Published in final edited form as:

Curr Opin Toxicol. 2017 February ; 2: 42–49. doi:10.1016/j.cotox.2017.01.011.

Molecular modeling of the AhR structure and interactions can shed light on ligand-dependent activation and transformation mechanisms

Laura Bonati*, Dario Corrada, Sara Giani Tagliabue, and Stefano Motta

Department of Earth and Environmental Sciences, University of Milano-Bicocca, Piazza della Scienza 1, 20126 Milan, Italy

Abstract

Molecular modeling has given important contributions to elucidation of the main stages in the AhR signal transduction pathway. Despite the lack of experimentally determined structures of the AhR functional domains, information derived from homologous systems has been exploited for modeling their structure and interactions. Homology models of the AhR PASB domain have provided information on the binding cavity and contributed to elucidate species-specific differences in ligand binding. Molecular Docking simulations of the ligand binding process have given insights into differences in binding of diverse agonists, antagonists, and selective AhR modulators, and their application to virtual screening of large databases of compounds have allowed identification of novel AhR ligands. Recently available structural information on protein-protein and protein-DNA complexes of other bHLH-PAS systems has opened the way for modeling the AhR:ARNT dimer structure and investigating the mechanisms of AhR transformation and DNA binding. Future research directions should include simulation of the protein dynamics to obtain a more reliable description of intermolecular interactions involved in signal transmission.

Keywords

molecular modeling; molecular docking; molecular dynamics; ligand-binding; protein-protein interactions; dimerization

1. Introduction

It is well known that the Ah receptor (AhR) responds to diverse ligands with the induction of expression of a large battery of genes and production of different biological and toxic effects [1–3]. In addition to the best-characterized high affinity ligands, including toxic halogenated aromatic hydrocarbons (HAHs), polycyclic aromatic hydrocarbons (PAHs) and PAH-like

*Corresponding Author: Department of Earth and Environmental Sciences, University of Milano-Bicocca, Piazza della Scienza 1, 20126 Milan, Italy. Phone: (+39) 0264482821. Fax (+39)0264482839. laura.bonati@unimib.it.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

chemicals [4–6], the AhR can bind and be activated or inhibited by a number of natural, endogenous and synthetic agonists and antagonists with diverse structures and physico-chemical characteristics [6,7]. Recently, it has been shown that this promiscuous receptor also plays a key regulatory role in a variety of endogenous developmental and immune response processes [8,9]. Although the mechanism of induction of gene expression (the ‘classical’ AhR signaling pathway, Figure 1) has been studied for many years [10], significant unanswered questions still remain. Major issues concern the molecular description of the key events regulating the three main stages in the AhR mechanism: binding and activation of the AhR by diverse ligands and the role played by the chaperone heat shock protein 90 (hsp90) in this process; ligand-dependent transformation of the AhR into its transcriptionally active form, related to dimerization with the AhR nuclear translocator (ARNT); binding of the ligand:AhR:ARNT complex to its specific DNA recognition site.

An in-depth molecular understanding of these events may not only provide insight into the mechanisms of activation and transformation of the AhR but also shed light on the reasons of ligand-dependent differences in AhR functionality. The lack of experimentally determined structures of the AhR has hampered such analyses for a long time. However, information derived from homologous systems has more recently been exploited for modeling the AhR structure and interactions.

2. Molecular insights

Structural information on the AhR N-terminal region, including the basic helix-loop-helix (bHLH) and the Per-ARNT-Sim (PAS) domains, is particularly relevant for the above molecular analyses. As indicated by extensive mutagenesis and deletion analyses [11–13] the PASB repeat is responsible both for ligand binding to the cytosolic AhR and, with the bHLH, for binding to hsp90 which role is to direct the AhR correct folding and maintain the ligand-binding conformation [14,15]. Both the bHLH and PAS domains are involved in AhR:ARNT dimerization in the cell nucleus, and the bHLH functions in DNA binding. However, no X-ray or NMR-determined structure of these domains has been reported to date, except for the AhR PASA homodimer, recently solved [16].

Also experimental structures of homologous proteins were not available until the late nineties, and the first structure depositions regarded distant homologous systems in the PAS superfamily [17–25]. From 2003 an increasing structural knowledge on homologous proteins belonging to the bHLH-PAS family has become available and has greatly contributed to advancements in molecular understanding of the AhR structure and interactions [26–34]. In addition to well-conserved N-terminal domain structures, the bHLH-PAS proteins show similarities in the mechanisms of action [35,36]. The AhR belongs to the class I bHLH-PAS proteins (along with the hypoxia inducible factor α , HIF-2 α , the circadian locomotor output cycles kaput, CLOCK, and some neuronal factors), which role is to sense environmental signals. Class II bHLH-PAS proteins (ARNT, ARNT2, and brain muscle ARNT-like, BMAL) have regulatory roles; they form dimers with one or more class I proteins and the resulting heterodimers bind to specific DNA sites to regulate target genes [35,36]. In particular, HIF-2 α has the highest degrees of sequence identity with

AhR among all the reported PAS structures and shows mechanistic commonalities with AhR. A number of experimental structures of the HIF-2 α PASB, both in the ligand free (apo) form [26,28,33] and in complexes with artificial ligands (holo form) [28–31,33,34], have become available, thus providing general information about the ligand binding cavity and the binding ability of this domain. Moreover, these depositions have given information about the putative dimerization mode of the isolated PASB domain, as well as the AhR PASA homodimer structure [16] into the PASA dimerization. Recently, elucidation on the whole architecture of bHLH-PASA-PASB dimers has been made possible for the first time by two X-ray structures, the murine CLOCK:BMAL1 [32] and the human HIF-2 α :ARNT [33] heterodimers, and also the interaction of these complexes with their specific DNA recognition site has been disclosed [33,37].

