quantify expression of inflammatory mediator genes,
260 *IL1B* and *EREG.* Additionally, FLS were pre-treated with 500 nM GNF351, an AHR antagonist, to
261 demonstrate the necessity of the AHR in synergistic expression of these genes. The IL1β and Trp Pool co-
262 treatment generated a comparable 185-fold induction in *IL1B* expression to the 192-fold increased
263 produced by the IL1β and TCDD. Expression of *EREG* significantly increased with both co-treatments
264 compared to IL1β alone, however, the IL1β and TCDD co-treatment produced a greater 27-fold induction
265 in *EREG* compared to a 17-fold change generated by IL1β and Trp Pool.

266 **Discussion**

267 Previous work from our group and others have quantified the abundance of potential AHR-
268 activating Trp metabolites from feces and cecal contents, however, it is not clear whether these represent
269 concentrations to which the intestinal epithelial lining is exposed or circulate in the bloodstream. The
270 extraction of water from waste material in the colon specifically concentrates hydrophobic metabolites,
271 therefore, the concentration of many nonpolar metabolites will likely be higher in this context than the
272 concentrations that intestinal epithelial cell are exposed to. Further, the exposure of intestinal epithelial
273 cells to metabolites will not be comparable to the cell exposure to these metabolites throughout the body.
274 Transport, metabolism, and compound clearance represent possible mechanisms that would reduce
275 circulating metabolite concentrations. To address these concerns, we chose to quantitatively profile
276 circulating metabolites to provide the foundation for investigating systemic or basal AHR activity. We
277 then focused on quantifying Trp metabolites that we previously demonstrated are AHR activators to
278 provide a measurable metabolite profile.[20] The promiscuity of the AHR ligand binding pocket means that
279 hitherto uncharacterized activators will inevitably be discovered and further complicate the dynamics
280 presented here. Nonetheless, the data provide a basis for more complex models that are necessary to
281 understand the nuances of physiological AHR activation.

282 To investigate systemic AHR activity, we used mouse models to test specific hypotheses,
283 however, we chose to minimize our reliance on mouse models because they do not accurately portray the
284 human AHR mediated activation by Trp metabolites. Similar to the human AHR, mouse AHR activity is
285 significantly induced in response to the mouse Trp Pool however, unlike human AHR, it is primarily

286 responsive to a single metabolite, KYN. This suggests the known evolutionary changes in the human
287 AHR binding pocket allow for greater recognition of KA and indole-containing compounds.[28,29]
288 Furthermore, the greater relative affinity of the endogenously produced KA for human AHR, may
289 represent an evolutionary change allowing for greater control over AHR activity. The differences in
290 receptor affinity highlight the importance of appreciating the structural differences between mouse and
291 human AHR and limit the use of mouse models to explore this topic. Beyond AHR structural differences,
292 the mouse microbiome is markedly different from human and this also influences the products of
293 metabolism.[30,31] Moreover, our lab has previously shown that, despite high expression of AHR in the
294 colon of conventional C57BL/6J mice and high concentrations of Trp metabolites, there is little to no
295 *Cyp1a1* expression.[32] Taken together, these observations imply that mouse AHR is relatively insensitive
296 to changes in Trp metabolites in the context of detectable serum concentrations. These distinctions
297 present the biological challenges limiting the accuracy of mouse models to provide specific insights into
298 physiologically induced human AHR activity.

299 Demonstrating AHR activation in response to a compound is often thought to be sufficient
300 evidence to assume that a specified compound is an AHR ligand. However, some compounds, such as
301 serotonin and butyrate, generate activity without directly interacting with the AHR.[33] The gold standard
302 for identifying an AHR ligand is a radioligand competition binding assay to demonstrate compound
303 interaction with the AHR binding pocket. We employed a cell-based radioligand competition assay to
304 evaluate all six metabolites for their ability to compete with the high affinity PAL. Compared to an *in*
305 *vitro* radioactive ligand competition assay, the cell-based method inherently considers transport, efflux,
306 and other cellular functions that may interfere with a compound from binding to the receptor, providing a
307 biologically relevant context. However, this assay is limited in that it allows further metabolism of
308 treatment compounds. We sought to address this limitation with both a short treatment time and with
309 supporting evidence from additional methods. We generated dose response curves for the six Trp
310 metabolites to characterize AHR responsiveness to each metabolite. Lastly, we used ligand binding
311 simulations to model the binding interaction of each metabolite with the AHR ligand binding pocket.
312 Simulated ligand docking similarly posed all six Trp metabolites and in the same plane as the well-known
313 AHR ligand, indirubin, supporting that the six Trp metabolites are indeed AHR ligands. Despite similar
314 binding energies and dissociation constants, suggesting little theoretical difference in activation potential
315 between the six metabolites, this is not reflected in cell-based activity assays. This discrepancy
316 underscores the importance of studying potential AHR ligands in a biological context that considers
317 metabolism and other cellular functions, because theoretical binding potential does not guarantee actual
318 receptor activation. Nonetheless, docking simulation analysis upholds the hypothesis that these six Trp

319 metabolites can act as AHR ligands. Support from these three methods provide evidence that the
320 examined metabolites are directly interacting with the AHR and are, indeed, genuine AHR ligands.

321 Research into Trp metabolism across taxonomic kingdoms has created confusion about the source
322 of specific metabolites, especially indole-containing compounds. It is often assumed that the production
323 of a Trp metabolite at relatively high concentrations by gut bacteria means that bacteria are the only or
324 primary source for the host. Our data suggest this is not always the case. IAA is often depicted as a
325 product of microbial metabolism, despite long recognition as an essential plant growth hormone. Further,
326 the data presented here from GF mice indicate that IAA is produced independent of bacteria and is
327 supported by previous publications reporting human enzymes that can produce IAA.[34] This confusion is
328 seen in previous reports of ILA or I3C as well, while our samples collected from GF mice indicate that
329 these compounds circulate in a host without the presence of bacteria.[35,36] Yet, we are confident that 2-
330 oxindole, indole-3-acrylic acid, and IPA are microbial products and represent metabolic products of the
331 microbiome that can circulate in a host. Not only are these below the limit of detection in GF mice, but
332 the published literature supports microbial production of these compounds.[14,37,38] Notably, the presence of
333 indole-3-acrylic acid and IPA are likely partially coupled because indole-3-acrylic acid is a known
334 intermediate in a microbial metabolic pathway that terminates in IPA.[14] Given that IPA circulates in
335 humans (HMDB ID: HMDB002302), it is possible that microbial metabolism may contribute to the
336 circulating pool of IAA, ILA, or I3C, but further work is needed to understand this dynamic.[39] Strikingly,
337 metabolites found in high concentration in feces, e.g. indole, were undetectable in serum.[20] This disparity
338 highlights the value of direct metabolite analysis of tissues or biofluids in addition to proxy
339 measurements, like feces, when measuring the concentrations of AHR ligands.

340 Study of AHR-activating Trp metabolites often leads to suggestions that changing the
341 composition of the gut microbiome may alter AHR activity, but this proposition may be more challenging
342 than it appears at first glance. Bacterial colonization of the gastrointestinal tract depends on multiple
343 competitive dynamics, thus, producing an environment that is difficult to control.[40,41] In contrast,
344 designing an individual's diet to achieve a clinical outcome is long standing practice. We wanted to
345 demonstrate the possibility that diet affecting Trp metabolism could alter potential clinical outcomes. We
346 observed that placing mice on a semi-purified diet does change Trp metabolite abundances. The two diets
347 produced differences in serum concentration of three of the nine metabolites, two of which are
348 microbially produced indole-3-acrylic acid and IPA. This suggests that the semi-purified diet can suppress
349 specific microbial metabolites, and these compounds appear particularly sensitive to change. The fact that
350 2-oxindole concentration showed no change suggests that altering diet may only affect specific microbial
351 pathways. Similarly, the significant reduction in IAA in mice fed a semi-purified diet indicates that host

352 Trp metabolism may be altered by diet**,** but changes in microbial production of IAA remains a possibility
353 as well. Although the removal of phytochemicals from the semi-purified diet may offer a possible
354 explanation, the cause of these metabolic changes is unclear and requires additional research. Whereas the
355 abundance of serum Trp metabolites can be modulated by non-pharmacological means, the influence this
356 would have on AHR activity is difficult to predict at this point.

357 We used a human AHR-dependent reporter system (HepG2 40/6) to specifically assess the
358 potential for each metabolite to activate the AHR at the identified mean concentration. This system is
359 limited in its capacity to dissect biological responses but is adequate for relative comparisons and
360 provides broad insights into how these metabolites activate AHR. This assay indicates some indoles, IPA,
361 I3C, and ILA, are not significantly contributing to AHR activation potential at physiological
362 concentrations. Measuring *CYP1A1* expression induced by the Trp Pool supports the notion that not all
363 Trp metabolites generate activity at homeostatic concentrations. Notably, KA is a downstream product of
364 KYN metabolism, coupling these two important AHR agonists.[42] However, this may be dependent on the
365 observed cell type, given that the expression profile of kynurenine pathway enzymes will differ between
366 cell types and physiological conditions. Determining the role of this pathway in producing AHR ligands
367 may be critical to understanding systemic AHR activity and warrants further investigation. The significant
368 positive correlation between KA and KYN presents another challenge of disentangling individual
369 metabolite activity due to possible continued metabolism within cells. Further, relative to the commonly
370 studied industrial pollutants (e.g., polychlorinated biphenyls) or dietary compounds (e.g., indole-3-
371 carbinol) that act as AHR ligands, these six Trp metabolites are low affinity ligands, and, as such, may
372 serve as a regulatory mechanism. Saturation of the AHR by Trp metabolites could prevent or hinder
373 AHR-mediated response to environmental toxins or immune challenge. Therefore, the low affinity of
374 these metabolites to the AHR may maintain basal homeostatic AHR activity while permitting a strong
375 response to environmental challenges.

376 The role of AHR in cancer is well documented, so we investigated whether the six identified
377 metabolites would generate AHR activity in a colorectal cancer model (Caco2).[43,44,45,46] Our results
378 indicate that the examined metabolites do activate AHR in Caco2 cells leading to a change in *CYP1A1*
379 expression**.** In some types of cancers this may provide a protective effect for a tumor by contributing to
380 immune system suppression. Host genetics, metabolic dysregulation, or diet may exacerbate this effect,
381 producing greater AHR activation and potentially influencing clinical outcomes in patients. A comparison
382 of activity produced in HepG2 and Caco2 cells also suggests that there may be cell type specific
383 responses to these metabolites.

384 To further examine this possibility, we exposed FLS to the Pool of Trp metabolites. Previously,
385 we have shown that AHR activation acts combinatorially with IL1β to induce inflammatory mediators,
386 including *IL1B* and *EREG*, in FLS.[4,47] Co-treatment with IL1β and the Trp Pool led to increased
387 expression of both biomarkers associated with worsened clinical outcomes. Regular or constant exposure
388 to the examined Trp metabolites may provide insights into understanding the development of rheumatoid
389 arthritis, an important AHR-linked disease.[48,49] The presence of Trp metabolites that activate the AHR
390 leads to the priming of the IL1B promoter facilitating enhanced transcriptional activity in the presence of
391 a pro-inflammatory response, such as NFkB activation.[4] This model represents potential future directions
392 for AHR-linked disease research. Indeed, Trp metabolites may exacerbate a disease state, or dysregulation
393 of Trp metabolism may lead to the development of disease.

394 Comparing the results of treating cells with the Trp Pool and individual metabolites demonstrate
395 the difficulty in interpreting the impact of a mixture of Trp metabolites on AHR activity. Our experiments
396 factor in the many cellular processes (metabolism, transport, and competition for receptor binding) that
397 likely act upon the metabolites without providing insight into how these processes are altering activity.
398 Transporters for KA and KYN have been identified, but the uptake mechanisms of the four indole-
399 containing compounds have been poorly studied.[50,51] The differential expression of transporter proteins in
400 distinct cell types may result in varied intracellular concentrations that affect AHR activity. Additionally,
401 the half-life of these metabolites is likely determined by further metabolism, host clearance and
402 compound stability. Furthermore, the rate of production of these metabolites may be affected by genetics,
403 microbiome composition, and environmental exposures that are unique to each host. This is further
404 complicated by the possibility that metabolite concentrations vary throughout the day as meals are
405 consumed and Trp availability changes.[52,53] These processes complicate interpretation and, to further
406 understand homeostatic AHR activity there is a need to develop more accurate models. Additionally, the
407 promiscuity of the AHR binding pocket and the diversity of metabolic compounds imply the inevitability
408 of identifying additional AHR activators. Novel activators should not be assumed to have physiological
409 relevance when assessed in isolation, considering the already significant abundances of endogenous
410 ligands. These studies revealed that circulating Trp metabolites capable of mediating AHR are produced
411 by the host and thus do not originate from the gut microbiome, in contrast to the prevailing dogma in the
412 literature. However, it is possible that homeostasis of circulating concentrations is modulated through
413 increased excretion and absorption of Trp metabolites. Future studies should turn to assessing the
414 important host metabolic pathways that result in the production of Trp metabolites that may regulate basal
415 AHR activity. Finally, it is possible that individual differences in Trp metabolite production capable of
416 mediating AHR activation may play a role in chronic diseases.

417 **Methods**

418 *Chemicals and reagents*

419 All chemicals, unless otherwise stated, were purchased from Sigma (St. Louis, MO, USA).
420 Indole-3-propionic acid and indole-3-lactic acid were from Alfa Aesar (Heysham, UK). AIN93G was
421 purchased from Dyets, Inc., (Bethlehem, PA). LC/MS grade solvents including, methanol and
422 acetonitrile, were purchased from Fisher Scientific (Hampton, NH, USA). The stock solutions of all
423 reference standards for liquid chromatography were prepared in 10% acetonitrile (v/v) containing 1 μM
424 chloropropamide (internal standard). The mixed standard solutions were gradually diluted with 10%
425 acetonitrile (v/v) containing 1 μM chloropropamide for generating the calibration curves. The AHR
426 photoaffinity ligand 2-azido-3-[$^{125}$I]iodo-7,8-dibromodibenzo-*p*-dioxin (PAL) was synthesized as
427 described.[26]

428 *Human serum samples*

429 Serum samples were collected as described.[25] Briefly, 40 selected trial participants were fed a
430 standardized isocaloric diet (representative of an average American macronutrient intake) to achieve
431 weight maintenance for four weeks At the end of the diet period and following a 12 h fast (and avoidance
432 of over-the-counter medication), serum was collected, stored at -80C and subjected to analysis by
433 metabolomics.

