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

Authors: Jakub Gruszczyk[*,1], Loïc 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], Aurélie Ancelin[1], Cherine Bechara[5,6], Alessandro Barducci[1], Patrick Balaguer[3], William Bourguet[*,1]

Affiliations:

[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.

[*]Corresponding authors. Email: jakub.gruszczyk@cbs.cnrs.fr (JG); william.bourguet@cbs.cnrs.fr (WB)

Abstract: Aryl hydrocarbon receptor (AHR) is a transcription factor recognizing numerous ligands including pollutants, natural products and metabolites, and therefore mediating a broad spectrum of (patho)physiological processes. However, the scarcity of structural data precludes understanding of the mechanisms governing its specificity and promiscuity. Our 2.85 Å cryo-EM structure of the indirubin-bound AHR complex with Hsp90 and XAP2 reveals a closed conformation Hsp90 dimer with AHR threaded through its lumen and XAP2 serving as a brace. We pinpoint residues crucial for the formation of the complex. We also provide the long-awaited structure of the human AHR PAS-B domain revealing the details of the interaction with its agonist. Our findings rationalize extensive prior biochemical data and provide a framework for future mechanistic studies and structure-guided drug design.

One-Sentence Summary: Structure of the key environmental and physiological sensor sheds light on the initial activation step.

Short title: Structure of the Hsp90-XAP2-AHR complex

Main Text: Aryl hydrocarbon receptor (AHR) is a ligand-dependent transcription factor involved in the regulation of genes governing many (patho)physiological processes including cell growth and differentiation, cell cycle and migration, apoptosis, hematopoiesis and carcinogenesis (*1–4*). Importantly, activation of AHR in the gastrointestinal tract participates in signaling between the enteric microflora and our immune system (*5–7*). Moreover, AHR is one of the crucial chemosensory proteins associated with the metabolism of various xenobiotics (*8*, *9*). AHR belongs to the basic helix-loop-helix (bHLH) PER-ARNT-SIM (PAS) protein family (*10*). The bHLH-PAS family, present ubiquitously in all kingdoms of life, is predominantly engaged in detection of such diverse signals like endogenous compounds, foreign chemicals, light and osmotic pressure, for example (*11*, *12*). The numerous cognate ligands of AHR exhibit distinctive chemical structures and include various compounds like polycyclic aromatic hydrocarbons (PAHs), halogenated aromatic hydrocarbons (HAHs), polychlorinated biphenyls (PCBs), indole derivatives, alkaloids, polyphenols and various pharmaceuticals, with dioxin (2,3,7,8-tetrachlorodibenzo-*p*-dioxin or TCDD) being the archetypal exogenous ligand of AHR (*8*). In the absence of ligand, AHR resides in the cytoplasm forming a complex with several other partners including heat shock protein 90 (Hsp90) and co-chaperones like X-associated protein 2 (XAP2) and p23. Upon ligand binding, AHR undergoes conformational changes leading to exposure of the nuclear localization signal and therefore triggering a translocation of the complex into the nucleus where AHR is released and interacts with the AHR nuclear translocator (ARNT). The newly formed heterodimer binds to specific DNA response elements and regulates the expression of target genes (*13*, *14*). So far, the structural basis of either the cytosolic complex organization or the AHR interaction with its ligands remains elusive.

To shed more light on these pivotal aspects of AHR physiology, we focused our attention on the cytosolic complex of AHR. We co-expressed a fragment of human AHR (residues 1-437) in

the presence of Hsp90, XAP2 and p23, and purified the Hsp90-XAP2-AHR complex (**fig. S1**; see Materials & Methods section for details). We combined the protein with indirubin and using cryo-electron microscopy (cryo-EM) we obtained a 2.85 Å cryo-EM reconstruction of the ligand-bound AHR cytosolic complex (**Fig. 1A**, **fig. S2, S3 and S4**, **table S1**). The structure reveals in its core a nucleotide-bound Hsp90 dimer (molecules denoted Hsp90A and Hsp90B, respectively) adopting a closed conformation (**Fig. 1B**). Hsp90 forms extensive interaction sites with the PAS-B domain of AHR and XAP2 that occupy the same side of Hsp90, reminiscent of the recently reported structure of the glucocorticoid receptor (GR) complex (Hsp90-p23-GR) (*15*). The structure of the bHLH and PAS-A domains of AHR could not be unambiguously resolved during image processing, indicating high dynamics of this region. PAS-B domain of AHR exhibits a canonical fold comprising a five-stranded antiparallel β-sheet (Aβ, Bβ, Gβ, Hβ, Iβ) flanked by four consecutive α-helices of variable lengths (Cα, Dα, Eα, Fα). Two additional short α-helices (Jα and Kα) are present at the carboxy-terminus of the domain (**fig. S5, A and B**). The PAS-B domain is surrounded by additional extensions at both ends. The amino-terminus includes an elongated 15-residue linker region interconnecting PAS-A and PAS-B domains threaded through the Hsp90 dimer lumen (**fig. S5, A and C**), and the 40 amino acid residues connecting PAS-B to the carboxy-terminal transactivation domain are forming a long loop that folds back onto the PAS-B domain and participates in the interface with XAP2 (**fig. S5, A and C**). XAP2 interacts with both Hsp90 and AHR (**Fig. 1A**). The amino-terminal domain of XAP2 is flexible and not well resolved in the general map. However, a focused refinement of the cryo-EM data yielded a 4.07 Å resolution map with an improved quality for the whole XAP2 protein and the carboxy-terminus of AHR (**fig. S2 and S3**).

