Cryo-EM structure of the agonist-bound Hsp90-XAP2-AHR cytosolic complex

Jakub Gruszczyk[1,*], Loic Grandvuillemin[1], Josephine Lai-Kee-Him[1], Matteo Paloni[1], Christos G. Savva[2], Pierre Germain[1], Marina Grimaldi[3], Abdelhay Boulahtouf[3], Hok-Sau Kwong[1], Julien Bous[4], Aurelie Ancelin[1], Cherine Bechara[5,6], Alessandro Barducci[1], Patrick Balaguer[3], William Bourguet[1,*]

[1]CBS (Centre de Biologie Structurale), Univ Montpellier, CNRS, Inserm; Montpellier, France.

[2]Leicester Institute of Structural & Chemical Biology and Department of Molecular & Cell Biology, University of Leicester; Leicester, UK.

[3]IRCM (Institut de Recherche en Cancérologie de Montpellier), Inserm U1194, Univ Montpellier, ICM; Montpellier, France.

[4]Section of Receptor Biology & Signaling, Department of Physiology & Pharmacology, Karolinska Institutet; Stockholm, Sweden.

[5]IGF, University of Montpellier, CNRS, INSERM; Montpellier, France.

[6]Institut Universitaire de France (IUF); Paris, France.

*Co-corresponding authors:

Jakub Gruszczyk

jakub.gruszczyk@cbs.cnrs.fr

William Bourguet

william.bourguet@cbs.cnrs.fr

Summary

Living organisms have developed protein sensors helping them to adapt to their environment[1]. The aryl hydrocarbon receptor (AHR) is an emblematic member of this class of proteins, and a ligand-dependent transcription factor that mediates a broad spectrum of (patho)physiological processes in response to numerous substances including pollutants, natural products and metabolites[2]. However, in the absence of high-resolution structural data, a molecular understanding of how AHR is activated by such diverse compounds is lacking. Here we present a 2.85 Å cryo-electron microscopy structure of the cytosolic complex comprising AHR bound to the ligand indirubin, the chaperone Hsp90 and the co-chaperone XAP2. The structure reveals a closed Hsp90 dimer with AHR threaded through its lumen. XAP2 directly interacts with Hsp90 and the AHR ligand-binding domain, thereby acting as a brace stabilizing the entire complex. Importantly, we provide the first experimental visualization of the AHR PAS-B domain bound to a ligand, revealing a unique organization of the ligand-binding pocket and the structural determinants of ligand-binding specificity and promiscuity of the receptor. By providing unprecedented structural details of the molecular initiating event leading to AHR activation, our study rationalizes prior biochemical data and provides a framework for future mechanistic studies and structure-guided drug design.

42 **Main text**

43 Also known as the dioxin receptor, the aryl hydrocarbon receptor (AHR) is a ligand-
44 dependent transcription factor involved in the regulation of genes governing many
45 (patho)physiological processes including cell growth and differentiation, cell cycle and
46 migration, apoptosis, haematopoiesis and carcinogenesis[3–6]. Importantly, activation of
47 AHR in the gastrointestinal tract participates in signalling between the enteric
48 microflora and our immune system[7–9]. A disruption of the intestinal homeostasis, for
49 example, has been shown to contribute in the aetiology of many disorders, including
50 inflammatory bowel disease, Crohn's disease, ulcerative colitis, diabetes and obesity[9].
51 Moreover, AHR is one of the crucial chemosensory proteins associated with the
52 metabolism of various xenobiotics, and sustained activation of the receptor can lead to
53 toxic outcomes[10,11]. AHR belongs to the basic helix-loop-helix (bHLH) PER-ARNT-
54 SIM (PAS) protein family[12]. The bHLH-PAS family, present ubiquitously in all
55 kingdoms of life, is predominantly engaged in detection of such diverse signals like
56 endogenous compounds, foreign chemicals, light and osmotic pressure, for example[2,13].
57 The numerous cognate ligands of AHR exhibit distinctive chemical structures and
58 include various compounds like polycyclic aromatic hydrocarbons (PAHs),
59 halogenated aromatic hydrocarbons (HAHs), polychlorinated biphenyls (PCBs), indole
60 derivatives, alkaloids, polyphenols and various pharmaceuticals, with dioxin (2,3,7,8-
61 tetrachlorodibenzo-*p*-dioxin or TCDD) being the archetypal exogenous ligand of
62 AHR[10]. As such, AHR is a key sensor which integrates numerous external and
63 endogenous chemical signals, and it is now considered a promising drug target[5].

64 In the absence of ligand, AHR resides in the cytoplasm forming a complex with
65 several other partners including heat shock protein 90 (Hsp90) and co-chaperones like
66 X-associated protein 2 (XAP2) and p23. Upon ligand binding, AHR undergoes

67 conformational changes leading to exposure of the nuclear localization signal (NLS)
68 and therefore triggering a translocation of the complex into the nucleus where AHR is
69 released and interacts with the AHR nuclear translocator (ARNT). The newly formed
70 heterodimer binds to so-called "dioxin-response element" (DRE) DNA sequences and
71 regulates the expression of target genes[14,15]. Previous structural work revealed how
72 AHR and ARNT heterodimerize and bind to DREs but provided no insights into the
73 AHR cytosolic complex organization or the structural basis of ligand-binding[16,17]. Here
74 we report the high-resolution cryo-electron microscopy (cryo-EM) structure of the
75 human ligand-bound AHR cytosolic complex as a snapshot of the first step of AHR
76 activation.
77
78 **Structure determination and architecture of the Hsp90-XAP2-AHR complex**
79 We co-expressed a fragment of human AHR (residues 1-437) in the presence of Hsp90,
80 XAP2 and p23, and purified the Hsp90-XAP2-AHR complex (Extended Data Fig. 1;
81 see Materials & Methods section for details). We combined the protein with indirubin
82 and using cryo-EM we obtained a 2.85 Å reconstruction of the ligand-bound AHR
83 cytosolic complex (Fig. 1a, Extended Data Fig. 2, 3 and Extended Data Table 1). The
84 structure reveals in its core a nucleotide-bound Hsp90 dimer (molecules denoted
85 Hsp90A and Hsp90B, respectively) adopting a closed conformation (Fig. 1b). Hsp90
86 forms extensive interaction sites with the PAS-B domain of AHR and XAP2 that
87 occupy the same side of Hsp90, reminiscent of the recently reported structure of the
88 glucocorticoid receptor (GR) complex (Hsp90-p23-GR)[18]. The structure of the bHLH
89 and PAS-A domains of AHR could not be unambiguously resolved during image
90 processing, indicating high dynamics of this region.