434 *Serum sample collection from conventional and GF mice*

435 C57BL/6J wild type mice were originally purchased from Jackson Laboratories (Bar Harbor, ME).
436 Germ-free (GF) C57BL/6J mouse experiments were conducted at the AAALAC-accredited Animal
437 Resource Center at Montana State University. GF mice were reared in standard, sterile (autoclaved cages
438 containing sterile (autoclaved) bedding inside a hermetically sealed isolator with HEPA-filtered airflow
439 and maintained on sterile (autoclaved) water and sterile (autoclaved) food (LabDiet® 5013, Land O'Lakes)
440 ad libitum. Surveillance of the GF condition was done with standard microbial cultivation and molecular
441 biology techniques[46]. Briefly, liquid 'bug' traps comprised of a mixture of food and drinking water were
442 left open to the air inside of isolators and observed daily for signs of microbial growth (i.e., turbidity). Stool
443 samples from mice were monitored prior to and throughout experiments for signs of growth on rich media
444 under anaerobic and regular atmospheric conditions (Mueller–Hinton broth and agar plates). Bulk DNA
445 was also extracted from stool samples (DNeasy PowerSoil Pro DNA isolation kit, Qiagen, Hilden,
446 Germany) to serve as a template for PCR using primers targeting the bacterial 16S rRNA encoding gene.
447 Evidence supporting the GF condition was the combination of a lack of observable microbial growth and

448 16S amplification by PCR. All GF mouse procedures were approved by Montana State University's
449 Institutional Animal Care and Use Committee. A subset of mice was switched onto a semi-purified AIN
450 93G from a chow diet for seven days prior to serum collection. Serum samples from conventional and GF
451 mice were collected and stored at -80°C. All animal protocols were reviewed and approved by the Animal
452 Care and Use Committee of Penn State University.

453 *Quantification of tryptophan metabolites in mouse and human serum*

454 Tryptophan metabolites were quantified by LC/MS based on established method with some
455 modifications.[20] Briefly, thawed on ice, 25 µL of serum sample, was mixed with 100 µL extraction solvent
456 of ice-cold methanol containing indole-3-acetic acid-d4 and kynurenic acid-d5. Mixture was vortexed and
457 incubated at −20 °C for 30 min. Following centrifugation at 12,000 × g for 15 min at 4°C, 90 µL supernatant
458 was collected, evaporated to dryness (Thermo Scientific, Waltham, MA) and dissolved in 45 µL of 10%
459 acetonitrile containing 1 µM chlorpropamide. After centrifugation at 12,000 × g for 15 min at 4°C,
460 supernatants were transferred to autosampler vials for LC-MS analysis. Quantitative analysis was
461 performed by reverse phase UHPLC using a Prominence 20 UFLCXR system (Shimadzu, C57Columbia,
462 MD) with a Waters (Milford, MA) BEH C18 column (2.1 × 100 mm, 1.7 µm particle size) maintained at 55
463 °C and a 20 min aqueous acetonitrile gradient, at a flow rate of 250 µL/min. Solvent A was water with 0.1%
464 formic acid and Solvent B was acetonitrile with 0.1% formic acid. The initial conditions were 97% A and
465 3 % B, increasing to 45% B at 10 min, 75% B at 12 min where it was held at 75% B until 17.5 min before
466 returning to the initial conditions. The eluate was delivered into a 5600 (QTOF) TripleTOF using a
467 Duospray™ ion source (SCIEX, Framingham, MA). Importantly, data was collected in multiple reaction
468 monitoring (MRM) mode, and the mass accuracy was calibrated by calibration solution purchased from
469 SCIEX (Framingham, MA).

470 *Cell culture conditions*

471 Caco2 cells were grown in αMEM (Sigma-Aldrich, St. Louis, MO, USA) supplemented with
472 20% (v/v) FBS (GeminiBio Products, Sacramento, CA, USA) and 1% (v/v) penicillin-streptomycin. Cells
473 were seeded into 6-well plates and grown seven days post-confluence prior to treatments to allow for
474 differentiation to occur. HepG2 40/6 cells were generated as previously described, and Hepa 1.1 cells
475 were a gift from Dr. Michael S. Denison.[54] HepG2 40/6 and Hepa 1.1 cells were grown in αMEM
476 supplemented with 8% (v/v) FBS and 1% (v/v) penicillin-streptomycin, and seeded into 12-well plates 24
477 h prior to treatment. Primary human fibroblast-like synoviocytes from healthy patients were purchased
478 from and maintained in Synoviocyte Growth Medium (Cell Application, Inc., San Diego, CA). HN30
479 cells were obtained and cultured as previously described.[55]

480 *Cell-based luciferase reporter assay*

481 Prior to treatment, culture media was aspirated from HepG2 40/6 cells, which were thoroughly
482 washed with PBS, then αMEM supplemented with 5 mg/mL BSA, 100 U/mL penicillin, and 100 μg/mL
483 streptomycin was added to each well. Cells were treated with TCDD or tryptophan metabolites in DMSO
484 (0.1% final concentration in cell culture). After 4 h incubation, cells were thoroughly washed with PBS,
485 lyzed using 100 μL of lysis buffer (25 mM Tris-phosphate pH 7.8, 2 mM dithiothreitol, 2 mM 1,2-
486 diaminocyclohexane-N,N,N',N'-tetraacetic acid, 10% (v/v) glycerol, 1% (v/v) Triton X-100), and stored
487 at -80°C until analysis. Luciferase activity was measured using a TD-20e luminometer (Turner
488 Biosystems Inc.) and luciferase substrate (Promega, Madison, WI, USA), per the manufacturer's
489 instructions.

490 *RNA isolation and RT-qPCR*

491 Prior to treatment, culture media was aspirated from HepG2 40/6 and Caco2 cells, which were
492 thoroughly washed with PBS, then αMEM supplemented with 5 mg/mL BSA and 1% (v/v)
493 penicillin/streptomycin was added to each well. HepG2 40/6 and Caco2 cells were exposed to Trp
494 metabolites for 4 h. FLS were treated in Synoviocyte Growth Medium for 24 h. RNA was isolated from
495 cells treated with tryptophan metabolites for quantitative real-time PCR (RT-PCR), performed as
496 previously described.[56] Primers are listed in supplementary table 3.

497 *Cell-based AHR ligand competition binding assay*

498 HN30 cells were seeded into 12-well plates at 50,000 cells/well and cultured for 48 h in
499 DMEM/F12 medium (Sigma) with 10% fetal bovine serum (Gemini BioProducts), 25 mM HEPES (pH
500 7.4), 100 U/mL penicillin, and 100 μg/mL streptomycin. The medium was removed, cells washed with
501 Dulbecco's phosphate buffered saline (PBS, Sigma). Cells in each well were incubated in 500 μl Hanks
502 balanced salt solution (Gibco), 25 mM HEPES pH 7.4, and 5 mg/ml bovine serum albumin for 30 min.
503 Next, cells were treated with AHR ligands, followed by the addition of 2 pmoles of PAL, cells were
504 placed back in the incubator for 30 min. Medium was removed, 500 μL of PBS was added, and cells were
505 exposed to UV light at >302 nm, at a distance of 8 cm, for 4 min using two 15-W UV lamps (Dazor Mfg.
506 Corp. St. Louis, MO). The PBS was removed and cells lyzed in 100 μL of 25 mM MOPS, 2 mM EDTA,
507 0.02% sodium azide, 10% glycerol and 1% NP40. Lysates were transferred to microfuge tubes and
508 centrifuged at 18,000 x g for 20 min. Supernatants were subjected to tricine SDS-PAGE, and
509 subsequently transferred to PVDF membrane. The radioactive AHR band was visualized with X-ray film
510 and densitometry analysis was performed with ImageJ v. 1.53t[57].

511 *Computational docking analysis*

512 The ligand binding affinity of 6 Trp-metabolites, including kynurenine, kynurenic acid, indole-3-
513 carboylate, indole-3-acetate, indole-3-lactate, and indole-3-propionate, were investigated using Autodock
514 Tools 4 and Autodock Vina, and a cryo-EM structure-based computational model of the human AHR
515 PAS B domain.[27,58,59,60] Docking models for each of the 6 Trp-related ligands were obtained from the
516 PubChem database and optimized for docking using Autodock Tools.[61] Autodock Vina was run using
517 standard settings, and Grid box parameters generated in Autogrid for a 30-60 Å3 docking grid centered on
518 the indirubin binding site, as described previously.[62] Docking results were analyzed using Autodock
519 Tools 4 and The PyMOL Molecular Graphics System, Version 2.52 Schrödinger, LLC.[63] Active site
520 cavity size was calculated using Caver 3.0.[64] AHR docking model quality was validated using an
521 Autodock Vina re-docking analysis with the structural coordinates of indirubin. RMSD re-docking error
522 was calculated using the program LigRMSD.[65]

523 *Statistical analysis*

524 All data were compared using one-way ANOVA in GraphPad Prism v. 9.4.1 (GraphPad
525 Software, San Diego, CA, USA) to determine statistical significance between groups. Values of $p < 0.05$
526 were considered statistically significant (*: $p<0.05$; **: $p<0.01$; ***: $p<0.001$; ****: $p<0.0001$).

527 **Acknowledgements**

528 We thank Marcia H. Perdew for critically reviewing the manuscript.

529 **Competing financial interest declaration**

530 The authors declare no conflict of interest.

531 **Disclosure of Potential Conflicts of Interests**

532 No potential conflicts of interest were disclosed.

533 **Funding**

534 This research was supported by National Institutes of Health under grants National Institutes of Health
535 Grants ES028244 (GHP), ES028288 (ADP) and T32DK120509 (EWM). In addition, this research was
536 supported by the USDA National Institute of Food and Federal Appropriations under Project PEN04607
537 and Accession number 1009993. The human research was supported, in part, by the McCormick Science
538 Institute and the National Center for Advancing Translational Sciences, National Institutes of Health,
539 through Grant UL1 TR002014 (PMK-E).

540 **References**


541 1. Beischlag TV, Morales JL, Hollingshead BD, Perdew GH. 2008. The Aryl Hydrocarbon
542 Receptor Complex and the Control of Gene Expression. Crit Rev Eukar Gene Expr.
543 18(3):207–250. doi:10.1615/CritRevEukarGeneExpr.v18.i3.20

544 2. Metidji A, Omenetti S, Crotta S, Li Y, Nye E, Ross E, Li V, Maradana MR, Schiering C,
545 Stockinger B. 2018. The Environmental Sensor AHR Protects from Inflammatory Damage by
546 Maintaining Intestinal Stem Cell Homeostasis and Barrier Integrity. Immunity. 49(2):353-
547 362.e5. doi:10.1016/j.immuni.2018.07.010

548 3. Gutiérrez-Vázquez C, Quintana FJ. 2018. Regulation of the Immune Response by the Aryl
549 Hydrocarbon Receptor. Immunity. 48(1):19–33. doi:10.1016/j.immuni.2017.12.012

550 4. Lahoti TS, John K, Hughes JM, Kusnadi A, Murray IA, Krishnegowda G, Amin S, Perdew
551 GH. 2013. Aryl hydrocarbon receptor antagonism mitigates cytokine-mediated inflammatory
552 signalling in primary human fibroblast-like synoviocytes. Annals of the Rheumatic Diseases.
553 72(10):1708–1716. doi:10.1136/annrheumdis-2012-202639

554 5. Quintana FJ, Basso AS, Iglesias AH, Korn T, Farez MF, Bettelli E, Caccamo M, Oukka M,
555 Weiner HL. 2008. Control of Treg and TH17 cell differentiation by the aryl hydrocarbon
556 receptor. Nature. 453(7191):65–71. doi:10.1038/nature06880

557 6. Nyström EEL, Martinez-Abad B, Arike L, Birchenough GMH, Nonnecke EB, Castillo PA,
558 Svensson F, Bevins CL, Hansson GC, Johansson MEV. 2021. An intercrypt subpopulation of
559 goblet cells is essential for colonic mucus barrier function. Science. 372(6539).
560 doi:10.1126/science.abb1590

561 7. Li J, Bhattacharya S, Zhou J, Phadnis-Moghe AS, Crawford RB, Kaminski NE. 2017. Aryl
562 Hydrocarbon Receptor Activation Suppresses EBF1 and PAX5 and Impairs Human B
563 Lymphopoiesis. The Journal of Immunology. 199(10):3504–3515.
564 doi:10.4049/jimmunol.1700289

565 8. Stockinger B, Shah K, Wincent E. 2021. AHR in the intestinal microenvironment:
566 safeguarding barrier function. Nat Rev Gastroenterol Hepatol. 18(8):559–570.
567 doi:10.1038/s41575-021-00430-8

568 9. Oweira H, Lahdou I, Mehrle S, Khajeh E, Nikbakhsh R, Ghamarnejad O, Terness P,
569 Reißfelder C, Sadeghi M, Ramouz A. 2022. Kynurenine Is the Main Metabolite of
570 Tryptophan Degradation by Tryptophan 2,3-Dioxygenase in HepG2 Tumor Cells. Journal of
571 Clinical Medicine. 11(16):4794. doi:10.3390/jcm11164794

572 10. Agus A, Planchais J, Sokol H. 2018. Gut Microbiota Regulation of Tryptophan
573 Metabolism in Health and Disease. Cell Host & Microbe. 23(6):716–724.
574 doi:10.1016/j.chom.2018.05.003

575 11. DiNatale BC, Murray IA, Schroeder JC, Flaveny CA, Lahoti TS, Laurenzana EM,
576 Omiecinski CJ, Perdew GH. 2010. Kynurenic Acid Is a Potent Endogenous Aryl Hydrocarbon

577 Receptor Ligand that Synergistically Induces Interleukin-6 in the Presence of Inflammatory
578 Signaling. Toxicological Sciences. 115(1):89–97. doi:10.1093/toxsci/kfq024

579 12. Manzella CR, Ackerman M, Singhal M, Ticho AL, Ceh J, Alrefai WA, Saksena S,
580 Dudeja PK, Gill RK. 2020. Serotonin Modulates AhR Activation by Interfering with
581 CYP1A1-Mediated Clearance of AhR Ligands | Cell Physiol Biochem. Cellular Physiology &
582 Biochemistry. 54(1):126–141.