All components of the complex are forming multiple interaction sites with each other (**Fig. 2A**). Interaction between AHR and Hsp90 involves two interfaces. The first interface

encompasses 731 $Å^2$ of buried surface area (BSA) and includes AHR residues from Aβ (I286-K290), Bβ (G299-D301), Gβ (M348), Hβ (Q364-N366) and Iβ (T382-R384) and Hsp90 residues W312, R338, A339, P340 and F341 from the middle domain of Hsp90A and Y596, T616, Y619 and M620 from the carboxy-terminal domain of Hsp90B (**Fig. 2B**). Note that residues localized in the β-strand Aβ were previously identified to be important for binding to Hsp90 (*16*). The second interface, confining 1,223 $Å^2$ of BSA, encompasses the 15-residue linker between the PAS-A and PAS-B domains (P271-F285) that adopts an extended conformation (**Fig. 2C and fig. S6A**). The linker is threaded through the Hsp90 lumen formed by two Hsp90 molecules, thereby positioning the two PAS domains on the opposite sides of the Hsp90 dimer (**fig. S5C**). The middle and carboxy-terminal domains from both Hsp90 chains contribute to this interface by providing a combination of hydrophobic and hydrophilic residues that are engaged in multiple contacts with the backbone and side chains of AHR (**Fig. 2C and fig. S6A**). We observed that AHR I277 and I280 occupy two small hydrophobic pockets within the Hsp90 lumen, similar to L525 and L528 in the Hsp90-p23-GR (*15*) and to V89 and V92 in the Hsp90-Cdc37-Cdk4 (*17*) structures (**fig. S6, B and C**). Hence, the aliphatic nature of residues at these positions might be a hallmark for client protein interaction with Hsp90. To validate these interactions, we mutated AHR residues P275 to T282 into alanine, and monitored the impact of each mutation on AHR-Hsp90 interaction using co-immunoprecipitation (Co-IP) assays. The results revealed that the most drastic effects are obtained for the mutation of polar/charged residues (S276, E279, R281), whereas the replacement of aliphatic residues (*e.g.*, I277, L278 and I280) by alanine has little or no impact on the stability of the complex (**Fig. 2D**). These data suggest that aliphatic residues at positions 277 and 280 are important for AHR recognition by Hsp90, whereas polar residues play a major role in the stability of the interaction.

Sodium molybdate is known to stabilize the AHR-Hsp90 interaction and to inhibit gene regulation by several nuclear receptors, including AHR (*18*), but its mode of action has remained elusive. Our map displays an unambiguous density for the nucleotide binding site in both Hsp90 chains (**fig. S7A**). We observed an unusually strong density at the expected position of ATP γ-phosphate, much stronger than either α- or β-phosphate (**fig. S7B**). We reasoned that, rather than the ATP γ-phosphate, this density could be that of a molybdate ion that was used in the buffer for complex preparation. We confirmed the presence of the molybdate inside the Hsp90 complex using mass spectrometry (**fig. S7B**). Therefore, we concluded that the nucleotide binding site contains one ADP molecule generated from ATP hydrolysis, one magnesium ion, and one molybdate ion that is also involved in the chelation of the magnesium ion (**fig. S7C**). Hence, by mimicking the γ-phosphate upon ATP hydrolysis, the molybdate ion would hold the Hsp90 dimer in the ATP-induced closed state, thereby preventing AHR release (*19*).

XAP2 is composed of an amino-terminal FKBP-type peptidyl-prolyl cis/trans isomerase (PPIase) domain and a carboxy-terminal tetratricopeptide repeat (TPR) helical domain (*20*, *21*) (**fig. S8A**). Our structure reveals the molecular basis of XAP2 interaction with AHR and Hsp90 suggesting that the co-chaperone serves as a brace interconnecting the other components of the complex (**Fig. 1**). The PPIase and TPR domains are connected by a 12-residue long hinge and adopt an open conformation without any direct inter-domain interaction (**fig. S8A**). The interaction between AHR and XAP2 (1,165 $Å^2$) is mainly conferred through helix Fα of PAS-B (I338, I341, K342) and the carboxy-terminal extension (F404, F406, G409-E415, P420-I423) on the AHR side, while the XAP2 regions in contact with AHR include the loop between strand βC' and helix αII (K66-K69) within the PPIase domain, helix α2 and the loop α2-α3 of the TPR domain (Y203, I206, A207, M214-P218, W223), and the linker region (W168) (**Fig. 2E**). Single amino acid substitutions of three AHR residues at the interface with XAP2 (F404A, F406A,

Y414A) had a deleterious effect on XAP2 binding and led to reduction in Hsp90 interaction, suggesting that XAP2 helps stabilize the whole complex (**Fig. 2F**). XAP2 also forms an extensive contact with Hsp90 (1,027 $Å^2$) mainly via the carboxy-terminal domain of the Hsp90B and the carboxy-terminal MEEVD motif, the latter previously described (*21*) (**Fig. 2G and fig. S8, B and C**). We mutated XAP2 residues at the interface with Hsp90 (Y248A in α4, W279A in loop α5-α6, or D317A, D320A, F324A, I327A, F328A, S329A in helix α7). All mutations except S329A abrogated or strongly reduced XAP2 binding and decreased the interaction between AHR and Hsp90 (**Fig. 2H**). Together, these data suggest that Hsp90 and XAP2 are interacting with AHR in a concerted fashion, XAP2 serving as a scaffolding protein that stabilizes the cytosolic complex.