Based on the knowledge of the three-dimensional structure of one (or more) homologous system/s (template/s) and on sequence alignment, homology modeling methods can be used for structure prediction of a target protein, providing reference models for mechanistic studies [38,39]. Moreover, a number of molecular modeling methods can be employed to characterize protein structures and properties related to their functions. For example, pockets and cavities, likely available for ligand binding, can be determined and characterized [40,41]. Molecular Docking methods can be employed to predict the binding geometry of ligand-protein complexes and to estimate their binding free energy [42,43]. Properties of domain surfaces and protein-protein interfaces can be characterized [44] and the structure of protein-protein complexes can be predicted by appropriate docking methods or by template-based modeling [45]. Given that proteins are dynamical objects that interconvert among a variety of conformations and that their dynamics plays a critical role in many biological processes, including ligand binding, protein-protein interactions and signaling mechanisms [46], several computational methods have been developed to study their motions. Among these, Molecular Dynamics (MD) simulations, that offer an atomistic description of such dynamical exchange, have directly contributed to the understanding of several biological mechanisms [47,48]. An extended description of structure prediction and molecular modeling methods available to date is out of the scope of this review, but a general knowledge can be gathered by the references indicated above.

3. Modeling the ligand binding domain and ligand binding

In the last years many studies have been focused on the development of structural models of the AhR ligand binding domain (LBD), in the PASB, to characterize the binding cavity and predict specific residues that play a critical role in ligand binding.

Since the first crystal structures of distant homologous proteins in the PAS superfamily became available [17–20], our group started developing homology models of the mouse AhR (mAHR) LBD [49], that provided initial information on the domain characteristics [49–51]. A more reliable structural proposal was reached when the PASB NMR structures of the bHLH-PAS HIF-2 α [26] and ARNT [27] proteins, with higher sequence identity to the AhR, were deposited. The mAHR model developed with these templates, associated to site-directed mutagenesis and functional analysis, confirmed the buried cavity in the core of the domain as the site involved in ligand binding [52]. Extension of modeling to several

mammalian AhR LBDs indicated that the physico-chemical characteristics of the residues in the binding cavity are conserved in all AhRs with high affinity for the 2,3,7,8-tetrachlorodibenzo-p-dioxin (TCDD) [53]. Mutagenesis of these residues allowed identification of the “TCDD-binding fingerprint”, *i.e.* of the group of residues associated to optimal TCDD binding [53].

Homology models of the LBD of AhRs of different species greatly contributed to rationalize species-specific differences in ligand binding to the AhR [54–57]. The comparison among the models of different isoforms of zebrafish AhRs (zfAhRs) allowed us to identify a dramatically reduced internal cavity in the zfAhR1a, which does not bind TCDD, with respect to the high affinity zfAhR2. Mutagenesis of the residues responsible for this difference restored the ability of zfAhR1a to bind TCDD and exhibit DNA-binding [54]. Similarly, residues responsible for low sensitivity of two amphibian AhRs (*Xenopus laevis* AhR1 β and *Ambystoma mexicanum* AhRs) to TCDD and to the endogenous 6-formylindolo[3,2-b]carbazole (FICZ) ligand were identified with the help of structural modeling [55,57]. Homology models of several avian AhR LBDs contributed to elucidate the importance of specific amino-acids in modulating the AhR sensitivity to TCDD and dioxin-like compounds [56].

Extensive experimental confirmations of the reliability of the AhR models prompted several groups to use them for Molecular Docking simulations of the ligand binding process [53,58–74].

Docking of the TCDD to the modeled wild-type mAhR and some mutants allowed us to interpret the role of residues belonging to the TCDD-binding fingerprint [53]. Homology models of the mAhR, zfAhR and human AhR (hAhR) LBDs, based on the same HIF-2 α PASB template structure [26], were developed by Bisson *et al.* and used for Molecular Docking of TCDD, FICZ and 2-(1'-H-indole-3'-carbonyl)-thiazole-4-carboxylic acid methyl ester (ITE) [59]. Their results allowed characterization of both the observed intra- and inter-species differences in binding of these ligands [59]. Moreover, for the first time, they proposed the use of docking for virtual screening of potential AhR ligands and their studies resulted in the identification of two flavonoids as new AhR agonists, in a database of about 500 natural compounds [59]. Several docking simulations to modeled AhR LBDs were performed [60–66,74] following the same computational protocol [59]. All these studies contributed to rationalize binding to AhR of some: agonists, like the anti-inflammatory drug leflunomide [60,63] or ligands showing anticancer properties (but not dioxin-like activities) [65,66]; antagonists, like GNF351 [62]; and selective AhR modulators (SAhRMs) [61].

All the above-cited AhR models were derived from apo template structures. However, the quality of the modeled binding site, that is critical for reliable predictions of the binding poses [75–78], can be greatly improved by the use of holo template structures, that take into account protein conformational rearrangements induced by the ligand (“induced fit” effects) [76]. The first X-ray structures of the HIF-2 α PASB in complex with artificial ligands [28,29] prompted us to develop new homology models of the AhR LBD, specifically aimed to improve the accuracy of the binding site description [58]. Also incorporating protein flexibility associated to ligand binding in Molecular Docking calculations is crucial for

accurate predictions and it represents one of the major challenges for this method [43,79,80]. To this aim, we proposed the use of the “ensemble-docking” technique (in which ligand is docked to multiple receptor conformations) followed by refinement of the docked poses by energy minimization [58]. The agreement of the TCDD:mAhR pose (Figure 1) with all the mutagenesis data available, as well as the satisfactory correlation between experimental and calculated binding affinities of a set of PCDDs confirmed the reliability of this new approach [58].