583 13. Lee J-H, Lee J. 2010. Indole as an intercellular signal in microbial communities. FEMS
584 Microbiology Reviews. 34(4):426–444. doi:https://doi.org/10.1111/j.1574-6976.2009.00204.x

585 14. Dodd D, Spitzer MH, Van Treuren W, Merrill BD, Hryckowian AJ, Higginbottom SK,
586 Le A, Cowan TM, Nolan GP, Fischbach MA, Sonnenburg JL. 2017. A gut bacterial pathway
587 metabolizes aromatic amino acids into nine circulating metabolites. Nature. 551(7682):648–
588 652. doi:10.1038/nature24661

589 15. Negatu DA, Yamada Y, Xi Y, Go ML, Zimmerman M, Ganapathy U, Dartois V,
590 Gengenbacher M, Dick T. 2019. Gut Microbiota Metabolite Indole Propionic Acid Targets
591 Tryptophan Biosynthesis in *Mycobacterium tuberculosis*. mBio. 10(2):e02781-18.
592 doi:10.1128/mBio.02781-18

593 16. Menni C, Hernandez MM, Vital M, Mohney RP, Spector TD, Valdes AM. 2019.
594 Circulating levels of the anti-oxidant indoleproprionic acid are associated with higher gut
595 microbiome diversity. Gut Microbes. 10(6):688–695. doi:10.1080/19490976.2019.1586038

596 17. Aoki R, Aoki-Yoshida A, Suzuki C, Takayama Y. 2018. Indole-3-Pyruvic Acid, an Aryl
597 Hydrocarbon Receptor Activator, Suppresses Experimental Colitis in Mice. J Immunol.
598 201(12):3683–3693. doi:10.4049/jimmunol.1701734

599 18. Ludwig-Müller J. 2011. Auxin conjugates: their role for plant development and in the
600 evolution of land plants. Journal of Experimental Botany. 62(6):1757–1773.
601 doi:10.1093/jxb/erq412

602 19. Rampey RA, LeClere S, Kowalczyk M, Ljung K, Sandberg G, Bartel B. 2004. A Family
603 of Auxin-Conjugate Hydrolases That Contributes to Free Indole-3-Acetic Acid Levels during
604 Arabidopsis Germination. Plant Physiology. 135(2):978–988. doi:10.1104/pp.104.039677

605 20. Dong F, Hao F, Murray IA, Smith PB, Koo I, Tindall AM, Kris-Etherton PM, Gowda K,
606 Amin SG, Patterson AD, Perdew GH. 2020. Intestinal microbiota-derived tryptophan
607 metabolites are predictive of Ah receptor activity. Gut Microbes. 12(1):1–24.
608 doi:10.1080/19490976.2020.1788899

609 21. Wikoff WR, Anfora AT, Liu J, Schultz PG, Lesley SA, Peters EC, Siuzdak G. 2009.
610 Metabolomics analysis reveals large effects of gut microflora on mammalian blood
611 metabolites. PNAS. 106(10):3698–3703. doi:10.1073/pnas.0812874106

612 22. Alexeev EE, Lanis JM, Kao DJ, Campbell EL, Kelly CJ, Battista KD, Gerich ME,
613 Jenkins BR, Walk ST, Kominsky DJ, Colgan SP. 2018. Microbiota-Derived Indole

614 Metabolites Promote Human and Murine Intestinal Homeostasis through Regulation of
615 Interleukin-10 Receptor. The American Journal of Pathology. 188(5):1183–1194.
616 doi:10.1016/j.ajpath.2018.01.011

617 23. Hubbard TD, Murray IA, Bisson WH, Lahoti TS, Gowda K, Amin SG, Patterson AD,
618 Perdew GH. 2015. Adaptation of the human aryl hydrocarbon receptor to sense microbiota-
619 derived indoles. Scientific Reports. 5:1–13. doi:10.1038/srep12689

620 24. Kumar P, Lee J-H, Lee J. 2021. Diverse roles of microbial indole compounds in
621 eukaryotic systems. Biological Reviews. 96(6):522–2545. doi:10.1111/brv.12765

622 25. Petersen KS, Davis KM, Rogers CJ, Proctor DN, West SG, Kris-Etherton PM. 2021.
623 Herbs and spices at a relatively high culinary dosage improves 24-hour ambulatory blood
624 pressure in adults at risk of cardiometabolic diseases: a randomized, crossover, controlled-
625 feeding study. Am J Clin Nutr. 114(6):1936–1948. doi:10.1093/ajcn/nqab291

626 26. Poland A, Glover E, Ebetino FH, Kende AS. 1986. Photoaffinity labeling of the Ah
627 receptor. Journal of Biological Chemistry. 261(14):6352–6365. doi:10.1016/S0021-
628 9258(19)84570-2

629 27. Gruszczyk J, Grandvuillemin L, Lai-Kee-Him J, Paloni M, Savva CG, Germain P,
630 Grimaldi M, Boulahtouf A, Kwong H-S, Bous J, et al. 2022. Cryo-EM structure of the
631 agonist-bound Hsp90-XAP2-AHR cytosolic complex. Nat Commun. 13(1):7010.
632 doi:10.1038/s41467-022-34773-w

633 31. Hahn ME. 2002. Aryl hydrocarbon receptors: diversity and evolution1. Chemico-
634 Biological Interactions. 141(1):131–160. doi:10.1016/S0009-2797(02)00070-4

635 29. Ramadoss P, Perdew GH. 2004. Use of 2-Azido-3-[125I]iodo-7,8-dibromodibenzo-p-
636 dioxin as a Probe to Determine the Relative Ligand Affinity of Human versus Mouse Aryl
637 Hydrocarbon Receptor in Cultured Cells. Mol Pharmacol. 66(1):129–136.
638 doi:10.1124/mol.66.1.129

639 30. Goodrich JK, Davenport ER, Waters JL, Clark AG, Ley RE. 2016. Cross-species
640 comparisons of host genetic associations with the microbiome. Science. 352(6285):532–535.
641 doi:10.1126/science.aad9379

642 31. Spor A, Koren O, Ley R. 2011. Unravelling the effects of the environment and host
643 genotype on the gut microbiome. Nat Rev Microbiol. 9(4):279–290. doi:10.1038/nrmicro2540

644 32. Zhou X, Chakraborty D, Murray IA, Coslo D, Kehs Z, Vijay A, Ton C, Desai D, Amin
645 SG, Patterson AD, Perdew GH. 2023. Aryl Hydrocarbon Receptor Activation Coordinates
646 Mouse Small Intestinal Epithelial Cell Programming. Laboratory Investigation.:100012.
647 doi:10.1016/j.labinv.2022.100012

648 33. Modoux M, Rolhion N, Lefevre JH, Oeuvray C, Nádvorník P, Illes P, Emond P, Parc Y,
649 Mani S, Dvorak Z, Sokol H. 2022. Butyrate acts through HDAC inhibition to enhance aryl

650 hydrocarbon receptor activation by gut microbiota-derived ligands. Gut Microbes.
651 14(1):2105637. doi:10.1080/19490976.2022.2105637

652 34. Zhang X, Gan M, Li J, Li H, Su M, Tan D, Wang S, Jia M, Zhang L, Chen G. 2020.
653 Endogenous Indole Pyruvate Pathway for Tryptophan Metabolism Mediated by IL4I1. J Agric
654 Food Chem. 68(39):10678–10684. doi:10.1021/acs.jafc.0c03735

655 35. Ehrlich AM, Pacheco AR, Henrick BM, Taft D, Xu G, Huda MN, Mishchuk D, Goodson
656 ML, Slupsky C, Barile D, et al. 2020. Indole-3-lactic acid associated with Bifidobacterium-
657 dominated microbiota significantly decreases inflammation in intestinal epithelial cells. BMC
658 Microbiol. 20(1):357. doi:10.1186/s12866-020-02023-y

659 36. Zelante T, Puccetti M, Giovagnoli S, Romani L. 2021. Regulation of host physiology and
660 immunity by microbial indole-3-aldehyde. Current Opinion in Immunology. 70:27–32.
661 doi:10.1016/j.coi.2020.12.004

662 37. Wlodarska M, Luo C, Kolde R, d'Hennezel E, Annand JW, Heim CE, Krastel P, Schmitt
663 EK, Omar AS, Creasey EA, et al. 2017. Indoleacrylic Acid Produced by Commensal
664 Peptostreptococcus Species Suppresses Inflammation. Cell Host & Microbe. 22(1):25-37.e6.
665 doi:10.1016/j.chom.2017.06.007

666 38. Zhang Y, Zou Y, Brock NL, Huang T, Lan Y, Wang X, Deng Z, Tang Y, Lin S. 2017.
667 Characterization of 2-Oxindole Forming Heme Enzyme MarE, Expanding the Functional
668 Diversity of the Tryptophan Dioxygenase Superfamily. J Am Chem Soc. 139(34):11887–
669 11894. doi:10.1021/jacs.7b05517

670 39. Wishart DS, Guo A, Oler E, Wang F, Anjum A, Peters H, Dizon R, Sayeeda Z, Tian S,
671 Lee BL, et al. 2022. HMDB 5.0: the Human Metabolome Database for 2022. Nucleic Acids
672 Research. 50(D1):D622–D631. doi:10.1093/nar/gkab1062

673 40. Shetty SA, Hugenholtz F, Lahti L, Smidt H, de Vos WM. 2017. Intestinal microbiome
674 landscaping: insight in community assemblage and implications for microbial modulation
675 strategies. FEMS Microbiology Reviews. 41(2):182–199. doi:10.1093/femsre/fuw045

676 41. Chen L, Wang D, Garmaeva S, Kurilshikov A, Vich Vila A, Gacesa R, Sinha T, Segal E,
677 Weersma RK, Wijmenga C, et al. 2021. The long-term genetic stability and individual
678 specificity of the human gut microbiome. Cell. 184(9):2302-2315.e12.
679 doi:10.1016/j.cell.2021.03.024

680 37. Wirthgen E, Hoeflich A, Rebl A, Günther J. 2018. Kynurenic Acid: The Janus-Faced
681 Role of an Immunomodulatory Tryptophan Metabolite and Its Link to Pathological
682 Conditions. Front Immunol. 8(1957). doi:10.3389/fimmu.2017.01957

683 43. Murray IA, Patterson AD, Perdew GH. 2014. Aryl hydrocarbon receptor ligands in
684 cancer: friend and foe. Nat Rev Cancer. 14(12):801–814. doi:10.1038/nrc3846

685 44. Kolluri SK, Jin U-H, Safe S. 2017. Role of the aryl hydrocarbon receptor in
686 carcinogenesis and potential as an anti-cancer drug target. Arch Toxicol. 91(7):2497–2513.
687 doi:10.1007/s00204-017-1981-2

688 45. Wu C, Yu S, Tan Q, Guo P, Liu H. 2018. Role of AhR in regulating cancer stem cell–like
689 characteristics in choriocarcinoma. Cell Cycle. 17(18):2309–2320.
690 doi:10.1080/15384101.2018.1535219

691 46. Safe S, Cheng Y, Jin U-H. 2017. The aryl hydrocarbon receptor (AhR) as a drug target
692 for cancer chemotherapy. Current Opinion in Toxicology. 2:24–29.
693 doi:10.1016/j.cotox.2017.01.012

694 47. Furman DP, Oshchepkova EA, Oshchepkov DYu, Shamanina MYu, Mordvinov VA.
695 2009. Promoters of the genes encoding the transcription factors regulating the cytokine gene
696 expression in macrophages contain putative binding sites for aryl hydrocarbon receptor.
697 Computational Biology and Chemistry. 33(6):465–468.
698 doi:10.1016/j.compbiolchem.2009.10.004

699 48. Ogura N, Akutsu M, Tobe M, Sakamaki H, Abiko Y, Kondoh T. 2007. Microarray
700 analysis of IL-1beta-stimulated chemokine genes in synovial fibroblasts from human TMJ. J
701 Oral Pathol Med. 36(4):223–228. doi:10.1111/j.1600-0714.2007.00515.x

702 49. Nguyen NT, Nakahama T, Nguyen CH, Tran TT, Le VS, Chu HH, Kishimoto T. 2015.
703 Aryl hydrocarbon receptor antagonism and its role in rheumatoid arthritis. J Exp Pharmacol.
704 7:29–35. doi:10.2147/JEP.S63549

705 50. Sinclair LV, Neyens D, Ramsay G, Taylor PM, Cantrell DA. 2018. Single cell analysis of
706 kynurenine and System L amino acid transport in T cells. Nat Commun. 9(1):1981.
707 doi:10.1038/s41467-018-04366-7

708 51. Uwai Y, Honjo H, Iwamoto K. 2012. Interaction and transport of kynurenic acid via
709 human organic anion transporters hOAT1 and hOAT3. Pharmacological Research. 65(2):254–
710 260. doi:10.1016/j.phrs.2011.11.003

711 52. Takahashi M, Ozaki M, Kang M-I, Sasaki H, Fukazawa M, Iwakami T, Lim PJ, Kim H-
712 K, Aoyama S, Shibata S. 2018. Effects of Meal Timing on Postprandial Glucose Metabolism
713 and Blood Metabolites in Healthy Adults. Nutrients. 10(11):1763. doi:10.3390/nu10111763

714 53. Shrestha A, Müllner E, Poutanen K, Mykkänen H, Moazzami AA. 2017. Metabolic
715 changes in serum metabolome in response to a meal. Eur J Nutr. 56(2):671–681.
716 doi:10.1007/s00394-015-1111-y

717 54. Long WP, Pray-Grant M, Tsai JC, Perdew GH. 1998. Protein Kinase C Activity Is
718 Required for Aryl Hydrocarbon Receptor Pathway-Mediated Signal Transduction. Mol
719 Pharmacol. 53(4):691–700. doi:10.1124/mol.53.4.691

720 55. John K, Lahoti TS, Wagner K, Hughes JM, Perdew GH. 2014. The Ah receptor regulates
721 growth factor expression in head and neck squamous cell carcinoma cell lines. Mol Carcinog.
722 53(10):765–776. doi:10.1002/mc.22032

723 56. Muku GE, Murray IA, Espín JC, Perdew GH. 2018. Urolithin A Is a Dietary Microbiota-
724 Derived Human Aryl Hydrocarbon Receptor Antagonist. Metabolites. 8(4):86.
725 doi:10.3390/metabo8040086

726 57. Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of
727 image analysis. Nat Methods. 9(7):671–675. doi:10.1038/nmeth.2089

728 58. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ.
729 2009. AutoDock4 and AutoDockTools4: Automated docking with selective receptor
730 flexibility. J Comput Chem. 30(16):2785–2791. doi:10.1002/jcc.21256

731 59. Trott O, Olson AJ. 2010. AutoDock Vina: improving the speed and accuracy of docking
732 with a new scoring function, efficient optimization, and multithreading. J Comput Chem.
733 31(2):455–461. doi:10.1002/jcc.21334

734 60. Eberhardt J, Santos-Martins D, Tillack AF, Forli S. 2021. AutoDock Vina 1.2.0: New
735 Docking Methods, Expanded Force Field, and Python Bindings. J Chem Inf Model.
736 61(8):3891–3898. doi:10.1021/acs.jcim.1c00203

737 61. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, Li Q, Shoemaker BA, Thiessen PA,
738 Yu B, et al. 2021. PubChem in 2021: new data content and improved web interfaces. Nucleic
739 Acids Res. 49(D1):D1388–D1395. doi:10.1093/nar/gkaa971

740 62. Annalora AJ, Goodin DB, Hong W-X, Zhang Q, Johnson EF, Stout CD. 2010. Crystal
741 structure of CYP24A1, a mitochondrial cytochrome P450 involved in vitamin D metabolism.
742 J Mol Biol. 396(2):441–451. doi:10.1016/j.jmb.2009.11.057

743 58. Schrödinger, L. & DeLano, W., 2020. PyMOL, Available at:
744 http://www.pymol.org/pymol.