Indirubin (**Fig. 3A**) is a dietary-derived endogenous ligand of AHR produced from tryptophan by intestinal bacteria (*22*). We confirmed that indirubin is a potent activator of AHR using cell-based assay (**Fig. 3B**). We also demonstrated that the purified Hsp90-XAP2-AHR complex is active and binds indirubin, and that the presence of p23 is not required for the interaction with the ligand (**Fig. 3C**). In the cryo-EM structure, the indirubin binding site is unambiguously resolved within the PAS-B pocket (**Fig. 3D**). The AHR ligand-binding pocket (LBP) takes a form of an elongated channel extending perpendicularly between helices Cα and Fα (**Fig. 3E**). Remarkably, all elements of the PAS-B secondary structure contribute to the LBP. The amino acids lining the cavity are essentially hydrophobic (71%), including 8 aromatic (26%) and 14 aliphatic (45%) residues. The LBP also comprises 9 polar amino acids (29%) but is essentially devoid of charged residues (**fig. S9A**). The ligand occupies a volume of approximately 220 $Å^3$ out of the total 682 $Å^3$ of the void volume as determined by CASTp (*23*). Accordingly, among 31 residues contributing the LBP, only 15 side chains interact with indirubin (4.2 Å distance cut-off), indicating that a significant portion of the cavity remains

unoccupied (**Fig. 3E and fig. S9B**). The amino acids involved in binding are located in the helix Fα region and involves residues originating mainly from Bβ, Eα, Fα, Gβ, Hβ and Iβ, whereas on the other side of the pocket, the volume delineated by Aβ, Cα, Dα and the loop Gβ-Hβ remains unoccupied (**Fig. 3F**). Indirubin is a planar molecule with an asymmetric double indole structure. The indole ring adjacent to helix Fα is engaged in many contacts with aliphatic and aromatic pocket residues, namely F295, Y322, I325, C333, H337, I349 and F351. The second indole ring is primarily involved in two hydrogen bonds with S365 and Q383 via its carbonyl and amine moieties, respectively, in addition to several van der Waals contacts with H291, P297, G321, F324, F351, L353 and V381. The hallmark of most AHR ligands is their planarity. Our structure reveals that indirubin intercalates between H291, F295 and P297 on one side of the ligand and I325, C333, F351 and L353 on the other side, the two layers of residues exerting a critical role in the selection of indirubin, and most likely other AHR ligands, based on its planar structure (**Fig. 3G**). Multiple microsecond-long molecular dynamics (MD) simulations were carried out to obtain an estimation of the relative energetic contributions of the LBP amino acids to the interaction between indirubin and the PAS-B domain. The ligand remains stable in the experimental pose and its presence stabilizes the bound conformation of the protein (**fig. S10A**), maintaining stable contacts with most of the residues observed in the cryo-EM structure (**fig. S10B**). Most of these amino acids form favorable Lennard-Jones interactions with indirubin (**Fig. 3H**, upper panel), in particular three residues with aromatic side chains form stable stacked or T-shaped π-interactions with the indole rings of indirubin, namely H291, F295, and Y322 (**Fig. 3G and fig. S10, C and D**). Our MD simulations confirms the substantial implication of S365 and Q383 that are involved in stable hydrogen bonds with the ligand (**Fig. 3H**, lower panel). Interestingly, the presence of indirubin reduces the overall mobility of the backbone, in particular

helices Cα, Dα, and the loop Gβ-Hβ (**Fig. 3, I and J**). The high dynamics of these two neighboring regions in the AHR apo form suggests a possible entry site for the ligand.

Together, these observations support the notion that the indirubin binding site adjacent to helix Fα is the primary anchor point of AHR ligands (**Fig. 3E**). This region of the LBP harbors all structural and molecular determinants controlling ligand-binding specificity, promiscuity and affinity by imposing strong geometrical constraints in order to fit flat hydrophobic ligands (*e.g.*, PAHs, HAHs, etc.), whilst providing polar amino acid residues with the potential to form hydrogen bonds with hydrophilic chemical groups (*e.g.*, indole derivatives such as indirubin). Moreover, our cryo-EM structure implies how AHR can accommodate also larger compounds. The binding pocket extension towards helix Cα, referred to as the secondary binding site (**Fig. 3E**), appears to be less geometrically constrained and contains a mix of aliphatic and polar side chains that can engage in diverse types of contacts with compounds featuring higher three-dimensional hindrance.

To experimentally validate the indirubin binding mode, we next mutated several interacting residues and tested the binding properties of the purified mutant proteins using nano-DSF (**fig. S11A**). We observed that substitution of H291, F295 and S365 by alanine residue strongly reduced interaction with indirubin. Previous reports revealed significant species differences in AHR response to indirubin treatment (*24*). Accordingly, we found that indirubin activates the human receptor more potently than the rat and zebrafish homologues (**fig. S11B**). In full support of these data, sequence (**fig. S11C**) and structural (**fig. S11D**) analyses revealed that important stabilizing interactions provided by I349 and V381 in human AHR are lost in the two other species which contain smaller residues at these positions (T and A, respectively). In the fish receptor, additional replacement of human S365 by an alanine residue is also very likely to

explain a decrease in indirubin response due to removal of a stabilizing hydrogen bond with the ligand.

As a whole, this study provides unprecedented structural details on the initiating event leading to AHR activation. By disclosing the atomic model for the AHR cytosolic complex and for its PAS-B domain bound to a ligand, our structure has enabled us to explain several prior biochemical observations and provides a strong rationale for future mechanistic studies and drug development.