Other computational strategies were proposed to both accurately model the AhR binding cavity and include local flexibility upon binding of different ligands. Perkins *et al.* performed docking to a hAhR model based on the HIF-2 α apo structure [26], but refined with the docked TCDD inside the cavity [67]. After validation on a test set of known agonists and antagonists, this approach was used for virtual screening of a large database of compounds and two novel agonists were identified [67]. Moreover, Molecular Dynamics (MD) simulations performed on the docked poses of both the TCDD and GNF315 ligands suggested a hypothesis about differential roles of agonists and antagonists in the rearrangement of the LBD that could be related to different effects on the activation/inhibition of the AhR [67]. Instead, Nuti *et al.* adopted an ensemble-docking protocol in which the different protein conformations were obtained by mono- and multi-template homology models of the mAHR based on different apo and holo HIF-2 α structures [68]. Docking of TCDD, FICZ, and the endogenous ligand L-kynurenine, followed by a MD simulation step, led to predict that diverse ligands fit different AhR LBD conformations and show specific interaction patterns with the binding site residues [68]. The obtained L-kynurenine binding mode was experimentally confirmed and it contributed to an in-depth study on the role of AhR in the induction of endotoxin tolerance [69]. Recently, molecular docking calculations on a hAhR LBD model based on a holo HIF-2 α structure [29], contributed to demonstrate the ability of AhR to sense bacterial virulence factors and control antibacterial responses [70]. Also ligand-, isoform-, and species-specific interactions in binding of both dioxin-like and endogenous ligands to some avian AhRs were rationalized by docking to homology models of AhR [71,72].

All the above results strongly support the increasing contribution of *in silico* modeling in studying ligand binding to the AhR and understanding the different roles of this receptor. However, the current PASB structural description has some limitations since it is related to domain conformations near to the ligand-bound state. It is conceivable that a PASB conformation with a more “open” binding cavity is adopted by the AhR when bound to hsp90 in the cytosolic multi-protein complex and that large conformational changes are associated with ligand binding, AhR activation, and partial hsp90 displacement. Therefore, a complete molecular description of such mechanistic step would require information about the initial hsp90-bound AhR conformation as well as the use of computational methods able to reliably simulate large protein motions. Moreover, it should be taken into account that computational prediction of the binding ability of AhR ligands is not sufficient by itself to completely answer key questions related to their biological and toxicological effects, *e.g.* prediction of differential AhR species-sensitivity to chemicals in ecological risk assessments, identification of novel AhR agonists or antagonists through virtual screening, design of new drugs able to stimulate/inhibit a certain biological effect without inducing

toxicity. In fact, the AhR diversity in responses to ligands is not only related to their initial binding but also to their role in the following steps of the AhR signaling pathway. To reach a better predictive ability, *in silico* modeling should move toward a molecular description of the entire process by which ligands differentially affect the AhR activation, stimulate AhR transformation and DNA binding, and activate AhR-dependent gene expression.

4. Modeling AhR dimerization and transformation

Understanding the AhR:ARNT dimerization mode would provide insights into the mechanisms of ligand-dependent transformation of the AhR into its transcriptionally active form. Dimerization occurs primarily through the HLH and PASA domains, and while these domains appear sufficient to allow transformation of the AhR:ARNT complex into its high affinity DNA binding form, the PASB has proven to be critically involved in upstream modulation of the dimerization process [12,81]. The high degree of structural conservation in the bHLH-PAS family of proteins [35,36] prompted us to develop models of the individual AhR:ARNT PAS domain dimers by homology modeling [82]. The models were based on the X-ray structures of the PASB and PASA dimers available at the time: the HIF-2 α :ARNT PASB and the CLOCK:BMAL1 bHLH-PASA-PASB heterodimers; the AhR PASA homodimer [16,28,32]. Due to the differences in the reciprocal orientation of the domains in the templates, particularly in the PASB dimers, alternative models were developed and the most reliable dimerization mode was identified through evaluations of the physico-chemical characteristics of the protein-protein interaction (PPI) interfaces [82]. On this basis, a set of mutagenesis experiments was designed in both the AhR and ARNT proteins and ligand-dependent DNA binding of the AhR:ARNT mutants was evaluated [82]. These studies suggested dimerization modes for the AhR and ARNT PAS domains similar to those observed in CLOCK:BMAL1 [32], where in the PASB interface the helical bundle of AhR interacts with the β -sheet of ARNT (Figure 1). The reliability of our prediction was confirmed by its agreement with previous mutagenesis studies on AhR or homologous systems [16,32,83]. Moreover, it was further supported by the comparison with the first dimer structures of HIF-2 α :ARNT including the entire bHLH-PASA-PASB region, recently published [33]. In fact, despite the different quaternary architectures of the N-terminal region, the CLOCK:BMAL1 and HIF-2 α :ARNT structures show very similar PASA:PASA and PASB:PASB interfaces [33].

These newly available structures of the entire N-terminal portion of two dimers in the bHLH-PAS family [32,33] have provided the opportunity to obtain a structural model of the AhR:ARNT complex encompassing the same region and, on this basis, to characterize the critical PPIs in its intricate network of inter-domain interfaces [84]. These studies will provide important new insights to start studying the mechanism of transformation of the AhR into its high-affinity DNA binding form.

5. Future perspectives

Some of the recent depositions of dimer structures belonging to bHLH-PAS family also contain a co-crystallized fragment of the DNA recognition site: CLOCK:BMAL1 with E-box DNA [37], and HIF-2 α :ARNT with the hypoxia response elements (HREs) [33] (Figure

1). This information opens new directions for modeling the AhR:ARNT dimer in complex with its specific DNA site, the Xenobiotic Responsive Element (XRE), and for characterizing the key protein-DNA interactions.

Despite structural models of AhR complexes (protein-ligand, protein-protein, protein-DNA) may provide a fundamental contribution to the comprehension of the main stages of the AhR mechanisms, they offer a static description. As pointed out before, it is now well accepted that protein dynamics plays a key role in protein interactions, signaling mechanisms and allosteric regulation [46–48]. Therefore, such models only represent the starting point for analyzing the processes by which ligands can affect the AhR activation, dimerization and transformation into its functional form, and DNA binding. Future research directions should include theoretical predictions of the above mechanisms based on MD simulations, validated by mutagenesis and functional analyses. These studies could give a more accurate description of signal propagation across the ligand:AhR:ARNT:DNA complex, highlighting the intermolecular interactions which determine ligand-specific pathways.

