Structural basis for aryl hydrocarbon receptor-mediated gene activation

Schulte, K.W., Green, E., Wilz, A., Platten, M., Daumke, O.

This is the final version of the accepted manuscript. The original article has been published in final edited form in:

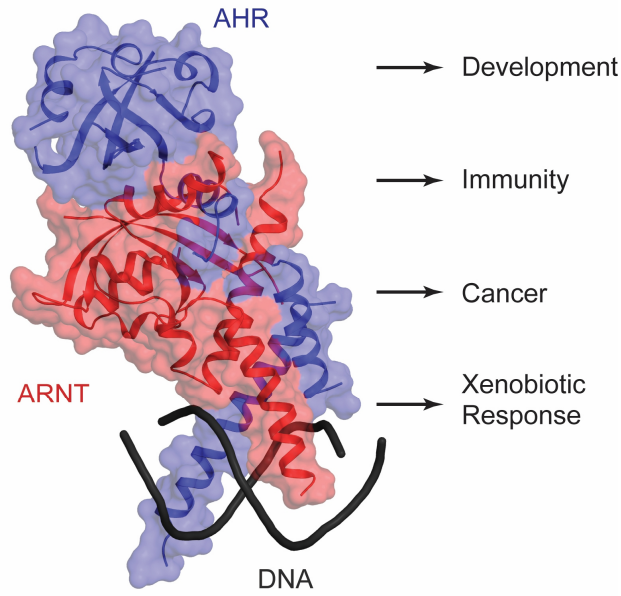
Structure
2017 JUL 05 ; 25(7): 1025-1033
2017 JUN 09 (first published online)
doi: 10.1016/j.str.2017.05.008

Publisher: Cell Press / Elsevier

Structural basis for aryl hydrocarbon receptor mediated gene activation

Kathrin Wiebke Schulte[1,2], Edward Green[3,4], Annabel Wilz[4], Michael Platten[3,4,5], Oliver Daumke[1,2*]

[1]Crystallography Department, Max-Delbrück-Center for Molecular Medicine, Robert-Rössle-Strasse 10, 13125 Berlin, Germany
[2]Institute of Chemistry and Biochemistry, Freie Universität Berlin, Takustrasse 6, 14195 Berlin, Germany
[3]Department of Neurology, University Hospital Heidelberg and National Center for Tumor Diseases, 69120 Heidelberg, Germany
[4]Clinical Cooperation Unit Neuroimmunology and Brain Tumor Immunology, German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), 69120 Heidelberg, Germany
[5]Department of Neurology, Universitätsmedizin Mannheim, Medical Faculty Mannheim, Heidelberg University, 68167 Mannheim, Germany



Correspondence: Oliver Daumke Tel: 0049-30-9406-3425; Fax: 0049-30-9406-3814; Email: oliver.daumke@mdc-berlin.de; Michael Platten, Tel: +49 621 383 2885; Fax: +49 621 383 3807; Email: michael.platten@dkfz.de.

***Lead contact**

Short title: Structure of the AHR-ARNT complex

Summary

The aryl hydrocarbon receptor (AHR) and the AHR nuclear translocator (ARNT) comprise a hetero-dimeric basic-helix-loop-helix-Per-ARNT-Sim (bHLH-PAS) domain-containing transcription factor with central functions in development and cellular homeostasis. AHR is activated by xenobiotics, notably dioxin, as well as by exogenous and endogenous metabolites. Modulation of AHR activity holds promise for the treatment of diseases featuring altered cellular homeostasis, such as cancer or autoimmune disorders. Here, we present the crystal structure of a heterodimeric AHR:ARNT complex containing the PAS A and bHLH domain bound to its target DNA. The structure provides insights into the DNA binding mode of AHR and elucidates how stable dimerization of AHR:ARNT is achieved through sophisticated domain interplay via three specific interfaces. Using mutational analyses, we prove the relevance of the observed interfaces for AHR-mediated gene activation. Thus, our work establishes the structural basis of AHR assembly and DNA interaction and provides a template for targeted drug design.

Keywords: Aryl hydrocarbon receptor; innate immunity; cancer; dioxin; X-ray structure analysis

Introduction

The aryl hydrocarbon receptor (AHR) and the AHR nuclear translocator (ARNT) belong to the family of basic helix-loop-helix-PER-ARNT-SIM (bHLH-PAS) domain containing transcription factors. Members of this family respond to various external stimuli in order to regulate diverse biological processes, including cellular differentiation (Gasiewicz et al., 2014), malignant transformation (Murray et al., 2014), immune responses (Stockinger et al., 2014) and circadian rhythmicity (Anderson et al., 2013). Proteins of the ARNT subfamily were grouped into the class II bHLH-PAS domain containing transcription factors, which localize to the nucleus. They can hetero-dimerize with members of the class I subfamily of bHLH-PAS proteins, such as AHR, the hypoxia inducible factors (HIF-1α/HIF-2α), the neuronal PAS domain proteins (NPAS) and single-minded proteins (SIM1/SIM2), in order to regulate the transcriptional response to xenobiotics, hypoxia and

neurogenesis, respectively (Denison et al., 2003; Kewley et al., 2004). In contrast to ARNT, AHR is a cytosolic signal sensor, which is guided into the nucleus upon binding of planar aromatic hydrocarbons. The best known high-affinity ligand for AHR is 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD), mediating its toxicity via activation of AHR (Poland et al., 1976; Perdew, 1992; Greenlee et al., 1979). Endogenous ligands of AHR include tryptophan-derived metabolites (Opitz et al., 2011; Denison et al., 2003; Veldhoen et al., 2009) and several dietary indoles, which are thought to modulate immune responses (Hubbard et al., 2015; Kiss et al., 2011; Li et al., 2011). More recently, bacterial metabolites from *Pseudomonas* and *Mycobacterium* strains were shown to bind and activate AHR to sense bacterial virulence and both activate antibacterial responses (Moura-Alves et al., 2014) and control sepsis (Bessede et al., 2014). In addition, AHR signaling is intricately involved in fine-tuning distinct immune responses, particularly at environmental interfaces where endogenous and exogenous AHR ligands accumulate (Zelante et al., 2013; Wincent et al., 2012; Kadow et al., 2011; Schiering et al., 2017). For instance, AHR activity is a key determinant of T helper (Th) cell differentiation (Gagliani et al., 2015; Di Meglio et al., 2014). Strikingly, ligands with differential AHR agonist activity may have opposing effects on Th cell differentiation, illustrating the delicate regulation of this transcription factor at the ligand-binding level (Quintana et al., 2008). Increased levels of AHR and its target genes, as well as constitutive localization of AHR in the nucleus, were also found in several aggressive tumors such as glioblastoma, skin tumors or non-small cell lung cancer (Opitz et al., 2011; Poland et al., 1982; Su et al., 2013; Liu et al., 2013). These observations render the AHR:ARNT complex an interesting target for the development of drugs targeting cancer (Opitz et al., 2011), and autoimmune conditions such as multiple sclerosis (Quintana et al., 2008; Veldhoen et al., 2008; Mascanfroni et al., 2015, Quintana et al., 2010).

AHR and ARNT are composed of an amino (N-)terminal bHLH domain required for DNA binding, followed by two PER-ARNT-SIM (PAS) domains (A and B) and a carboxy (C-)terminal trans-activation domain (TAD) (Fukunaga et al., 1995; Crews et al., 1999) (Fig. 1A). The bHLH domains of AHR dimerize to form a four-helical bundle (Pongratz et al., 1998), which can bind to distinct DNA motifs leading to transcriptional activation of target genes (Reyes et al., 1992). The dioxin response element (DRE, 5'- NGCGTG -3', in which N stands for any nucleotide) was identified as the minimal

DNA sequence recognized by the bHLH domains of the AHR:ARNT complex (Yao et al., 1992; Swanson et al., 1995). It can be subdivided into a 5' part (NGC) recognized by AHR, and a 3' part (GTG) recognized by ARNT.

Biochemical and mutagenesis studies indicate that the PAS A domains of AHR and ARNT also dimerize with each other, thereby controlling dimerization specificity, enhancing complex stability and strengthening DNA binding (Pongratz et al., 1998; Erbel et al., 2003; Lindebro et al., 1995; Chapman-Smith et al., 2004). Crystal structures of the hetero-dimeric CLOCK:BMAL complex (Huang et al., 2012), a PAS A domain dimer of AHR (Wu et al., 2013), the hetero-dimeric HIF:ARNT complex (Wu et al., 2015) and, most recently, the NPAS1:ARNT and NPAS3:ARNT heterodimers (Wu et al., 2016) provided molecular insights into the dimerization and DNA binding modes of bHLH-PAS domain proteins. The 'connector helix' of the PAS A domain was shown to play a major role in its dimerization. The PAS B domains of AHR and ARNT differ from the PAS A domains by a three loop insertion, as initially found in the crystal structure of the tandem PAS domain containing PERIOD protein (Yildiz et al., 2005). They have also been suggested to contribute to hetero-dimerization (Erbel et al., 2003; Huang et al., 2012; Wu et al., 2015). In AHR, the PAS B domain constitutes the ligand binding domain (Poland et al., 1976; Poland et al., 1982; Burbach et al., 1992). Interestingly, an AHR construct lacking the PAS B domain forms a constitutively active dimer in mice (Andersson et al., 2002), which suggests that ligand binding is only required for AHR translocation into the nucleus, and not for hetero-dimerization and trans-activation.

Here, we present the crystal structure of an AHR:ARNT transcription factor complex containing the bHLH and PAS A domains bound to a 12mer double strand (ds) DNA. The structure provides the framework to understand AHR-based gene activation and serves as a template for the design of AHR targeting drugs.

Results

Structure determination of an AHR:ARNT complex bound to DNA

To obtain structural insights into AHR-mediated gene activation, we sought to establish a purification protocol for the AHR:ARNT complex. Since all attempts to produce a soluble AHR construct including the PAS B domain failed, we co-expressed truncated human AHR and mouse ARNT constructs without the PAS B and transactivation domain (TAD) (Fig. 1A, referred to as AHR' and ARNT' from hereon). Deleting a predicted internal loop in the PAS A domain of ARNT (Δ274-301) and mutating a reactive cysteine (C256S, in the ARNT'Δ construct) increased protein yields and allowed the successful purification of the complex (Fig. 1B, Methods).

Assembly of the purified complex was analyzed by analytical gel filtration, combined with a right-angle laser scattering (RALS) analysis. In these experiments, a stable heterodimeric complex of about 60 kDa was observed, comprising one molecule of AHR' (26 kDa) and ARNT'Δ (29 kDa) each (Fig. 1C). Addition of a 12mer dsDNA oligomer containing the DRE motif in its center slightly decreased the retention time in gel filtration analysis but increased the OD_{260}/OD_{280} ratio of the peak fraction, indicating stable DNA binding. In contrast, purified ARNT protein was monomeric and formed homo-dimers only upon binding to a DNA sequence containing a central palindromic 5'-CAC-GTG-3' motif corresponding to the ARNT binding sequence (Sogawa et al., 1995) (Fig. S1).

To quantitate DNA binding of the AHR':ARNT'Δ complex, the affinities towards dsDNA of various lengths, containing a central DRE motif, were determined by isothermal titration calorimetry (ITC) experiments. dsDNA oligomers of 16, 14 and 12 nucleotides in length bound the AHR:ARNT complex with an affinity of approximately 100 nM (Figs. 1D, E), indicating that bases flanking the DRE element beyond the 12mer sequence do not influence the binding affinity. Further shortening of the dsDNA to a 10mer sequence led to an 8-fold reduction in the binding affinity (K_D = 790 nM). Thus, although not directly involved in DNA interactions (see below), residues in the immediate vicinity of the DRE motif may influence DNA binding, for example by altering DNA stability or structure.