References and Notes

1. L. Larigot, L. Benoit, M. Koual, C. Tomkiewicz, R. Barouki, X. Coumoul, Aryl Hydrocarbon Receptor and Its Diverse Ligands and Functions: An Exposome Receptor. *Annu. Rev. Pharmacol. Toxicol.* **62**, 383–404 (2022).

2. C. Esser, B. P. Lawrence, D. H. Sherr, G. H. Perdew, A. Puga, R. Barouki, X. Coumoul, Old Receptor, New Tricks-The Ever-Expanding Universe of Aryl Hydrocarbon Receptor Functions. Report from the 4th AHR Meeting, 29¯31 August 2018 in Paris, France. *Int. J. Mol. Sci.* **19**, E3603 (2018).

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

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

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

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

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

8. M. S. Denison, S. R. Nagy, Activation of the aryl hydrocarbon receptor by structurally diverse exogenous and endogenous chemicals. *Annu. Rev. Pharmacol. Toxicol.* **43**, 309–334 (2003).

9. A. B. Okey, An aryl hydrocarbon receptor odyssey to the shores of toxicology: the Deichmann Lecture, International Congress of Toxicology-XI. *Toxicol. Sci. Off. J. Soc. Toxicol.* **98**, 5–38 (2007).

10. D. W. Nebert, Aryl hydrocarbon receptor (AHR): "pioneer member" of the basic-helix/loop/helix per-Arnt-sim (bHLH/PAS) family of "sensors" of foreign and endogenous signals. *Prog. Lipid Res.* **67**, 38–57 (2017).

11. D. Wu, F. Rastinejad, Structural characterization of mammalian bHLH-PAS transcription factors. *Curr. Opin. Struct. Biol.* **43**, 1–9 (2017).

12. E. Vazquez-Rivera, B. Rojas, J. C. Parrott, A. L. Shen, Y. Xing, P. R. Carney, C. A. Bradfield, The aryl hydrocarbon receptor as a model PAS sensor. *Toxicol. Rep.* **9**, 1–11 (2022).

13. T. V. Beischlag, J. Luis Morales, B. D. Hollingshead, G. H. Perdew, The aryl hydrocarbon receptor complex and the control of gene expression. *Crit. Rev. Eukaryot. Gene Expr.* **18**, 207–250 (2008).

14. M. S. Denison, A. A. Soshilov, G. He, D. E. DeGroot, B. Zhao, Exactly the same but different: promiscuity and diversity in the molecular mechanisms of action of the aryl hydrocarbon (dioxin) receptor. *Toxicol. Sci. Off. J. Soc. Toxicol.* **124**, 1–22 (2011).

15. C. M. Noddings, R. Y.-R. Wang, J. L. Johnson, D. A. Agard, Structure of Hsp90-p23-GR reveals the Hsp90 client-remodelling mechanism. *Nature*. **601**, 465–469 (2022).

16. A. Soshilov, M. S. Denison, Ligand displaces heat shock protein 90 from overlapping binding sites within the aryl hydrocarbon receptor ligand-binding domain. *J. Biol. Chem.* **286**, 35275–35282 (2011).

17. K. A. Verba, R. Y.-R. Wang, A. Arakawa, Y. Liu, M. Shirouzu, S. Yokoyama, D. A. Agard, Atomic structure of Hsp90-Cdc37-Cdk4 reveals that Hsp90 traps and stabilizes an unfolded kinase. *Science*. **352**, 1542–1547 (2016).

18. S. E. Heid, R. S. Pollenz, H. I. Swanson, Role of heat shock protein 90 dissociation in mediating agonist-induced activation of the aryl hydrocarbon receptor. *Mol. Pharmacol.* **57**, 82–92 (2000).

19. A. A. Soshilov, S. Motta, L. Bonati, M. S. Denison, Transitional States in Ligand-Dependent Transformation of the Aryl Hydrocarbon Receptor into Its DNA-Binding Form. *Int. J. Mol. Sci.* **21**, E2474 (2020).

20. M. Linnert, Y.-J. Lin, A. Manns, K. Haupt, A.-K. Paschke, G. Fischer, M. Weiwad, C. Lücke, The FKBP-type domain of the human aryl hydrocarbon receptor-interacting protein reveals an unusual Hsp90 interaction. *Biochemistry*. **52**, 2097–2107 (2013).

21. R. M. L. Morgan, L. C. Hernández-Ramírez, G. Trivellin, L. Zhou, S. M. Roe, M. Korbonits, C. Prodromou, Structure of the TPR domain of AIP: lack of client protein interaction with the C-terminal α-7 helix of the TPR domain of AIP is sufficient for pituitary adenoma predisposition. *PloS One*. **7**, e53339 (2012).

22. T. D. Hubbard, I. A. Murray, G. H. Perdew, Indole and Tryptophan Metabolism: Endogenous and Dietary Routes to Ah Receptor Activation. *Drug Metab. Dispos. Biol. Fate Chem.* **43**, 1522–1535 (2015).

23. W. Tian, C. Chen, X. Lei, J. Zhao, J. Liang, CASTp 3.0: computed atlas of surface topography of proteins. *Nucleic Acids Res.* **46**, W363–W367 (2018).