The complete structural and dynamical molecular description of the AhR mechanism will greatly improve the ability of *in silico* modeling to predict ligand-dependent differences in AhR functionality, with high impact on both human/ecological risk assessment and drug design efforts.

Acknowledgments

This work was supported by the National Institute of Environmental Health Sciences US [grant number R01ES07685].

References

- Schmidt JV, Bradfield CA. Ah receptor signaling pathways. *Annu Rev Cell Dev Biol.* 1996; 12:55–89. [PubMed: 8970722]
- Ma Q. Induction of CYP1A1. The AhR/DRE paradigm: transcription, receptor regulation, and expanding biological roles. *Curr Drug Metab.* 2001; 2:149–164. [PubMed: 11469723]
- Denison MS, Soshilov AA, He G, DeGroot DE, Zhao B. Exactly the same but different: Promiscuity and diversity in the molecular mechanisms of action of the aryl hydrocarbon (dioxin) receptor. *Toxicol Sci.* 2011; 124:1–22. This is a comprehensive review on the classical and non-classical mechanisms of the AhR that can contribute to the different sensitivity and diversity in responses observed upon ligand-dependent activation of the receptor. [PubMed: 21908767]
- Poland A, Knutson JC. 2,3,7,8-tetrachlorodibenzo-p-dioxin and related halogenated aromatic hydrocarbons: examination of the mechanism of toxicity. *Annu Rev Pharmacol Toxicol.* 1982; 22:517–554. [PubMed: 6282188]
- Safe S. Polychlorinated biphenyls (PCBs), dibenzo-p-dioxins (PCDDs), dibenzofurans (PCDFs), and related compounds: environmental and mechanistic considerations which support the development of toxic equivalency factors (TEFs). *Crit Rev Toxicol.* 1990; 21:51–88. [PubMed: 2124811]
- Denison, MS., Seidel, SD., Rogers, WJ., Ziccardi, M., Winter, GM., Heath-Pagliuso, S. Natural and synthetic ligands for the Ah receptor. In: Puga, A., Wallace, KB., editors. *Molecular biology approaches to toxicology*. Taylor & Francis; Philadelphia PA: 1998. p. 393-410.
- Denison MS, Nagy SR. Activation of the aryl hydrocarbon receptor by structurally diverse exogenous and endogenous chemicals. *Annu Rev Pharmacol Toxicol.* 2003; 43:309–34. [PubMed: 12540743]

8. Murray IA, Patterson AD, Perdew GH. Aryl hydrocarbon receptor ligands in cancer: friend and foe. *Nat Rev Cancer*. 2014; 14:801–814. [PubMed: 25568920]
9. Esser C, Rannug A. The aryl hydrocarbon receptor in barrier organ physiology, immunology, and toxicology. *Pharmacol Rev*. 2015; 67:259–279. [PubMed: 25657351]
10. Gasiewicz, TA., Henry, EC. History of research on the AHR. In: Pohjanvirta, R., editor. *The Ah Receptor in Biology and Toxicology*. Wiley; Hoboken NJ: 2011. p. 3-32.
11. Coumailleau P, Poellinger L, Gustafsson JA, Whitelaw ML. Definition of a minimal domain of the dioxin receptor that is associated with Hsp90 and maintains wild type ligand binding affinity and specificity. *J Biol Chem*. 1995; 270:25291–25300. [PubMed: 7559670]
12. Fukunaga BN, Probst MR, Reisz-Porszasz S, Hankinson O. Identification of functional domains of the aryl hydrocarbon receptor. *J Biol Chem*. 1995; 270:29270–29278. [PubMed: 7493958]
13. McGuire J, Okamoto K, Whitelaw ML, Tanaka H, Poellinger L. Definition of a dioxin receptor mutant that is a constitutive activator of transcription: Delineation of overlapping repression and ligand binding functions within the PAS domain. *J Biol Chem*. 2001; 276:41841–41849. [PubMed: 11551926]
14. Carver LA, Jackiw V, Bradfield CA. The 90-kDa heat shock protein is essential for Ah receptor signaling in a yeast expression system. *J Biol Chem*. 1994; 269:30109–30112. [PubMed: 7982913]
15. Pongratz I, Mason GG, Poellinger L. Dual roles of the 90-kDa heat shock protein hsp90 in modulating functional activities of the dioxin receptor. Evidence that the dioxin receptor functionally belongs to a subclass of nuclear receptors which require hsp90 both for ligand binding activity and repression of intrinsic DNA binding activity. *J Biol Chem*. 1992; 267:13728–13734. [PubMed: 1320028]
16. Wu D, Potluri N, Kim Y, Rastinejad F. Structure and dimerization properties of the aryl hydrocarbon receptor PAS-A domain. *Mol Cell Biol*. 2013; 33:4346–56. The first crystal structure of an AhR domain, the PASA homodimer (PDB ID: 4M4X), is presented and the residues important for dimerization are highlighted by mutational analysis. [PubMed: 24001774]
17. Borgstahl GE, Williams DR, Getzoff ED. 1.4 Å structure of photoactive yellow protein, a cytosolic photoreceptor: unusual fold, active site, and chromophore. *Biochemistry*. 1995; 34:6278–6287. [PubMed: 7756254]
18. Dux P, Rubinstenn G, Vuister GW, Boelens R, Mulder FAA, Hård K, Hoff WD, Kroon AR, Crielaard W, Hellingwerf KJ, Kaptein R. Solution structure and backbone dynamics of the photoactive yellow protein. *Biochemistry*. 1998; 37:12689–12699. [PubMed: 9737845]
19. Morais Cabral JH, Lee A, Cohen SL, Chait BT, Li M, Mackinnon R. Crystal structure and functional analysis of the HERG potassium channel N terminus: a eukaryotic PAS domain. *Cell*. 1998; 95:649–655. [PubMed: 9845367]
20. Gong W, Hao B, Mansy SS, Gonzalez G, Gilles-Gonzalez MA, Chan MK. Structure of a biological oxygen sensor: a new mechanism for heme-driven signal transduction. *Proc Natl Acad Sci U S A*. 1998; 95:15177–15182. [PubMed: 9860942]
21. Crosson S, Moffat K. Structure of a flavin-binding plant photoreceptor domain: Insights into light-mediated signal transduction. *Proc Natl Acad Sci*. 2001; 98:2995–3000. [PubMed: 11248020]
22. Amezcua CA, Harper SM, Rutter J, Gardner KH. Structure and interactions of PAS Kinase N-terminal PAS domain: Model for intramolecular kinase regulation. *Structure*. 2002; 10:1349–1361. [PubMed: 12377121]
23. Getzoff ED, Gutwin KN, Genick UK. Anticipatory active-site motions and chromophore distortion prime photoreceptor PYP for light activation. *Nat Struct Biol*. 2003; 10:663–668. [PubMed: 12872160]
24. Razeto A, Ramakrishnan V, Litterst CM, Giller K, Griesinger C, Carlomagno T, Lakomek N, Heimburg T, Lodrini M, Pfitzner E, Becker S. Structure of the NCoA-1/SRC-1 PAS-B domain bound to the LXXLL motif of the STAT6 transactivation domain. *J Mol Biol*. 2004; 336:319–329. [PubMed: 14757047]
25. Yildiz Ö, Doi M, Yujnovsky I, Cardone L, Berndt A, Hennig S, Schulze S, Urbanke C, Sassone-Corsi P, Wolf E. Crystal structure and interactions of the PAS repeat region of the *Drosophila* clock protein PERIOD. *Mol Cell*. 2005; 17:69–82. [PubMed: 15629718]