91 The PAS-B domain of AHR exhibits a canonical fold comprising a five-
92 stranded antiparallel β-sheet (Aβ, Bβ, Gβ, Hβ, Iβ) flanked by four consecutive α-
93 helices of variable lengths (Cα, Dα, Eα, Fα). Two additional short α-helices (Jα and
94 Kα) are present at the carboxy-terminus of the domain (Extended Data Fig. 4a, b). The
95 PAS-B domain is surrounded by additional extensions at both ends. The amino-
96 terminus includes an elongated 15-residue linker region interconnecting PAS-A and
97 PAS-B domains threaded through the Hsp90 dimer lumen (Extended Data Fig. 4a, c),
98 and the 40 amino acid residues connecting PAS-B to the carboxy-terminal
99 transactivation domain are forming a long loop that folds back onto the PAS-B domain
100 and participates in the interface with XAP2 (Extended Data Fig. 4a, c).
101 XAP2 interacts with both Hsp90 and AHR (Fig. 1a). The amino-terminal
102 domain of XAP2 is flexible and not well resolved in the general map. However, a
103 focused refinement of the cryo-EM data yielded a 4.07 Å resolution map with an
104 improved quality for the whole XAP2 protein and the carboxy-terminus of AHR
105 (Extended Data Fig. 2, 3). This is, to our knowledge, the first report of the full-length
106 XAP2 structure.
107
108 **Hsp90 forms a core of the complex**
109 All components of the complex are forming multiple interaction sites with each other
110 (Fig. 2a). Interaction between AHR and Hsp90 involves two interfaces. The first
111 interface encompasses 731 $Å^2$ of buried surface area (BSA) and includes AHR residues
112 from Aβ (I286-K290), Bβ (G299-D301), Gβ (M348), Hβ (Q364-N366) and Iβ (T382-
113 R384) and Hsp90 residues W312, R338, A339, P340 and F341 from the middle domain
114 of Hsp90A and Y596, T616, Y619 and M620 from the carboxy-terminal domain of

115 Hsp90B (Fig. 2b). Note that residues localized in the β-strand Aβ were previously
116 identified to be important for binding to Hsp90[19].

117 The second interface, confining 1,223 Å^2 of BSA, encompasses the 15-residue
118 linker between the PAS-A and PAS-B domains (P271-F285) that adopts an extended
119 conformation (Fig. 2c and Extended Data Fig. 5a). The linker is threaded through the
120 Hsp90 lumen formed by two Hsp90 molecules, thereby positioning the two PAS
121 domains on the opposite sides of the Hsp90 dimer (Extended Data Fig. 4c). The middle
122 and carboxy-terminal domains from both Hsp90 chains contribute to this interface by
123 providing a combination of hydrophobic and hydrophilic residues that are engaged in
124 multiple contacts with the backbone and side chains of AHR (Fig. 2c and Extended
125 Data Fig. 5a). We observed that AHR I277 and I280 occupy two small hydrophobic
126 pockets within the Hsp90 lumen, similar to L525 and L528 in the Hsp90-p23-GR[18] and
127 to V89 and V92 in the Hsp90-Cdc37-Cdk4[20] structures (Extended Data Fig. 5b, c).
128 Hence, the aliphatic nature of residues at these positions might be a hallmark for client
129 protein interaction with Hsp90. To validate these interactions, we mutated AHR
130 residues P275 to T282 into alanine, and monitored the impact of each mutation on
131 AHR-Hsp90 interaction using co-immunoprecipitation (Co-IP) assays. The results
132 revealed that the most drastic effects are obtained for the mutation of polar/charged
133 residues (S276, E279, R281), whereas the replacement of aliphatic residues (*e.g.*, I277,
134 L278 and I280) by alanine has little or no impact on the stability of the complex (Fig.
135 2d). These data suggest that aliphatic residues at positions 277 and 280 are important
136 for AHR recognition by Hsp90, whereas polar residues play a major role in the stability
137 of the interaction.
138
139

140 **Sodium molybdate mimics ATP γ-phosphate**

141 Sodium molybdate is known to stabilize the client-Hsp90 interaction and to inhibit gene
142 regulation by several nuclear receptors, including AHR[21], but its mode of action has
143 remained elusive. Our map displays an unambiguous density for the nucleotide binding
144 site in both Hsp90 chains (Extended Data Fig. 6a). We observed an unusually strong
145 density at the expected position of ATP γ-phosphate, much stronger than either α- or
146 β-phosphate (Extended Data Fig. 6b). We reasoned that, rather than the ATP γ-
147 phosphate, this density could be that of a molybdate ion that was used in the buffer for
148 complex preparation. We confirmed the presence of the molybdate inside the Hsp90
149 complex using mass spectrometry (Extended Data Fig. 6b). Therefore, we concluded
150 that the nucleotide binding site contains one ADP molecule generated from ATP
151 hydrolysis, one magnesium ion, and one molybdate ion that is also involved in the
152 chelation of the magnesium ion (Extended Data Fig. 6c). Hence, by mimicking the γ-
153 phosphate upon ATP hydrolysis, the molybdate ion would hold the Hsp90 dimer in the
154 ATP-induced closed state, thereby preventing AHR release[22] (Extended Data Fig. 6d).