24. S. C. Faber, A. A. Soshilov, S. Giani Tagliabue, L. Bonati, M. S. Denison, Comparative In Vitro and In Silico Analysis of the Selectivity of Indirubin as a Human Ah Receptor Agonist. *Int. J. Mol. Sci.* **19**, E2692 (2018).

25. S. H. W. Scheres, RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).

26. S. Q. Zheng, E. Palovcak, J.-P. Armache, K. A. Verba, Y. Cheng, D. A. Agard, MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods*. **14**, 331–332 (2017).

27. A. Rohou, N. Grigorieff, CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).

28. P. B. Rosenthal, R. Henderson, Optimal determination of particle orientation, absolute hand, and contrast loss in single-particle electron cryomicroscopy. *J. Mol. Biol.* **333**, 721–745 (2003).

29. S. Chen, G. McMullan, A. R. Faruqi, G. N. Murshudov, J. M. Short, S. H. W. Scheres, R. Henderson, High-resolution noise substitution to measure overfitting and validate resolution in 3D structure determination by single particle electron cryomicroscopy. *Ultramicroscopy*. **135**, 24-35 (2013).

30. D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L.-W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams and P. D. Adams Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. *Acta Cryst.* **D75**, 861-877 (2019).

31. Y. Z. Tan, P. R. Baldwin, J. H. Davis, J. R. Williamson, C. S. Potter, B. Carragher, D. Lyumkis, Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat. Methods*. **14**, 793–796 (2017).

32. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera--a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

33. J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson, Y. Zhang, The I-TASSER Suite: protein structure and function prediction. *Nat. Methods*. **12**, 7–8 (2015).

34. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).

35. C. J. Williams, J. J. Headd, N. W. Moriarty, M. G. Prisant, L. L. Videau, L. N. Deis, V. Verma, D. A. Keedy, B. J. Hintze, V. B. Chen, S. Jain, S. M. Lewis, W. B. Arendall 3rd, J. Snoeyink, P. D. Adams, S. C. Lovell, J. S. Richardson, D. C. Richardson, MolProbity: More and better reference data for improved all-atom structure validation. *Protein Sci.* **27(1)**, 293–315 (2018).

36. T. H. Pham, S. Lecomte, R. Le Guevel, A. Lardenois, B. Evrard, F. Chalmel, F. Ferriere, P. Balaguer, T. Efstathiou, F. Pakdel, Characterization of Glyceollins as Novel Aryl Hydrocarbon Receptor Ligands and Their Role in Cell Migration. *Int. J. Mol. Sci.* **21**, E1368 (2020).

37. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*. **1–2**, 19–25 (2015).

38. C. Tian, K. Kasavajhala, K. A. A. Belfon, L. Raguette, H. Huang, A. N. Migues, J. Bickel, Y. Wang, J. Pincay, Q. Wu, C. Simmerling, ff19SB: Amino-Acid-Specific Protein Backbone Parameters Trained against Quantum Mechanics Energy Surfaces in Solution. *J. Chem. Theory Comput.* **16**, 528–552 (2020).

39. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman, D. A. Case, Development and testing of a general amber force field. *J. Comput. Chem.* **25**, 1157–1174 (2004).

40. S. Izadi, R. Anandakrishnan, A. V. Onufriev, Building Water Models: A Different Approach. *J. Phys. Chem. Lett.* **5**, 3863–3871 (2014).

41. D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, R. J. Woods, The Amber biomolecular simulation programs. *J. Comput. Chem.* **26**, 1668–1688 (2005).

42. G. Bussi, D. Donadio, M. Parrinello, Canonical sampling through velocity rescaling. *J. Chem. Phys.* **126**, 014101 (2007).

43. M. Parrinello, A. Rahman, Polymorphic transitions in single crystals: A new molecular dynamics method. *J. Appl. Phys.* **52**, 7182–7190 (1981).

44. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, L. G. Pedersen, A smooth particle mesh Ewald method. *J. Chem. Phys.* **103**, 8577–8593 (1995).

45. B. Hess, P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *J. Chem. Theory Comput.* **4**, 116–122 (2008).

46. T. D. Goddard, C. C. Huang, E. C. Meng, E. F. Pettersen, G. S. Couch, J. H. Morris, T. E. Ferrin, UCSF ChimeraX: Meeting modern challenges in visualization and analysis. *Protein Sci. Publ. Protein Soc.* **27**, 14–25 (2018).

47. A. Stivala, M. Wybrow, A. Wirth, J. C. Whisstock, P. J. Stuckey, Automatic generation of protein structure cartoons with Pro-origami. *Bioinforma. Oxf. Engl.* **27**, 3315–3316 (2011).

48. F. Madeira, Y. M. Park, J. Lee, N. Buso, T. Gur, N. Madhusoodanan, P. Basutkar, A. R. N. Tivey, S. C. Potter, R. D. Finn, R. Lopez, The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **47**, W636–W641 (2019).

49. X. Robert, P. Gouet, Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* **42**, W320-324 (2014).

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.

Funding:

European Union's Horizon 2020 research and innovation program grant No. 825489 (WB)

ATIP-Avenir 2020 grant No. R20059SP (JG)

ANR (the French National Research Agency "Investissements d'avenir" programme reference No. ANR-16-IDEX-0006 (JG)

Author contributions:

JG, WB initiated and supervised project

JG, LG purified protein samples

JLKH, AA prepared EM grids

JG, JLKH, CGS, JB, AA collected and analyzed EM data

JG, HSK, WB performed model building and refinement

JG, WB analyzed the structure

LG, PG carried out biochemical analysis

MG, AbBo, PB performed cell-based assays

MP, AlBa performed molecular dynamics simulations

CB collected and analyzed MS data

JG, WB wrote the paper

Competing interests: Authors declare that they have no competing interests

Data availability: The atomic coordinates for the indirubin-bound Hsp90-XAP2-AHR complex have been deposited in the Protein Data Bank (PDB) under the accession code 7ZUB. The cryo-EM map obtained in this study has been deposited in the Electron Microscopy Data Bank (EMDB) under the accession code EMD-14971. Focused map used for model refinement has been deposited in EMDB under the accession code EMD-14972.