To explore the structural basis of AHR-mediated gene transcription, we crystallized the AHR':ARNT'Δ complex bound to the 12mer dsDNA. Crystals of this complex diffracted to 3.3 Å

resolution (Table 1). Experimental phases were calculated using erbium ions as anomalous scatterers resulting in a clearly interpretable electron density map. The final R_{work}/R_{free} of the model was 29.2%/33.3%.

The asymmetric unit contained two hetero-dimeric AHR':ARNT'Δ complexes (complex 1 and 2), each bound to a dsDNA (Fig. S2A). These two complexes showed a similar architecture and orientation of their domains. They could be superimposed with a root mean square deviation (rmsd) of 0.9 Å for 252 Cα atoms (Fig. S2B). Due to missing crystal contacts, the PAS A domains of AHR' and ARNT'Δ in complex 2 were not well resolved in electron density and could be modelled only incompletely (see Star Methods). Therefore, the following discussion refers to complex 1.

DNA binding of AHR':ARNT'Δ

The AHR':ARNT'Δ complex has an intertwined architecture, with all four domains participating in the assembly (Fig. 2A). The bHLH domains show a prototypical structure comprising two helices and a flexible connector. Similar to other bHLH domain containing proteins, the DRE motif (underlined in the sequence of the 12mer oligomer ^{1}GGTTGCGTGACC^{12}) is bound in between the two N-terminal α1-helices via an interface of 890 Å^2 (Fig. 2B, C, S3, S4). Similar to the Hif-2α:ARNT complex (Wu et al., 2015), His94 and Arg102 of ARNT interact with the terminal guanine base in the DRE motif of one DNA strand whereas Glu98 binds to the adenine and cytosine base of the complementary strand. AHR interacts via Ser36, His39 and Arg40 with the other half-site of the DRE motif. His39 contacts the phosphate backbone of the thymine base whereas Ser36 and Arg40 form hydrogen bonds to the cytosine and guanine of the complementary DNA strand. In agreement with our structure, His39 and Arg40 in human AHR and His79, Asp83, Arg86 and Arg87 in human ARNT (corresponding to His94, Glu98, Arg102, Arg103 in mouse ARNT) were previously shown to be crucial for DRE binding (Reisz-Porszasz et al., 1994; Dolwick et al., 1993; Bacsi et al., 1995; Dong et al., 1996). Overall, the DNA binding mode in bHLH-PAS transcription factors is related although the detailed interactions differ allowing specific sequence recognition (Fig. 2C).

Assembly of AHR':ARNT'Δ

Dimerization of AHR' and ARNT'Δ is mediated by an intricate domain interplay, involving interaction interfaces in both the bHLH and PAS A domains (Fig. 3A-F). The bHLH domains interact with each other via two hydrophobic interaction sites covering 980 Å^2 involving helices α1 and α2 (from here on referred to as interface 1 according to (Wu et al., 2016)). For example, Ile109 and Leu112 in α1 of ARNT interacts with Val74 and Leu70 in α2 of AHR (Fig. 3A, Interface 1a). A second contact site is provided by a hydrophobic surface between α1 of AHR and α2 of ARNT, in which Leu47, Leu50 and Leu53 of AHR interact with Leu132, Val136 and Met139 of ARNT (Fig. 3B, Interface 1b).

As previously described (Wu et al., 2013; Wu et al., 2015; Wu et al., 2016), the PAS A domains of AHR and ARNT have a canonical PAS domain fold containing a five stranded antiparallel β-sheet flanked by four helices at one side. They form an extensive dimerization interface of 1030 Å^2 to which we refer as interface 2. At the N-terminus of the PAS A domains, the helical connectors (A'α) display a number of mainly hydrophobic contacts with each other and with residues in the opposing PAS A domain. For example, hydrophobic interactions are mediated by Leu118, Leu119 and L122 in A'α of AHR with Leu164 and Ile168 in A'α of ARNT, whereas Asp114 in AHR forms a salt bridge to Lys165 in ARNT (Fig. 3C, Interface 2a). Furthermore, Phe158, Leu159, Leu167, Ile168 and Ala172 N-terminally of ARNT A'α form a hydrophobic interface with Tyr137, Val126, Phe266, Ile268 in the PAS A domain of AHR (Fig. 3D, Interface 2b). Vice versa, Phe117, Leu118, Ala121 and Leu122 in A'α of AHR interact with Leu176, Ile340, Thr309 and Arg260 in the PAS A domain of ARNT (Fig. 3E, Interface 2c).

A third, mostly intramolecular interface of 960 Å^2 is formed between the bHLH and the PAS A domain of AHR. Phe56, Val60, Leu72, Ala79, Phe82 in the bHLH domain form a hydrophobic network to Phe136, Ser151, Ile154, and Leu246 in the PAS A domain. Furthermore, Pro55, Phe56, Tyr76, Ala79 and Phe83 in the AHR bHLH domain interact with Pro156, Ser157, Phe158, Leu159 in the N-terminal region of the helical connector and with Glu163 in the helical connector of ARNT (Fig. 3F, Interface 3).

The identified interfaces explain previously described mutations affecting assembly of AHR and ARNT. Based on the crystal structure of the PAS A domain dimer of AHR, Wu and co-workers found that single mutations of Glu112, Phe115, Leu116, Ala119, Leu120, Val124, Phe260 and Ile262 in mouse AHR (corresponding to Glu114, Phe117, Leu118, Ala121, Leu122, Val126, Phe266, Ile268 in human AHR) disrupt formation of the AHR:ARNT complex, leading to reduced activation of a reporter gene (Wu et al., 2013). These residues directly participate (Glu114, Leu118, Ala121, Leu122, Phe266, Ile268) or are close to (Phe117, Val126) interface 2. In addition, mutations of interface 2 residues Leu167, Ile168, Ala171 in mouse ARNT were shown to disrupt dimerization (Wu et al., 2013). Two further studies found that mutations of Glu112 in mouse AHR (corresponding to Glu114 in human AHR) and of Glu163, Ala339 and Gly341 in mouse ARNT interfere with dimerization, in agreement with their direct involvement in interface 2 or 3 (Numayama-Tsuruta et al., 1997; Hao et al., 2011).

To further evaluate the AHR:ARNT domain interfaces, we used a reporter gene assay, in which GFP expression was driven under the control of an DRE-dependent promotor. Co-expression of full length ARNT and a constitutively active AHR (AHR.cA) with an internal deletion of the PAS B domain (Andersson et al., 2002) yielded robust, ligand-independent activation of the GFP reporter gene (Fig. 3G). Subsequently, we tested the effects of replacing single residues in AHR.cA and ARNT with negatively charged aspartate on gene transcription (Fig. 3G). In agreement with the structural data, mutations of the DNA binding residues Arg40 in AHR.cA and Arg102 in ARNT decreased reporter gene activation. Also mutations in interface 1 (L50D, V74D in AHR, L112D, M139D in ARNT), interface 2 (L118D and L122D in AHR.cA) and interface 3 (A79D, F82D, F136D, I154D in AHR.cA) interfered with gene activation. In contrast, the L164D mutation in ARNT in interface 2 did not have an effect. Leu164 locates to the periphery of interface 2 and its replacement by aspartate apparently does not interfere with a functional AHR.cA:ARNT complex. Overall, these data are in agreement with crucial roles of the observed interfaces in AHR-dependent gene activation.

Discussion

bHLH-PAS transcription factors are central regulators of cellular homeostasis. Given their crucial functions in health and disease, it is astonishing that structural data reporting their DNA binding properties and assembly modes only recently became available. The CLOCK-BMAL structure revealed an asymmetric heterodimer, in which both the bHLH, PAS A and PAS B domains contributed to dimerization (Huang et al., 2012). Despite the close relationship, the overall topology is clearly different compared to our structure of the AHR:ARNT complex (Fig. 4A). More recently, the structure of HIF-1α and 2α-ARNT heterodimers bound to DNA revealed an alternate conformation of bHLH-PAS transcription factors in which the two PAS A domains were rotated with respect to the bHLH domain (Wu et al., 2015). Consequently, it was unclear which conformation is adopted by AHR:ARNT dimers, and modeling studies yielded conflicting results (Corrada et al., 2016). Our current structural analysis indicates that AHR:ARNT shows a similar conformation compared to HIF-2α:ARNT bound to DNA (Fig. 4B). Very recently, the structures of NPAS1- and NPAS2:ARNT bound to DNA were also demonstrated to have a similar overall topology as the AHR:ARNT dimers (Fig. 4C) (Wu et al., 2016). The structural similarities of the AHR:ARNT, HIF-2α:ARNT and NPAS:ARNT complexes may be related to the common ARNT subunit and its absence in the CLOCK:BMAL complex. Interestingly, when comparing the two PAS A domains in our structure to those of the AHR PAS A domain dimer (Wu et al., 2013), an almost 70° rotation of the second subunit was observed (Fig. 4D), pointing to a non-physiological domain assembly of the PAS A domain dimer.

DNA binding of AHR:ARNT is mediated by the bHLH domains which interact via positively charged residues with the DNA. Similar to the HIF-2α:ARNT complex, ARNT in the AHR complex recognizes its target sequence GTG via the conserved His94, Glu98 and Arg102. Specificity of DNA recognition of the AHR:ARNT complex is mediated by Ser36, His39 and Arg40 of AHR which contact the first three bases of the DRE motif. In contrast, HIF-2α uses Ala23 and Arg27 to recognize a TAC sequence. Interestingly, in HIF-α, also the PAS A domain participates in DNA binding via a loop structure (Wu et al., 2015). In contrast to HIF-2α, we did not observe such a contact in the AHR PAS A domain. The corresponding loop is 8 amino acids shorter in AHR and CLOCK compared to HIF-2α

and may therefore not reach the DNA, whereas it has a similar size as HIF-2α in NPAS proteins (Fig. 1D, S3).

In the published bHLH-PAS structures, the PAS B domain is intimately involved in dimerization. However, we found that AHR:ARNT can stably dimerize without the PAS B domain. Furthermore, despite the absence of the PAS B domain, an active, DNA binding conformation was adopted. These observations point to a regulatory function of the PAS B domain in gene activation and are in agreement with previous deletion analysis (Andersson et al., 2002). Thus, ligand binding to the PAS B of AHR domain may release auto-inhibitory constraints, thereby allowing nuclear localization of AHR and dimerization with ARNT via the described interfaces. Whether PAS B domain interactions in the AHR:ARNT complex have a functional relevance needs to be further examined.

bHLH-PAS domain proteins, and in particular the AHR transcription factor, comprise interesting targets for the development of therapeutic agents (Scheuermann et al., 2013). For example, it was shown that AHR activation by the endogenous ligand kynurenine promotes the development of glioblastoma (Opitz et al., 2011), thus an AHR inhibitor is expected to suppress tumor growth and metastasis and enhance anti-tumor immunity. Our structure may guide rational drug design approaches: for example, inhibitors targeting the DNA binding site of the bHLH domains may specifically interfere with transcription of the AHR:ARNT complex without affecting related bHLH PAS domain transcription factor complexes, while the assembly interfaces 1-3 may also be targeted. Sequence comparisons indicate that residues in the assembly interfaces are highly conserved within the class II bHLH transcription factor family, but deviate to a larger extent in the class I family (Fig. 4E, F, S3, S4). For example, all ARNT residues referred to in the text are completely invariant in the ARNT family, whereas AHR residues Leu47, Leu53, Val74 in interface 1, Val126, Phe136, Ile154, Leu246, Phe266, Ile268 in interface 2 and Val60, Ala79, Phe82, Phe83 in interface 3 are only partially conserved in class I transcription factors. Furthermore, published biochemical data indicate differential importance of interface residues for the assembly. For example, mutations F100D and L160E in NPAS1 were reported not to disturb assembly with ARNT (Wu et al., 2016), whereas mutations in corresponding positions in AHR (F82D, F136D) interfere with AHR-mediated gene transcription. Thus, even closely related

transcription factors complexes, such HIF-2α:ARNT, NPAS:ARNT and AHR:ARNT may offer unique surfaces in their PAS A and bHLH domains to specifically target assembly and function by small molecule inhibitors. Our study also indicates that small molecule inhibitors of the PAS B domain will likely not interfere with dimerization, but could alternatively affect the function of the AHR:ARNT complex, for example by modifying translocation of AHR into the nucleus.