155

156 **XAP2 stabilizes the complex**

157 The XAP2 protein is composed of an amino-terminal FKBP-type peptidyl-prolyl
158 cis/trans isomerase (PPIase) domain and a carboxy-terminal tetratricopeptide repeat
159 (TPR) helical domain[23,24] (Extended Data Fig. 7a). Our structure reveals the molecular
160 basis of XAP2 interaction with AHR and Hsp90 suggesting that the co-chaperone
161 serves as a brace interconnecting the other components of the complex (Fig. 1). The
162 PPIase and TPR domains are connected by a 12-residue long hinge and adopt an open
163 conformation without any direct inter-domain interaction (Extended Data Fig. 7a). The
164 interaction between AHR and XAP2 (1,165 $Å^2$) is mainly conferred through helix Fα

165 of PAS-B (I338, I341, K342) and the carboxy-terminal extension (F404, F406, G409-
166 E415, P420-I423) on the AHR side, while the XAP2 regions in contact with AHR
167 include the loop between strand βC' and helix αII (K66-K69) within the PPIase
168 domain, helix α2 and the loop α2-α3 of the TPR domain (Y203, I206, A207, M214-
169 P218, W223), and the linker region (W168) (Fig. 2e). Single amino acid substitutions
170 of three AHR residues at the interface with XAP2 (F404A, F406A, Y414A) had a
171 deleterious effect on XAP2 binding and led to reduction in Hsp90 interaction,
172 suggesting that XAP2 helps stabilize the whole complex (Fig. 2f).

173 XAP2 also forms an extensive contact with Hsp90 (1,027 $Å^2$) mainly via the
174 carboxy-terminal domain of the Hsp90B and the carboxy-terminal MEEVD motif, the
175 latter previously described[24] (Fig. 2g and Extended Data Fig. 7b, c). We mutated XAP2
176 residues at the interface with Hsp90 (Y248A in α4, W279A in loop α5-α6, or D317A,
177 D320A, F324A, I327A, F328A, S329A in helix α7). All mutations except S329A
178 abrogated or strongly reduced XAP2 binding and decreased the interaction between
179 AHR and Hsp90 (Fig. 2h). Together, these data suggest that Hsp90 and XAP2 are
180 interacting with AHR in a concerted fashion, XAP2 serving as a scaffolding protein
181 that stabilizes the cytosolic complex.

182

183 **Exploring AHR ligand-binding mechanism**

184 Indirubin (Fig. 3a) is a dietary-derived endogenous ligand of AHR produced from
185 tryptophan by intestinal bacteria[25]. We confirmed that indirubin is a potent activator of
186 AHR using cell-based assay (Fig. 3b). We also demonstrated that the purified Hsp90-
187 XAP2-AHR complex is active and binds indirubin, and that the presence of p23 is not
188 required for the interaction with the ligand (Fig. 3c). In the cryo-EM structure, the
189 indirubin binding site is unambiguously resolved within the PAS-B pocket (Fig. 3d).

190 The AHR ligand-binding pocket (LBP) takes a form of an elongated channel extending
191 perpendicularly between helices Cα and Fα (Fig. 3e). Remarkably, all elements of the
192 PAS-B secondary structure contribute to the LBP. The amino acids lining the cavity
193 are essentially hydrophobic (71%), including 8 aromatic (26%) and 14 aliphatic (45%)
194 residues. The LBP also comprises 9 polar amino acids (29%) but is essentially devoid
195 of charged residues (Extended Data Fig. 8a). The ligand occupies a volume of
196 approximately 220 $Å^3$ out of the total 682 $Å^3$ of the void volume as determined by
197 CASTp[26]. Accordingly, among 31 residues contributing the LBP, only 15 side chains
198 interact with indirubin (4.2 Å distance cut-off), indicating that a significant portion of
199 the cavity remains unoccupied (Fig. 3e and Extended Data Fig. 8b). The amino acids
200 involved in binding are located in the helix Fα region and involves residues originating
201 mainly from Bβ, Eα, Fα, Gβ, Hβ and Iβ, whereas on the other side of the pocket, the
202 volume delineated by Ab, Ca, Da and the loop Gβ-Hβ remains unoccupied (Fig. 3f).
203 Indirubin is a planar molecule with an asymmetric double indole structure. The indole
204 ring adjacent to helix Fα is engaged in many contacts with aliphatic and aromatic
205 pocket residues, namely F295, Y322, I325, C333, H337, I349 and F351. The second
206 indole ring is primarily involved in two hydrogen bonds with S365 and Q383 via its
207 carbonyl and amine moieties, respectively, in addition to several van der Waals contacts
208 with H291, P297, G321, F324, F351, L353 and V381. The hallmark of most AHR
209 ligands is their planarity. Our structure reveals that indirubin intercalates between
210 H291, F295 and P297 on one side of the ligand and I325, C333, F351 and L353 on the
211 other side, the two layers of residues exerting a critical role in the selection of indirubin,
212 and most likely other AHR ligands, based on its planar structure (Fig. 3g).
213 Multiple microsecond-long molecular dynamics (MD) simulations were carried
214 out to obtain an estimation of the relative energetic contributions of the LBP amino

215 acids to the interaction between indirubin and the PAS-B domain. The ligand remains
216 stable in the experimental pose and its presence stabilizes the bound conformation of
217 the protein (Extended Data Fig. 9a), maintaining stable contacts with most of the
218 residues observed in the cryo-EM structure (Extended Data Fig. 9b). Most of these
219 amino acids form favourable Lennard-Jones interactions with indirubin (Fig. 3h, upper
220 panel), in particular three residues with aromatic side chains form stable stacked or T-
221 shaped π-interactions with the indole rings of indirubin, namely H291, F295, and Y322
222 (Fig. 3g and Extended Data Fig. 9c, d). Our MD simulations confirms the substantial
223 implication of S365 and Q383 that are involved in stable hydrogen bonds with the
224 ligand (Fig. 3h, lower panel). Interestingly, the presence of indirubin reduces the overall
225 mobility of the backbone, in particular helices Cα, Dα, and the loop Gβ-Hβ (Fig. 3i, j).
226 The high dynamics of these two neighbouring regions in the AHR apo form suggests a
227 possible entry site for the ligand.