745 64. Chovancova E, Pavelka A, Benes P, Strnad O, Brezovsky J, Kozlikova B, Gora A, Sustr
746 V, Klvana M, Medek P, et al. 2012. CAVER 3.0: A Tool for the Analysis of Transport
747 Pathways in Dynamic Protein Structures. PLOS Computational Biology. 8(10):e1002708.
748 doi:10.1371/journal.pcbi.1002708

749 65. Velázquez-Libera JL, Durán-Verdugo F, Valdés-Jiménez A, Núñez-Vivanco G,
750 Caballero J. 2020. LigRMSD: a web server for automatic structure matching and RMSD
751 calculations among identical and similar compounds in protein-ligand docking.
752 Bioinformatics. 36(9):2912–2914. doi:10.1093/bioinformatics/btaa018

753

754 **List of abbreviations:**

| | |
|---|---|
| 755 AHR | Aryl hydrocarbon receptor |
| 756 CoV | Coefficient of variance |
| 757 GC-MS | Gas chromatography paired with mass spectrometry |
| 758 GF | Germ-free |
| 759 FLS | Fibroblast-like synoviocytes |
| 760 IAA | Indole-3-acetic acid |
| 761 IAcA | Indole-3-acrylic acid |
| 762 I3C | Indole-3-carboxaldehyde |
| 763 ILA | Indole-3-lactic acid |
| 764 IPA | Indole-3-propionic acid |
| 765 LC-MS/MS | Liquid chromatography paired with tandem mass spectrometry |
| 766 KA | Kynurenic acid |
| 767 KYN | Kynurenine |
| 768 PAL | Photoaffinity ligand |
| 769 SD | Standard deviation |
| 770 SEM | Standard error of the mean |
| 771 TCDD | 2,3,7,8-tetrachlorodibenzo-*p*-dioxin |
| 772 Trp | Tryptophan |

773 **Figure legends**

774 **Figure 1.** Comparison of tryptophan metabolites in serum of GF mice and conventional mice.
775 LC-MS analysis identified nine Trp metabolites in serum of conventional and GF mice fed a
776 standard chow diet, *ad libitum*. Statistical comparisons were made using Mann-Whitney and
777 unpaired *t*-test. Each box represents the median value with Q1 and Q3 range, and whiskers
778 describe the maximal and minimal values. Conv, conventional mice; GF, germ-free mice; <LOD,
779 below the limit of detection. (***): p-value = 0.0005; (****): p-value < 0.0001.

780 **Figure 2.** Comparison of tryptophan metabolites in serum of conventional mice fed a standard
781 chow or semi-purified diet. Conventional mice were fed a standard chow diet or semi-purified
782 diet (AIN-93G), *ad libitum*, and the nine Trp metabolites were quantified by LC-MS analysis.
783 Statistical comparisons were made using Mann-Whitney or unpaired *t*-test. Each box represents
784 the median value with Q1 and Q3 range, and whiskers describe the maximal and minimal values.
785 (***): p-value = 0.0005; (****): p-value < 0.0001.

786 **Figure 3.** Mouse AHR activity generated from circulating Trp metabolites. Mean concentrations
787 of Trp metabolites detected in mouse sera were assessed for their ability to activate the mouse
788 AHR using a stably transfected luciferase reporter construct in the Hepa1.1 cell line. Nine Trp
789 metabolites and the Trp metabolite Pool were used to treat Hepa1.1 cells for 4 h. Metabolite
790 activity was compared to VEH treatment. Metabolite concentrations were based on serum
791 samples collected from C57BL/6 mice fed a standard laboratory chow diet shown in Table 2.
792 Statistical significance was determined using one-way ANOVA. The data are mean ± SEM. *: p-
793 value < 0.05; **: p-value < 0.005; ***: p-value < 0.0005; ****: p-value < 0.0001. ns: not
794 significant.

795 **Figure 4.** Concentrations of six tryptophan metabolites in human circulating blood serum**.** LC-
796 MS analysis was used to quantitate tryptophan metabolites (a) Concentrations of six individual
797 Trp metabolites were identified in circulating blood serum collected from study participants
798 provided a controlled diet. (b) Total concentration of Trp metabolites. Each violin plot displays
799 the median value (---) and Q1 and Q3 (⋯). n = 40.

800 **Figure 5.** Circulating tryptophan metabolites generate AHR activity at physiological
801 concentrations. Trp metabolites and Pools were assessed for their ability to activate the human
802 AHR using a stably transfected luciferase reporter construct or *CYP1A1* induction in the HepG2
803 40/6 cell line. (a) The composition of the serum metabolite "Pool" determined by the mean of
804 each tryptophan metabolite. (b) Six tryptophan metabolites at the mean concentration detected in
805 serum and the Pool were used to treat HepG2 40/6 cells for 4 h. Metabolite-generated activity
806 was compared to VEH control. (c) Specific metabolites were excluded from the Pool and used to
807 treat HepG2 40/6 cells for 4 h. Activity from VEH control was subtracted from experimental
808 treatments and AHR activity was compared to the complete Pool treatment. AHR activity was
809 determined using a luciferase reporter system. (d) HepG2 40/6 cells were treated for 4 h with Trp
810 metabolite profiles and AHR activity was measured by the relative expression of *CYP1A1* to

811 *ACTB.* Data are normalized to a saturating dose of TCDD (10 nM) and VEH (0.1% DMSO, final
812 concentration) treatment, and presented as a percentage of maximal AHR activity. (e) Metabolite
813 concentrations representing the Trp metabolite profiles from six individual trial participants,
814 grouped into "High" and "Low" groups based on total metabolite abundance. Statistical
815 significance was determined using one-way ANOVA or unpaired *t*-test. The data are mean ±
816 SEM. *: p-value < 0.05; **: p-value < 0.005 ; ***: p-value < 0.0005; ****: p-value < 0.0001.

817 **Figure 6**. AHR activity in response to multiple treatment doses of tryptophan metabolites. Trp
818 metabolites were assessed for the ability to induce AHR activity across a range of treatment
819 concentrations using a stably transfected luciferase reporter construct in the HepG2 40/6 cell
820 line. AHR activity was determined using a luciferase reporter system following a 4 h treatment
821 time. Data are normalized to a saturating dose of TCDD (10 nM) and VEH (0.1% DMSO, final
822 concentration) treatment, and presented as a percentage of maximal AHR activity. $EC_{25}$ values
823 were generated by nonlinear regression and represented in the figure by a dotted line (⋯). The
824 range of quantified Trp metabolites from analyzed human serum are represented by dashed lines
825 (- - -), and the range is depicted by the yellow fill.

826 **Figure 7.** Serum tryptophan metabolites prevent binding of radioactive photoaffinity ligand. Trp
827 metabolites were assessed as AHR ligands by binding to the ligand binding pocket of AHR and
828 preventing the binding of the high affinity photoaffinity ligand 2-azido-3-[$^{125}$I]iodo-7,8-dibromodibenzo-
829 *p*-dioxin (PAL). HN30 cells were with treated 10 μM Trp metabolite and the metabolite Pool (9.26 μM,
830 total Trp metabolite concentration; italicized) prior to the addition of 2 pmoles PAL. Cell lysates were
831 processed and subjected to tricine SDS-PAGE, transferred to PVDF membranes, and AHR band intensity
832 was quantified using densitometry. Kynurenic acid (KA) is a known AHR ligands. Samples were
833 normalized to AHR abundance. (a) PAL competition assay with the Trp Pool and individual metabolites
834 metabolites; all reduced radioligand binding. (b) Quantitative graphic of the results from panel a. The
835 data are mean ± SEM. *: p-value < 0.05; **: p-value < 0.005.

836 **Figure 8. Trp metabolite docking in the cryo-EM model of the human AHR PAS B domain.**
837 The substrate binding properties of kynurenine, kynurenic acid and four tryptophan-related
838 metabolites (Indole-3-acetate, Indole-3-carboxylate, Indole-3-lactate, Indole-3-propionate) were
839 studied using the cryo-EM structure of the AhR (7ZUB) and Autodock Vina (see Supplementary
840 Table 1; see Methods). (a) Kynurenine (KYN) docks deeply into the primary, ligand binding
841 domain of the AHR with low micromolar affinity (-7.5 kcal/mol; KD = 3.07 μM; grey stick).
842 KYN engages the primary, indirubin binding site in close contact to key active site residues:
843 F295, Y322, C333, I349, F351, L353, S365, V381, N383 and H391. Electrostatic and
844 hydrophobic bond distances are shown (yellow dashes). (b) Kynurenic Acid (KA) displayed over
845 3-fold higher affinity to the PAS B domain as KYN (-8.2 kcal/mol; KD = 0.95 μM; orange
846 stick). KA docking revealed balanced interactions among the ligand and both sides of the closed,
847 active site, including a strong electrostatic contact with Fα helix residue S336. (c) Docking
848 results for Indole-3-carboxylate (I3C) revealed over 4-fold lower affinity for the AHR compared
849 to KA (-7.3 kcal/mol, KD = 4.29 μM; yellow stick). I3C preferentially docked in the distal
850 pocket as well, but showed reduced interactions with L353, N383 and H391. (d) Indole-3-

851 acetate (IAA) displayed nearly 2-fold higher affinity for the AHR than I3C (-7.6 kcal/mol, KD =
852 2.60 μM; dark blue stick), docking the left side of the distal binding pocket via strong
853 interactions with key residues G321, C333, Y322 and H391. (e) Indole-3-lactate (ILA) docked
854 the AHR with high nanomolar affinity (-8.3 kcal/mol; KD = 0.81 μM; light green stick),
855 displaying the highest affinity of all 6 Trp-related ligands tested. ILA binds deeply into the
856 primary pocket, forming close contacts with all key active site residues shown, except for G321,
857 Y322 and N383. (f) Indole-3-propionate (IPA) bound the AHR with similar affinity as ILA and
858 KA (-8.1 kcal/mol; KD = 1.13 μM; purple stick), but deeper in the distal pocket, forming strong
859 contacts with F295 and S336, but not L353, which may be a strong determinant of proper ligand
860 recognition in the AHR, as needed for efficient receptor transactivation.

861 **Figure 9.** Circulating tryptophan metabolites influence AHR-mediated gene expression in
862 human disease models. (a) AHR target gene, *CYP1A1,* induction in response to 4 h treatment
863 with Trp metabolites in human colorectal adenocarcinoma cells (Caco2) compared to VEH or
864 Pool treatment. Data presented as a percentage of maximal AHR activity (10 nM TCDD). (b)
865 Primary human fibroblast-like synoviocytes isolated from heathy patients were co-treated for 24
866 h with IL-1β and the Trp pool or pretreated with AHR antagonist (GNF351) prior to treatments.
867 Induction of inflammatory mediators, *IL1B* and *EREG*, was measured. Gene induction was
868 determined by real-time qPCR. Data are normalized to a saturating dose of TCDD (10 nM) and
869 VEH control, and presented as relative to *ACTB* expression. Statistical significance was
870 determined using one-way ANOVA. The data are mean ± SEM.

871

872 **Figure S1.** Serum concentrations of hippuric acid were used as a control to demonstrate depletion
873 of bacteria in germ free mice. A statistical comparison was made using a Mann-Whitney test. Each
874 box represents the median value with Q1 and Q3 range, and whiskers describe the maximal and
875 minimal values. Conv, conventional mice; GF, germ-free mice. (****): p-value < 0.0001.

876 **Figure S2.** Heatmap of Pearson's correlations between tryptophan metabolites detected in human
877 serum. Numbers and color indicate Pearson's r.* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

878 **Figure S3.** Trp metabolite cytotoxicity assay in Caco2 cells. Cytotoxicity at was assessed using
879 Cell Counting Kit-8 (SIGMA-ALDRICH, Co., St. Louis, Missouri, USA), per the
880 manufacturer's recommendations, following a 4 h exposure of Caco2 cells to mean serum
881 concentrations of Trp metabolites. Statistical comparisons were determined by one-way
882 ANOVA.