Supplementary Materials

Materials and Methods

Figs. S1 to S11

Table S1

References (25–49)

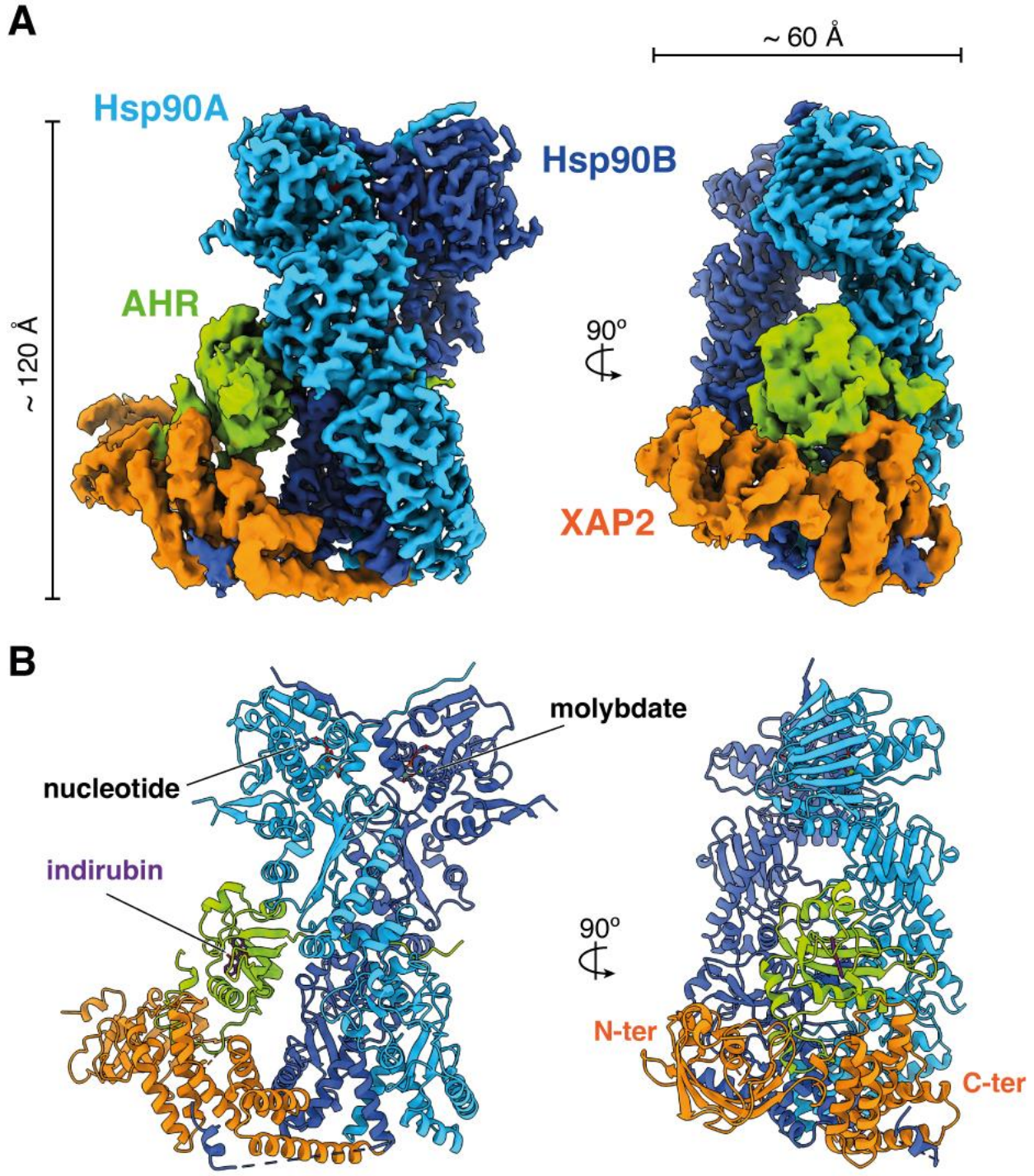

Fig. 1. Overall architecture of the agonist-bound cytosolic complex of AHR. (**A**) A composite map of the Hsp90-XAP2-AHR complex in two orthogonal views. Hsp90A (light blue), Hsp90B (dark blue), XAP2 (orange), AHR (green). The same color scheme is used

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.