228 To experimentally validate the indirubin binding mode, we next mutated several
229 interacting residues and tested the binding properties of the purified mutant proteins
230 using nano-DSF (Extended Data Fig. 10a). We observed that substitution of H291,
231 F295 and S365 by alanine residue strongly reduced interaction with indirubin. Previous
232 reports revealed significant species differences in AHR response to indirubin
233 treatment[27]. Accordingly, we found that indirubin activates the human receptor more
234 potently than the rat and zebrafish homologues (Extended Data Fig. 10b). In full
235 support of these data, sequence (Extended Data Fig. 10c) and structural (Extended Data
236 Fig. 10d) analyses revealed that important stabilizing interactions provided by I349 and
237 V381 in human AHR are lost in the two other species which contain smaller residues
238 at these positions (T and A, respectively). In the fish receptor, additional replacement

239 of human S365 by an alanine residue is also very likely to explain a decrease in
240 indirubin response due to removal of a stabilizing hydrogen bond with the ligand.
241 Together, these observations support the notion that the indirubin binding site
242 adjacent to helix Fα is the primary anchor point of AHR ligands (Fig. 3e). This region
243 of the LBP harbours all structural and molecular determinants controlling ligand-
244 binding specificity, promiscuity and affinity by imposing strong geometrical
245 constraints in order to fit flat hydrophobic ligands (*e.g.*, PAHs, HAHs, etc.), whilst
246 providing polar amino acid residues with the potential to form hydrogen bonds with
247 hydrophilic chemical groups (*e.g.*, indole derivatives such as indirubin). Moreover, our
248 cryo-EM structure implies how AHR can accommodate also larger compounds. The
249 binding pocket extension towards helix Cα, referred to as the secondary binding site
250 (Fig. 3e), appears to be less geometrically constrained and contains a mix of aliphatic
251 and polar side chains that can engage in diverse types of contacts with compounds
252 featuring higher three-dimensional hindrance.
253
254 **Discussion**
255 Attempts to express, purify and solve the structure of the AHR ligand-binding domain,
256 either in the context of the full-length protein or as an isolated domain have failed for
257 many years. In our approach we reasoned that the reconstitution of the cytosolic
258 complex by co-expressing AHR with its main chaperone and co-chaperone would help
259 structurally and functionally stabilize the receptor. Using the Sf9 insect cell expression
260 system we were able to obtain a stable protein sample and solve the cryo-EM structure
261 of the indirubin-bound Hsp90-XAP2-AHR complex, at a resolution of 2.85 Å. This
262 study provides unprecedented structural insight into the details of the complex
263 assembly. AHR is forming extensive interaction sites with both protein partners,

264 partially explaining why all previous efforts to obtain the isolated PAS-B domain of
265 human AHR have failed. The structure reveals how Hsp90 and XAP2 are interacting
266 with AHR and maintaining the protein in the active form. The interactions involve not
267 only the linker interconnecting AHR PAS-A and B domains and PAS-B itself but also
268 the mainly disordered C-terminal extension of the protein. Our data shows that Hsp90
269 and XAP2 play indispensable role in the formation of the complex and our mutational
270 analysis confirmed that both proteins are working in a concerted fashion highlighting
271 the importance of using an entire protein complex to study the molecular mechanism
272 behind AHR-dependent signalling pathway. Our cryo-EM structure of the complex also
273 provides insight into a general mechanism of Hsp90-client recognition. Reminiscent of
274 the GR and Cdk4 structures[18,20], the AHR protein is threaded through the Hsp90 dimer
275 lumen, suggesting that the mechanism of client recognition is conserved between
276 different client proteins. Particularly, the position of key hydrophobic residues within
277 the extended linker region that interact with Hsp90 seems to be conserved across all
278 three available to date structures.

279 The structure explains previous observations suggesting that AHR Hsp90-
280 binding and ligand-binding sites overlap[19]. Indeed, residues localized within PAS-B
281 strands Aβ and Bβ are involved in the interactions with Hsp90 and the ligand. The
282 proposed mechanism of AHR activation includes ligand binding and induction of
283 conformational changes in AHR that could be directly sensed by Hsp90. Subsequently,
284 the signal would be transferred to the amino-terminal part of the receptor leading to
285 exposure of the NLS and triggering translocation to the nucleus. In case of AHR, the
286 NLS sequence is predicted to be localized within the first 60 amino acid residues.
287 Interestingly, as our complex was prepared in the presence of a ligand, the PAS-A
288 domain and the N-terminal part of AHR could not be convincingly placed in the

289 electron density suggesting high flexibility of this region. This part of the protein is
290 therefore exposed to the solvent and can interact freely with the nuclear transport
291 machinery including the importin complex.