Our structural data on AHR:ARNT lay the foundation for understanding AHR-dependent gene activation and provide a template for structure-guided drug design.

Note added in proof: While our manuscript was in review, a related AHR-ARNT structure bound to DNA was reported (Seok et al., 2017).

Author contributions

KS purified the AHR':ARNT'Δ complex, performed the biochemical experiments and determined the structure of the complex. EG and AW performed the activation assays. KS, EG, MP, OD designed research and wrote the manuscript.

Acknowledgements

This project was supported by a grant of the Leibnitz Graduate School to KS/OD. We would like to thank Katja Fälber and Manuel Hessenberger for critically reading the manuscript, and Manfred Weiss, Uwe Müller (from the Berlin Synchrotron beamline 14), Yvette Roske and Katja Fälber for their support during data collection.

Supplemental Information

Supplemental Information includes three figure and one supplemental table and can be found with this article online at xxx.

References

Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010). PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D. Biol. Crystallogr. 66**,** 213-221.

Anderson, G., Beischlag, T. V., Vinciguerra, M., and Mazzoccoli, G. (2013). The circadian clock circuitry and the AHR signaling pathway in physiology and pathology. Biochem. Pharmacol. 85**,** 1405-1416.

Andersson, P., McGuire, J., Rubio, C., Gradin, K., Whitelaw, M. L., Pettersson, S., Hanberg, A., and Poellinger, L. (2002). A constitutively active dioxin/aryl hydrocarbon receptor induces stomach tumors. Proc. Natl. Acad. Sci. U. S. A 99**,** 9990-9995.

Bacsi, S. G., Reisz-Porszasz, S., and Hankinson, O. (1995). Orientation of the heterodimeric aryl hydrocarbon (dioxin) receptor complex on its asymmetric DNA recognition sequence. Mol. Pharmacol. 47**,** 432-438.

Bessede, A., Gargaro, M., Pallotta, M. T., Matino, D., Servillo, G., Brunacci, C., Bicciato, S., Mazza, E. M., Macchiarulo, A., and Vacca, C. (2014). Aryl hydrocarbon receptor control of a disease tolerance defence pathway. Nature 511**,** 184-190.

Burbach, K. M., Poland, A., and Bradfield, C. A. (1992). Cloning of the Ah-receptor cDNA reveals a distinctive ligand-activated transcription factor. Proc. Natl. Acad. Sci. U. S. A 89**,** 8185-8189.

Chapman-Smith, A., Lutwyche, J. K., and Whitelaw, M. L. (2004). Contribution of the Per/Arnt/Sim (PAS) domains to DNA binding by the basic helix-loop-helix PAS transcriptional regulators. J Biol. Chem. 279**,** 5353-5362.

Chen, V. B., Arendall, W. B., III, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr 66**,** 12-21.

Corrada, D., Soshilov, A. A., Denison, M. S., and Bonati, L. (2016). Deciphering Dimerization Modes of PAS Domains: Computational and Experimental Analyses of the AhR:ARNT Complex Reveal New Insights Into the Mechanisms of AhR Transformation. PLoS. Comput. Biol. 12**,** e1004981.

Crews, S. T., and Fan, C. M. (1999). Remembrance of things PAS: regulation of development by bHLH-PAS proteins. Curr. Opin. Genet. Dev. 9**,** 580-587.

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

Di Meglio, P., Duarte, J. H., Ahlfors, H., Owens, N. D., Li, Y., Villanova, F., Tosi, I., Hirota, K., Nestle, F. O., and Mrowietz, U. (2014). Activation of the aryl hydrocarbon receptor dampens the severity of inflammatory skin conditions. Immunity 40**,** 989-1001.

Dolwick, K. M., Swanson, H. I., and Bradfield, C. A. (1993). In vitro analysis of Ah receptor domains involved in ligand-activated DNA recognition. Proc. Natl. Acad. Sci. U. S. A 90**,** 8566-8570.

Dong, L., Ma, Q., and Whitlock, J. P., Jr. (1996). DNA binding by the heterodimeric Ah receptor. Relationship to dioxin-induced CYP1A1 transcription in vivo. J Biol. Chem. 271**,** 7942-7948.

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

Erbel, P. J., Card, P. B., Karakuzu, O., Bruick, R. K., and Gardner, K. H. (2003). Structural basis for PAS domain heterodimerization in the basic helix--loop--helix-PAS transcription factor hypoxia-inducible factor. Proc. Natl. Acad. Sci. U. S. A 100**,** 15504-15509.

Fukunaga, B. N., Probst, M. R., Reisz-Porszasz, S., and Hankinson, O. (1995). Identification of functional domains of the aryl hydrocarbon receptor. J Biol. Chem. 270**,** 29270-29278.

Gagliani, N., Vesely, M. C. A., Iseppon, A., Brockmann, L., Xu, H., Palm, N. W., de Zoete, M. R., Licona-Lim+¦n, P., Paiva, R. S., and Ching, T. (2015). Th17 cells transdifferentiate into regulatory T cells during resolution of inflammation. Nature 523**,** 221-225.

Gasiewicz, T. A., Singh, K. P., and Bennett, J. A. (2014). The Ah receptor in stem cell cycling, regulation, and quiescence. Ann. N. Y. Acad. Sci. 1310**,** 44-50.

Greenlee, W. F., and Poland, A. (1979). Nuclear uptake of 2,3,7,8-tetrachlorodibenzo-p-dioxin in C57BL/6J and DBA/2J mice. Role of the hepatic cytosol receptor protein. J Biol. Chem. 254**,** 9814-9821.

Hao, N., Whitelaw, M. L., Shearwin, K. E., Dodd, I. B., and Chapman-Smith, A. (2011). Identification of residues in the N-terminal PAS domains important for dimerization of Arnt and AhR. Nucleic Acids Res. 39**,** 3695-3709.

Huang, N., Chelliah, Y., Shan, Y., Taylor, C. A., Yoo, S. H., Partch, C., Green, C. B., Zhang, H., and Takahashi, J. S. (2012). Crystal structure of the heterodimeric CLOCK:BMAL1 transcriptional activator complex. Science 337**,** 189-194.

Hubbard, T. D., Murray, I. A., Bisson, W. H., Lahoti, T. S., Gowda, K., Amin, S. G., Patterson, A. D., and Perdew, G. H. (2015). Adaptation of the human aryl hydrocarbon receptor to sense microbiota-derived indoles. Scientific reports 5:12689.

Kabsch, W. (2010). XDS. Acta Crystallogr D Biol Crystallogr 66**,** 125-132.

Kadow, S., Jux, B., Zahner, S. P., Wingerath, B., Chmill, S., Clausen, B. E., Hengstler, J., and Esser, C. (2011). Aryl hydrocarbon receptor is critical for homeostasis of invariant +¦+¦ T cells in the murine epidermis. The Journal of Immunology 187**,** 3104-3110.

Kewley, R. J., Whitelaw, M. L., and Chapman-Smith, A. (2004). The mammalian basic helix-loop-helix/PAS family of transcriptional regulators. Int. J Biochem. Cell Biol. 36**,** 189-204.

Kiss, E. A., Vonarbourg, C., Kopfmann, S., Hobeika, E., Finke, D., Esser, C., and Diefenbach, A. (2011). Natural aryl hydrocarbon receptor ligands control organogenesis of intestinal lymphoid follicles. Science 334**,** 1561-1565.

Krissinel, E., and Henrick, K. (2007). Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372**,** 774-797.

Leslie, A. G. W. (2006). The integration of macromolecular diffraction data. Acta Crystallogr D Biol Crystallogr 62**,** 48-57.

Li, Y., Innocentin, S., Withers, D. R., Roberts, N. A., Gallagher, A. R., Grigorieva, E. F., Wilhelm, C., and Veldhoen, M. (2011). Exogenous stimuli maintain intraepithelial lymphocytes via aryl hydrocarbon receptor activation. Cell 147**,** 629-640.

Lindebro, M. C., Poellinger, L., and Whitelaw, M. L. (1995). Protein-protein interaction via PAS domains: role of the PAS domain in positive and negative regulation of the bHLH/PAS dioxin receptor-Arnt transcription factor complex. EMBO J 14**,** 3528-3539.

Liu, Z., Wu, X., Zhang, F., Han, L., Bao, G., He, X., and Xu, Z. (2013). AhR expression is increased in hepatocellular carcinoma. J Mol. Histol. 44**,** 455-461.

Mascanfroni, I. D., Takenaka, M. C., Yeste, A., Patel, B., Wu, Y., Kenison, J. E., Siddiqui, S., Basso, A. S., Otterbein, L. E., and Pardoll, D. M. (2015). Metabolic control of type 1 regulatory T cell differentiation by AHR and HIF1-[alpha]. Nat. Med. 21**,** 638-646.

Moura-Alves, P., Fae, K., Houthuys, E., Dorhoi, A., Kreuchwig, A., Furkert, J., Barison, N., Diehl, A., Munder, A., Constant, P., Skrahina, T., Guhlich-Bornhof, U., Klemm, M., Koehler, A. B., Bandermann, S., Goosmann, C., Mollenkopf, H. J., Hurwitz, R., Brinkmann, V., Fillatreau, S., Daffe, M., Tummler, B., Kolbe, M., Oschkinat, H., Krause, G., and Kaufmann, S. H. (2014). AhR sensing of bacterial pigments regulates antibacterial defence. Nature 512**,** 387-392.

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

Numayama-Tsuruta, K., Kobayashi, A., Sogawa, K., and Fujii-Kuriyama, Y. (1997). A point mutation responsible for defective function of the aryl-hydrocarbon-receptor nuclear translocator in mutant Hepa-1c1c7 cells. Eur. J. Biochem. 246**,** 486-495.

Opitz, C. A., Litzenburger, U. M., Sahm, F., Ott, M., Tritschler, I., Trump, S., Schumacher, T., Jestaedt, L., Schrenk, D., Weller, M., Jugold, M., Guillemin, G. J., Miller, C. L., Lutz, C., Radlwimmer, B., Lehmann, I., von, D. A., Wick, W., and Platten, M. (2011). An endogenous tumour-promoting ligand of the human aryl hydrocarbon receptor. Nature 478**,** 197-203.

Perdew, G. H. (1992). Chemical cross-linking of the cytosolic and nuclear forms of the Ah receptor in hepatoma cell line 1c1c7. Biochem. Biophys. Res. Commun. 182**,** 55-62.

Poland, A., Glover, E., and Kende, A. S. (1976). Stereospecific, high affinity binding of 2,3,7,8-tetrachlorodibenzo-p-dioxin by hepatic cytosol. Evidence that the binding species is receptor for induction of aryl hydrocarbon hydroxylase. J Biol. Chem. 251**,** 4936-4946.

Poland, A., Palen, D., and Glover, E. (1982). Tumour promotion by TCDD in skin of HRS/J hairless mice. Nature 300**,** 271-273.

Pongratz, I., Antonsson, C., Whitelaw, M. L., and Poellinger, L. (1998). Role of the PAS domain in regulation of dimerization and DNA binding specificity of the dioxin receptor. Mol. Cell Biol. 18**,** 4079-4088.

Quintana, F. J., Basso, A. S., Iglesias, A. H., Korn, T., Farez, M. F., Bettelli, E., Caccamo, M., Oukka, M., and Weiner, H. L. (2008). Control of Treg and TH17 cell differentiation by the aryl hydrocarbon receptor. Nature 453**,** 65-71.

