Structural basis for the ligand-dependent activation of heterodimeric AHR-ARNT complex

Xiaotong Diao[1,#], Qinghong Shang[1,#], Mengqi Guo[1], Yubin Huang[1], Meina Zhang[1], Xiaoyu Chen[1], Yinping Liang[1], Xiangnan Sun[1], Fan Zhou[2], Jingjing Zhuang[1], Shuang-Jiang Liu[1], Christoph F. A. Vogel[3], Fraydoon Rastinejad[4], Dalei Wu[1,*]

[1]State Key Laboratory of Microbial Technology, Shandong University, Qingdao 266237, China

[2]Shanghai Zelixir Biotech, Shanghai 200030, China

[3]Center for Health and the Environment, University of California, One Shields Avenue, Davis, CA 95616, USA

[4]Target Discovery Institute, NDM Research Building, University of Oxford, Old Road Campus, Oxford OX3 7FZ, UK

[#]These authors contributed equally: Xiaotong Diao, Qinghong Shang

*e-mail: dlwu@sdu.edu.cn

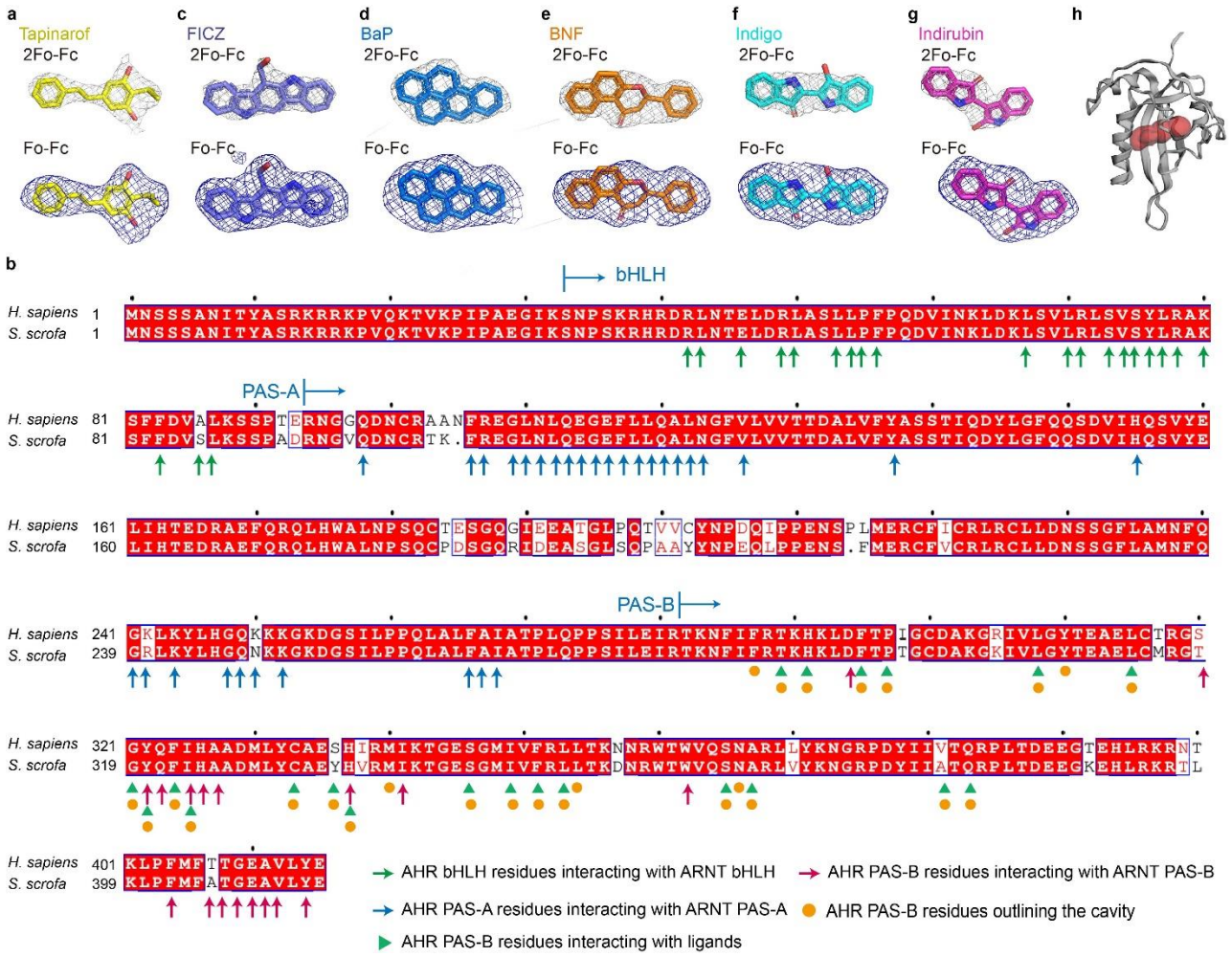
Inventory of Supporting Information

1. Supplementary Figures

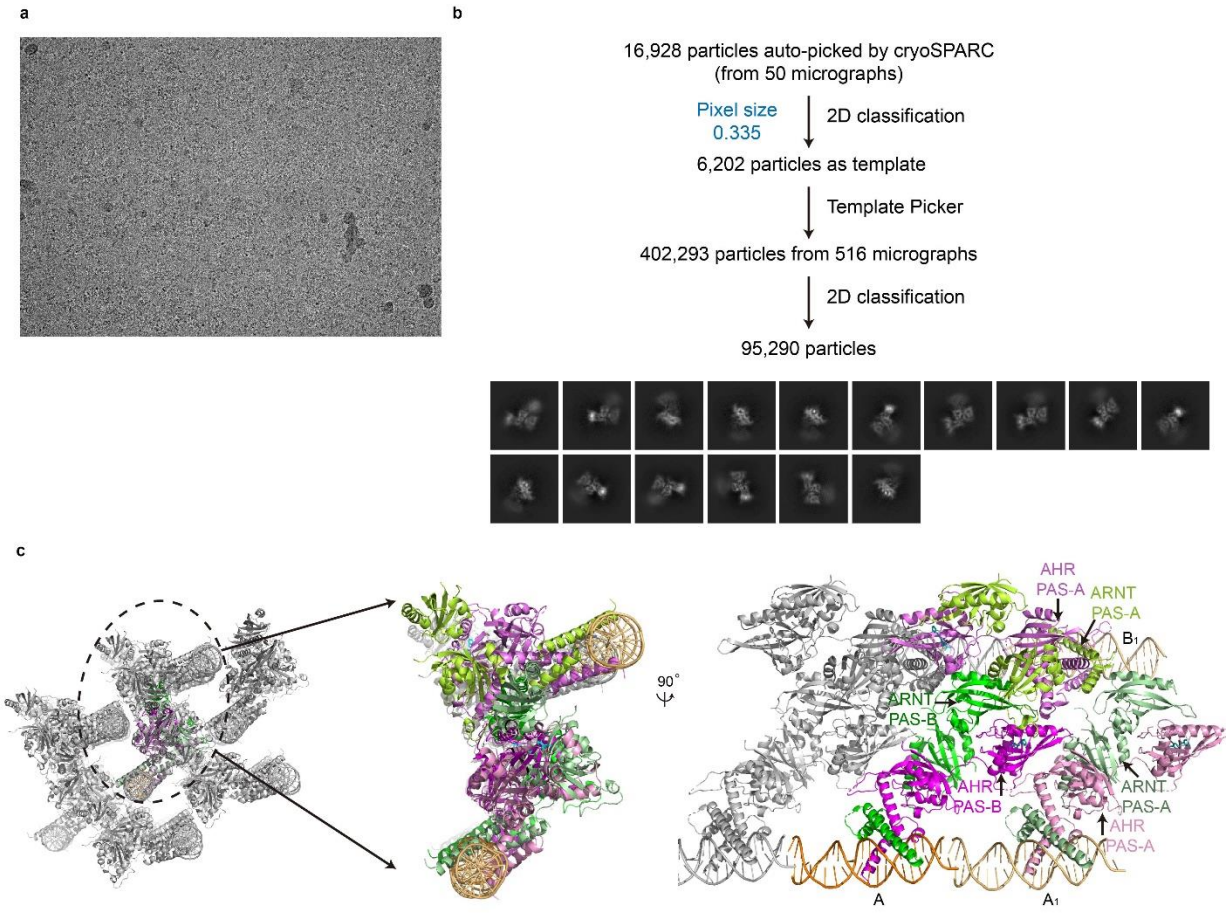
2. Supplementary Table

3. Supplementary Movies

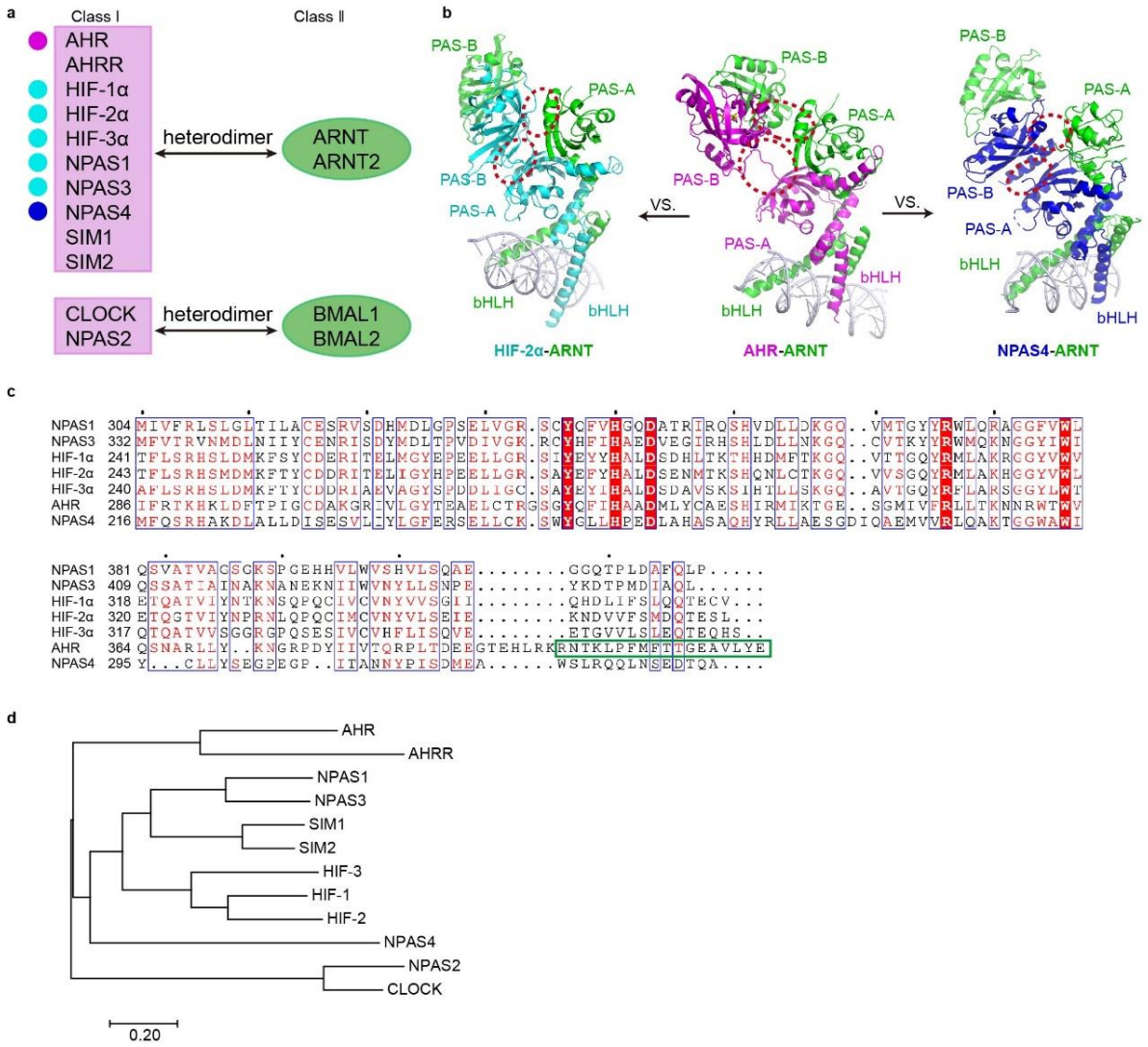
4. Supplementary Data



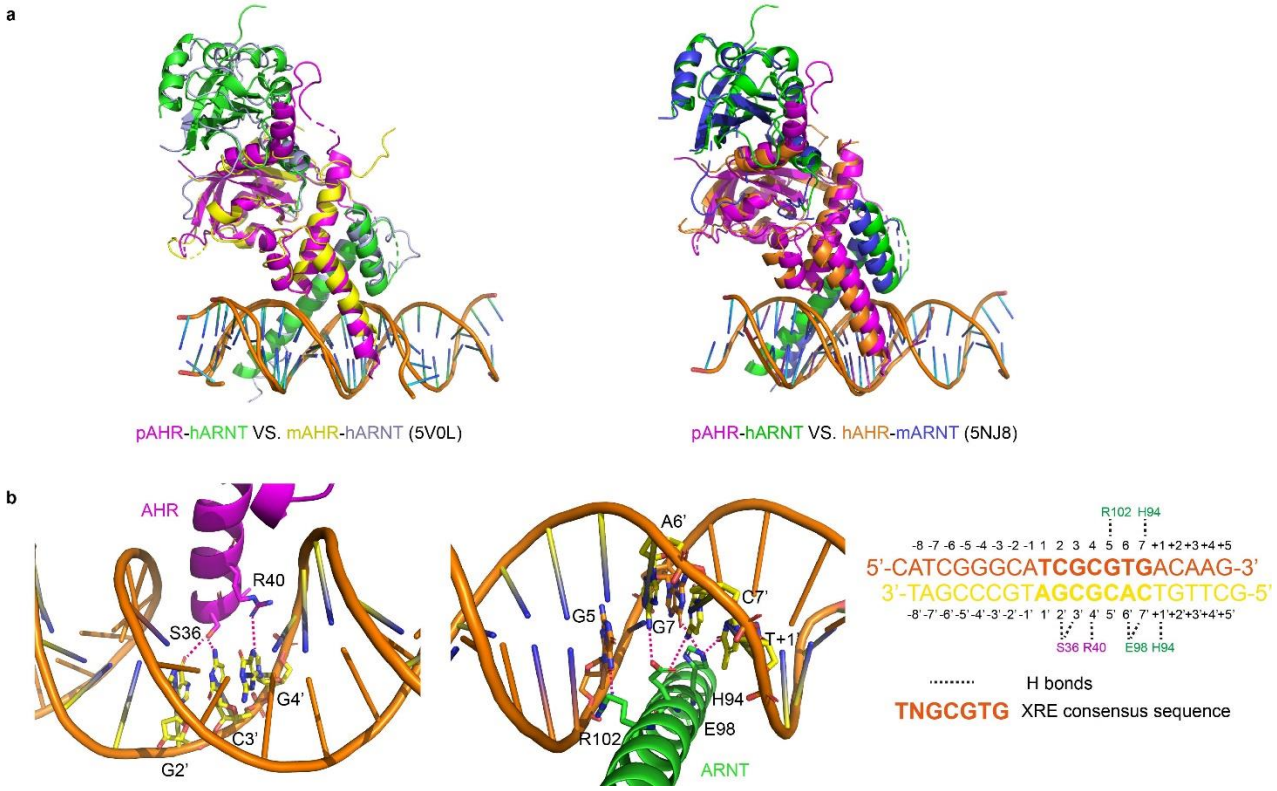
Supplementary Fig. 1 Electron density maps of ligands within the AHR-ARNT-DNA complex structures and sequence alignment of porcine with human AHR. a, The 2Fo-Fc maps (contour level = 1.0 σ, shown as the grey mesh) and the Fo-Fc omit maps (contour level = 3.0 σ, shown as the blue mesh) of Tapinarof. **b,** Sequence alignment of porcine and human AHR proteins. The residues involved in the interactions with ARNT and/or the ligands are marked by arrows or triangles, and the residues outlining the cavity are marked by circles. **c-g**, The 2Fo-Fc maps (contour level = 1.0 σ, shown as the grey mesh) and the Fo-Fc omit maps (contour level = 3.0 σ, shown as the blue mesh) of FICZ (**c**), BaP (**d**), BNF (**e**), Indigo (**f**), and Indirubin (**g**) in their corresponding complex structures, respectively. **h**, The size and location analysis of the AHR PAS-B cavity using CASTp algorithm.



