Academic Sciences International Journal of Pharmacy and Pharmaceutical Sciences

ISSN- 0975-1491 Vol 5, Issue 2, 2013

Research Article

HOMOLOGY MODELING OF ARYL HYDROCARBON RECEPTOR AND DOCKING OF AGONISTS AND ANTAGONISTS

MANOJ KUMAR GADHWAL[1], SWATI PATIL[1], PRISCILLA D'MELLO[2] ANDURMILA JOSHI*[1]

[1]Department of Pharmaceutical Chemistry and[2] Department of Pharmacognosy, Prin. K. M. Kundnani College of Pharmacy, Mumbai 400005. Email: urmilajjoshi@hotmail.com

Received: 02 Oct 2012, Revised and Accepted: 05 Feb 2013

ABSTRACT

The aryl hydrocarbon receptor (Ah receptor or AhR) is a nuclear receptor, located in the cytoplasm. Because of its association with carcinogenicity, the AhR has become a critical receptor for identifying the unknown endocrine disruptors. The experimental 3D structure of the receptor is not available. This calls for the creation of the homology model of the Ligand-binding Domain (LBD) for the purpose of application of the structure-based methods. High affinity heterodimer of HIF2 alpha and ARNT C-terminal PAS domain was used as a template for the development of Homology model. The generated model was evaluated stereochemically and validated further by docking a set of known agonists and antagonists into the modeled LBD of mouse AhR. An attempt has been made to explain the observed experimental binding affinities and the site-directed mutagenesis data as reported in the literature with the results of docking. The results of docking of antagonists indicate that these form distinct H-bonds with the receptor as against the agonists where hydrophobic interactions have predominated. The model can be used for the screening of ligands for AhR binding activity.

Keywords: Aryl Hydrocarbon Receptor (Ah receptor or AhR), Homology Modeling, Docking.

INTRODUCTION

The aryl hydrocarbon receptor (Ah receptor or AhR) is a ligand-activated transcription factor belonging to the basic-helix-loop-helix (bHLH) PAS family[1-5]. It is a nuclear receptor, located in the cytoplasm and exists as one component of the complex[6]; the other components being two molecules of heat shock protein (hsp90), an X-associated protein and a co-chaperone protein[7]. When agonists bind to the receptor, hsp90 dissociates from the complex; the complex translocates to the nucleus and dimerises with AhR nuclear translocator protein (ARNT)[8,9]. The AhR-ARNT heterodimer acts as a transcriptional activator by binding to specific DNA sequences[10], mediating the upregulation of the target genes. The AhR is constitutively expressed in a large number of mammalian tissues, with the highest amounts of mRNA found in liver, kidney, lung, heart, thymus, and placenta. Genes regulated by the AhR include those encoding cytochromes P450 CYPlAl, CYPlA2, and CYPlB1, as well as Phase II enzymes, such as UDP-glucuronosyl transferase UGTlA6, and other growth factors and proteins[11].

AhR is composed of multiple functional domains[8,9]. In the N-terminal end, the AhR contains a basic-helix-loop-helix (bHLH) region that is involved in DNA binding, dimerization with its nuclear partner Arnt and association with heat shock protein 90 (hsp90). The N-terminal of this region also contains nuclear localization (NLS) and export (NES) domains. The bHLH domain is required for heterodimerization, reorganisation and binding of these factors to dioxin response element upstream of the target genes. C-terminal to the bHLH domain is PAS domain, which composes of two imperfect repeats of 50 amino acids, PAS-A and PAS-B. PAS-B domain is reported to be involved in ligand binding[12-14]. Therefore a detailed understanding of the functions of AhR requires structural information about PAS-B domain. In absence of availability of an experimentally determined structure of AhR PAS-B domain, a 3D model needs to be developed by Homology Modeling techniques.

Several attempts of development of homology models are reported in the literature. The initial models[15-17] were based on the reported NMR structure of the C-Terminal PAS domain of Human HIF-2α (PDB code 1P97). PYP has also been used as a template for the homology modeling[18]. The structural similarity between AhR and the template has been close to 25% in all these cases, probably leading to loss of accuracy of docking and subsequent virtual screening. Attempts have also been made to model nuclear receptors[19] including the AhR using the crystal structure of hERα. All these models have docked TCDD to prove the utility of the models. These models have defined the ligand binding cavity in terms of the agonists and stressed the importance of the planar geometry of the agonists for the binding to the AhR. No amino acid interactions have been discussed except for the possible involvement of Phe in the π-π stacking interactions.

With the availability of the crystal structure of ARNT, the latest model published[20] used a combination of structures of both HIF-2α and ARNT. However this combination also could not improve the identity beyond 30%. The ligand - amino acid interactions have been discussed in a greater detail for the agonists in the ligand-binding domain of the AhR. None of the models have discussed the antagonist binding as studied by docking the antagonists in the active site of the receptor. In view of these finding, we decided to build a homology model of mouse AhR and dock agonists as well as the antagonists into the model.