Quintana, F. J., Murugaiyan, G., Farez, M. F., Mitsdoerffer, M., Tukpah, A. M., Burns, E. J., and Weiner, H. L. (2010). An endogenous aryl hydrocarbon receptor ligand acts on dendritic cells and T cells to suppress experimental autoimmune encephalomyelitis. Proceedings of the National Academy of Sciences 107**,** 20768-20773.

Reisz-Porszasz, S., Probst, M. R., Fukunaga, B. N., and Hankinson, O. (1994). Identification of functional domains of the aryl hydrocarbon receptor nuclear translocator protein (ARNT). Mol. Cell Biol. 14**,** 6075-6086.

Reyes, H., Reisz-Porszasz, S., and Hankinson, O. (1992). Identification of the Ah receptor nuclear translocator protein (Arnt) as a component of the DNA binding form of the Ah receptor. Science 256**,** 1193-1195.

Scheuermann, T. H., Li, Q., Ma, H. W., Key, J., Zhang, L., Chen, R., Garcia, J. A., Naidoo, J., Longgood, J., Frantz, D. E., Tambar, U. K., Gardner, K. H., and Bruick, R. K. (2013). Allosteric inhibition of hypoxia inducible factor-2 with small molecules. Nat. Chem. Biol. 9**,** 271-276.

Schiering, C., Wincent, E., Metidji, A., Iseppon, A., Li, Y., Potocnik, A. J., Omenetti, S., Henderson, C. J., Wolf, C. R., Nebert, D. W., and Stockinger, B. (2017). Feedback control of AHR signalling regulates intestinal immunity. Nature 542**,** 242-245.

Seok, S. H., Lee, W., Jiang, L., Molugu, K., Zheng, A., Li, Y., Park, S., Bradfield, C. A., and Xing, Y. (2017). Structural hierarchy controlling dimerization and target DNA recognition in the AHR transcriptional complex. Proc. Natl. Acad. Sci. U. S. A., https://doi.org/10.1073/pnas.161703511

Sogawa, K., Nakano, R., Kobayashi, A., Kikuchi, Y., Ohe, N., Matsushita, N., and Fujii-Kuriyama, Y. (1995). Possible function of Ah receptor nuclear translocator (Arnt) homodimer in transcriptional regulation. Proc. Natl. Acad. Sci. U. S. A 92**,** 1936-1940.

Stockinger, B., Meglio, P. D., Gialitakis, M., and Duarte, J. H. (2014). The aryl hydrocarbon receptor: multitasking in the immune system. Annu. Rev. Immunol. 32**,** 403-432.

Su, J. M., Lin, P., and Chang, H. (2013). Prognostic value of nuclear translocation of aryl hydrocarbon receptor for non-small cell lung cancer. Anticancer Res. 33**,** 3953-3961.

Swanson, H. I., Chan, W. K., and Bradfield, C. A. (1995). DNA binding specificities and pairing rules of the Ah receptor, ARNT, and SIM proteins. J Biol. Chem. 270**,** 26292-26302.

Terwilliger, T. C., Adams, P. D., Read, R. J., McCoy, A. J., Moriarty, N. W., Grosse-Kunstleve, R. W., Afonine, P. V., Zwart, P. H., and Hung, L. W. (2009). Decision-making in structure solution using Bayesian estimates of map quality: the PHENIX AutoSol wizard. Acta Crystallogr. D. Biol. Crystallogr. 65**,** 582-601.

Veldhoen, M., Hirota, K., Christensen, J., O'Garra, A., and Stockinger, B. (2009). Natural agonists for aryl hydrocarbon receptor in culture medium are essential for optimal differentiation of Th17 T cells. J. Exp. Med. 206**,** 43-49.

Veldhoen, M., Hirota, K., Westendorf, A. M., Buer, J., Dumoutier, L., Renauld, J. C., and Stockinger, B. (2008). The aryl hydrocarbon receptor links TH17-cell-mediated autoimmunity to environmental toxins. Nature 453**,** 106-109.

Wang, Z., Wu, Y., Li, L., and Su, X. D. (2013). Intermolecular recognition revealed by the complex structure of human CLOCK-BMAL1 basic helix-loop-helix domains with E-box DNA. Cell Res. 23**,** 213-224.

Wincent, E., Bengtsson, J., Bardbori, A. M., Alsberg, T., Luecke, S., Rannug, U., and Rannug, A. (2012). Inhibition of cytochrome P4501-dependent clearance of the endogenous agonist FICZ as a mechanism for activation of the aryl hydrocarbon receptor. Proceedings of the National Academy of Sciences 109**,** 4479-4484.

Winn, M. D., Ballard, C. C., Cowtan, K. D., Dodson, E. J., Emsley, P., Evans, P. R., Keegan, R. M., Krissinel, E. B., Leslie, A. G., McCoy, A., McNicholas, S. J., Murshudov, G. N., Pannu, N. S., Potterton, E. A., Powell, H. R., Read, R. J., Vagin, A., and Wilson, K. S. (2011). Overview of the CCP4 suite and current developments. Acta Crystallogr. D. Biol. Crystallogr. 67**,** 235-242.

Wu, D., Potluri, N., Kim, Y., and Rastinejad, F. (2013). Structure and dimerization properties of the aryl hydrocarbon receptor PAS-A domain. Mol. Cell Biol. 33**,** 4346-4356.

Wu, D., Potluri, N., Lu, J., Kim, Y., and Rastinejad, F. (2015). Structural integration in hypoxia-inducible factors. Nature 524**,** 303-308.

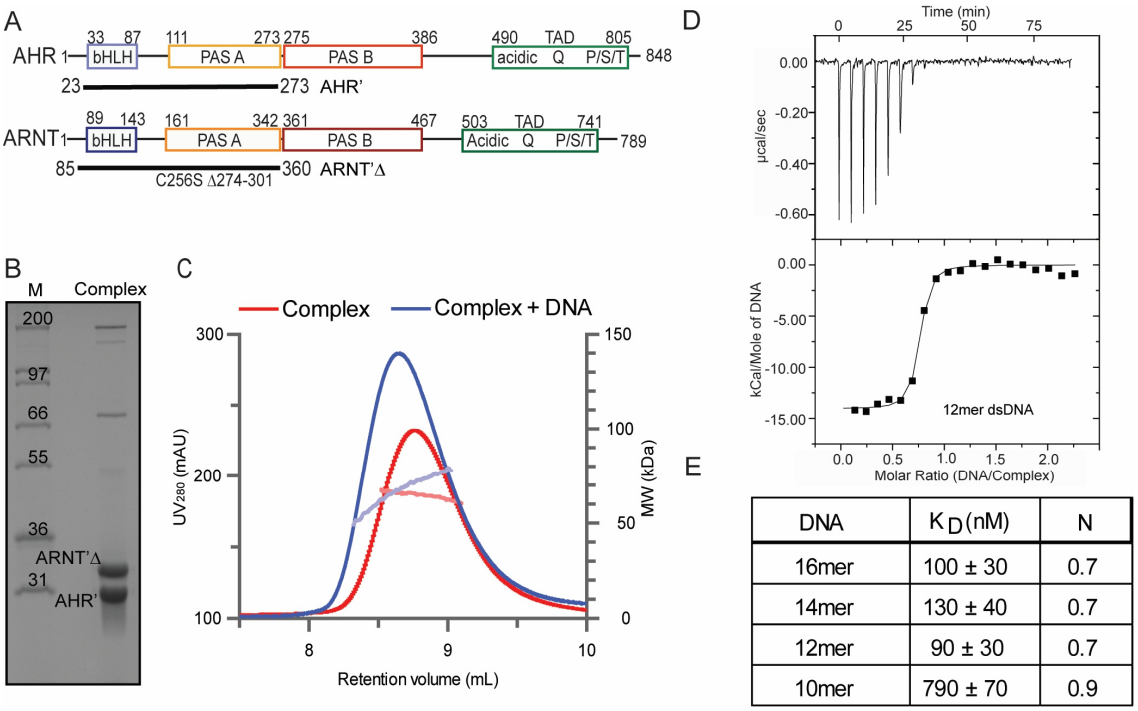
Wu, D., Su, X., Potluri, N., Kim, Y., and Rastinejad, F. (2016). NPAS1-ARNT and NPAS3-ARNT crystal structures implicate the bHLH-PAS family as multi-ligand binding transcription factors. Elife. 5**,** 1-15.

Yao, E. F., and Denison, M. S. (1992). DNA sequence determinants for binding of transformed Ah receptor to a dioxin-responsive enhancer. Biochemistry (Mosc). 31**,** 5060-5067.

Yildiz, O., Doi, M., Yujnovsky, I., Cardone, L., Berndt, A., Hennig, S., Schulze, S., Urbanke, C., Sassone-Corsi, P., and Wolf, E. (2005). Crystal structure and interactions of the PAS repeat region of the Drosophila clock protein PERIOD. Mol. Cell 17**,** 69-82.

Zelante, T., Iannitti, R. G., Cunha, C., De Luca, A., Giovannini, G., Pieraccini, G., Zecchi, R., D`Angelo, C., Massi-Benedetti, C., and Fallarino, F. (2013). Tryptophan catabolites from microbiota engage aryl hydrocarbon receptor and balance mucosal reactivity via interleukin-22. Immunity 39**,** 372-385.

Figures:



DNA	K_D (nM)	N
16mer	100 ± 30	0.7
14mer	130 ± 40	0.7
12mer	90 ± 30	0.7
10mer	790 ± 70	0.9

Figure 1: The dimeric AHR':ARNT'Δ complex binds with high affinity to DNA

A) Domain architecture of AHR and ARNT. Numbers at the domain boundaries refer to the amino acids of the human proteins. The crystallized constructs are indicated below. B) SDS gel showing the purified AHR':ARNT'Δ complex. M: marker. C) Analytical gelfiltration analysis coupled to RALS using 150 µg of the AHR':ARNT'Δ complex in the absence (red) and presence (blue) of the 12mer dsDNA. UV absorbance is shown in dark colors and refers to the left y-axis, while the absolute molecular mass as determined by RALS is shown in light colors and refers to the right y-axis. D) Top: A DNA solution containing the 12mer dsDNA with the central DRE motif was titrated to the AHR':ARNT'Δ complex at 20 °C, and the time-dependent resulting heat changes were monitored by ITC. Bottom: Binding isotherms were derived by integration of the heat peaks and fitted to a quadratic binding equation. E) Table showing the equilibrium dissociation constants (K_D; middle panel) of the 16mer, 14mer, 12mer and 10mer ds DNA (left panel) to the AHR':ARNT'Δ complex determined by ITC measurements. Errors represent standard errors of the fit. The substoichiometric binding numbers N may be explained by incomplete annealing of the titrated dsDNA oligomers. See also Fig. S1.