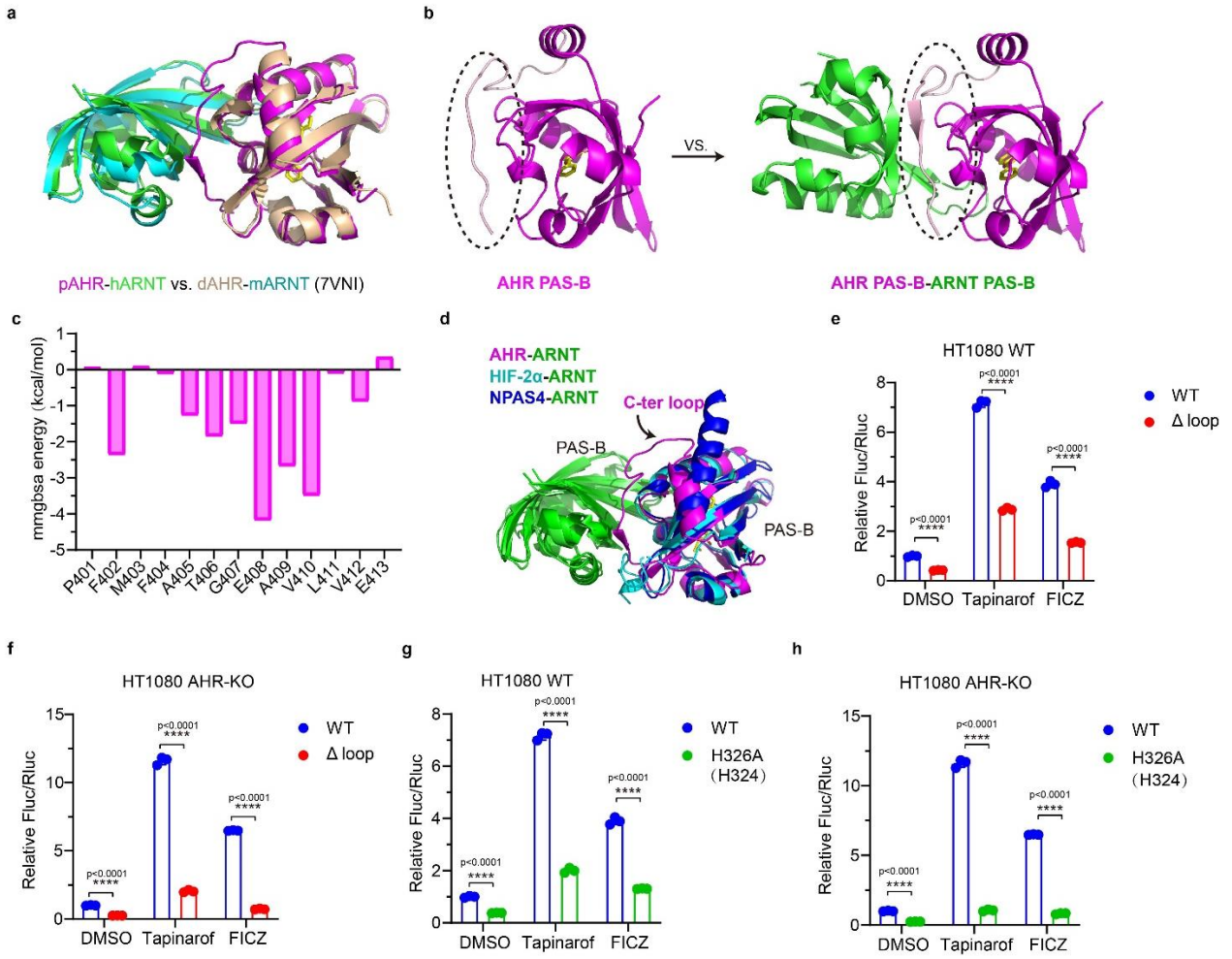
Supplementary Fig. 2 Details of cryo-EM data 2D class processing for the AHR-ARNT-DNA complex and the crystal packing pattern of the complex. a, Representative micrograph of the sample from 516 movies selected for particle picking. **b**, Flowchart for Cryo-EM data processing of AHR-ARNT-DNA and the representative 2D class averages. **c**, Packing arrangement between AHR-ARNT-DNA molecules within the crystal. The PAS-B domains of the A molecule form interaction interfaces with the PAS-A domains of the A_1 and B_1 molecules.