- 26••. Erbel PJ, Card PB, Karakuzu O, Bruick RK, Gardner KH. 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*. 2003; 100:15504–15509. The first experimental (solution NMR) structure of the HIF-2 α PASB domain (PDB ID: 1P97) is described and point mutations are used to investigate the putative dimerization interface with ARNT. [PubMed: 14668441]
27. Card PB, Erbel PJ, Gardner KH. Structural basis of ARNT PAS-B dimerization: Use of a common beta-sheet interface for hetero- and homodimerization. *J Mol Biol*. 2005; 353:664–677. [PubMed: 16181639]
- 28••. Scheuermann TH, Tomchick DR, Machius M, Guo Y, Bruick RK, Gardner KH. Artificial ligand binding within the HIF2 α PAS-B domain of the HIF2 transcription factor. *Proc Natl Acad Sci U S A*. 2009; 106:450–455. The paper reports the first crystal structure of the HIF-2 α PASB domain in complex with an artificial ligand (PDB ID: 3F1O), in addition to apo structures of the same domain, and gives insights into the PASB internal cavity characteristics. [PubMed: 19129502]
29. Key J, Scheuermann TH, Anderson PC, Daggett V, Gardner KH. Principles of ligand binding within a completely buried cavity in HIF2 α PAS-B. *J Am Chem Soc*. 2009; 131:17647–17654. [PubMed: 19950993]
30. Rogers JL, Bayeh L, Scheuermann TH, Longgood J, Key J, Naidoo J, Melito L, Shokri C, Frantz DE, Bruick RK, Gardner KH, MacMillan JB, Tambar UK. Development of inhibitors of the PAS-B domain of the HIF-2 α transcription factor. *J Med Chem*. 2013; 56:1739–1747. [PubMed: 23363003]
31. Scheuermann TH, Li Q, Ma HW, Key J, Zhang L, Chen R, Garcia JA, Naidoo J, Longgood J, Frantz DE, Tambar UK, Gardner KH, Bruick RK. Allosteric inhibition of hypoxia inducible factor-2 with small molecules. *Nat Chem Biol*. 2013; 9:271–276. [PubMed: 23434853]
- 32••. Huang N, Chelliah Y, Shan Y, Taylor CA, Yoo SH, Partch C, Green CB, Zhang H, Takahashi JS. Crystal structure of the heterodimeric CLOCK:BMAL1 transcriptional activator complex. *Science*. 2012; 337:189–94. The crystal structure of the CLOCK:BMAL1 heterodimer is presented (PDB ID: 3F3L) and the dimerization interfaces are analyzed. This represents the first experimental structure of dimers of the bHLH-PAS family of proteins including the entire bHLH-PASA-PASB region. [PubMed: 22653727]
- 33••. Wu D, Potluri N, Lu J, Kim Y, Rastinejad F. Structural integration in hypoxia-inducible factors. *Nature*. 2015; 524:303–308. The first crystal structure of the HIF-2 α :ARNT heterodimer including the entire bHLH-PASA-PASB region is presented. In addition to the apo form (PDB ID: 4ZP4), ligand- and DNA-bound structures and a structure of the HIF-1 α :ARNT dimer are reported. These depositions, along with the CLOCK:BMAL1 one [32], give insights into the quaternary architecture of the bHLH-PAS multi-domain complexes. [PubMed: 26245371]
34. Scheuermann TH, Stroud D, Sleet CE, Bayeh L, Shokri C, Wang H, Caldwell CG, Longgood J, MacMillan JB, Bruick RK, Gardner KH, Tambar UK. Isoform-Selective and Stereoselective Inhibition of Hypoxia Inducible Factor-2. *J Med Chem*. 2015; 58:5930–5941. [PubMed: 26226049]
- 35••. Bersten DC, Sullivan AE, Peet DJ, Whitelaw ML. bHLH-PAS proteins in cancer. *Nat Rev Cancer*. 2013; 13:827–841. This is an up-to-date review on the bHLH-PAS proteins, dealing with their structures, mechanisms, and functions, and highlighting their role in several pro-tumor and anti-tumor pathways. [PubMed: 24263188]
36. Kewley RJ, Whitelaw ML, Chapman-Smith A. The mammalian basic helix-loop-helix/PAS family of transcriptional regulators. *Int J Biochem Cell Biol*. 2004; 36:189–204. [PubMed: 14643885]
- 37•. Wang Z, Wu Y, Li L, Su XD. Intermolecular recognition revealed by the complex structure of human CLOCK-BMAL1 basic helix-loop-helix domains with E-box DNA. *Cell Res*. 2013; 23:213–224. The crystal structure of CLOCK:BMAL1 bHLH domains bound to a canonical E-box DNA (PDB ID: 4H10) is presented. This deposition, along with the HIF-2 α :ARNT DNA-bound structure [33], sheds light on DNA binding of bHLH-PAS proteins. [PubMed: 23229515]
38. Cavasotto CN, Phatak SS. Homology modeling in drug discovery: current trends and applications. *Drug Discov Today*. 2009; 14:676–683. [PubMed: 19422931]
39. Schmidt T, Bergner A, Schwede T. Modelling three-dimensional protein structures for applications in drug design. *Drug Discov Today*. 2014; 19:890–897. [PubMed: 24216321]