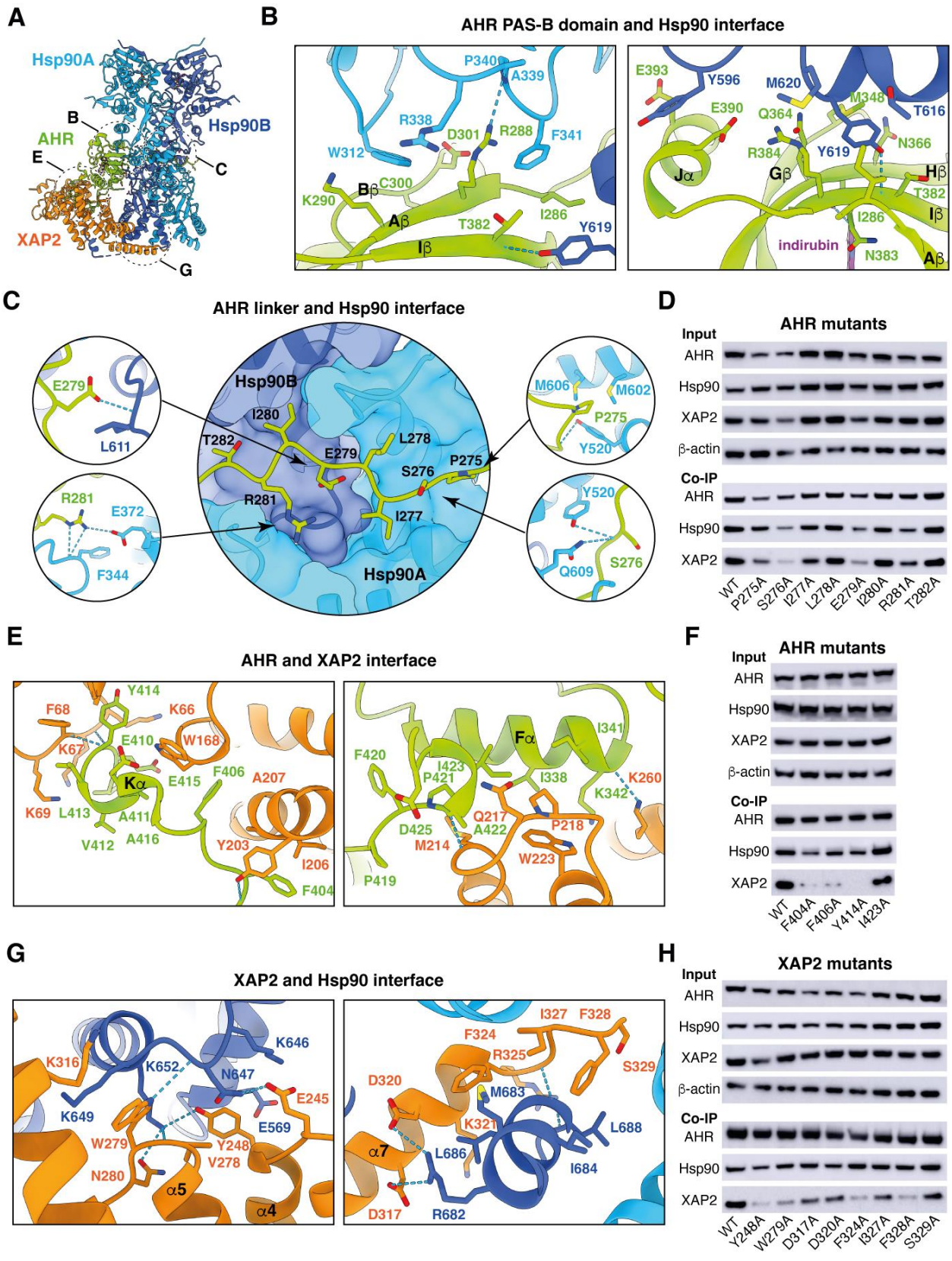
A
Hsp90A
Hsp90B
AHR
XAP2
B
C
E
G
B
AHR PAS-B domain and Hsp90 interface
P340
A339
R338
D301
R288
F341
W312
Bβ
C300
K290
Aβ
T382
I286
Y619
Iβ
E393
Y596
M620
E390
Q364
M348
T616
N366
R384
Y619
Jα
Gβ
Hβ
T382
Iβ
I286
indirubin
N383
Aβ
C
AHR linker and Hsp90 interface
Hsp90B
Hsp90A
E279
L611
I280
T282
E279
L278
S276
P275
R281
I277
R281
E372
F344
M606
M602
P275
Y520
Y520
Q609
S276
D
AHR mutants
Input
AHR
Hsp90
XAP2
β-actin
Co-IP
AHR
Hsp90
XAP2
WT
P275A
S276A
I277A
L278A
E279A
I280A
R281A
T282A
E
AHR and XAP2 interface
Y414
F68
K66
K67
E410
W168
Kα
E415
F406
A207
K69
L413
A411
V412
A416
Y203
I206
F404
Fα
I341
F420
I423
P421
I338
K260
K342
D425
Q217
A422
P218
M214
W223
P419
F
AHR mutants
Input
AHR
Hsp90
XAP2
β-actin
Co-IP
AHR
Hsp90
XAP2
WT
F404A
F406A
Y414A
I423A
G
XAP2 and Hsp90 interface
K316
K652
K646
N647
E245
K649
E569
W279
Y248
V278
N280
α5
α4
I327
F328
F324
R325
S329
D320
M683
K321
L688
α7
L686
I684
D317
R682
H
XAP2 mutants
Input
AHR
Hsp90
XAP2
β-actin
Co-IP
AHR
Hsp90
XAP2
WT
Y248A
W279A
D317A
D320A
F324A
I327A
F328A
S329A

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.