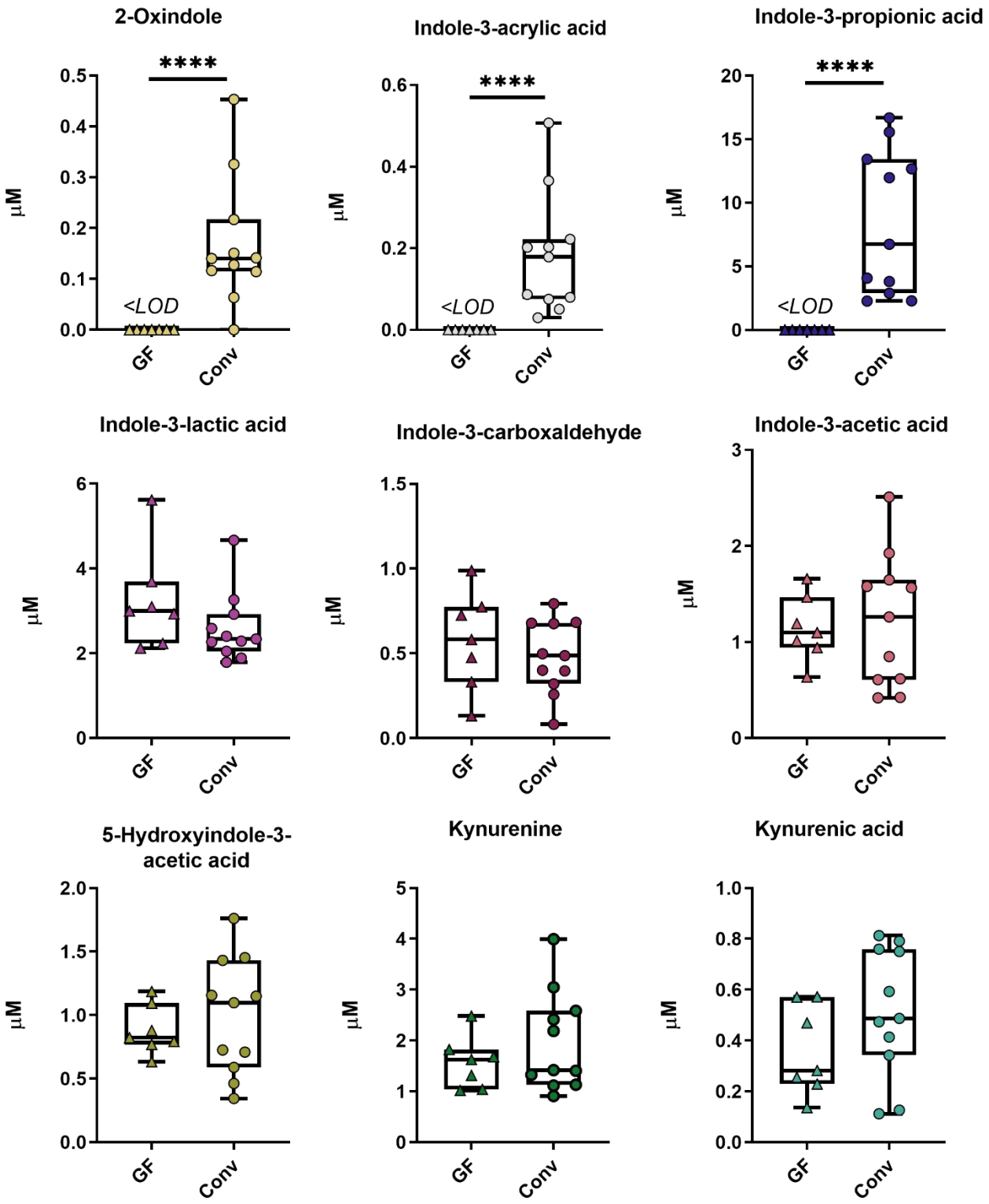
883 **Figure S4. Trp metabolite docking analysis in the cryo-EM structure of the AHR.** Ligand
884 binding properties were assessed using Autodock Vina and the PAS B domain of the indirubin-
885 bound Hsp90-XAP2-AHR complex (PDB: 7ZUB [55]; see methods). A.) The PAS B domain of
886 the AHR is shown in cartoon format with α-helical elements C, D and F (in cyan) and β-sheet
887 domains A, G, H, and I (in magenta) highlighted. Key amino acid residues that define the

888 primary ligand binding site are shown (line format), including H291, F295, P297, I325, C333,
889 H337, I349, F351, L353, S365, A367, V381, and Q383. The structural coordinates for indirubin
890 (transparent green stick) are shown superimposed on the low energy, indirubin redocking result
891 (-12.8 kcal/mol; $K_D$ = 0.44 nM; see Supplemental Table 5). B.) Autodock Vina redocking
892 results for indirubin in the AHR active site demonstrated high precision in recapitulating ligand
893 recognition (RMSD = 0.498 Å, see Supplemental Table 1), validating the cryo-EM AHR PAS B
894 domain structure as a suitable model for the comparative ligand binding analysis conducted here.
895 While individual ligands may contribute different enthalpic cues to facilitate closure and
896 transactivation of the receptor, in situ, the indirubin-bound active site provides a tangible, closed
897 conformation for exploring structure-activity relationships that are independent of structural
898 artifacts induced via MD-simulation based ligand docking approaches.

899 **Figure S5. Trp metabolite docking in the Indirubin-bound model of the human AHR.** As
900 shown in Figure 1, the substrate binding properties of (a) Kynurenine; (b) Kynurenic acid, and
901 four tryptophan-related metabolites, (c) Indole-3-carboxylate; (d) Indole-3-acetate; (e) Indole-3-
902 lactate, and (f) Indole-3-propionate were explored using the cryo-EM structure of the AHR and
903 Autodock Vina (see Table 1; see Methods). The results highlighted in Figure 1, are shown here
904 with respect to the crystal structure coordinates of indirubin (shown in transparent green stick) in
905 each panel, with respect to key active site residues: F295, G321, Y322, S336, S365, V381, N383
906 and H391. Key electrostatic and hydrophobic bond distances are shown (yellow dashes). High
907 nanomolar affinity binding observed for indole-3-lactate, indole-3-propionate, and kynurenic acid
908 suggest that a combination of discrete hydrophobic contacts (with F295, Y322, L349, F351, L353
909 and V381) and electrostatic contacts (with S336, S365, N383 and H391) are needed to fully engage
910 the primary ligand binding domain of the human AHR.