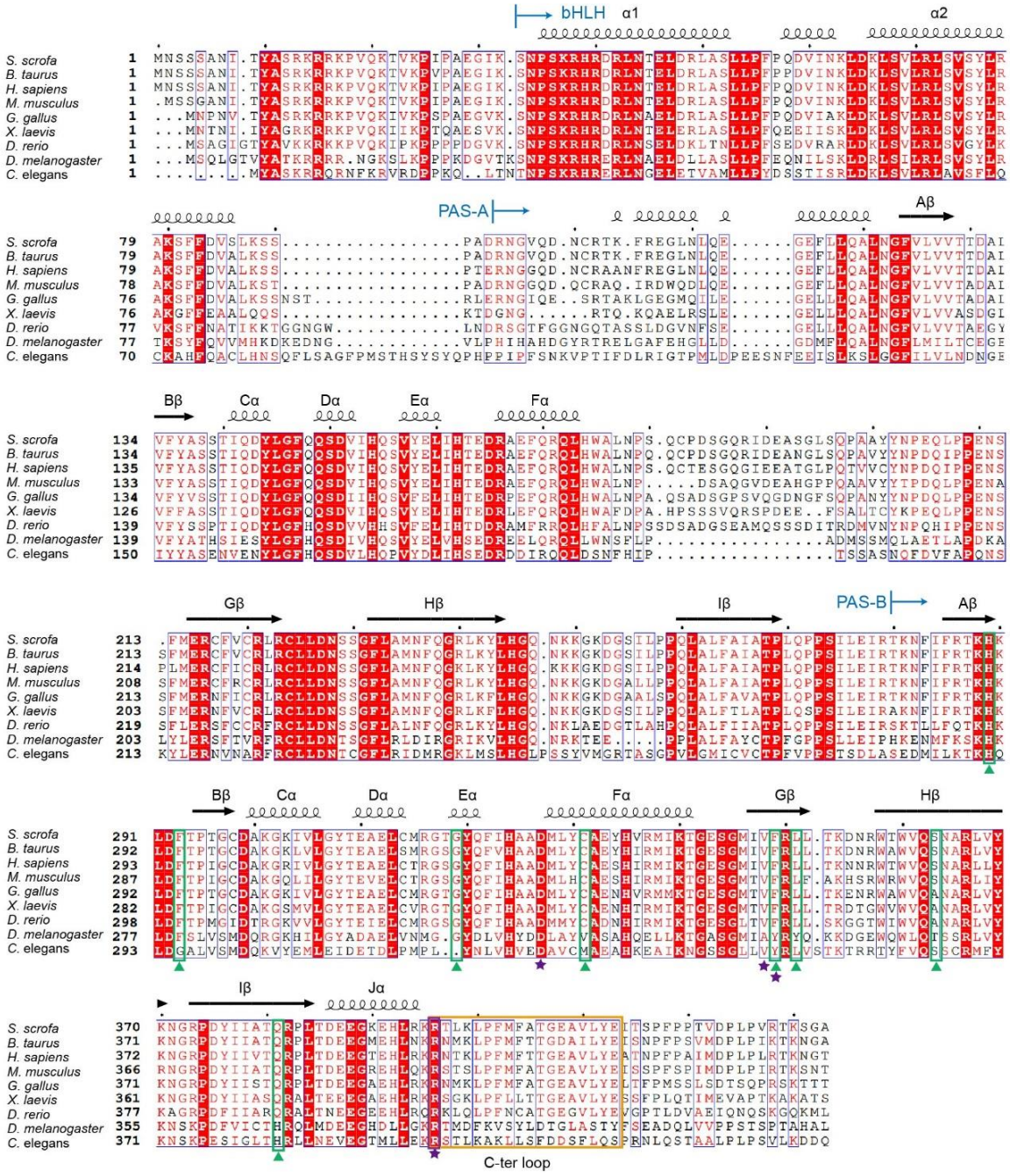
Supplementary Fig. 3 The mammalian bHLH-PAS family of transcription factors. a, Diagram showing the partner choices for heterodimerization between different members of bHLH-PAS family. The magenta, cyan and blue dots represent different dimerization modes. **b**, The overall structure comparison of AHR-ARNT with HIF-2α-ARNT (4ZPK) and NPAS4-ARNT (7XI4). The interfaces between the PAS-B domains of HIF-2α or NPAS4 with their own PAS-A domain and the PAS-A domain of ARNT are circled with dashed lines, and the corresponding regions in the AHR-ARNT heterodimer structure are also circled. AHR, HIF-2α, NPAS4, and ARNT are colored in magenta, cyan, blue, and green, respectively. **c**, Protein sequence alignment (figure generated by ESPript 3.0) of the PAS-B domains in members of the bHLH-PAS family, including NPAS1, NPAS3, HIF-1α, HIF-2α, HIF-3α, AHR, and NPAS4. The C-ter loop of AHR PAS-B is marked with a green box. **d**, Phylogenetic tree of human bHLH-PAS family Class I members based on their protein sequences.



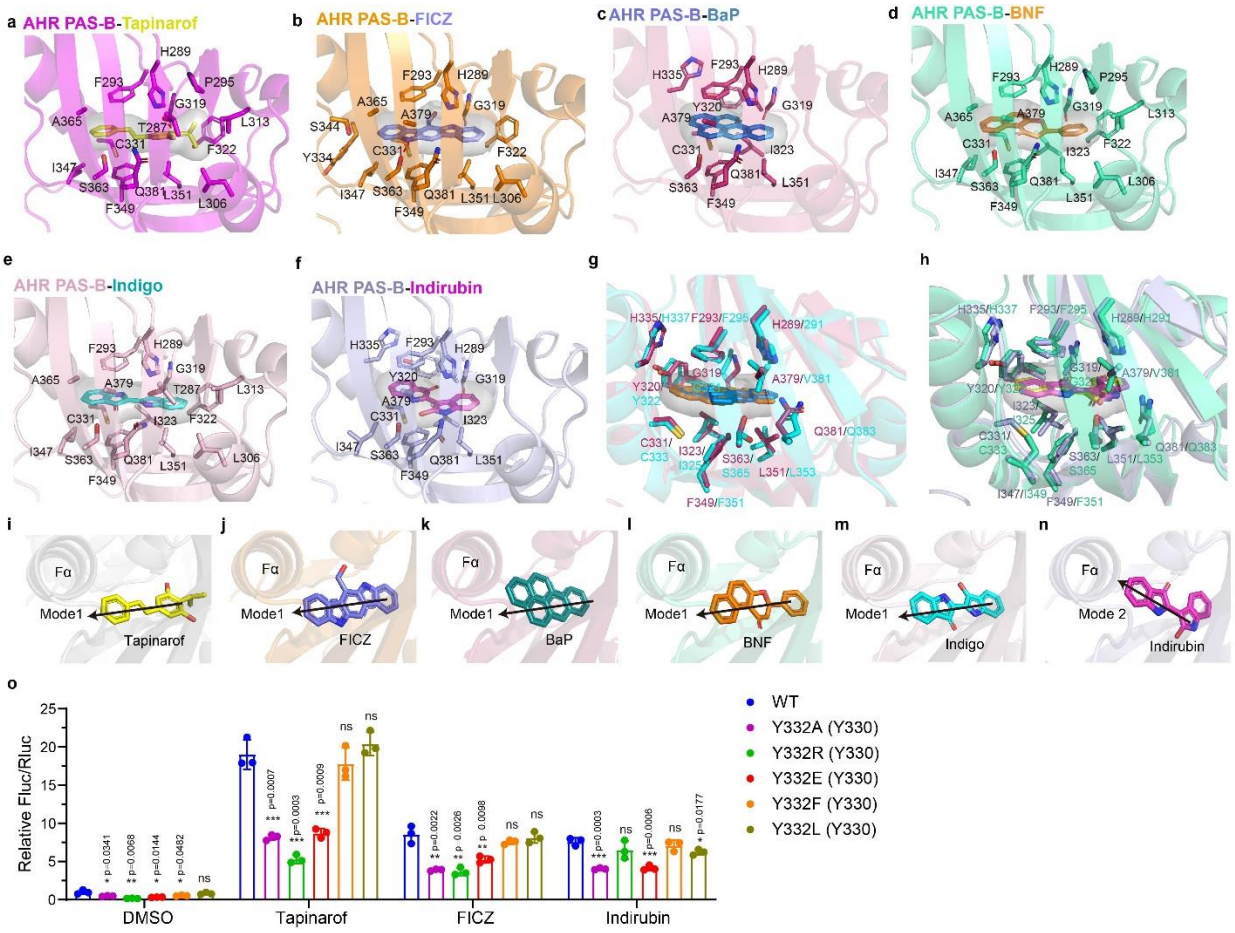
Supplementary Fig. 4 DNA binding pattern of AHR-ARNT heterodimers. a, Superimposition of the bHLH and PAS-A domain structures from pAHR-hARNT heterodimer with those from mAHR-hARNT (5V0L, left) or hAHR-mARNT (5NJ8, right). **b**, Recognition sites of AHR and ARNT bHLH domains directly involved in DNA binding. Schematic recognition diagram of AHR-ARNT to the XRE sequence on DNA (right). The dotted lines represent hydrogen bonds. AHR and ARNT are colored in magenta and green, respectively.