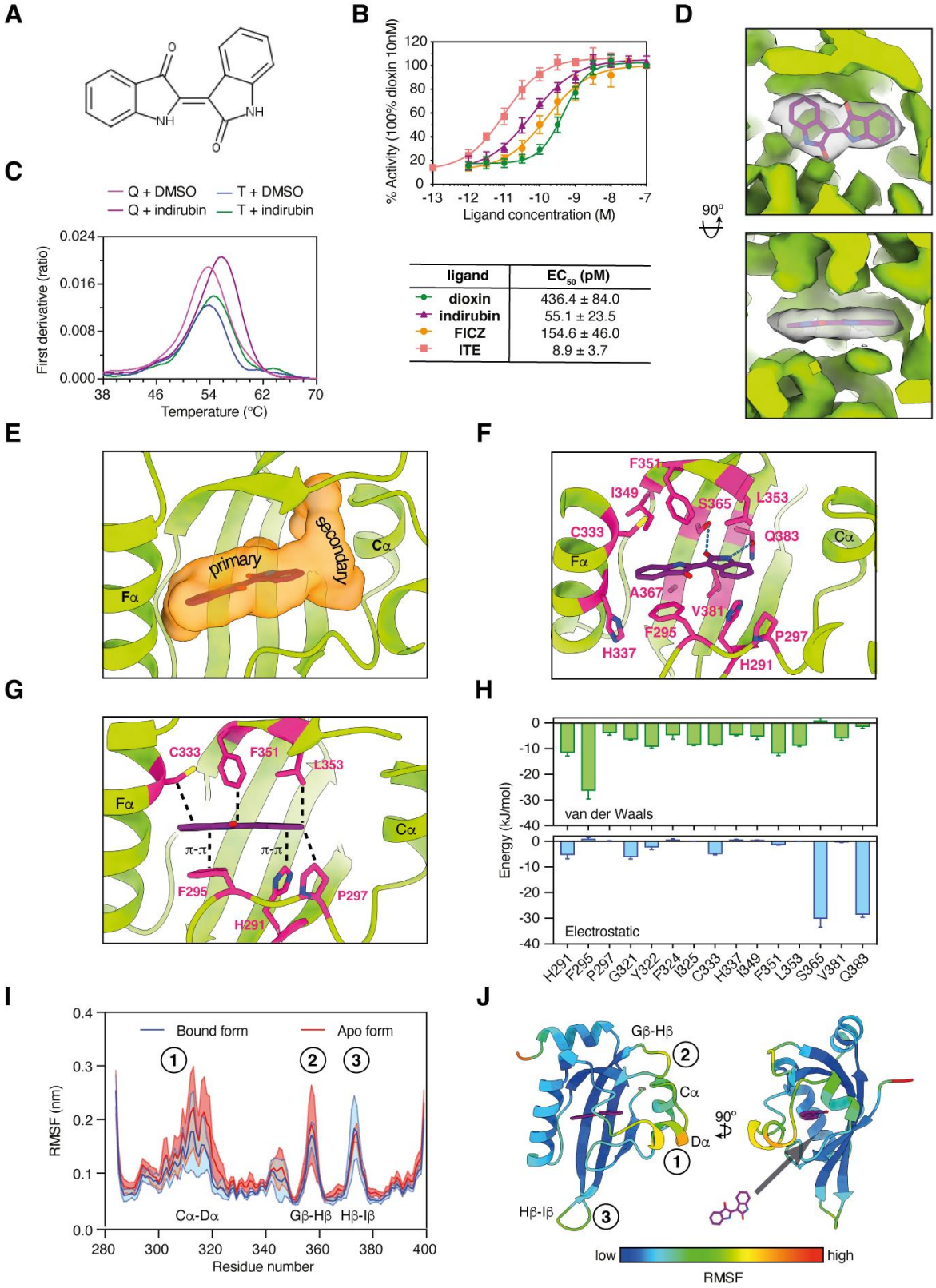
A
B
C
D
E
F
G
H
I
J
Q + DMSO
T + DMSO
Q + indirubin
T + indirubin
First derivative (ratio)
Temperature (°C)
% Activity (100% dioxin 10nM)
Ligand concentration (M)
ligand
EC50 (pM)
dioxin 436.4 ± 84.0
indirubin 55.1 ± 23.5
FICZ 154.6 ± 46.0
ITE 8.9 ± 3.7
90°
primary
secondary
Fα
Cα
F351
I349
S365
L353
C333
Q383
A367
V381
F295
P297
H337
H291
π-π
van der Waals
Electrostatic
Energy (kJ/mol)
H291 F295 P297 G321 Y322 F324 I325 C333 H337 I349 F351 L353 S365 V381 Q383
Bound form
Apo form
RMSF (nm)
Residue number
Cα-Dα
Gβ-Hβ
Hβ-Iβ
Dα
low
high
RMSF

Fig. 3. Characterization of the interaction between indirubin and AHR. (**A**), Structure of indirubin. (**B**) Results of cell-based activity assays for AHR ligands: dioxin, indirubin, FICZ and ITE. The values of measured EC_{50} are indicated. 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 first derivative of the ratio between intrinsic fluorescence at 350 nm and 330 nm is plotted against the temperature. The results are representative of independent biological replicates (n=3). (**E**) Close-up view of the ligand binding pocket (LBP) of AHR shown as an orange surface. The LBP is composed of two distinct parts. Indirubin within the LBP is shown in magenta sticks. (**F**) Close-up view of the indirubin binding site. Residues interacting with the ligand are shown as sticks and colored in pink. Hydrogen bonds are indicated as dashed blue lines. (**G**) Two layers of LBP residues (magenta sticks) determine AHR ligand selectivity. π-π and dashed lines denote Pi and van der Waals interactions. (**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, while 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, while 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.