292 One of the hallmarks of AHR is its remarkable promiscuity as the protein can
293 bind many different ligands. Our cryo-EM structure allowed us to explain the
294 mechanism behind selectivity towards certain types of ligands. Amino acid residues
295 involved in the formation of the primary binding site are responsible for the selectivity
296 towards molecules exhibiting a planar shape. Importantly, mutational analyses
297 confirmed the importance of residues involved in the formation of π-interactions with
298 the ligand. The structure also provides the explanation for receptor promiscuity. The
299 ligand-binding pocket with its relatively large interior cavity can accommodate
300 molecules of different sizes. Additionally, MD simulations provide insight into the
301 potential entry site of the ligand. The region localized between helices Cα and Dα,
302 exposed to the solvent and not obscured by the other components of the complex,
303 exhibits an elevated degree of flexibility in the simulations. It is also worth noting that
304 despite striking similarity in the overall structure, the PAS-A domain of AHR does not
305 possess any internal cavity that could serve as a potential binding pocket. Instead, the
306 interior of the domain is occluded by a presence of several bulky aromatic amino acid
307 residues supporting the hypothesis that in the course of its evolution the PAS-A domain
308 has been selected to serve primarily in protein-protein interactions.

309 Interestingly, previous comparison studies between human and rodent proteins
310 also revealed intriguing interspecies differences concerning AHR specificity. Despite
311 relatively high sequence homology between human and mouse AHR receptors (around
312 85%), the same chemical molecule can interact with the orthologous proteins with
313 drastically different affinities. For example, mouse AHR binds dioxin with a tenfold

314 higher affinity than its human counterpart[28]. In contrast, human AHR binds indirubin
315 and indoxyl sulphate with much higher affinity compared with the mouse receptor[29,30].
316 Intriguingly, human AHR protein appears to have evolved towards better recognition
317 of compounds belonging to indole family, which is most likely directly connected with
318 its involvement in the signalling in the human digestive track[31]. Despite high
319 promiscuity of the PAS-B domain, its specificity can still be finely tuned by small
320 variations in the protein sequence. Indeed, our analyses of AHR polymorphism between
321 human, rat and zebrafish combined with structural data explain the interspecies
322 differences observed in the binding affinity for indirubin. Our structure demonstrates
323 how microbiome-derived compound can interact with a human protein and modulate
324 host signalling pathways. Therefore, it provides an insight into the molecular
325 mechanism that emerged as a result of coevolution between human body and its
326 microbiome.

327 By disclosing the first atomic model for the AHR cytosolic complex and for its
328 PAS-B domain bound to a ligand, our structure has enabled us to explain many prior
329 biochemical observations. Moreover, it provides a strong rationale for future
330 mechanistic studies and development of novel pharmaceuticals that could find
331 applications in the treatment of many diseases, including irritable bowel disease,
332 allergic responses and cancer.

333 **Main References**

334 1. Gu, Y. Z., Hogenesch, J. B. & Bradfield, C. A. The PAS superfamily: sensors
335 of environmental and developmental signals. *Annu. Rev. Pharmacol. Toxicol.* **40**,
336 519–561 (2000).

337 2. Vazquez-Rivera, E. *et al.* The aryl hydrocarbon receptor as a model PAS sensor.
338 *Toxicol. Rep.* **9**, 1–11 (2022).

339 3. Larigot, L. *et al.* Aryl Hydrocarbon Receptor and Its Diverse Ligands and
340 Functions: An Exposome Receptor. *Annu. Rev. Pharmacol. Toxicol.* **62**, 383–404
341 (2022).

342 4. Esser, C. *et al.* Old Receptor, New Tricks-The Ever-Expanding Universe of
343 Aryl Hydrocarbon Receptor Functions. Report from the 4th AHR Meeting, 29–31
344 August 2018 in Paris, France. *Int. J. Mol. Sci.* **19**, E3603 (2018).

345 5. Safe, S., Cheng, Y. & Jin, U.-H. The Aryl Hydrocarbon Receptor (AhR) as a
346 Drug Target for Cancer Chemotherapy. *Curr. Opin. Toxicol.* **2**, 24–29 (2017).

347 6. Murray, I. A., Patterson, A. D. & Perdew, G. H. Aryl hydrocarbon receptor
348 ligands in cancer: friend and foe. *Nat. Rev. Cancer* **14**, 801–814 (2014).

349 7. Esser, C. & Rannug, A. The aryl hydrocarbon receptor in barrier organ
350 physiology, immunology, and toxicology. *Pharmacol. Rev.* **67**, 259–279 (2015).

351 8. Rothhammer, V. & Quintana, F. J. The aryl hydrocarbon receptor: an
352 environmental sensor integrating immune responses in health and disease. *Nat. Rev.*
353 *Immunol.* **19**, 184–197 (2019).

354 9. Dong, F. & Perdew, G. H. The aryl hydrocarbon receptor as a mediator of host-
355 microbiota interplay. *Gut Microbes* **12**, 1859812 (2020).

356 10. Denison, M. S. & Nagy, S. R. Activation of the aryl hydrocarbon receptor by
357 structurally diverse exogenous and endogenous chemicals. *Annu. Rev. Pharmacol.*
358 *Toxicol.* **43**, 309–334 (2003).
359 11. Okey, A. B. An aryl hydrocarbon receptor odyssey to the shores of toxicology:
360 the Deichmann Lecture, International Congress of Toxicology-XI. *Toxicol. Sci. Off.*
361 *J. Soc. Toxicol.* **98**, 5–38 (2007).
362 12. Nebert, D. W. Aryl hydrocarbon receptor (AHR): 'pioneer member' of the
363 basic-helix/loop/helix per-Arnt-sim (bHLH/PAS) family of 'sensors' of foreign and
364 endogenous signals. *Prog. Lipid Res.* **67**, 38–57 (2017).
365 13. Wu, D. & Rastinejad, F. Structural characterization of mammalian bHLH-PAS
366 transcription factors. *Curr. Opin. Struct. Biol.* **43**, 1–9 (2017).
367 14. Beischlag, T. V., Luis Morales, J., Hollingshead, B. D. & Perdew, G. H. The
368 aryl hydrocarbon receptor complex and the control of gene expression. *Crit. Rev.*
369 *Eukaryot. Gene Expr.* **18**, 207–250 (2008).
370 15. Denison, M. S., Soshilov, A. A., He, G., DeGroot, D. E. & Zhao, B. Exactly the
371 same but different: promiscuity and diversity in the molecular mechanisms of action
372 of the aryl hydrocarbon (dioxin) receptor. *Toxicol. Sci. Off. J. Soc. Toxicol.* **124**, 1–
373 22 (2011).
374 16. Seok, S.-H. *et al.* Structural hierarchy controlling dimerization and target DNA
375 recognition in the AHR transcriptional complex. *Proc. Natl. Acad. Sci. U. S. A.* **114**,
376 5431–5436 (2017).
377 17. Schulte, K. W., Green, E., Wilz, A., Platten, M. & Daumke, O. Structural Basis
378 for Aryl Hydrocarbon Receptor-Mediated Gene Activation. *Structure* **25**, 1025-
379 1033.e3 (2017).