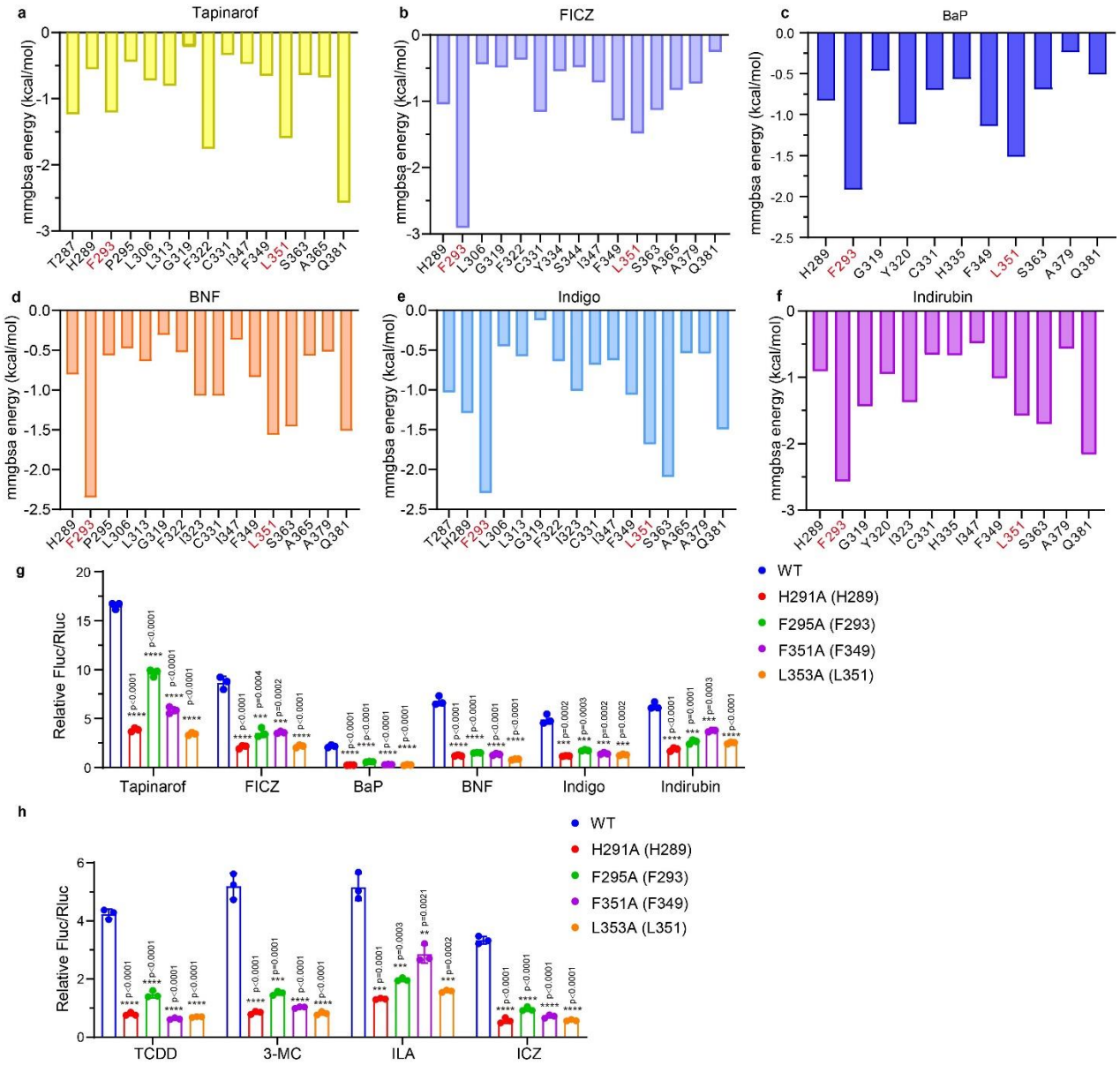
Supplementary Fig. 5 The C-ter loop of AHR PAS-B domain mediating its interaction with ARNT. a, Superimposition of PAS-B domain structure from pAHR-hARNT heterodimer with those from dAHR-mARNT (7VNI). **b**, The middle region of C-ter loop (407-409 aa) of AHR PAS-B forms a small β-sheet together with the nearby β-strand (Bβ) of ARNT PAS-B domain. **c**, A per-residue decomposition study of relative binding energy for the C-ter loop, based on MD simulations using only the PAS-B domains of AHR-ARNT. **d**, Superimposition of the dimeric PAS-B domain structure from pAHR-hARNT heterodimer with those from HIF-2α-ARNT (4ZPK) and NPAS4-ARNT (7XI4). **e-h**, XRE reporter assays evaluating the effects of C-ter loop deletion and H326A mutation on AHR transcriptional activity in HT1080 (**e,g**) and HT1080 AHR-KO (**f,h**) cells. Error bars, mean ± s.d.; n=3 (biological replicates); statistical significance: **** $p < 0.0001$. P values (shown on the charts) were calculated using unpaired two-tailed t-test versus the WT group.