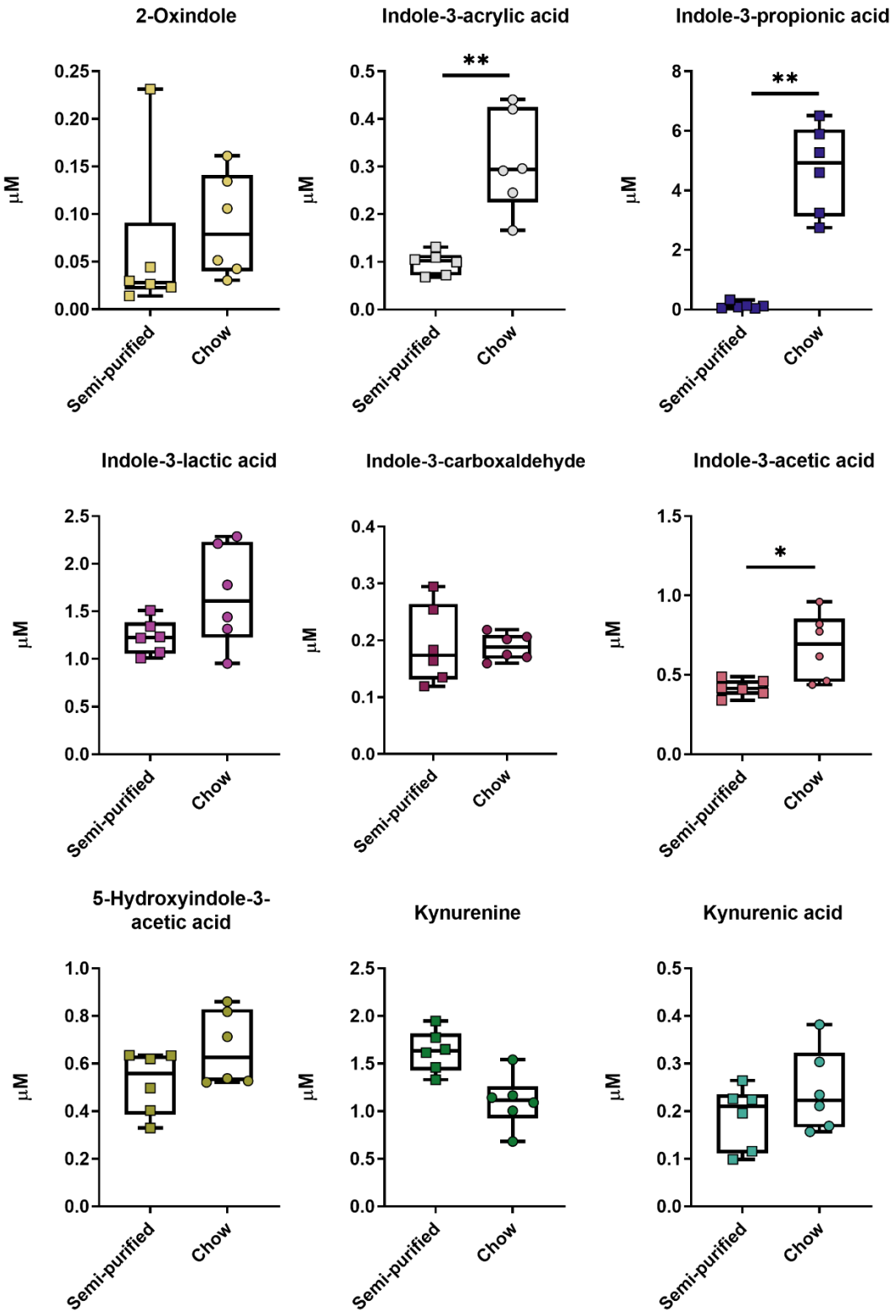
911

912 **Figure 1**




913

914 **Figure 2**

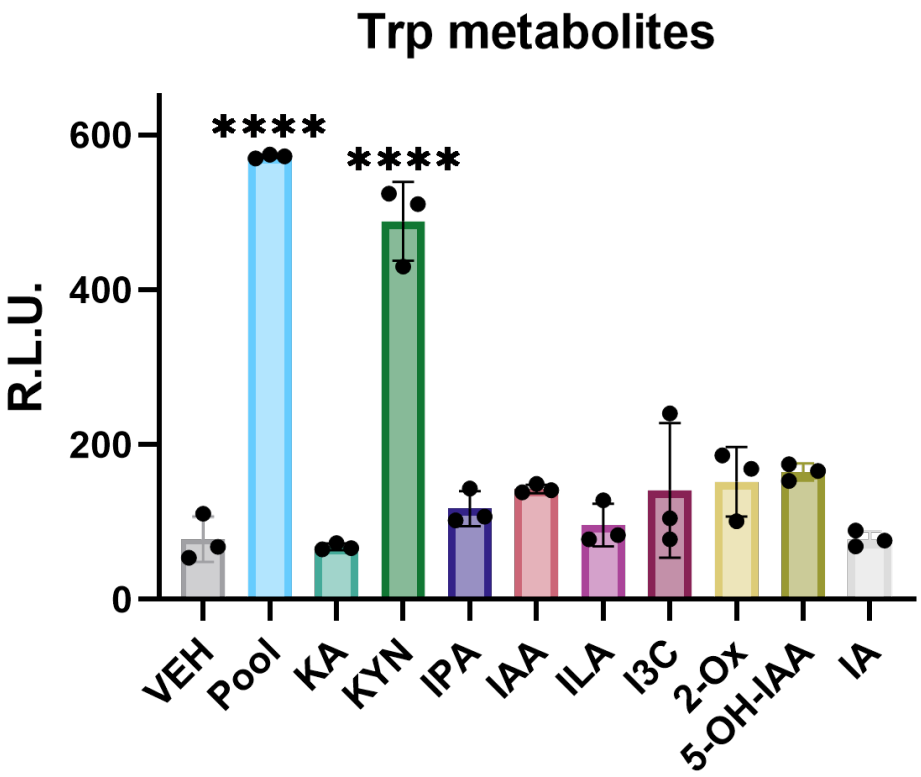



915

916

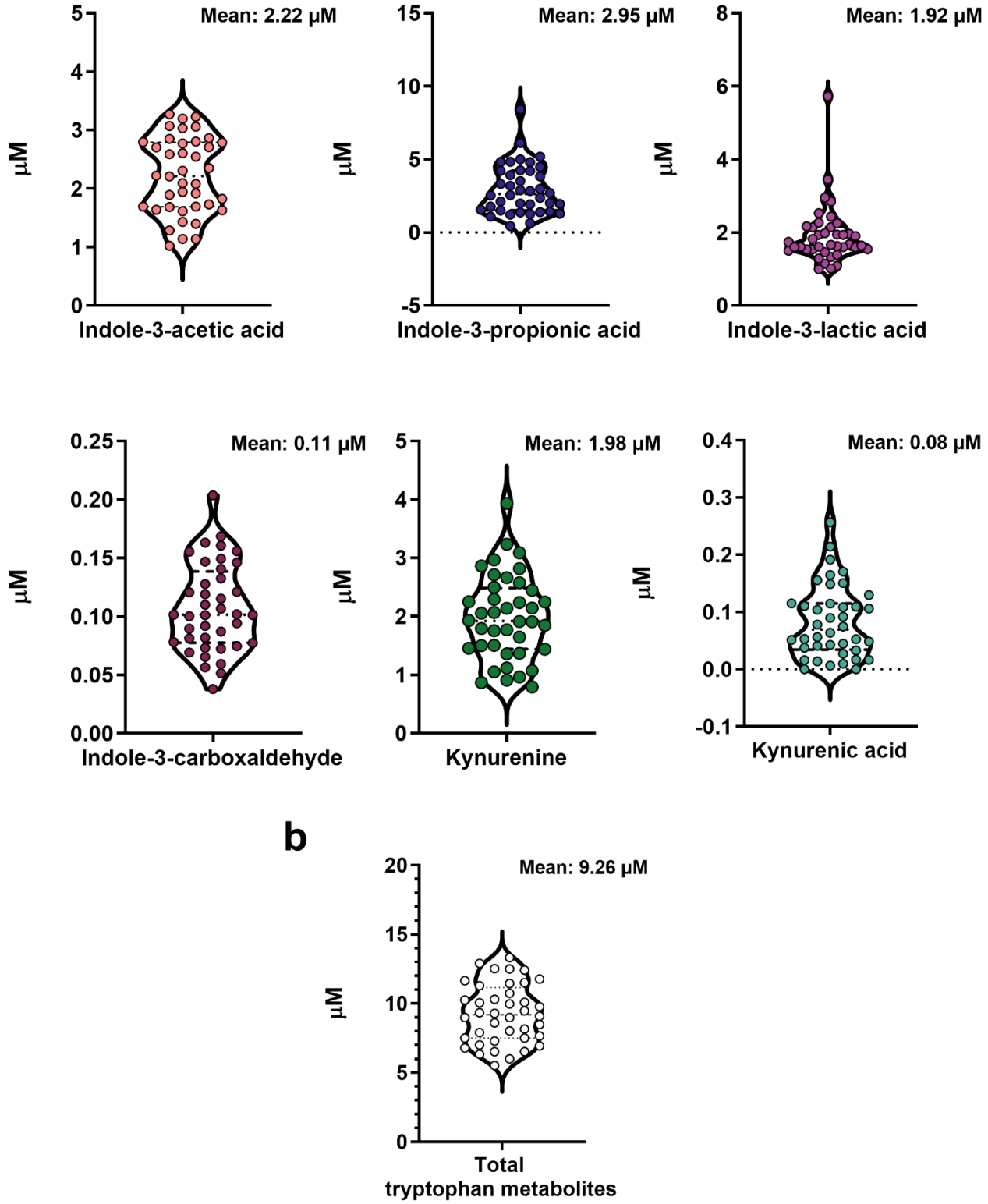
917 **Figure 3**

918

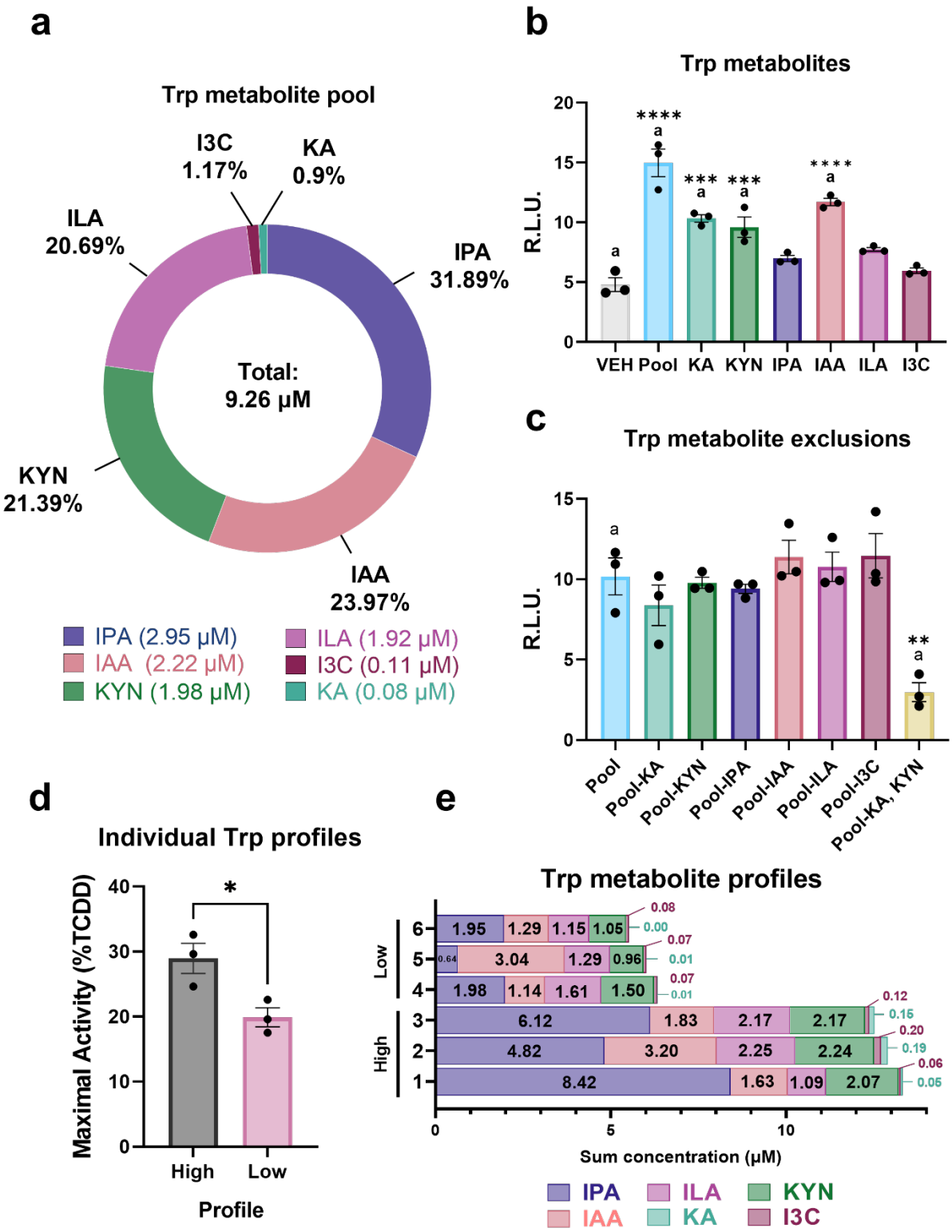



919

920 **Figure 4**

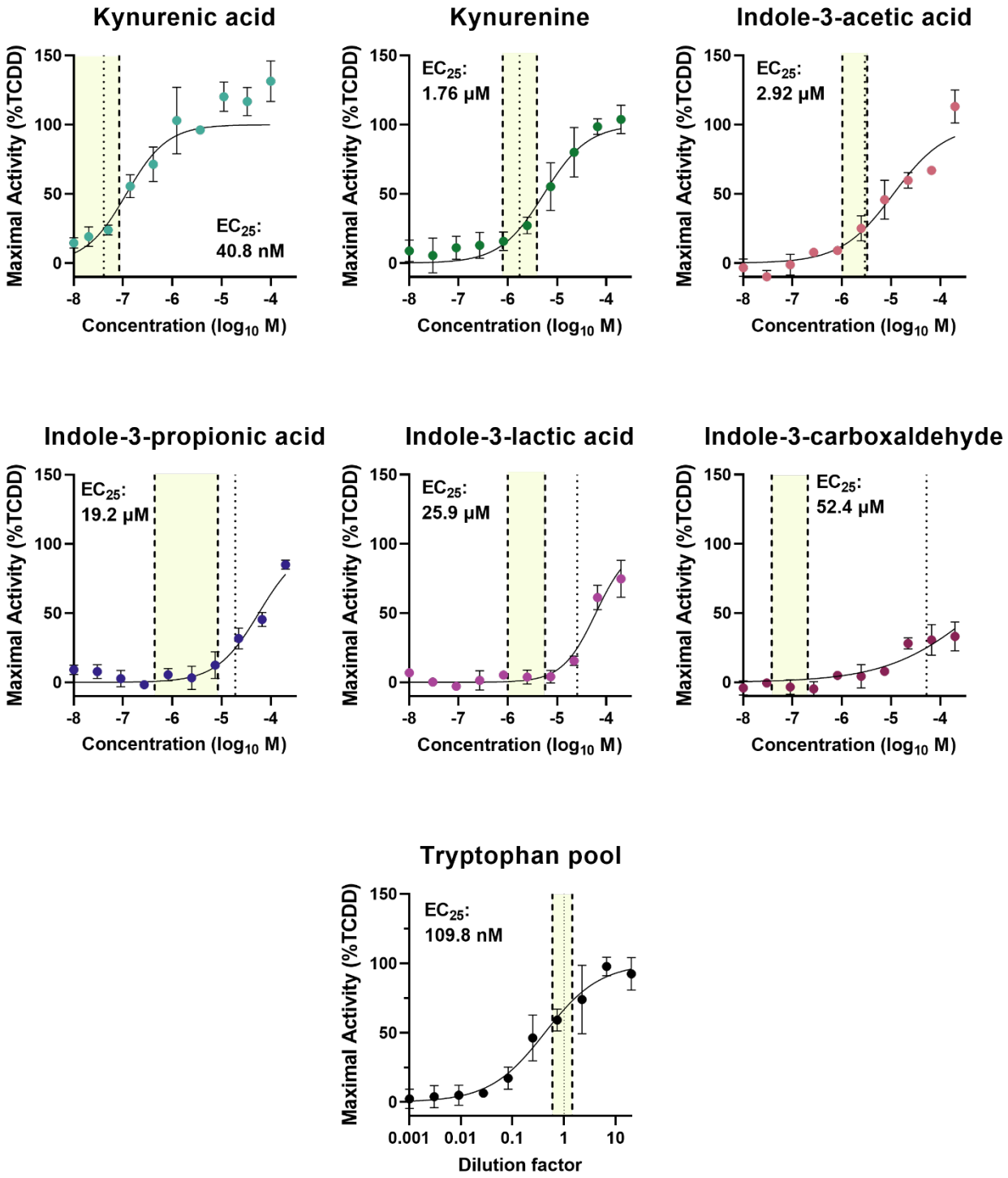

Mean: 2.22 µM
µM
Indole-3-acetic acid
Mean: 2.95 µM
µM
Indole-3-propionic acid
Mean: 1.92 µM
µM
Indole-3-lactic acid
Mean: 0.11 µM
µM
Indole-3-carboxaldehyde
Mean: 1.98 µM
µM
Kynurenine
Mean: 0.08 µM
µM
Kynurenic acid
b
Mean: 9.26 µM
µM
Total
tryptophan metabolites