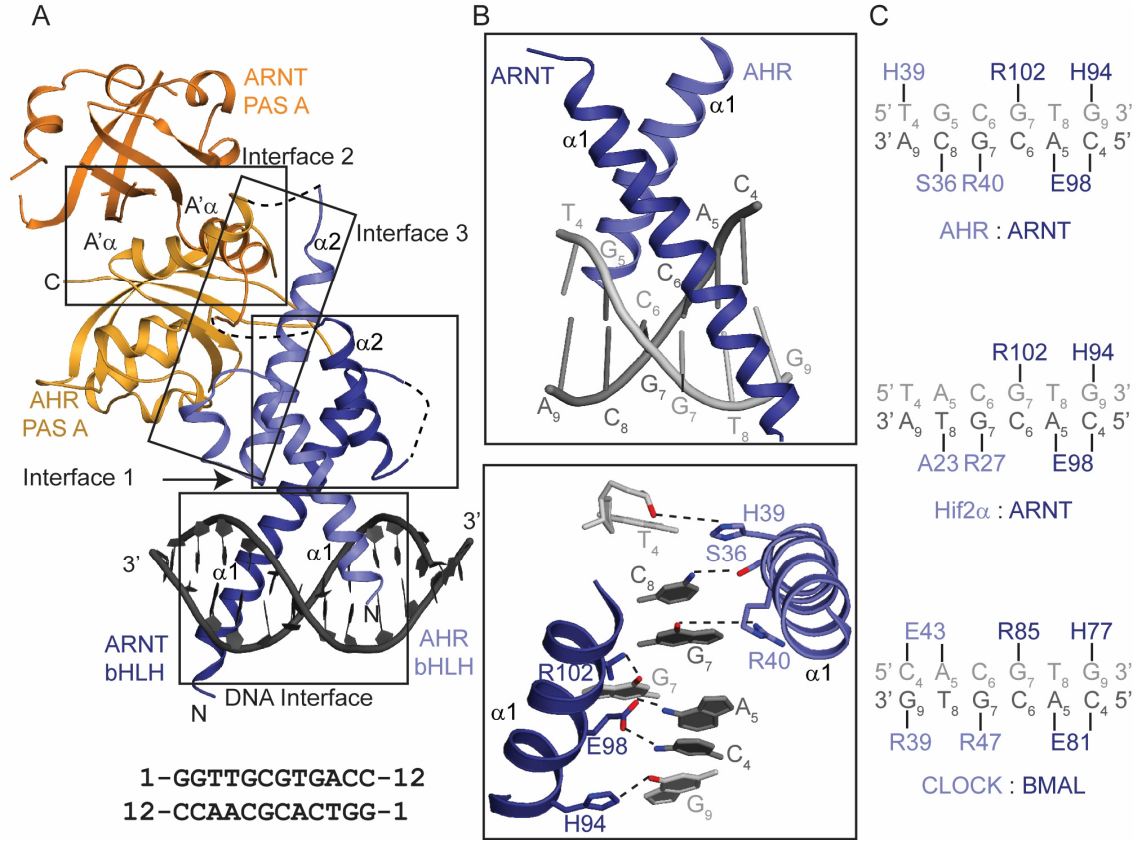


Figure 2: Structural basis of DNA binding.

A) Crystal structure of the AHR':ARNT'Δ complex bound to DNA (domain colors as in Fig. 1A). Distinct interfaces in the AHR:ARNT complex are boxed and shown in higher magnification in Fig 2B and 3. The nucleotide sequence of the 12mer dsDNA oligomer is shown at the bottom. N- and C-termini are marked. B) Structural details of DNA binding by the AHR':ARNT'Δ complex. C) Schematic interaction diagram of various bHLH-PAS complexes with DNA. The HIF-2α:ARNT and CLOCK:BMAL interactions were derived from pdb entry 4ZPK (Wu et al., 2015) and 4H10 (Wang et al., 2013), respectively. See also Fig. S2.

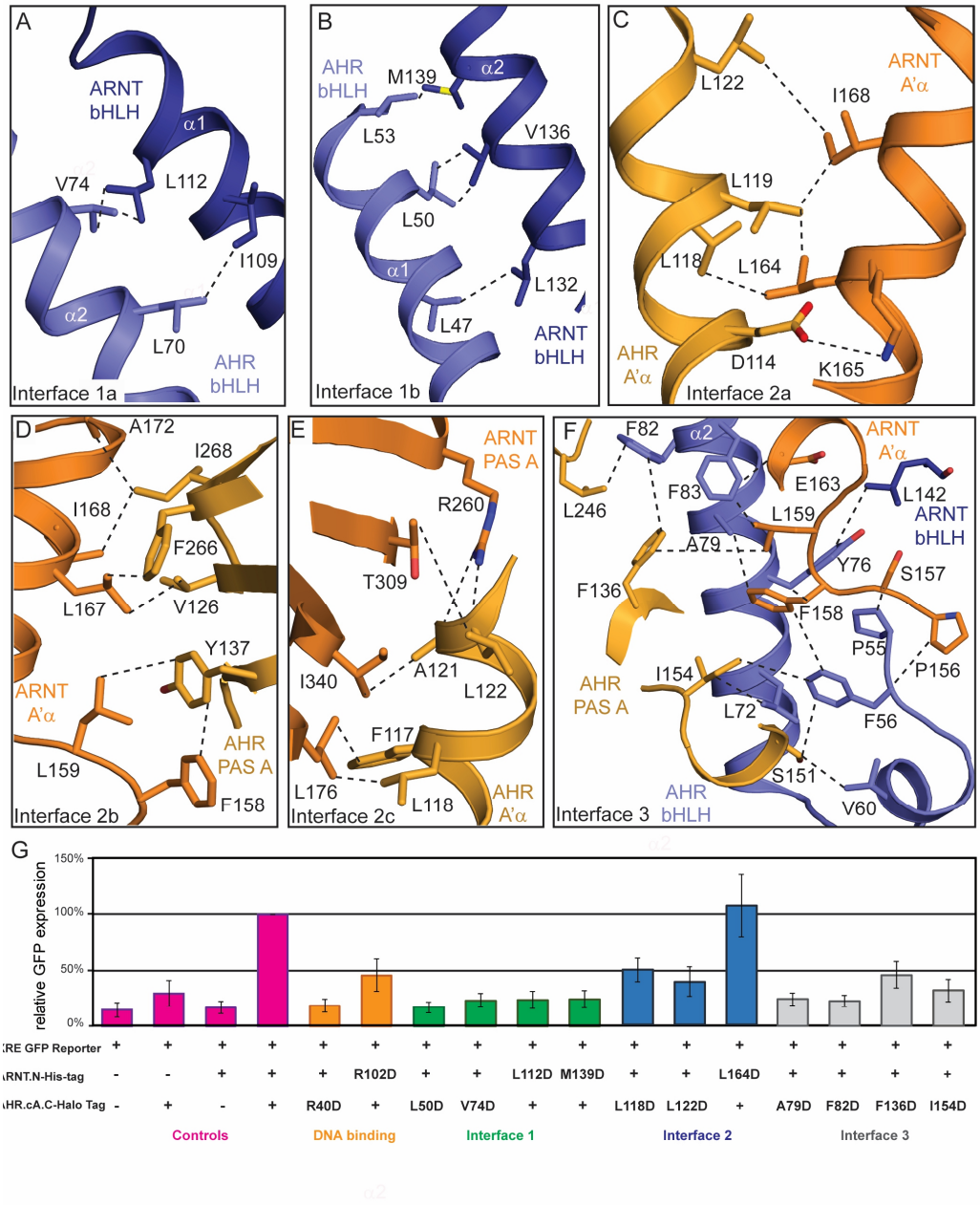


Figure 3: Structural determinants of oligomerization.

A-F) Structural details of the indicated interfaces. Selected residues involved in interface formation are shown in stick representation. Dotted lines represent hydrophobic or polar contacts.
G) Probing the effect of AHR:ARNT interface mutations on DRE reporter expression containing repeats of the 21bp motif TGAGTTCTCACGCTAGCAGAT in HEK293T cells. Publicly available RNAseq data (www.proteinatlas.org) suggested that HEK293T cells express endogenous AHR and ARNT at low levels, resulting in low background reporter activity. Overexpression of AHR.cA or ARNT alone has limited effects on reporter activity without cotransfection of the second subunit, suggesting both are limiting in HEK293T cells. All tested mutants, except L164D, show significantly decreased DRE reporter activity relative to cotransfection of non-mutagenized forms. Data represent the mean +/- SEM of three biological replicates normalized to the highest intensity result, where within each biological replicate each condition was repeated in triplicate. See also Fig. S3 and S4.

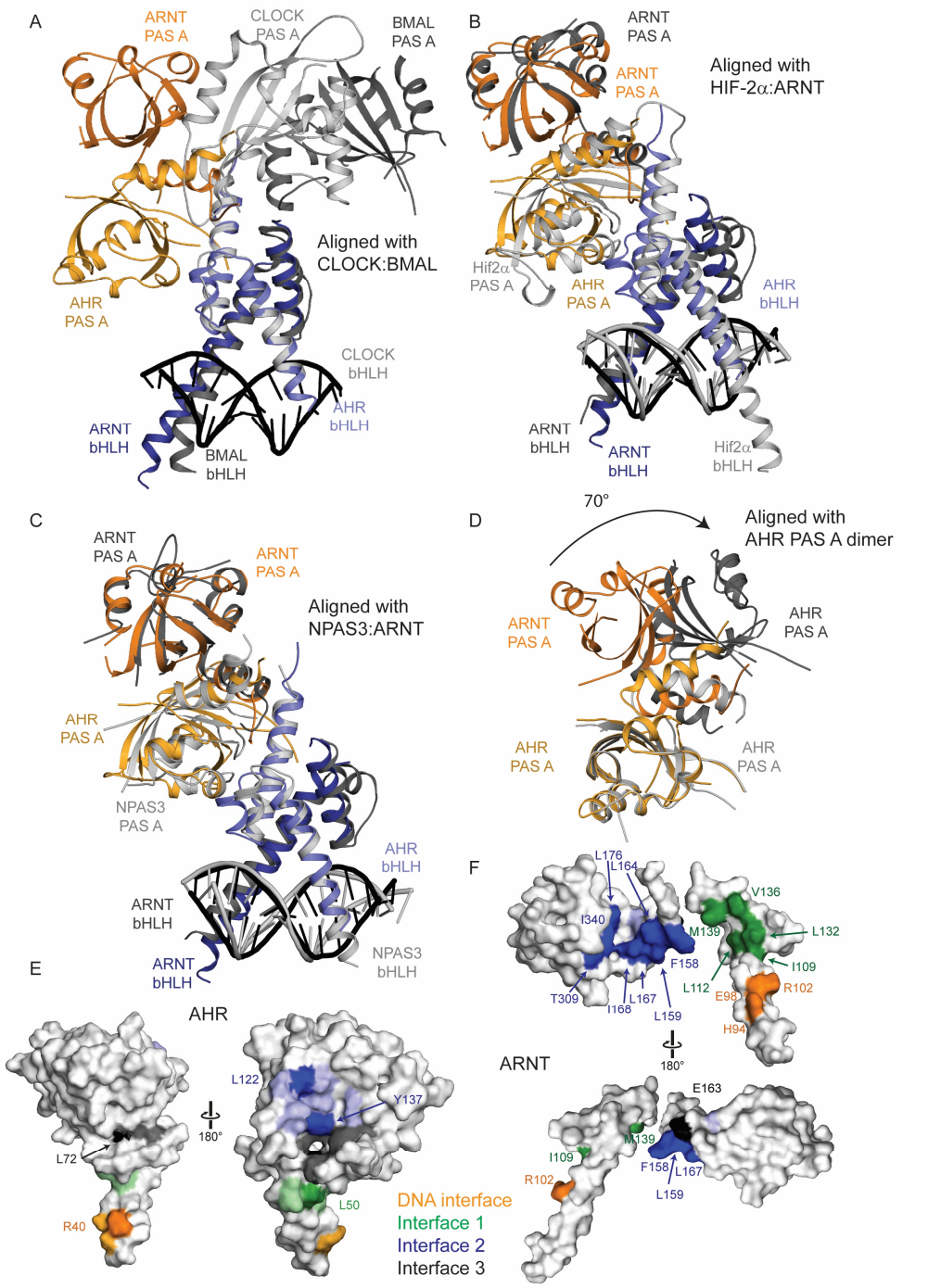

A
ARNT PAS A
CLOCK PAS A
BMAL PAS A
Aligned with CLOCK:BMAL
AHR PAS A
CLOCK bHLH
AHR bHLH
ARNT bHLH
BMAL bHLH
B
ARNT PAS A
ARNT PAS A
Aligned with HIF-2α:ARNT
Hif2α PAS A
AHR PAS A
AHR bHLH
ARNT bHLH
ARNT bHLH
Hif2α bHLH
C
ARNT PAS A
ARNT PAS A
Aligned with NPAS3:ARNT
AHR PAS A
NPAS3 PAS A
AHR bHLH
ARNT bHLH
ARNT bHLH
NPAS3 bHLH
D
70°
Aligned with AHR PAS A dimer
ARNT PAS A
AHR PAS A
AHR PAS A
AHR PAS A
E
AHR
L122
Y137
180°
L72
R40
L50
F
L176
L164
V136
I340
M139
L132
F158
I109
L112
E98
R102
T309
I168
L167
L159
H94
180°
ARNT
E163
M139
I109
F158
L167
R102
L159
DNA interface
Interface 1
Interface 2
Interface 3

Figure 4: Structural comparisons to other bHLH transcription factors.