40. Dundas J, Ouyang Z, Tseng J, Binkowski A, Turpaz Y, Liang J. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.* 2006; 34:W116–W118. [PubMed: 16844972]
41. An J, Totrov M, Abagyan R. Pocketome via comprehensive identification and classification of ligand binding envelopes. *Mol Cell Proteomics.* 2005; 4:752–761. [PubMed: 15757999]
42. Kitchen DB, Decornez H, Furr JR, Bajorath J. Docking and scoring in virtual screening for drug discovery: methods and applications. *Nat Rev Drug Discov.* 2004; 3:935–949. [PubMed: 15520816]
43. Cozzini P, Kellogg GE, Spyraakis F, Abraham DJ, Costantino G, Emerson A, Fanelli F, Gohlke H, Kuhn LA, Morris GM, Orozco M, Pertinhez TA, Rizzi M, Sottriffer CA. Target flexibility: an emerging consideration in drug discovery and design. *J Med Chem.* 2008; 51:6237–6255. [PubMed: 18785728]
44. Tuncbag N, Kar G, Keskin O, Gursoy A, Nussinov R. A survey of available tools and web servers for analysis of protein-protein interactions and interfaces. *Brief Bioinform.* 2009; 10:217–232. [PubMed: 19240123]
45. Szilagyi A, Zhang Y. Template-based structure modeling of protein-protein interactions. *Curr Opin Struct Biol.* 2014; 24:10–23. [PubMed: 24721449]
46. Henzler-Wildman K, Kern D. Dynamic personalities of proteins. *Nature.* 2007; 450:964–972. [PubMed: 18075575]
47. Karplus M, Kuriyan J. Molecular dynamics and protein function. *Proc Natl Acad Sci.* 2005; 102:6679–6685. [PubMed: 15870208]
48. Dror RO, Dirks RM, Grossman JP, Xu H, Shaw DE. Biomolecular Simulation: A computational microscope for molecular biology. *Annu Rev Biophys.* 2012; 41:429–452. [PubMed: 22577825]
49. Procopio M, Lahm A, Tramontano A, Bonati L, Pitea D. A model for recognition of polychlorinated dibenzo-p-dioxins by the aryl hydrocarbon receptor. *Eur J Biochem.* 2002; 269:13–18. [PubMed: 11784293]
50. Denison MS, Pandini A, Nagy SR, Baldwin EP, Bonati L. Ligand binding and activation of the Ah receptor. *Chem-Biol Interact.* 2002; 141:3–24. [PubMed: 12213382]
51. Murray IA, Reen RK, Leathery N, Ramadoss P, Bonati L, Gonzalez FJ, Peters JM, Perdew GH. Evidence that ligand binding is a key determinant of Ah receptor-mediated transcriptional activity. *Arch Biochem Biophys.* 2005; 442:59–71. [PubMed: 16137638]
- 52•. Pandini A, Denison MS, Song Y, Soshilov AA, Bonati L. Structural and functional characterization of the aryl hydrocarbon receptor ligand binding domain by homology modeling and mutational analysis. *Biochemistry.* 2007; 46:696–708. The paper presents the first homology model of the AhR LBD based on bHLH-PAS template structures (HIF-2 α and ARNT) and validated by site-directed mutagenesis and AhR functional analysis. [PubMed: 17223691]
- 53•. Pandini A, Soshilov AA, Song Y, Zhao J, Bonati L, Denison MS. Detection of the TCDD binding-fingerprint within the Ah receptor ligand binding domain by structurally driven mutagenesis and functional analysis. *Biochemistry.* 2009; 48:5972–5983. In this paper the group of residues responsible for TCDD binding in high-affinity mammalian AhRs is revealed on the basis of amino-acid conservation in the modeled ligand binding cavities and of mutagenesis studies. [PubMed: 19456125]
- 54•. Fraccalvieri D, Soshilov AA, Karchner SI, Franks DG, Pandini A, Bonati L, Hahn ME, Denison MS. Comparative analysis of homology models of the Ah receptor ligand binding domain: verification of structure–function predictions by site-directed mutagenesis of a nonfunctional receptor. *Biochemistry.* 2013; 52:714–725. In this paper comparison of several AhR homology models sheds light on the binding cavity characteristics in mammalian, avian and fish AhRs, and allows to predict the amino-acids responsible for the different TCDD sensitivity of zebrafish AhR isoforms. [PubMed: 23286227]
55. Odio C, Holzman SA, Denison MS, Fraccalvieri D, Bonati L, Franks DG, Hahn ME, Powell WH. Specific ligand binding domain residues confer low dioxin responsiveness to AHR1 β of *xenopus laevis*. *Biochemistry.* 2013; 52:1746–1754. [PubMed: 23394719]
56. Farmahin R, Manning GE, Crump D, Wu D, Mundy LJ, Jones SP, Hahn ME, Karchner SI, Giesy JP, Bursian SJ, Zwiernik MJ, Fredricks TB, Kennedy SW. Amino acid sequence of the ligand-