921 **Figure 5**




922

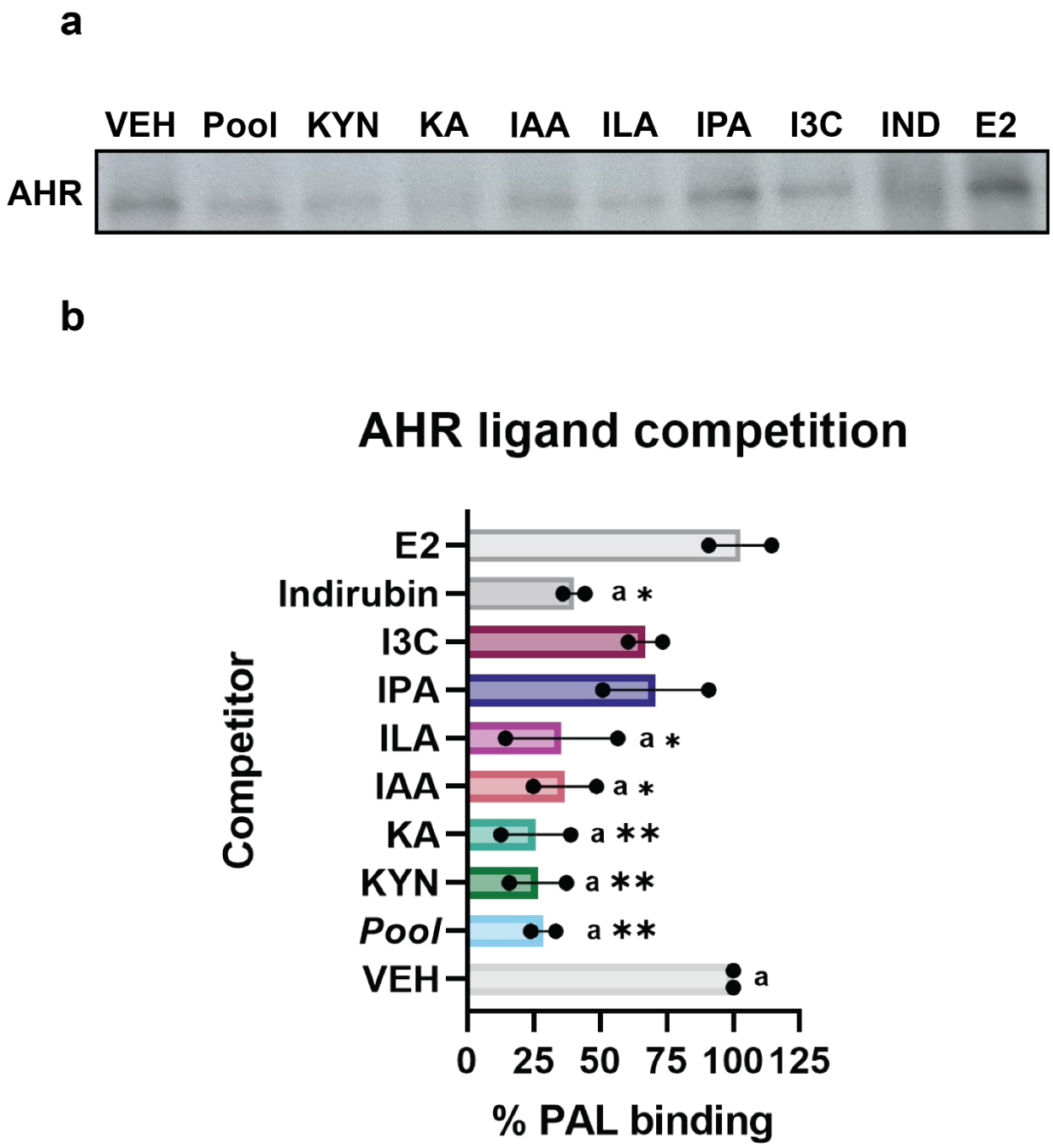
923 **Figure 6**




924

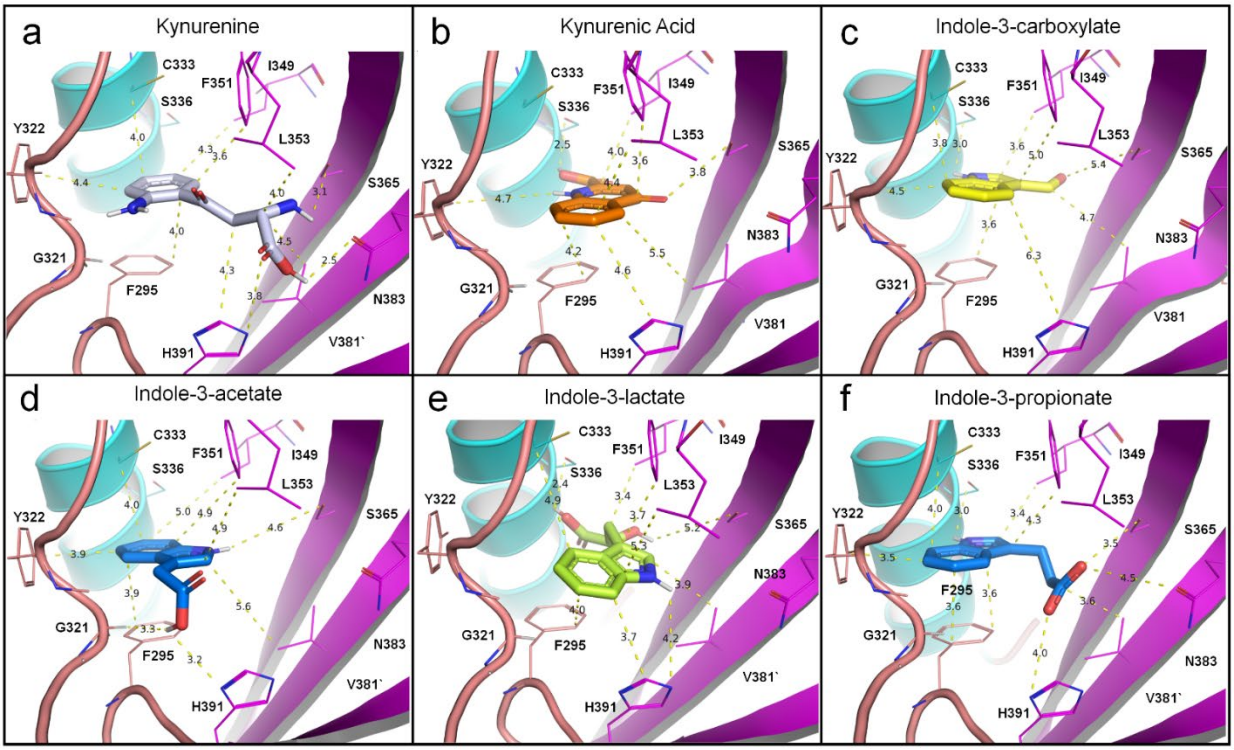
925

926 **Figure 7**




927

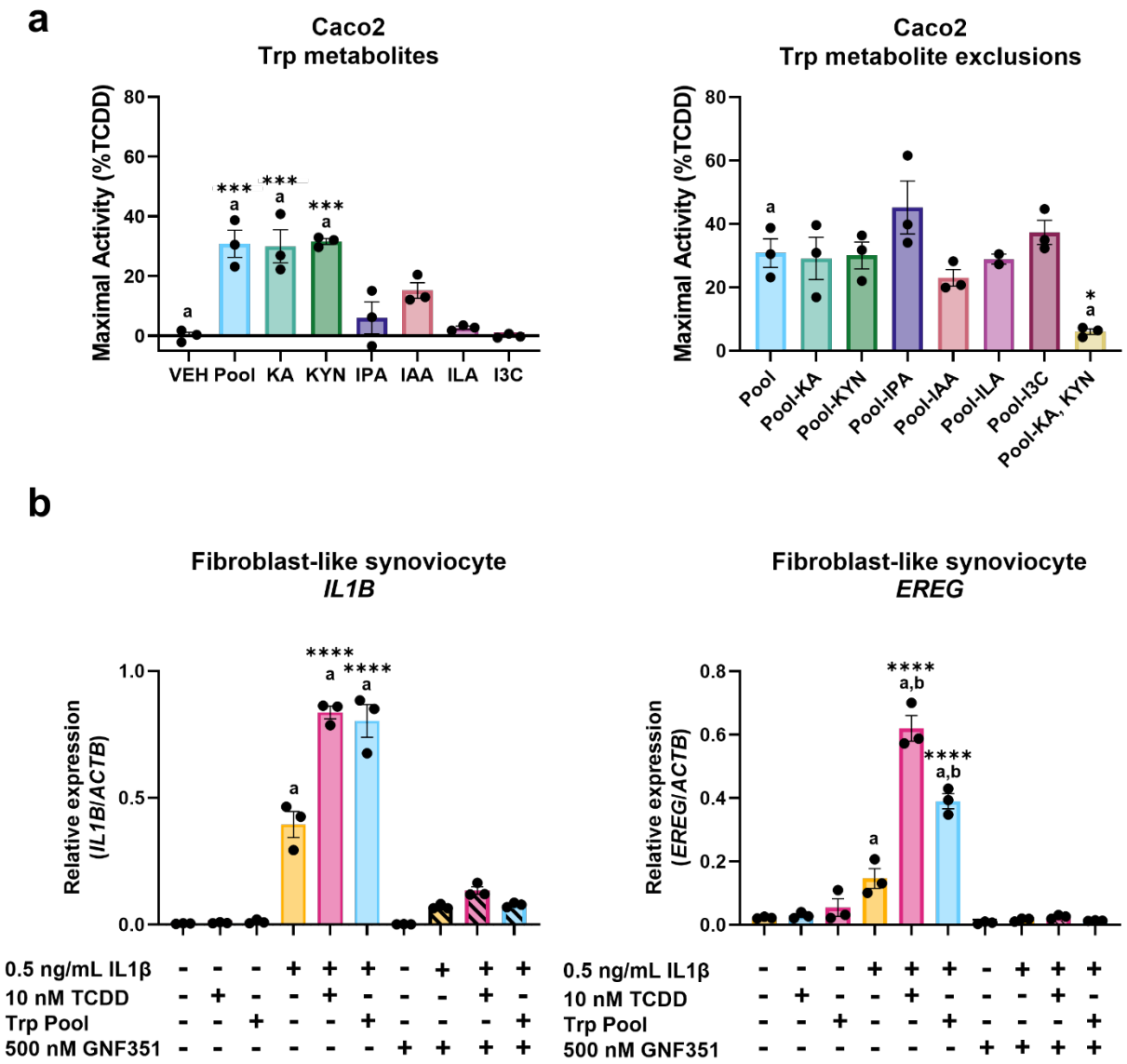
928 **Figure 8**




929

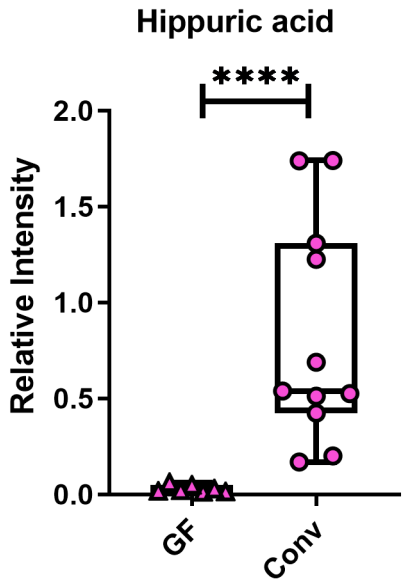
930

931 **Figure 9**



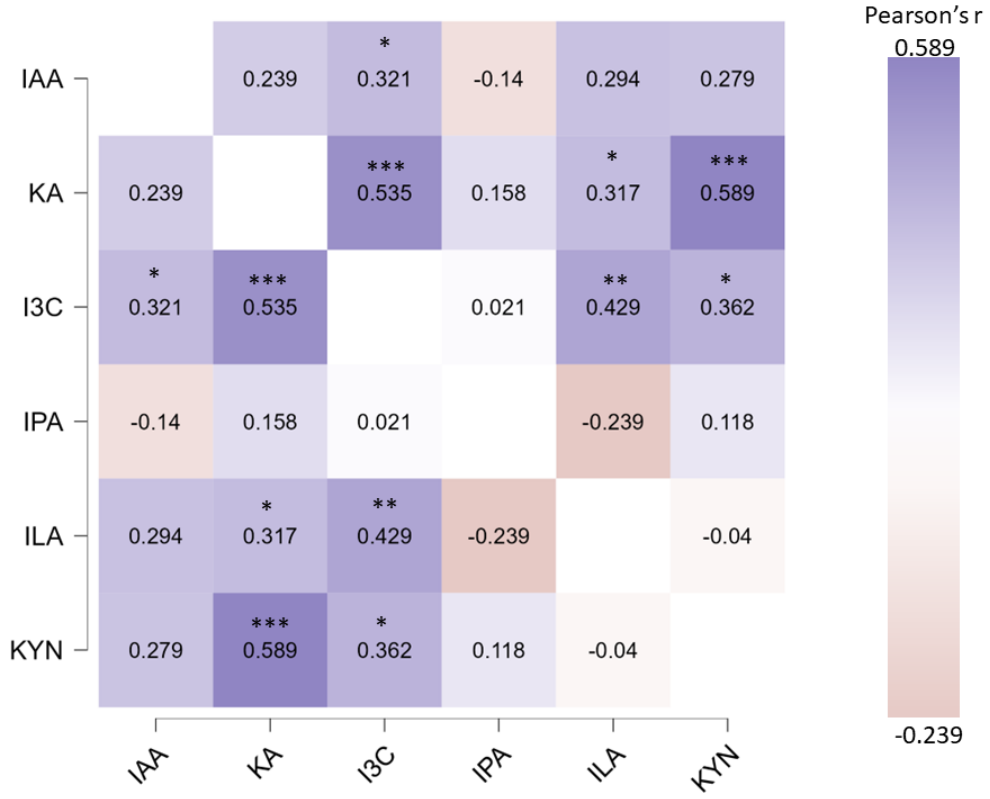

932

933 **Figure S1**

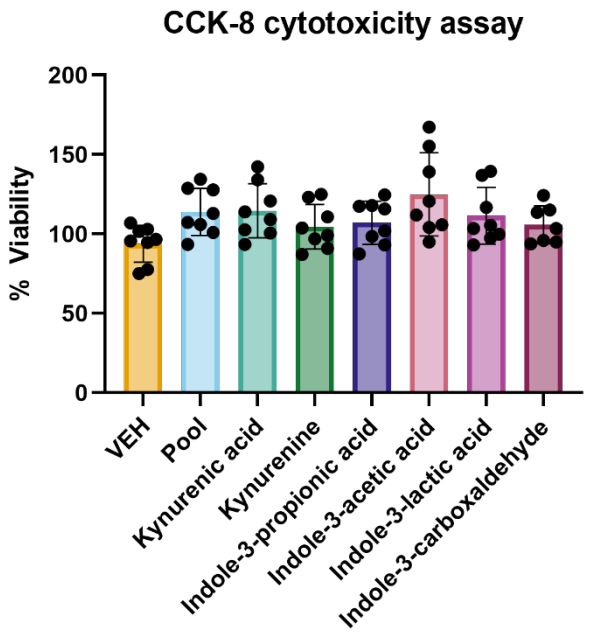



934

935 **Figure S2**

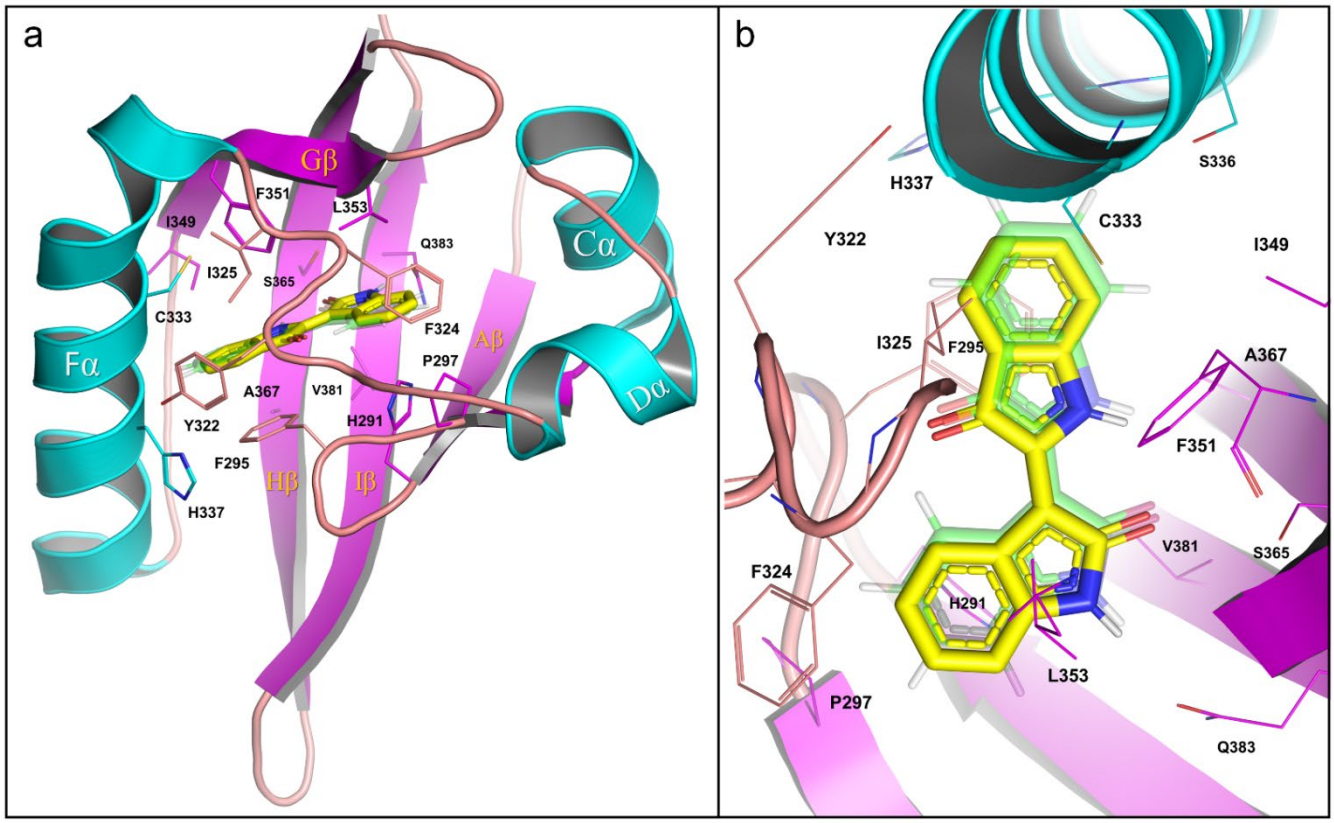


937 **Figure S3**



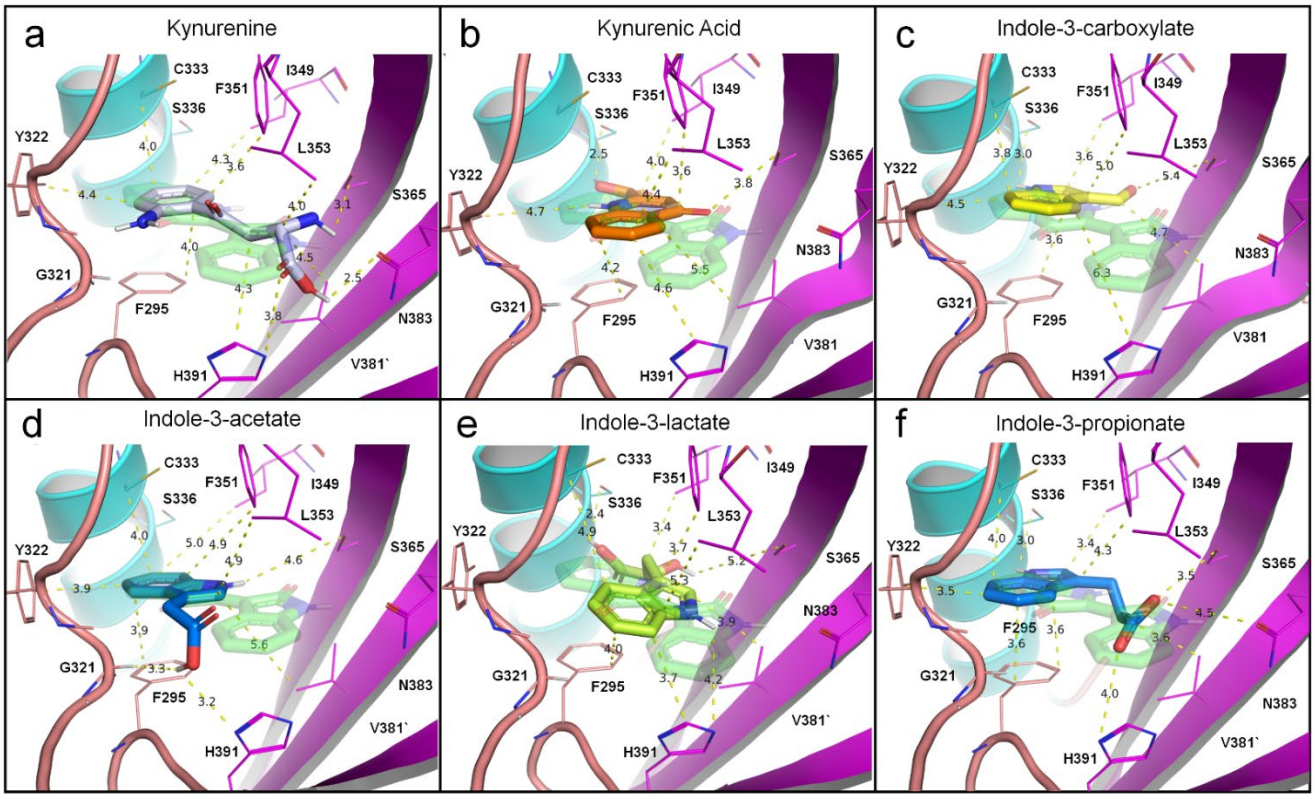

938

939 **Figure S4**




940

941 **Figure S5**




942

943

944 **Tables**

945 **Table 1.** Summary of profiled tryptophan metabolites in mouse serum and human sera[a].

| Metabolite | Mouse sera | Human sera | LOD (nM)[b] |
|---|---|---|---|
| Indole-3-acetic acid | Y | Y | 3 |
| Indole-3-carboxaldehyde | Y | Y | 3 |
| Indole-3-lactic acid | Y | Y | 5 |
| Indole-3-propionic acid | Y | Y | 2 |
| Kynurenic acid | Y | Y | 2 |
| Kynurenine | Y | Y | 7 |
| 5-Hydroxyindole-3-acetic acid | Y | N.D. | 22 |
| Indole-3-acrylic acid | Y | N.D. | 4 |
| 2-Oxindole | Y | N.D. | 3 |
| 2,8-Dihydroxyquinoline | N.D. | N.D. | 2 |
| Indole | N.D. | N.D. | 587 |
| Indole-3-carboxylic acid | N.D. | N.D. | 7 |
| 3-Methylindole | N.D. | N.D. | 60 |
| 3-Methyl-2-oxindole | N.D. | N.D. | 2 |
| Tryptamine | N.D. | N.D. | 7 |
| Xanthurenic acid | N.D. | N.D. | 25 |

946 a: Y presents the metabolites that were quantified in the biological samples; N.D.presents the
947 concentration of metabolites that were below the limit of detection. b: LOD was estimated by the
948 signal to noise ratio (S/N) of 3.

949 **Table 2.** Comparison of serum Trp metabolite concentrations in conventional mice and GF mice

| | **Mean abundance ± S.D. (μM)** | |
|---|---|---|
| **Metabolite** | **Conventional mice** | **GF mice** |
| **5-hydroxyindole-3-acetic acid** | 0.99 ±0.46 | 0.88 ± 0.19 |
| **Indole-3-acetic acid** | 1.22 ± 0.69 | 1.14 ± 0.34 |
| **Indole-3-acrylic acid** | 0.18 ± 0.15 | <LOD |
| **Indole-3-carboxaldehyde** | 0.48 ± 0.22 | 0.57 ± 0.29 |
| **Indole-3-lactic acid** | 2.59 ± 0.81 | 3.24 ± 1.18 |
| **Indole-3-propionic acid** | 8.4 ± 5.69 | <LOD |
| **Kynurenic acid** | 0.51 ± 0.25 | 0.36 ± 0.18 |
| **Kynurenine** | 1.96 ± 0.97 | 1.57 ± 0.51 |
| **2-Oxindole** | 0.17 ± 0.13 | <LOD |

950

951 **Table 3.** Comparison of serum Trp metabolite concentrations in conventional mice fed semi-
952 purified and standard chow diets.

| | **Mean abundance ± S.D. (μM)** | |
|---|---|---|
| **Metabolite** | **Semi-purified diet** | **Standard chow diet** |
| **5-hydroxyindole-3-acetic acid** | 0.52 ± 0.13 | 0.66 ± 0.15 |
| **Indole-3-acetic acid** | 0.42 ± 0.05[a] | 0.68 ± 0.21 |
| **Indole-3-acrylic acid** | 0.10 ± 0.02[a] | 0.31 ± 0.10 |
| **Indole-3-carboxaldehyde** | 0.19 ± 0.07 | 0.19 ± 0.02 |
| **Indole-3-lactic acid** | 1.12 ± 0.18 | 1.66 ± 0.53 |
| **Indole-3-propionic acid** | 0.12 ± 0.10[a] | 4.71 ± 1.48 |
| **Kynurenic acid** | 0.19 ± 0.07 | 0.24 ± 0.09 |
| **Kynurenine** | 1.63 ± 0.22 | 1.10 ± 0.28 |
| **2-Oxindole** | 0.06 ± 0.08 | 0.09 ± 0.05 |

953

954 **Table 4.** Descriptive statistics of human serum metabolites

| Metabolite | Mean ± S.D. (µM) | Range (µM) | Coefficient of Variation | Skewness | Std. Error of Skewness |
|---|---|---|---|---|---|
| **Indole-3-propionic acid** | 2.954 ± 1.686 | 0.443-8.417 | 0.571 | 0.979 | 0.374 |
| **Indole-3-acetic acid** | 2.222 ± 0.655 | 1.025-3.274 | 0.295 | -0.09 | 0.374 |
| **Kynurenine** | 1.981 ± 0.727 | 0.794-3.934 | 0.367 | 0.4 | 0.374 |
| **Indole-3-lactic acid** | 1.917 ± 0.808 | 0.998-5.731 | 0.421 | 2.963 | 0.374 |
| **Indole-3-carboxaldehyde** | 0.107 ± 0.038 | 0.038-0.204 | 0.35 | 0.43 | 0.374 |
| **Kynurenic acid** | 0.083 ± 0.063 | 0-0.257 | 0.763 | 0.782 | 0.374 |

955

956 **Table S1.** Parameters of AHR-related compounds quantified in serum of pathogen-free mice on a
957 chow diet[a]..

| Compound | LOQ (nM) | Mean concentration (µM) | Recovery (%) | Matrix effect (%) |
|---|---|---|---|---|
| Indole-3-acetic acid | 10 | 1.75 ± 0.18 | 93.2 ± 6.8 | 86.7 ± 3.2 |
| Indole-3-carboxaldehyde | 10 | 0.64 ± 0.05 | 95.9 ± 6.7 | 96.0 ± 7.0 |
| Indole-3-lactic acid | 17 | 2.23 ± 0.11 | 89.6 ± 7.4 | 92.8 ± 6.1 |
| Indole-3-propionic acid | 7 | 12.84 ± 1.42 | 93.6 ± 2.6 | 75.3 ± 1.6 |