(A-D) Superposition of the AHR':ARNT'Δ DNA complex (colors as in Fig. 2A) with (A) the bHLH-PAS A domains of the BMAL:CLOCK structure (pdb 4F3L), (B) the bHLH and PAS A domains of the HIF-2α:ARNT-DNA complex (pdb 4ZPK), (C) the bHLH and PAS A domains of the NPAS3:ARNT-DNA complex (pdb 5SY7), (D) the dimeric PAS A domain of human AHR (pdb 4M4X). For the latter, two AHR domains were superimposed, indicating a 70° rotation of the second AHR versus the ARNT PAS A domain. (E,F) The conservation of human class I (Fig. S3) and class II (Fig. S4) bHLH transcription factor sequences was plotted onto the four interaction interfaces of the AHR':ARNT'Δ complex. Shown are two views on the AHR (E) and ARNT (F) molecules. 80% type conserved residues in AHR and 100% type conserved residues in ARNT are shown in darker colours, non-conserved residues in lighter colours (see legend for the colour scheme of the interfaces). See also Fig. S3 and S4.

Data collection	**Er^{3+} data set**
Beamline	BESSY II BL 14.1
Wavelength [Å]	1.476
Space group	$P6_522$
Cell dimensions	
a, b, c [Å]	91.32, 91.32, 461.91
α, β, γ [°]	90.00, 90.00, 120.00
Resolution [Å][a]	50.00 (3.49)- 3.3
R_{meas} [%]	12.6 (63.7)
I/σ	17.48 (4.5)
Completeness [%]	99.6 (98.5)
Redundancy	20.2 (18.6)
Refinement	
Resolution [Å]	50.00 (3.47) - 3.3
No. reflections	18,323 (2,352)
R_{work}/R_{free} [%][b]	29.2/33.3 (31.6/41.5)
Complexes/ASU	2
No. protein/DNA atoms/ASU	5,836
Avg. B-factor [$Å^2$]	106
No. Er^{3+} atoms	12
Avg. B-factor [$Å^2$]	116
R.m.s. deviations	
Bond length [Å]	0.005
Angles [°]	0.678
Ramachandran favored [%]	91
Ramachandran outlier [%]	0.5

Table 1: **Data collection and refinement statistics of the Er(III) containing AHR'/ARNT'Δ complex bound to 12mer dsDNA.**

[a] Numbers in parentheses apply to the highest resolution shell.
[b] 5% of the reflections were excluded from refinement to calculate R_{free}. The relatively high R factors may be explained by the two flexible PAS A domains of complex 2, which could be modelled only incompletely.
Complexes/ASU – AHR'-ARNT'Δ–DNA complexes per asymmetric unit.

STAR methods

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Oliver Daumke (oliver.daumke@mdc-berlin.de).

METHOD DETAILS

Protein expression and purification: A codon-optimized cDNA of human (hs) AHR encoding amino acid residues 23-273 was commercially synthesized (MWG Eurofins) and cloned into pET32-EK/LIC (Novagen). Mouse ARNT (residues 85-345), Δ274-301, C256S was cloned into pET30-EK/LIC (Novagen). Both constructs were co-expressed as N-terminal Hexa-His-fusion proteins in BL21 Rosetta2(DE3) cells, followed by a PreScission cleavage site. Bacteria were grown to an OD_{600} of 0.5 in TB medium when protein expression was induced with 200 µM Isopropyl-β-D-thiogalactopyranosid (IPTG), followed by overnight expression at 18 °C. Cells were resuspended in ice-cold lysis buffer (20 mM HEPES/NaOH pH 8.0, 200 mM NaCl, 100 µM Pefabloc SC (Roth), 1 µM DNAse (Roth), 5 mM β-mercaptoethanol (βME) and lyzed using a microfluidizer (Microfluidics, Newton, USA). After centrifugation at 100,000 x g for 45 min at 4 °C, the supernatant was applied to a Ni-nitroloacetic acid (NTA) column pre-equilibrated with equilibration buffer (20 mM HEPES/NaOH pH 8.0, 200 mM NaCl, 5 mM βME), and afterwards extensively washed with the same buffer. The protein was eluted using equilibration buffer containing 200 mM imidazole. PreScission protease was added in a 1:200 ratio (w/w) to cleave the N-terminal His-tag of AHR and ARNT, while dialyzing overnight against 20 mM HEPES/NaOH pH 8.0, 200 mM NaCl, 5 mM dithiothreitole (DTT) at 4 °C. Cleaved proteins were concentrated using Amicon concentrators (Millipore) and further purified by SEC on a Superdex 200 16/60 column using 20 mM HEPES/NaOH pH 8.0, 200 mM NaCl, 5 mM DTT as running buffer. Protein containing fractions were pooled and concentrated to 10 mg/mL. The protein complex was aliquoted, snap frozen and stored at -80 °C.

Mouse ARNT (residues 85-345), Δ274-301, C256S, cloned into pET 30EK-LIC (Novagen), was expressed as N-terminal Hexa-His-fusion proteins in BL21 Rosetta2(DE3) cells and purified as described above for the hetero-dimeric complex.

DNA oligomer annealing: Equal concentrations (typically 800 µM) of DNA strand A and its complementary strand B were incubated at 90 °C for 5 min, followed by a slow cooling phase to room temperature over a time range of 6 h to allow for strand annealing. For simplicity, the sequence of only one strand is listed in this manuscript.

RALS Analysis: A RALS-refractive index detector (Malvern) coupled to an analytical SEC column (Superdex 75 10/300) was used to determine absolute molecular masses of the applied samples. In each run, 100 µL of a 1.5 mg/mL protein solution was applied. Data were analyzed with the provided software OmniSec. The running buffer contained 20 mM HEPES pH 8.0, 200 mM NaCl, 5 mM DTT. For testing complex formation, a 12mer ds oligomer with the sequence 5‘-GGTTGCGTGACC-3‘ was used. For ARNT homo-dimerization experiments, an 18mer ds DNA with sequence 5’-GCATAACACGTGTTATGC-3‘ was employed (presumed binding sites underlined).

Isothermal titration calorimetry: ITC experiments were carried out at 20 °C in an ITC-200 device (GE Healthcare). dsDNA oligomers of various length (16mer: 5‘-AAGGTTGCGTGACCTT-3‘, 14mer: 5‘-AGGTTGCGTGACCT-3‘, 12mer: 5‘-GGTTGCGTGACC-3‘, 10mer: 5‘-GTTGCGTGAC-3‘) in 20 mM HEPES/NaOH pH 8.0, 200 mM NaCl, 5 mM DTT at a concentration of 400-550 µM was titrated into the reaction cell containing the AHR’:ARNT’Δ complex (30 µM). Binding isotherms were fitted and equilibrium dissociation constants calculated using the Microcal ORIGIN software.

Crystallization and structure determination: For crystallization trials, the protein complex was incubated with a 1.2-fold molar excess of the 12mer dsDNA (12mer: 5‘-GGTTGCGTGACC-3‘) and purified by SEC using a Superdex 200 10/300 column. Fractions containing the AHR’:ARNT’Δ-DNA complex were pooled, concentrated to 8.5 mg/mL and crystallized in a 24 well format by using the hanging drop vapor diffusion technique. 1 µl of the protein:DNA solution was mixed with 1.5 µL of the reservoir solution containing 10 mM HEPES/NaOH pH 6.8, 18% PEG3350, 200 mM ammonium formate. Plate-like crystals appeared after 5 days at 4 °C and reached their final size after 8 days (~ 0.4

mm x 0.1 mm x 0.025 mm). For crystal dehydration and derivatization with an anomalous scatterer, crystals were transferred for 1.5 h into a solution containing 35% PEG3350, 25 mM HEPES/NaOH pH 8.0, 100 mM NaCl, 2 mM DTT, 230 mM ammonium formate and 5 mM Erbium-(III)-acetatedihydrate at 4 °C and afterwards directly flash cooled in liquid nitrogen. A full data set from a single crystal was recorded at beamline BL 14.1 at BESSY II, using a PILATUS 6M detector. Calculation of an optimal data collection strategy was done with the Mosflm software (Leslie, 2006). The SAD dataset was processed using the XDS program suite (Kabsch, 2010). The phase problem was solved by using the anomalous signal of 12 Er^{3+} atoms in the asymmetric unit using AutoSol (Terwilliger et al., 2009), which resulted in an unambiguous electron density map. Atomic model building was done by C*oot* (Emsley et al., 2010) and aided by the known bHLH domain of CLOCK-BMAL and the PAS-A domain of AHR. Iterative refinement was done using PHENIX (Adams et al., 2010) at a maximum resolution of 3.3 Å including the experimental phases and secondary structure restraints throughout refinement. TLS (translation, libration, screw axis rotation) refinement was done with one TLS group per domain or DNA oligomer. For the erbium atoms, anomalous B factors were refined. Chain A contains residues 34-88, 109-250, 261-272, chain B residues 85-119, 126-143, 152-228, 259-269, 302-316, 324-329, 334-345, chain C residues 32-88, 107-173, 211-227, 234-251, 255-271, chain D residues 84-144, 148-177, 188-196, 199-203, 207-211, 259-269, 307-314, 337-342, chain E and G the DNA sequence GGTCACGCAACC, chain F and H the DNA sequence GGTTGCGTGACC. Figures were prepared using PyMOL Molecular Graphics System, Version 1.3 Schrödinger, LLC. Evaluation of atom contacts and geometry of the atomic model was done by the Molprobity server (Chen et al., 2010) and ccp4 (Winn et al., 2011). Interface sizes were calculated with the PISA server (Krissinel et al., 2007).

Plasmid mutagenesis: hsARNT and hsAHR were PCR amplified using Phusion polymerase (NEB) and cloned into pDONR201 using Gateway® cloning (Invitrogen) following the manufacturer's instructions. Site directed mutagenesis was performed using primers (Table S2) designed with the NEBBaseChanger software (http://nebasechanger.neb.com/) and the Q5 SDM kit (NEB) following the manufacturer's instructions. Mutagenized ARNT clones were shuttled into the Gateway® pDEST™26 vector (Invitrogen) expressing N-terminally His tagged proteins, while AHR mutants were shuttled into

a C-terminally Halo tagged vector (gift of S. Pusch). Plasmid DNA was prepared for transfection using the Qiagen Plasmid Midi Kit.

HEK cell transformation: HEK293T cells were transfected in 96 well plates with the indicated combination of AHR, ARNT and Cignal xenobiotic response element (XRE)-GFP reporter (Qiagen) plasmids using FuGENE®HD (Promega) following the manufacturer's instructions (FuGENE®HD:DNA ratio 3:1). After 30 hours, the activity of the XRE-GFP reporter was measured using a Pherastar FS plate reader (BMG Labtech). Data represent the mean of three biological replicates normalized to the highest intensity result, where within each biological replicate each condition was repeated in triplicate.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data fitting including baseline corrections, normalization, calculation of mean and error (SEM), and statistical tests were carried out in GraphPad Prism (version 5). The number of biological and technical replicates as well as the entity plotted are indicated in the figure legends. Data baseline correction and normalization, where applied, were indicated in the corresponding method section and in the axis labels. No explicit power analysis was used.