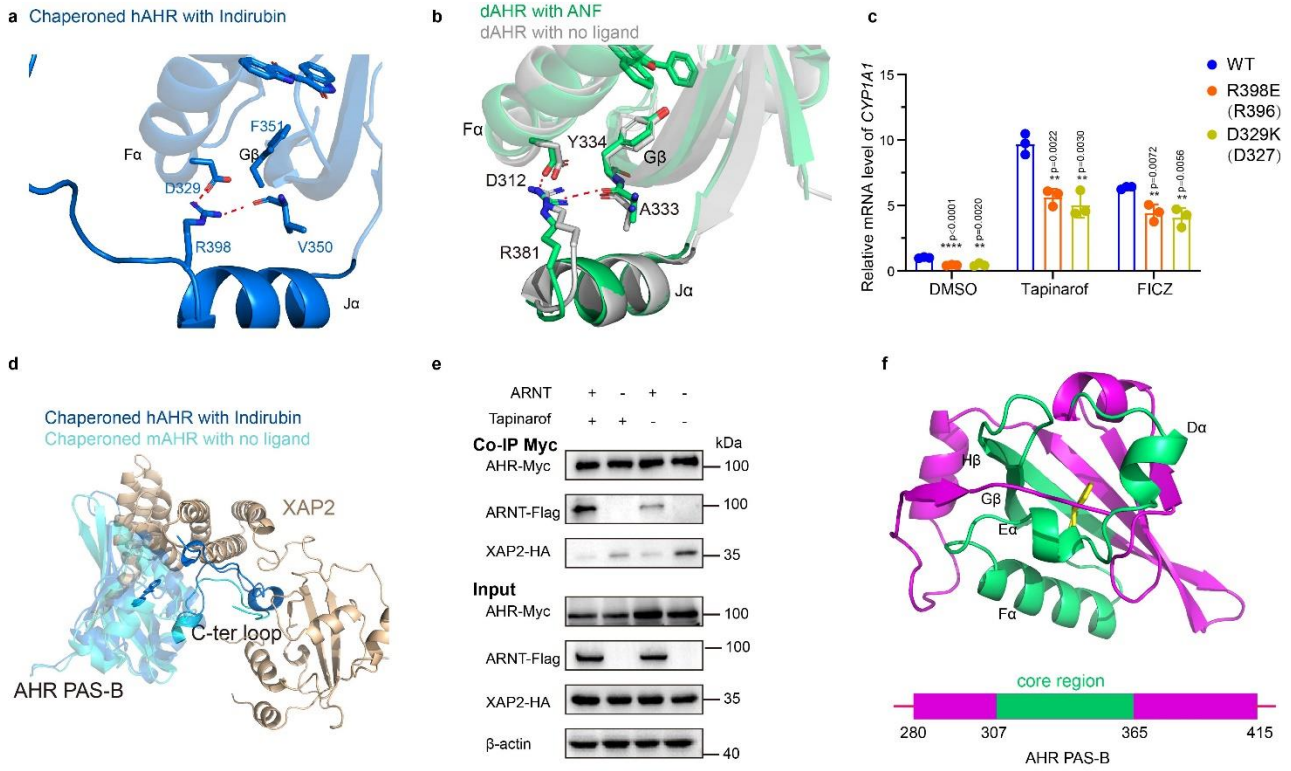
Supplementary Fig. 6 Sequence alignment of the N-terminal portions of AHR proteins from different species. These species include *Sus scrofa*, *Bos taurus*, *Homo sapiens*, *Mus musculus*, *Gallus gallus*, *Xenopus laevis*, *Danio rerio*, *Drosophila melanogaster*, *Caenorhabditis elegans*. The C-ter loop segment of PAS-B is marked with an orange box. Eight residues directly interacting with the six ligands are marked with green boxes and triangles. The residues (D327, V348, F349, and R396 of pAHR) forming the network that mediates allosteric effects linking ligand binding to the conformations of Jα and the C-ter loop, are marked with purple asterisks.



Supplementary Fig. 7 The complex structures of AHR-ARNT bound with different ligands. a-f, Interactions between Tapinarof (yellow, **a**), FICZ (purple, **b**), BaP (blue, **c**), BNF (orange, **d**), Indigo (cyan, **e**), or Indirubin (magenta, **f**) and surrounding residues in the AHR PAS-B pocket. **g**,**h**, Structural superimposition of AHR PAS-B domains bound with BaP (**g**) or Indirubin (**h**) in the AHR-ARNT heterodimers and those bound with BaP (PDB code: 8QMO) or Indirubin (PDB code 7ZUB) in the chaperoned HSP90-XAP2-AHR complexes. **i-n**, Comparison of binding modes of Tapinarof (**i**), FICZ (**j**), BaP (**k**), BNF (**l**), Indigo (**m**), and Indirubin (**n**) in the AHR PAS-B pocket. **o**, XRE reporter assay evaluating the effects of hAHR Y332 mutations (to alanine, arginine, glutamate, phenylalanine, and leucine) on the transactivation in the presence of Tapinarof, FICZ, Indirubin in the AHR-KO HT1080 cell line. Error bars, mean ± s.d.; n=3 (biological replicates); * $p < 0.05$,** $p < 0.01$, *** $p < 0.001$, ns $p>0.05$. P values (shown on the charts) were calculated using unpaired two-tailed t-test versus the WT group.