MATERIALS AND METHODS

Homology modeling is a method of constructing and predicting an atomic resolution model of the target protein from its amino acid sequence based on an experimentally determined 3D structure of a related homologous protein called the template protein.

The four basic steps in homology modeling are: (1) identifying the template structure sequence, (2) aligning the query sequence with the template structure sequence, (3) building the model structure of the query based on the information from the template structure and (4) evaluating the predicted model. Homology modeling is therefore a useful methodology in predicting undetermined protein structures like the Ah receptor[21]. Homology modeling technique was used for predicting the LBD structures of mouse aryl hydrocarbon receptor. All computational and molecular modeling of mouse aryl hydrocarbon receptor were carried out on Maestro (version 8.5**,** of **Schrödinger,** LLC, 2008 software). The amino acids sequence for mouse (P30561) AhR consisting of 848 amino acids residues, was obtained from swissprot database. Template identification was performed using PSI-BLAST to search the nonredundant PDB database[22]. The X-ray crystalline structure of the high affinity heterodimer of HIF2 alpha and ARNT C-terminal PAS domains with the artificial ligand THS017 (PDB Id 3H7W)[23] showed detectable degree of similarity with the query sequence and was therefore used for the model generation. The coordinates were obtained from Protein Data Bank. The first step consisted of aligning the template and the target. The alignment**s** were sorted by score, expectation value, identities, positives, and gaps and evaluated statistically so as to select the best alignment as shown in **Fig. 1**. The homology models were generated using the Prime Module. The co-ordinates of

the conserved residues were used as the basis for modeling the LBD of AhR. The side chain coordinates for all non-identical residues were predicted using PRIME. Loop refinement of LBD of mouse AhR (8 loops) was carried out and multiple loop conformations were constructed using Prime functionality. Scoring of these conformations was done by side-chain predictions and all-atom minimizations. After completion of model building calculations, the model was further optimized and minimized. The insertions, deletions, and template transitions were built. These are cases in which the backbone itself needs to be reconstructed, either due to gaps in the alignment or template transitions that produce gaps in the 3D structure. These gaps were closed by reconstruction of the affected region ab initio, using a backbone dihedral library. Gap reconstruction was done by finding a single loop conformation that closes the structure and is physically reasonable. Side-chain were predicted and then optimized. The non-template regions were minimized. The generated models were refined further to remove the steric clashes. Bond length, bond angle, side chain dihedrals and chiralities were adjusted and the final model was energy minimized with a truncated-Newton energy minimization using OPLS_2000 all-atom force field (protein-optimized). Ramchandran plot of the final model as shown in **Fig. 2** indicated that 98% of the residues were in the core and allowed region. The ribbon representation of the modeled LBD of mouse AhR is shown in **Fig. 3**.

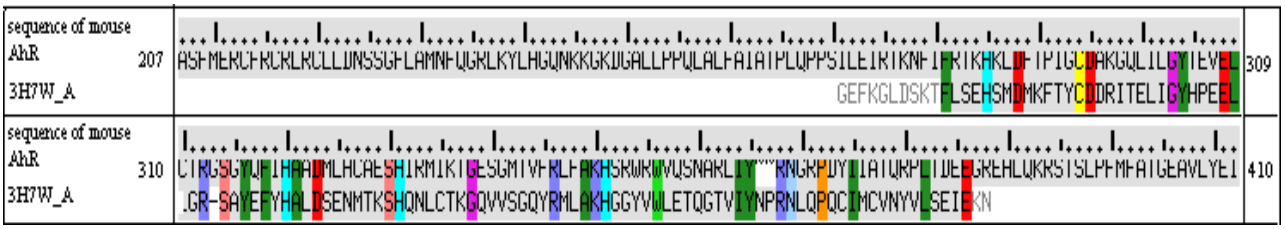

Fig. 1: Alignment of mouse AhR and 3H7W (X-ray crystalline structure of the high affinity heterodimer of HIF2 alpha and ARNT C-terminal PAS domains).