DATA AND SOFTWARE AVAILABILITY

Coordinates and structure factors of the AHR:ARNT complex bound to DNA have been deposited in the protein database (PDB) under accession code 5NJ8.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
E. coli strain BL21Rosetta2(DE3)	Novagen	
Chemicals, Peptides, and Recombinant Proteins		
Pefabloc SC	Carl Roth	
DNAse I	Carl Roth	
Critical Commercial Assays		
Q5 SDM kit	NEB	
FuGENE®HD	Promega	
Deposited Data		
Coordinates of DNA-bound AHR-ARNT	PDB database	5NJ8
Experimental Models: Cell Lines		
HEK293T	ATCC	
Oligonucleotides		
Codon optimized cDNA of human AHR	MWG Eurofins	
Primers	Sigma	see Table S1
Recombinant DNA		
pET32-EK/LIC	Novagen	
pET32-EK/LIC hsAHR (residues 23-273)	This manuscript	
pET30-EK/LIC	Novagen	
pET30-EK/LIC mmARNT (residues 85-345), Δ274-301, C256S	This manuscript	
pDONR201	Invitrogen	
pDONR201-hsAHR	This manuscript	
pDONR201-hsARNT	This manuscript	
Gateway® pDEST™26	Invitrogen	
Gateway® pDEST™26 hsARNT	This manuscript	
pHalo	Gift S. Pusch	
pHalo hsAHR	This manuscript	
Cignal XRE-GFP	Qiagen	
Software and Algorithms		
Mosflm	Leslie, 2006	
XDS	Kabsch, 2010	
Autosol	Terwilliger…., 2009	
C*oot*	Emsely et al., 2010	
PHENIX	Adams et al, 2010	
PyMOL Molecular Graphics System, 1.3	Schrödinger, LLC	
Molprobity	Chen et al., 2010	
CCP4	Winn et al., 2011	
PISA server	Krissinel et al., 2007	
NEBBaseChanger	NEB	http://nebasechanger.neb.com/
GraphPad Prism, version 5	GraphPad	https://www.graphpad.com/

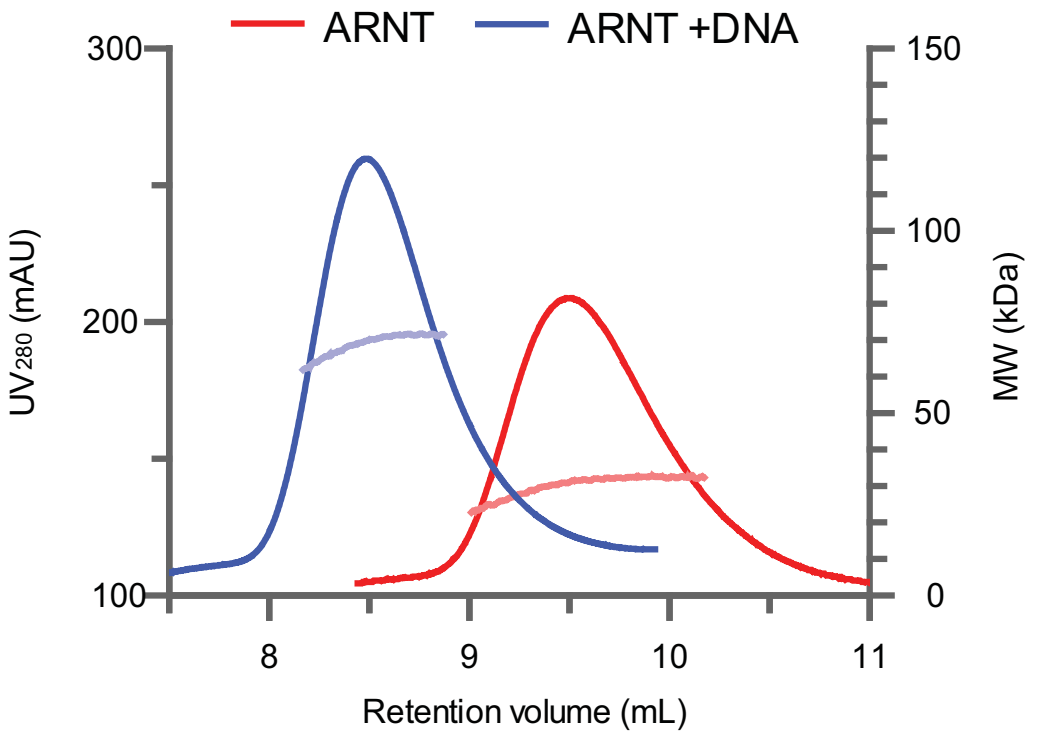


Figure S1: RALS analysis of ARNT'Δ, related to Fig. 1.
Chromatogram of the analytical gel filtration coupled to RALS of the ARNT protein in the absence and presence of a cognate 18mer dsDNA (sequence 5'-GCATAACACGTGTTATGC-3', ARNT binding site is underlined). UV absorbance is shown in dark blue/dark red and refers to the left y-axis, while the absolute molecular mass determined by RALS is shown in light blue/light red and refers to the right y-axis. The retention volume is plotted on the x-axis.

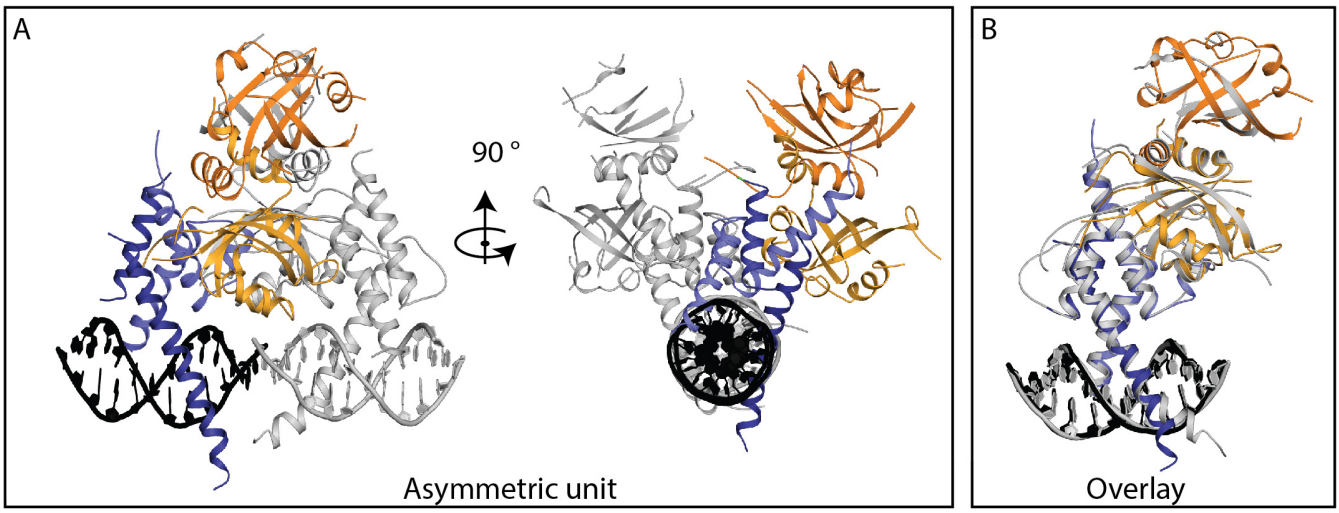
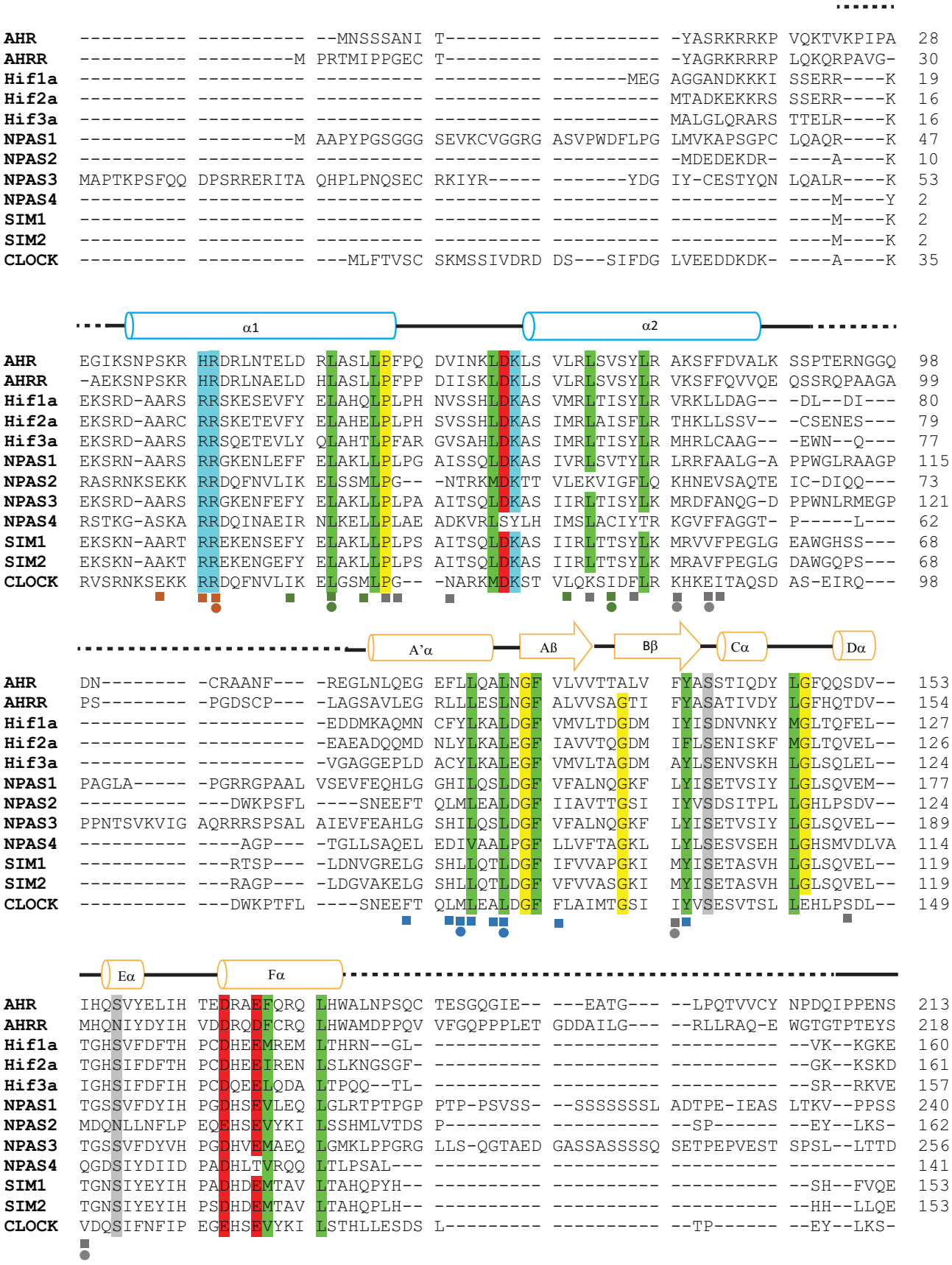


Figure S2: Content of the asymmetric unit, related to Fig. 2.
A) Ribbon type representation of the two AHR/ARNT complexes bound to the 12mer dsDNA within the asymmetric unit. The DNA oligomers assembled into a continuous DNA double strand throughout the crystal. Domains of complex 1 are color coded as in Figure 2, complex 2 is colored in gray. The PAS domains of complex 2 are not stabilized by crystal contacts, resulting in disorder, a poor electron density and an incomplete model. B) Overlay of complex 1 and 2 of the asymmetric unit. The two complexes overlaid with a root mean square deviation of 1.2 Å.