| Kynurenic acid | 7 | 0.34 ± 0.08 | 94.1 ± 0.5 | 96.9 ± 3.7 |
| Kynurenine | 24 | 2.59 ± 0.36 | 92.6 ± 3.1 | 91.7 ± 2.8 |
| 5-Hydroxyindole-3-acetic acid | 74 | 1.34 ± 0.10 | 81.8 ± 1.2 | 114.6 ± 8.6 |
| Indole-3-acrylic acid | 14 | 0.16 ± 0.05 | 94.2 ± 4.5 | 99.3 ± 3.4 |
| 2-Oxindole | 10 | 0.14 ± 0.02 | 96.4 ± 25.0 | 64.6 ± 8.4 |

958 [a]: Concentration of each compound was corrected by recovery and matrix effect. Data was
959 shown as mean ± SEM.

960

961 **Table S2.** Parameters of AHR-related compounds quantified in human sera[a].

| Compound | LOQ (nM) | Mean concentration (μM) | Recovery (%) | Matrix effect (%) |
|---|---|---|---|---|
| Indole-3-acetic acid | 10 | 2.22 ± 0.10 | 92.5 ± 4.6 | 92.6 ± 3.6 |
| Indole-3-carboxaldehyde | 10 | 0.11 ± 0.01 | 73.0 ± 3.6 | 146.1 ± 18.4 |
| Indole-3-lactic acid | 17 | 1.92 ± 0.13 | 86.5 ± 8.5 | 85.2 ± 9.3 |
| Indole-3-propionic acid | 7 | 2.95 ± 0.27 | 98.5 ± 7.3 | 77.5 ± 7.0 |
| Kynurenic acid | 7 | 0.08 ± 0.01 | 80.0 ± 7.5 | 114.9 ± 8.3 |
| Kynurenine | 24 | 1.98 ± 0.11 | 80.2 ± 5.0 | 100.1 ± 4.0 |

962 [a]: Concentration of each compound was corrected by recovery and matrix effect. Data was
963 shown as mean ± SEM.

964

965 **Table S3.** RT-qPCR primer sequences used in this study.

| Gene | Forward Primer | Reverse Primer |
|---|---|---|
| *ACTB* | CACCATTGGCAATGAGCGGTTC | AGGTCTTTGCGGATGTCCACGT |
| *Actb* | CATTGCTGACAGGATGCAGAAGG | TGCTGGAAGGTGGACAGTGAGG |
| *CYP1A1* | TCTTCCTTCGTCCCCTTCAC | TGGTTGATCTGCCACTGGTT |
| *Cyp1a1* | CATCACAAGTGCCCTGTTCAAGC | AATGCTCCAGGTGATGGCTGTG |
| *EREG* | CTTATCACAGTCGTCGGTTCCAC | GCCATTCAGACTTGCGGCAACT |
| *IL1B* | CAGGCTGCTCTGGGATTCTC | GTCCTGGAAGGAGCACTTCAT |

966

967

968

969

970

971

972

**Table S4**. Pearson's Correlations between tryptophan metabolites detected in human serum.

| **Pearson's Correlations** | | | | | | | |
|---|---|---|---|---|---|---|---|
| **Metabolite** | | **IAA** | **KA** | **I3C** | **IPA** | **ILA** | **KYN** |
| 1. IAA | Pearson's r | — | | | | | |
| | p-value | — | | | | | |
| 2. KA | Pearson's r | 0.239 | — | | | | |
| | p-value | 0.137 | — | | | | |
| 3. I3C | Pearson's r | 0.321 | 0.535 | — | | | |
| | p-value | 0.043 | < .001 | — | | | |
| 4. IPA | Pearson's r | -0.140 | 0.158 | 0.021 | — | | |
| | p-value | 0.388 | 0.332 | 0.898 | — | | |
| 5. ILA | Pearson's r | 0.294 | 0.317 | 0.429 | -0.239 | — | |
| | p-value | 0.066 | 0.046 | 0.006 | 0.138 | — | |
| 6. KYN | Pearson's r | 0.279 | 0.589 | 0.362 | 0.118 | -0.040 | — |
| | p-value | 0.082 | < .001 | 0.022 | 0.468 | 0.808 | — |

985 **Table S5.** Computational Docking Analysis of Tryptophan Metabolites in the Human AHR
986 PAS-B domain.

| Substrate | | Autodock Vina | |
|---|---|---|---|
| | | ***Structure-based Homology Model*** | |
| | | **Human AhR PAS B domain** | |
| | | **Binding Energy** [a] (kcal/mole) | **Dissociation Constant** [b] *[$K_D$] $\mu M$* |
| **Kynurenine** | *Max* [c] | **-7.5** | **3.07** |
| | *Avg* [d] | *-7.05 ± 0.28* | *7.20 ± 3.08* |
| **Kynurenic Acid** | *Max* | **-8.2** | **0.95** |
| | *Avg* | *-7.75 ± 0.35* | *2.42 ± 1.69* |
| **Indole-3-carboxylate** | *Max* | **-7.3** | **4.29** |
| | *Avg* | *-6.89 ± 0.30* | *9.98 ± 7.33* |
| **Indole-3-acetate** | *Max* | **-7.6** | **2.60** |
| | *Avg* | *-7.13 ± 0.32* | *6.64 ± 4.05* |
| **Indole-3-lactate** | *Max* | **-8.3** | **0.81** |
| | *Avg* | *-7.59 ± 0.39* | *3.20 ± 1.83* |
| **Indole-3-propionate** | *Max* | **-8.1** | **1.13** |
| | *Avg* | *-7.68 ± 0.21* | *2.47 ± 0.96* |

987 [a] Substrate Binding Energies were derived computationally using Autodock Vina.

988 [b] Dissociation binding constants ($K_D$) (in the micromolar ($\mu$M) scale), were derived
989 computationally from Autodock Vina data using the Autodock 4.2 inhibitory constant conversion
990 scale, based on the equation: $y = 0.5982\ln(x) - 12.304$.

991 [c] Max refers to the low-energy Vina docking solution (in kcal/mole), with the highest predicted
992 binding affinity [$K_D$] for the AhR PAS B domain. If multiple results had the same binding
993 energy, the pose with the closest contact distance to the heme center was selected.

994 [d] Avg refers to the average or cumulative dissociation constant and binding energy (± the
995 standard deviation) for the Vina docking poses obtained for each individual substrate:model
996 docking combination (N < 10).