380 18. Noddings, C. M., Wang, R. Y.-R., Johnson, J. L. & Agard, D. A. Structure of
381 Hsp90-p23-GR reveals the Hsp90 client-remodelling mechanism. *Nature* **601**, 465–
382 469 (2022).
383 19. Soshilov, A. & Denison, M. S. Ligand displaces heat shock protein 90 from
384 overlapping binding sites within the aryl hydrocarbon receptor ligand-binding
385 domain. *J. Biol. Chem.* **286**, 35275–35282 (2011).
386 20. Verba, K. A. *et al.* Atomic structure of Hsp90-Cdc37-Cdk4 reveals that Hsp90
387 traps and stabilizes an unfolded kinase. *Science* **352**, 1542–1547 (2016).
388 21. Heid, S. E., Pollenz, R. S. & Swanson, H. I. Role of heat shock protein 90
389 dissociation in mediating agonist-induced activation of the aryl hydrocarbon
390 receptor. *Mol. Pharmacol.* **57**, 82–92 (2000).
391 22. Soshilov, A. A., Motta, S., Bonati, L. & Denison, M. S. Transitional States in
392 Ligand-Dependent Transformation of the Aryl Hydrocarbon Receptor into Its DNA-
393 Binding Form. *Int. J. Mol. Sci.* **21**, E2474 (2020).
394 23. Linnert, M. *et al.* The FKBP-type domain of the human aryl hydrocarbon
395 receptor-interacting protein reveals an unusual Hsp90 interaction. *Biochemistry* **52**,
396 2097–2107 (2013).
397 24. Morgan, R. M. L. *et al.* Structure of the TPR domain of AIP: lack of client
398 protein interaction with the C-terminal α-7 helix of the TPR domain of AIP is
399 sufficient for pituitary adenoma predisposition. *PloS One* **7**, e53339 (2012).
400 25. Hubbard, T. D., Murray, I. A. & Perdew, G. H. Indole and Tryptophan
401 Metabolism: Endogenous and Dietary Routes to Ah Receptor Activation. *Drug
402 Metab. Dispos. Biol. Fate Chem.* **43**, 1522–1535 (2015).
403 26. Tian, W., Chen, C., Lei, X., Zhao, J. & Liang, J. CASTp 3.0: computed atlas of
404 surface topography of proteins. *Nucleic Acids Res.* **46**, W363–W367 (2018).

405 27. Faber, S. C., Soshilov, A. A., Giani Tagliabue, S., Bonati, L. & Denison, M. S.
406 Comparative In Vitro and In Silico Analysis of the Selectivity of Indirubin as a
407 Human Ah Receptor Agonist. *Int. J. Mol. Sci.* **19**, E2692 (2018).
408 28. Ramadoss, P. & Perdew, G. H. Use of 2-azido-3-[125I]iodo-7,8-
409 dibromodibenzo-p-dioxin as a probe to determine the relative ligand affinity of
410 human versus mouse aryl hydrocarbon receptor in cultured cells. *Mol. Pharmacol.*
411 **66**, 129–136 (2004).
412 29. Flaveny, C. A., Murray, I. A., Chiaro, C. R. & Perdew, G. H. Ligand selectivity
413 and gene regulation by the human aryl hydrocarbon receptor in transgenic mice. *Mol.*
414 *Pharmacol.* **75**, 1412–1420 (2009).
415 30. Schroeder, J. C. *et al.* The uremic toxin 3-indoxyl sulfate is a potent endogenous
416 agonist for the human aryl hydrocarbon receptor. *Biochemistry* **49**, 393–400 (2010).
417 31. Hubbard, T. D. *et al.* Adaptation of the human aryl hydrocarbon receptor to
418 sense microbiota-derived indoles. *Sci. Rep.* **5**, 12689 (2015).

Main Figures

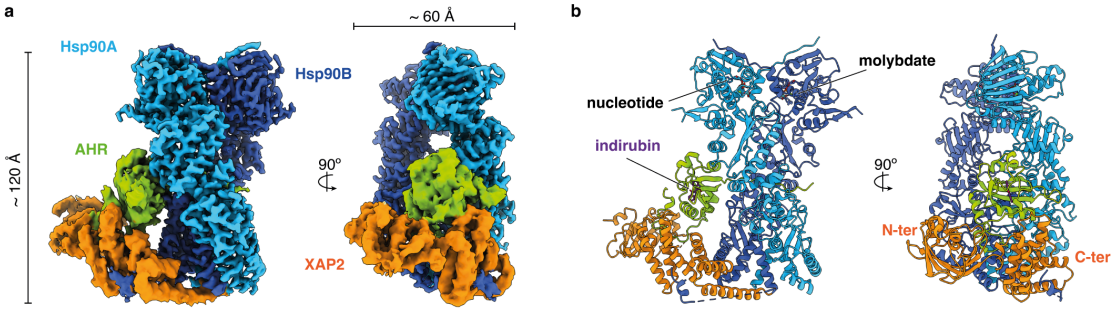

Fig. 1 | Overall architecture of the agonist-bound cytosolic complex of AHR. a, A composite map of the Hsp90-XAP2-AHR complex in two orthogonal views. Hsp90A (light blue), Hsp90B (dark blue), XAP2 (orange), AHR (green). The same color scheme is used throughout the manuscript unless stated otherwise. **b**, The atomic model of the complex in cartoon representation (the orientation of the molecule same as in a). Nucleotide molecule, molybdate ion and indirubin ligand are shown in sticks. The corresponding binding sites are indicated.