α1
α2
A'α
Aß
Bβ
Cα
Dα
Eα
Fα

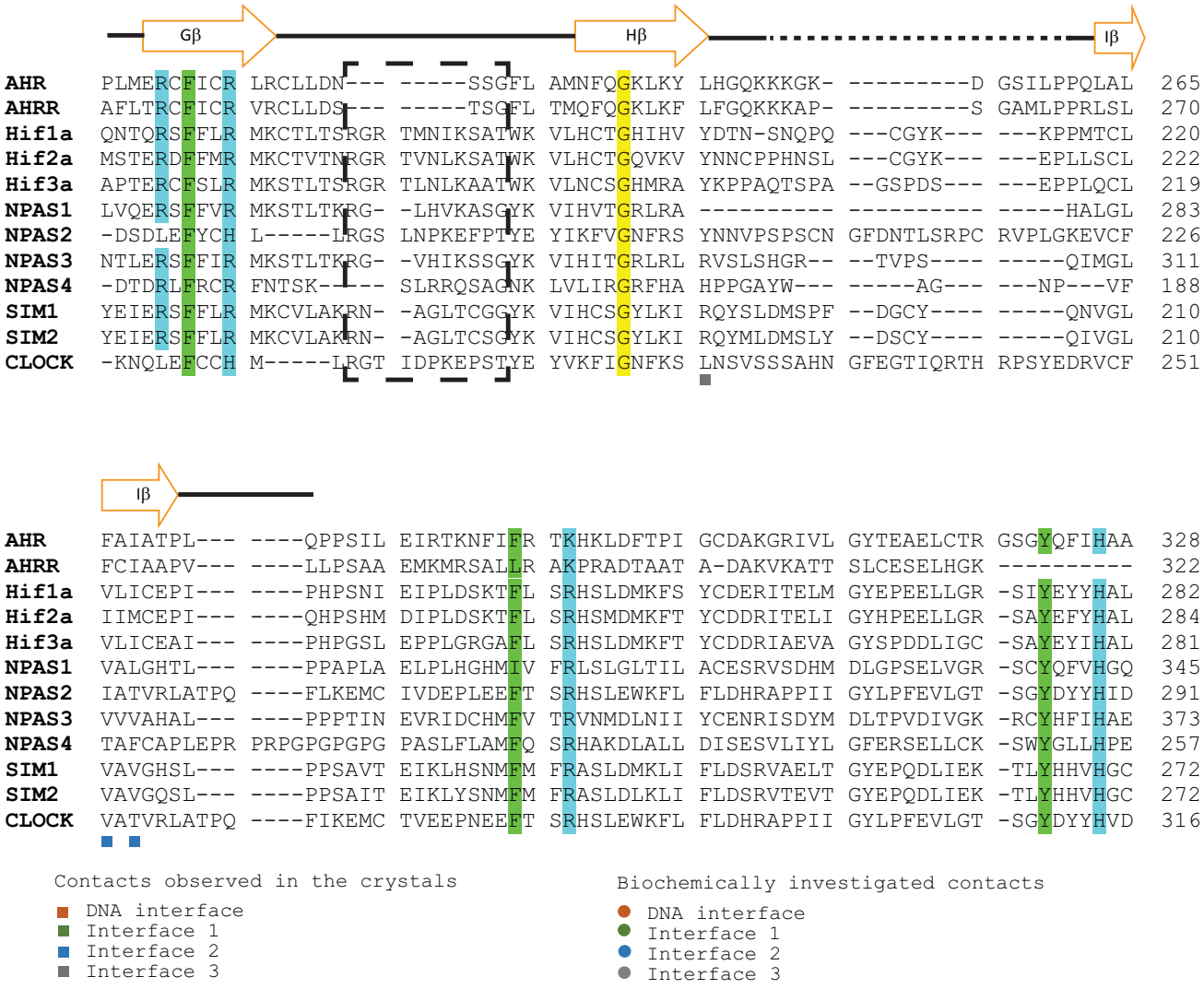


Figure S3: Sequence alignment of human class I bHLH-PAS domain transcription factors, related to Fig. 3 and 4.

bHLH and PAS A domains of the following human sequences were aligned with ClustalW (Larkin et al., 2007): AHR (expasy accession code P35869), AHR repressor (A9YTQ3), Hif-1α (Q16665), HIF-2α (Q99814), HIF3α (Q9Y2N7), NPAS1 (Q99742), NPAS2 (Q99743), NPAS3 (Q8IXF0), NPAS4 (Q8IUM7), SIM1 (P81133), SIM2 (Q14190), CLOCK (O15516). The secondary structure as observed in the AHR'-ARNT'Δ complex is indicated on top. Residues with a type conservation of greater than 80% are color-coded (D,E in red; R,K,H in blue; N,Q,S,T in grey; L, I, V, F, Y, W, M, C in green, A,G,P in yellow). Interface residues, as observed in our structure, are marked with a square. Residues which have previously been shown to be crucial for dimerization (Wu et al., 2013; Wu et al., 2015; Wu et al., 2016) are indicated with a circle. The HIF-2α loop shown to be involved in DNA binding is boxed.

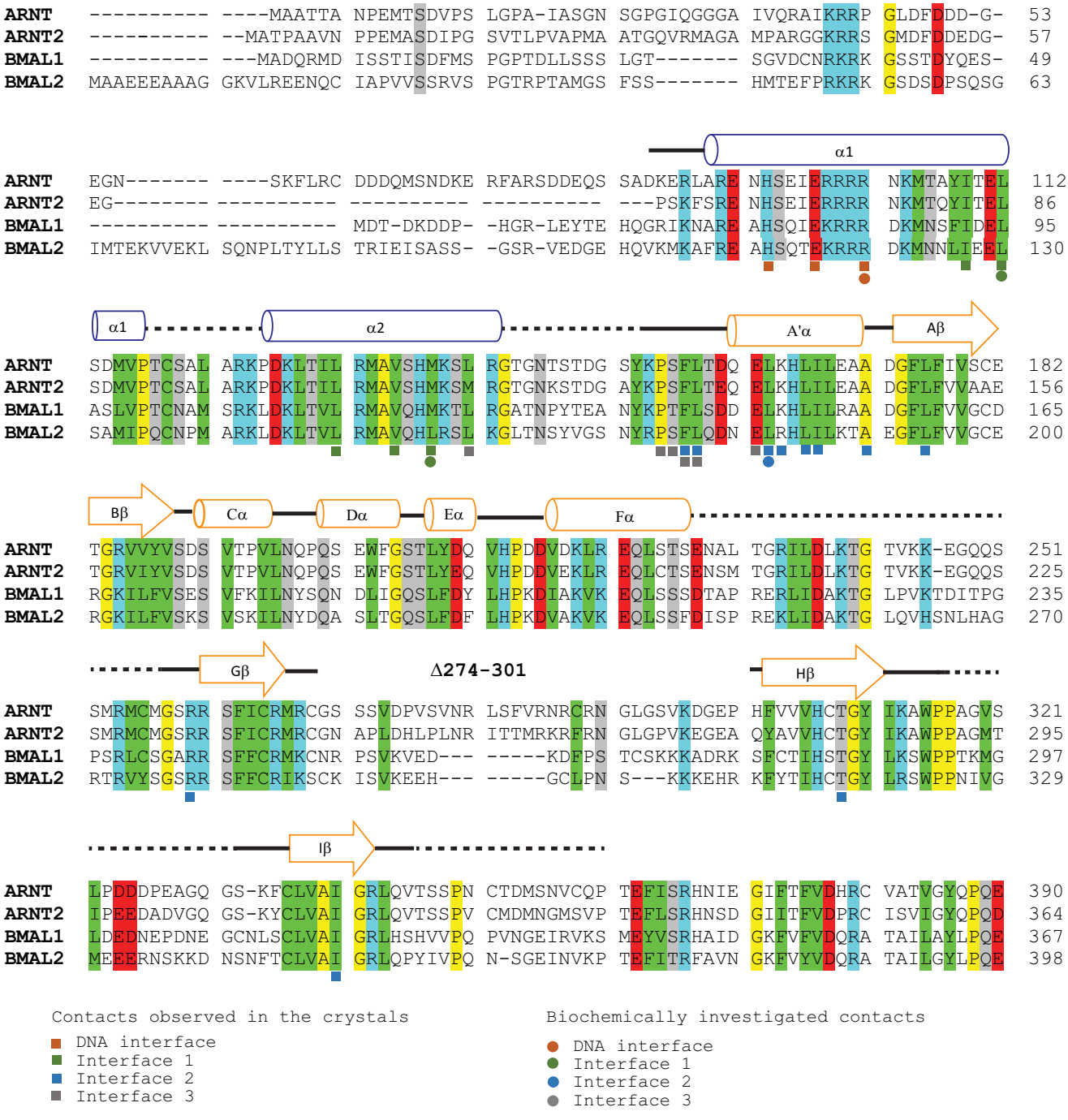


Figure S4: Sequence alignment of human class II bHLH-PAS domain transcription factors, related to Fig. 3 and 4.

bHLH and PAS A domains of the following human sequences were aligned with ClustalW: ARNT (expasy accession code P27540), ARNT2 (Q9HBZ2), BMAL1 (O00327) and BMAL2 (Q8WYA1). The secondary structure as observed in the AHR'-ARNT'Δ complex is indicated on top. Residues with a type conservation of 100% are color-coded (D,E in red; R,K,H in blue; N,Q,S,T in grey; L, I, V, F, Y, W, M, C in green, A,G,P in yellow). Interface residues, as observed in our structure, are marked with a square. Residues which have previously been shown to be crucial for dimerization (Wu et al., 2013; Wu et al., 2015; Wu et al., 2016) are indicated with a circle.

Table S1: Primers used for site directed mutagenesis, related to Fig. 3.

hsAHR R40D SDM F Tm58	CAAGCGGCATgatGACCGACTTAATAC
hsAHR R40D SDM R Tm58	GAAGGATTTGACTTGATTCC
hsAHR L50D SDM F Tm59	GTTGGACCGTgatGCTAGCCTGC
hsAHR L50D SDM R Tm59	TCTGTATTAAGTCGGTCTC
hsAHR V74D SDM F Tm61	TAGGCTCAGCgatAGTTACCTGAG
hsAHR V74D SDM R Tm61	AGAACTGAAAGTTTGTCCAAC
hsAHR A79D SDM F Tm62	TTACCTGAGAgatAAGAGCTTCTTTGATG
hsAHR F79D SDM R Tm61	CTGACGCTGAGCCTAAGAAC
hsAHR F81D SDM F Tm61	AGCCAAGAGCgatTTTGATGTTGCATTAAAATC
hsAHR A81D SDM R Tm62	CTCAGGTAACTGACGCTG
hsAHR F136D SDM F Tm59	TGCTTTGGTCgatTATGCTTCTTCTAC
hsAHR F136D SDM R Tm59	TCTGTAGTGACAACTAATACAAAG
hsAHR I154D SDM F Tm59	GTCTGATGTCgatCATCAGAGTGTATATGAACTTATC
hsAHR I154D SDM R Tm59	TGCTGAAACCCTAGATAATC
hsAHR Δ288-421 SDM F Tm59	gCCATAATGGATCCCTTACC
hsAHR Δ288-421 SDM R Tm59	tCTAAAGATAAAATTTTTGGTCCG
hsAHR R102D SDM F Tm58	ACGGCGGCGAgatAACAAGATGAC
hsAHR R102D SDM R Tm58	TCAATTTCACTGTGATTTTCC
hsAHR L112D SDM F Tm59	CATCACAGAAgatTCAGATATGGTACCCACCTG
hsAHR L112D SDM R Tm59	TAGGCTGTCATCTTGTTC
hsAHR M139D SDM F Tm61	AGTTTCTCACgatAAGTCCTTGCGGGGAAC
hsAHR M139D SDM R Tm61	GCCATGCGTAAGATGGTTAG
hsAHR L164D SDM F Tm59	TGATCAGGAAgatAAACATTTGATCTTGGAGG
hsAHR L164D SDM R Tm59	GTGAGGAAAGACGGCTTATAG