- binding domain of the aryl hydrocarbon receptor 1 predicts sensitivity of wild birds to effects of dioxin-like compounds. *Toxicol Sci.* 2013; 131:139–152. [PubMed: 22923492]
57. Shoots J, Fraccalvieri D, Franks DG, Denison MS, Hahn ME, Bonati L, Powell WH. An aryl hydrocarbon receptor from the salamander *ambystoma mexicanum* exhibits low sensitivity to 2,3,7,8-tetrachlorodibenzo-p-dioxin. *Environ Sci Technol.* 2015; 49:6993–7001. [PubMed: 25941739]
58. Motto I, Bordogna A, Soshilov AA, Denison MS, Bonati L. New aryl hydrocarbon receptor homology model targeted to improve docking reliability. *J Chem Inf Model.* 2011; 51:2868–2881. The paper shows the first homology models of the AhR LBD based on holo template structures, developed to improve the binding site description, and suggests an appropriate computational protocol for Molecular Docking applications. [PubMed: 21981577]
59. Bisson WH, Koch DC, O'Donnell EF, Khalil SM, Kerkvliet NI, Tanguay RL, Abagyan R, Kolluri SK. Modeling of the aryl hydrocarbon receptor (AhR) ligand binding domain and its utility in virtual ligand screening to predict new AhR ligands. *J Med Chem.* 2009; 52:5635–5641. In this paper homology models of AhR LBDs are used both for Molecular Docking applications and, for the first time, for virtual screening of a large database of compounds to identify new AhR agonists. [PubMed: 19719119]
60. O'Donnell EF, Saili KS, Koch DC, Kopparapu PR, Farrer D, Bisson WH, Mathew LK, Sengupta S, Kerkvliet NI, Tanguay RL, Kolluri SK. The anti-inflammatory drug leflunomide is an agonist of the aryl hydrocarbon receptor. *PLoS One.* 2010; 5:e13128. [PubMed: 20957046]
61. Murray IA, Flaveny CA, Chiaro CR, Sharma AK, Tanos RS, Schroeder JC, Amin SG, Bisson WH, Kolluri SK, Perdew GH. Suppression of cytokine-mediated complement factor gene expression through selective activation of the Ah receptor with 3',4'-dimethoxy- α -naphthoflavone. *Mol Pharmacol.* 2011; 79:508–519. [PubMed: 21127131]
62. Smith KJ, Murray IA, Tanos R, Tellev J, Boitano AE, Bisson WH, Kolluri SK, Cooke MP, Perdew GH. Identification of a high-affinity ligand that exhibits complete aryl hydrocarbon receptor antagonism. *J Pharmacol Exp Ther.* 2011; 338:318–327. [PubMed: 21493753]
63. Goodale BC, La Du JK, Bisson WH, Janszen DB, Waters KM, Tanguay RL. AHR2 mutant reveals functional diversity of aryl hydrocarbon receptors in zebrafish. *PLoS One.* 2012; 7:e29346. [PubMed: 22242167]
64. Gerlach CV, Das SR, Volz DC, Bisson WH, Kolluri SK, Tanguay RL. Mono-substituted isopropylated triaryl phosphate, a major component of Firemaster 550, is an AHR agonist that exhibits AHR-independent cardiotoxicity in zebrafish. *Aquat Toxicol.* 2014; 154:71–79. [PubMed: 24865613]
65. O'Donnell EF, Koch DC, Bisson WH, Jang HS, Kolluri SK. The aryl hydrocarbon receptor mediates raloxifene-induced apoptosis in estrogen receptor-negative hepatoma and breast cancer cells. *Cell Death Dis.* 2014; 5:e1038. [PubMed: 24481452]
66. Koch DC, Jang HS, O'Donnell EF, Punj S, Kopparapu PR, Bisson WH, Kerkvliet NI, Kolluri SK. Anti-androgen flutamide suppresses hepatocellular carcinoma cell proliferation via the aryl hydrocarbon receptor mediated induction of transforming growth factor- β 1. *Oncogene.* 2015:6092–6104. [PubMed: 25867062]
67. Perkins A, Phillips JL, Kerkvliet NI, Tanguay RL, Perdew GH, Kolluri SK, Bisson WH. A structural switch between agonist and antagonist bound conformations for a ligand-optimized model of the human aryl hydrocarbon receptor ligand binding domain. *Biology (Basel).* 2014; 3:645–669. In the paper, an optimized homology-model for the human AhR LBD is proposed to identify new AhR agonists on the basis of virtual screening of a large database of compounds. [PubMed: 25329374]
68. Nuti R, Gargaro M, Martino D, Dolciami D, Grohmann U, Puccetti P, Fallarino F, Macchiarulo A. Ligand Binding and functional selectivity of L-tryptophan metabolites at the mouse aryl hydrocarbon receptor (mAHR). *J Chem Inf Model.* 2014; 54:3373–3383. In this paper a combination of homology modeling, molecular docking and molecular dynamics is applied to the AhR LBD to detect specific receptor conformations and interactions selected by L-tryptophan metabolites. [PubMed: 25402742]
69. Bessede A, Gargaro M, Pallotta MT, Martino D, Servillo G, Brunacci C, Biciato S, Mazza EMC, Macchiarulo A, Vacca C, Iannitti R, Tissi L, Volpi C, Belladonna ML, Orabona C, Bianchi R,

Lanz TV, Platten M, Della Fazio MA, Piobbico D, Zelante T, Funakoshi H, Nakamura T, Gilot D, Denison MS, Guillemin GJ, DuHadaway JB, Prendergast GC, Metz R, Geffard M, Boon L, Pirro M, Iorio A, Veyret B, Romani L, Grohmann U, Fallarino F, Puccetti P. Aryl hydrocarbon receptor control of a disease tolerance defence pathway. *Nature*. 2014; 511:184–190. This paper reports an important study on the role of AhR in the induction of endotoxin tolerance; in addition to several experimental techniques employed, prediction of the L-kynurenine binding mode by molecular docking to the AhR homology model contributes to the study. [PubMed: 24930766]