Supplementary Fig. 8 Effects of key residues in AHR PAS-B pocket on the binding and transactivation by various ligands. a-f, A per-residue decomposition study of relative binding energy between Tapinarof (**a**), FICZ (**b**), BaP (**c**), BNF (**d**), Indigo (**e**), or Indirubin (**f**) and interacting residues, based on MD simulations using only the PAS-B domains of AHR-ARNT. **g,h,** XRE reporter assay evaluating the effects of mutations of hAHR H291, F295, F351, and L353 residues on the transactivation in the presence of Tapinarof, FICZ, BaP, BNF, Indigo, Indirubin (**g**), or TCDD, 3-MC, ILA, ICZ (**h**). Error bars, mean ± s.d.; n=3 (biological replicates); ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$. P values (shown on the charts) were calculated using unpaired two-tailed t-test versus the WT group.



Supplementary Fig. 9 Comparison of AHR PAS-B domains in heterodimeric and in chaperoned complexes. a, A view of chaperoned hAHR bound with Indirubin (7ZUB, marine blue). D329 in Fα, F351 and V350 in Gβ, and R398 in Jα are shown as sticks. Dotted lines represent the salt bridge and hydrogen bond. **b**, Superimposition of dAHR PAS-B structures in ligand-free (7VNI, grey) and ANF-bound (7VNH, green) states. D312 in Fα, Y334 and A333 in Gβ, and R381 in Jα are shown as sticks **c**, qPCR assay evaluating the effects of hAHR R398E mutation (corresponding to R396 in pAHR) and hAHR D329K mutation (corresponding to D327 in pAHR) on transactivation. **d**, C-ter loop of AHR PAS-B domain interacts with XAP2 in the chaperoned complexes before (8H77) and after ligand binding (7ZUB). **e**, Co-IP experiments showing interactions between AHR and XAP2 in the presence or absence of ARNT. **f**, The core region of AHR PAS-B domain (from Dα to Hβ) defining the ligand-responsiveness is shown in green. Error bars, mean ± s.d.; n=3 (biological replicates); statistical significance: ** $p < 0.01$, **** $p < 0.0001$. P values (shown on the charts) were calculated using unpaired two-tailed t-test versus the WT group.

Supplementary Table 1 Crystallographic data collection and refinement statistics

	AHR-ARNT-DNA-Tapinarof	AHR-ARNT-DNA-FICZ	AHR-ARNT-DNA-BaP	AHR-ARNT-DNA-BNF	AHR-ARNT-DNA-Indigo	AHR-ARNT-DNA-Indirubin
PDB ID	8XS6	8XS7	8XS8	8XS9	8XSA	8XSB
Data collection						
Space group	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$	$P2_1$
Cell dimensions						
a, *b*, *c* (Å)	67.95, 98.01, 79.45	68.36, 98.52, 80.30	68.32, 99.15, 80.37	67.99, 98.51, 79.99	68.01, 98.83, 80.30	67.89, 99.45, 79.77
α, β, γ (°)	90.00, 90.93, 90.00	90.00, 90.50, 90.00	90.00, 90.12, 90.00	90.00, 90.65, 90.00	90.00, 90.41, 90.00	90.00, 90.48, 90.00
Resolution (Å)	50.0-2.95 (3.0-2.95) *	34.18-2.77 (2.84-2.77)	68.32-3.11 (3.28-3.11)	50.0-2.80 (2.85-2.80)	50.0-2.60 (2.64-2.60)	50.0-3.06 (3.24-3.06)
R_{sym} or R_{merge}	15.4 (85.3)	14.4 (73.3)	15.7 (69.2)	13.7 (80.0)	8.4 (58.1)	13.7 (56.5)
I / σI	12.4 (2.0)	9.5 (2.2)	4.3 (1.6)	14.2 (2.0)	17.7 (2.3)	8.8 (2.1)
Completeness (%)	96.6 (82.5)	99.6 (98.5)	90.3 (92.9)	93.9 (70.8)	94.1 (69.1)	98.6 (97.1)
Redundancy	5.8 (5.9)	6.7 (6.8)	2.9 (2.8)	6.4 (6.0)	3.5 (3.5)	5.1 (4.8)
Refinement						
Resolution (Å)	34.05-2.95 (3.05-2.95)	34.16-2.77 (2.87-2.77)	68.32-3.11 (3.22-3.11)	39.89-2.80 (2.90-2.80)	30.48-2.60 (2.69-2.60)	36.00-3.06 (3.17-3.06)
No. reflections	21447 (1828)	27166 (2669)	17002 (1500)	24516 (1843)	30953 (2256)	20122 (1981)
R_{work} / R_{free}	0.215/0.271	0.207/0.266	0.182/0.229	0.203/0.263	0.195/0.248	0.200/0.240
No. atoms						
Protein	5877	5864	5829	5894	5904	5869
Ligand/ion	19	30	20	29	28	20
Water	31	50	26	39	144	23
B-factors						
Protein	62.3	57.4	66.0	48.3	48.3	61.6
Ligand/ion	53.0	60.1	50.5	51.0	53.6	42.0
Water	47.3	49.0	46.5	36.7	41.5	44.5
R.m.s. deviations						
Bond lengths (Å)	0.005	0.003	0.004	0.004	0.004	0.003
Bond angles (°)	0.83	0.68	0.61	0.66	0.69	0.53

*Values in parentheses are for highest-resolution shell; one crystal for each structure.