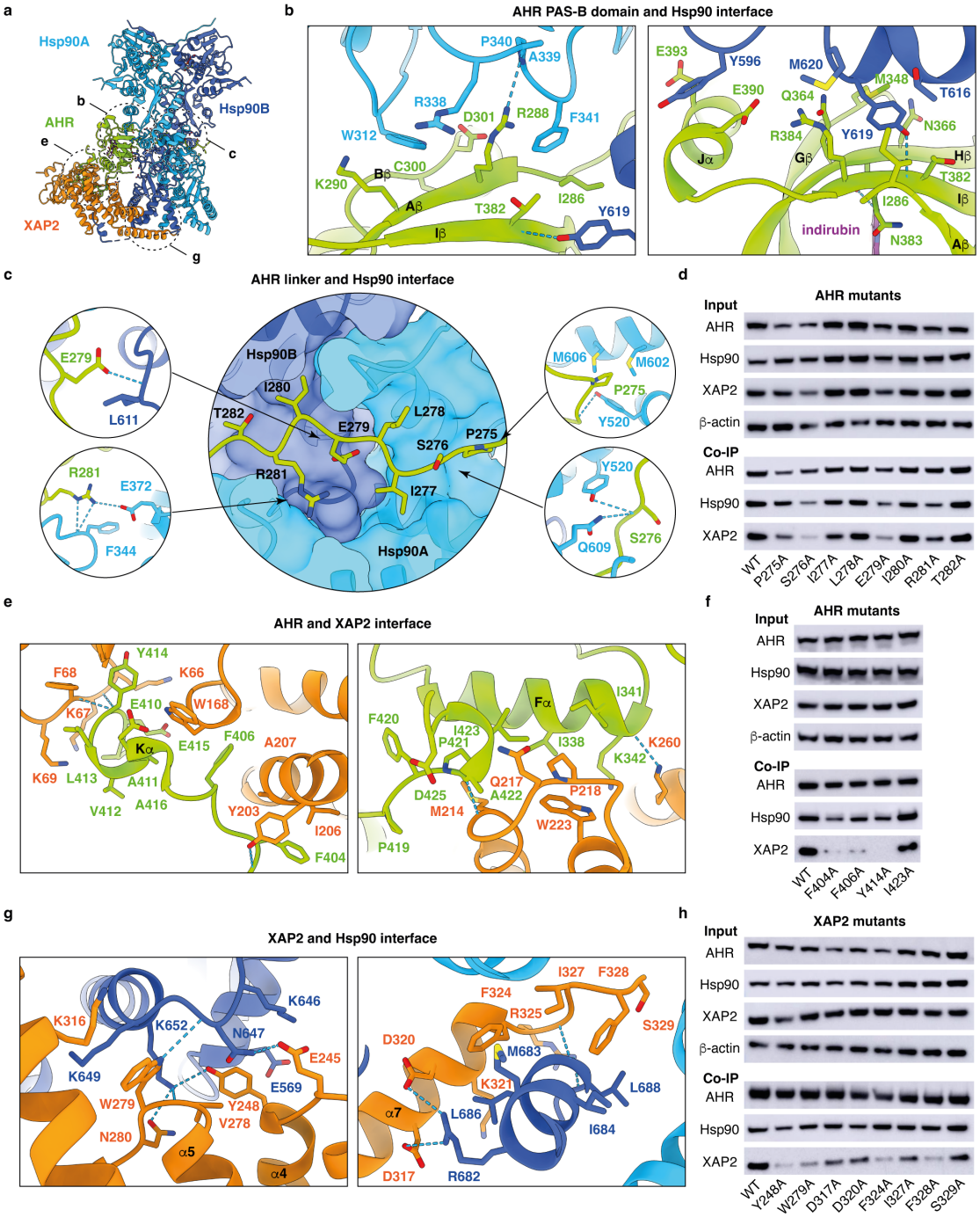

Fig. 2 | Details of the Hsp90-XAP2-AHR complex organization. a, An overall view of the complex with indicated location of the interaction sites between the proteins. The dashed circles indicate the contacts described in the following panels. **b**, Two principal interaction sites between Hsp90 and PAS-B domains. **c**, Close-up view of the AHR linker that is threaded through the Hsp90 lumen and interacts with both molecules. The residues for which we observed the strongest effect of the mutation on the complex stability in Co-IP assay are shown in small circles. **d**, Results of Co-IP for the mutation of the AHR linker. **e**, Close-up view of the interaction sites between AHR and XAP2. **f**, Co-IP results for the AHR mutants at the interface with XAP2. **g**, Interactions between XAP2 and Hsp90. **h**, Co-IP results for the XAP2 mutants at the interface with Hsp90. Co-IP experiments were successfully reproduced at least three times.