70. Moura-Alves P, Fae K, Houthuys E, Dorhoi A, Kreuchwig A, Furkert J, Barison N, Diehl A, Munder A, Constant P, Skrahina T, Gühlich-Bornhof U, Klemm M, Koehler AB, Bandermann S, Goosmann C, Mollenkopf HJ, Hurwitz R, Brinkmann V, Fillatreau S, Daffe M, Tummeler B, Kolbe M, Oschkinat H, Krause G, Kaufmann SH. AhR sensing of bacterial pigments regulates antibacterial defence. *Nature*. 2014; 512:387–392. This paper presents an important research on the ability of AhR to regulate antibacterial defence. Molecular docking of ligands of bacterial origin to AhR homology models contributes to elucidate the AhR role. [PubMed: 25119038]
71. Hirano M, Hwang JH, Park HJ, Bak SM, Iwata H, Kim EY. In silico analysis of the interaction of avian aryl hydrocarbon receptors and dioxins to decipher isoform-, ligand-, and species-specific activations. *Environ. Sci Technol*. 2015; 49:3795–3804.
72. Kim IS, Hwang JH, Hirano M, Iwata H, Kim EY. In vitro and in silico evaluation of transactivation potencies of avian AHR1 and AHR2 by endogenous ligands: Implications for the physiological role of avian AHR2. *Comp. Biochem Physiol C Toxicol Pharmacol*. 2016; 187:1–9.
73. Gutierrez MA, Davis SS, Rosko A, Nguyen SM, Mitchell KP, Mateen S, Neves J, Garcia TY, Mooney S, Perdew GH, Hubbard TD, Lamba DA, Ramanathan A. A novel AhR ligand, 2AI, protects the retina from environmental stress. *Sci Rep*. 2016; 6:29025. [PubMed: 27364765]
74. Hubbard TD, Murray IA, Bisson WH, Lahoti TS, Gowda K, Amin SG, Patterson AD, Perdew GH. Adaptation of the human aryl hydrocarbon receptor to sense microbiota-derived indoles. *Sci Rep*. 2015; 5:12689. [PubMed: 26235394]
75. Ferrara P, Jacoby E. Evaluation of the utility of homology models in high throughput docking. *J Mol Model*. 2007; 13:897–905. [PubMed: 17487515]
76. McGovern SL, Shoichet BK. Information decay in molecular docking screens against holo, apo, and modeled conformations of enzymes. *J Med Chem*. 2003; 46:2895–2907. [PubMed: 12825931]
77. Fan H, Irwin JJ, Webb BM, Klebe G, Shoichet BK, Sali A. Molecular docking screens using comparative models of proteins. *J Chem Inf Model*. 2009; 49:2512–2527. [PubMed: 19845314]
78. Bordogna A, Pandini A, Bonati L. Predicting the accuracy of protein-ligand docking on homology models. *J Comput Chem*. 2011; 32:81–98. [PubMed: 20607693]
79. Lill MA. Efficient incorporation of protein flexibility and dynamics into molecular docking simulations. *Biochemistry*. 2011; 50:6157–6169. [PubMed: 21678954]
80. Feixas F, Lindert S, Sinko W, McCammon JA. Exploring the role of receptor flexibility in structure-based drug discovery. *Biophys Chem*. 2014; 186:31–45. [PubMed: 24332165]
81. Reisz-Porszasz S, Probst MR, Fukunaga BN, Hankinson O. Identification of functional domains of the aryl hydrocarbon receptor nuclear translocator protein (ARNT). *Mol Cell Biol*. 1994; 14:6075–6086. [PubMed: 8065341]
82. Corrada D, Soshilov AA, Denison MS, Bonati L. 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*. 2016; 12:e1004981. This paper presents the first homology models of the AhR:ARNT PASA and PASB dimers; the most reliable dimerization modes, among the ones shown by different templates, are identified through computational analyses of the dimerization interfaces and mutagenesis experiments. [PubMed: 27295348]
83. Hao N, Whitelaw ML, Shearwin KE, Dodd IB, Chapman-Smith A. Identification of residues in the N-terminal PAS domains important for dimerization of Arnt and AhR. *Nucleic Acids Res*. 2011; 39:3695–3709. [PubMed: 21245039]
84. Corrada, D., Denison, MS., Bonati, L. Structural modeling of the AhR: ARNT complex in the bHLH-PASA-PASB region elucidates the key determinants of dimerization. Submitted for publication

Highlights

The contribution of molecular modeling in elucidating the AhR mechanism is reviewed

The structure of the AhR ligand binding domain can be predicted by Homology Modeling

The ligand binding process can be described by Molecular Docking

The AhR: ARNT dimer structure and interactions can be modeled

Future perspectives: modeling DNA binding; considering the role of protein dynamics

Figure 1.

Contributions of molecular modeling in elucidating the key mechanistic stages in the AhR induction of gene expression. Central panel: schematic representation of the three main stages in the AhR mechanism: ligand binding to the cytosolic AhR; dimerization with ARNT in the cell nucleus; binding to the DNA recognition site (XRE). In the upper left: homology model of the AhR LBD with the TCDD pose obtained by Molecular Docking [58]. At the right side: homology models of the PASA and PASB AhR:ARNT dimers (AhR in green, with the TCDD in the PASB; ARNT in orange) [82] superimposed to the X-ray structure of the bHLH-PASA-PASB HIF-2α:ARNT dimer [33] (HIF-2α in pale green, ARNT in pale orange). In the lower left: detail of the X-ray structure of the HIF-2α:ARNT dimer [33] bound to the DNA recognition site (HIF-2α in pale green, ARNT in pale orange, the DNA HRE in grey).