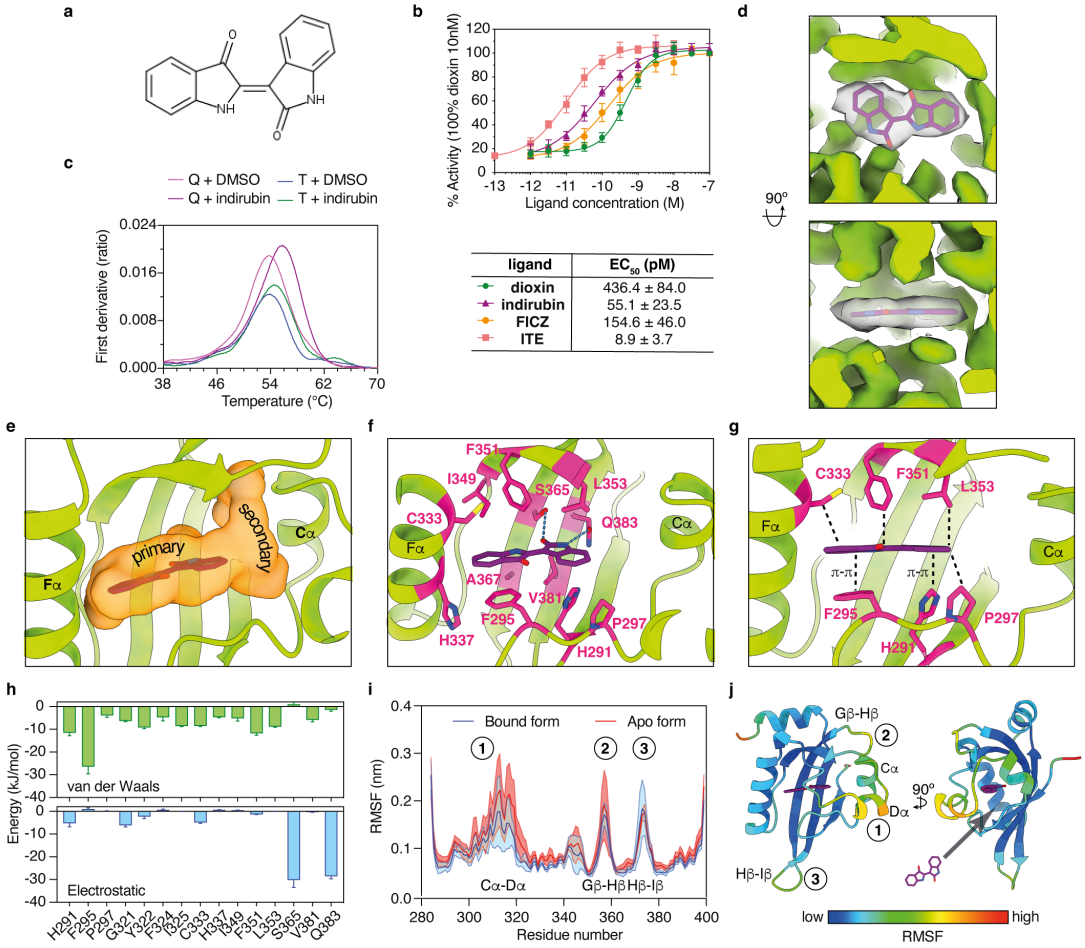

Fig. 3 | Characterization of the interaction with indirubin. a, Structure of indirubin. **b**, Results of cell-based activity assays for four AHR ligands. The results are representative of independent biological replicates (n=4). **c**, Two orthogonal views of the cryo-EM map for the region corresponding to indirubin. **d**, Nano-DSF analysis of the interaction between Hsp90-XAP2-AHR (T, ternary) and Hsp90-XAP2-p23-AHR (Q, quaternary) complexes and indirubin. The results are representative of independent biological replicates (n=3). **e**, Close-up view of the ligand binding pocket (LBP) shown as an orange surface. **f**, Close-up view of the indirubin binding site. Residues interacting with the ligand are shown as magenta sticks. Hydrogen bonds are indicated as dashed blue lines. **g**, Two layers of LBP residues determine AHR ligand selectivity. **h**, Van der Waals (upper panel) and electrostatic (lower panel) interaction energies between indirubin and the residues of the protein. Values represent averages between 8 MD simulations; error bars are standard deviation between the replicas. **i**, Root mean square fluctuations (RMSF) of the protein backbone in bound (blue) and apo (red) conformations. Values are averages over the MD simulations; shaded areas represent the standard deviation between the replicas. **j**, Two views of the PAS-B domain colored according to the RMSF. The location of a potential ligand entry site to LBP is indicated.

Acknowledgments We acknowledge The Midlands Regional CryoEM Facility at the Leicester Institute of Structural and Chemical Biology (LISCB), major funding from MRC (MC_PC_17136). We thank Rémi Freydier and Léa Causse (AETE-ISO platform, OSU-OREME/Université de Montpellier) for the determination of molybdate concentrations in our samples. Centre de Biologie Structurale is a member of the French Infrastructure for Integrated Structural Biology (FRISBI), a national infrastructure supported by the French National Research Agency (ANR-10-INBS-04-01). Native MS experiments were carried out at the Montpellier Proteomics Platform (PPM, Biocampus Montpellier) supported by the regional funds FEDER/Région Occitanie, MUSE and Labex EpiGenMed. The work was supported by funding from European Union's Horizon 2020 research and innovation program grant No. 825489 (W.B.), ATIP-Avenir 2020 grant No. R20059SP (J.G.) and the ANR (the French National Research Agency) "Investissements d'avenir" programme reference No. ANR-16-IDEX-0006 (J.G.).

Author Contributions J.G. and W.B. initiated and supervised project. J.G. and L.G. purified protein samples. J.L.K.H. and A.A. prepared EM grids. J.G., J.L.K.H., C.G.S., J.B. and A.A. collected and analyzed EM data. J.G., H.S.K. and W.B. performed model building and refinement. J.G. and W.B. analyzed the structure. L.G. and P.G. carried out biochemical analysis. M.G., Ab.Bo. and P.B. performed cell-based assays. M.P. and Al.Ba. performed molecular dynamics simulations. C.B. collected and analyzed MS data. J.G. and W.B. principally wrote the manuscript which was finalised with input from all authors.

Competing interests The authors declare no competing financial interests.

Correspondence and requests for materials should be addressed to Jakub Gruszczyk and William Bourguet.