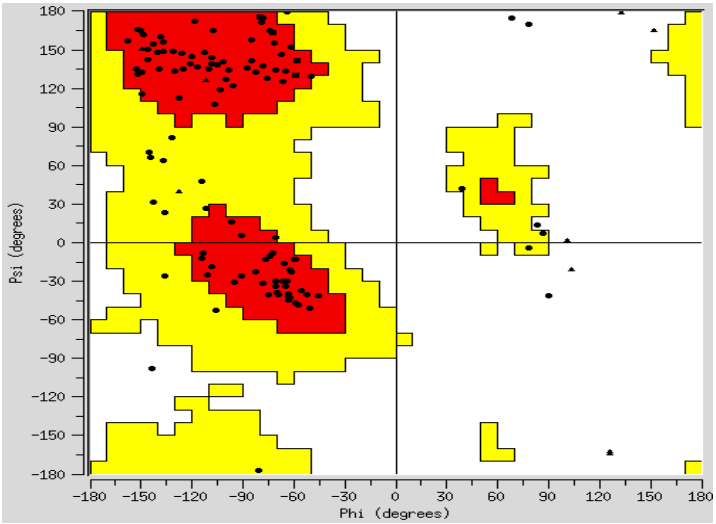

Fig. 2: Ramachandran Plot for the modeled LBD of mouse AhR after refinement. The plot is organized as follows: Glycine, proline and all other residues are plotted as triangles, squares, and circles respectively. The red, yellow and white regions represent the favoured, allowed and the disallowed regions respectively.

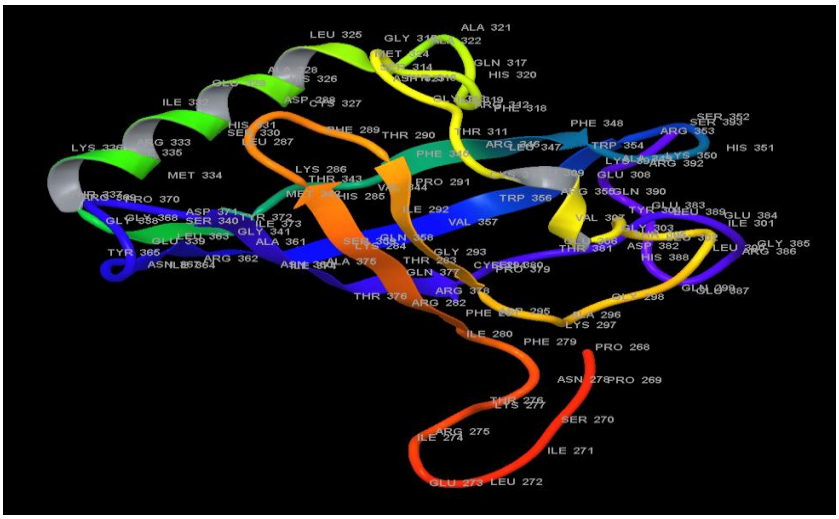

Fig. 3: A ribbon representation of modeled LBD of mouse AhR

To validate the model, some known agonist and antagonist were docked into the modeled ligand-binding domain of the mouse aryl hydrocarbon receptor. **Fig. 4** and **Fig. 5** shows the agonists of AhR: TCDD and E-80 which were docked into the modeled LBD of AhR, and **Fig. 6** and **Fig. 7** show the antagonists of AhR: flavones and ellipticines respectively; which were docked into the modeled LBD of mouse AhR. The Docking studies were carried out by using GLIDE (Maestro, version 8.5**, Schrödinger,** LLC, 2008) software. Protein preparation utilities in Maestro were used to assign the charge state of ionizable residues, add hydrogens, and carry out a highly constrained minimization of the generated model. The ligand set was prepared for docking using the LigPrep utility to define the charge state and enumerate the stereoisomers for each ligand. The ligands were geometry minimized using the OPLS_2005 force field and the Truncated Newton Conjugate Gradient to a gradient RMSD below 0.01 kJ/Å[24].

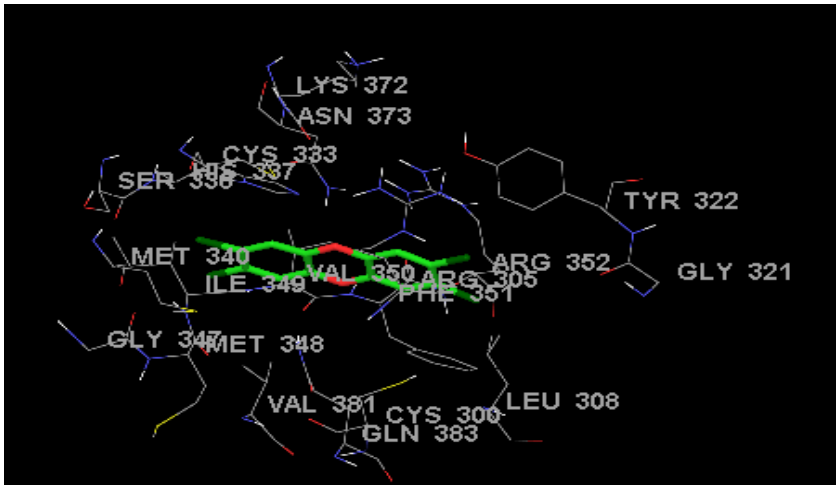

Fig. 4: Docked image of agonist TCDD into the homology model of mAhR.

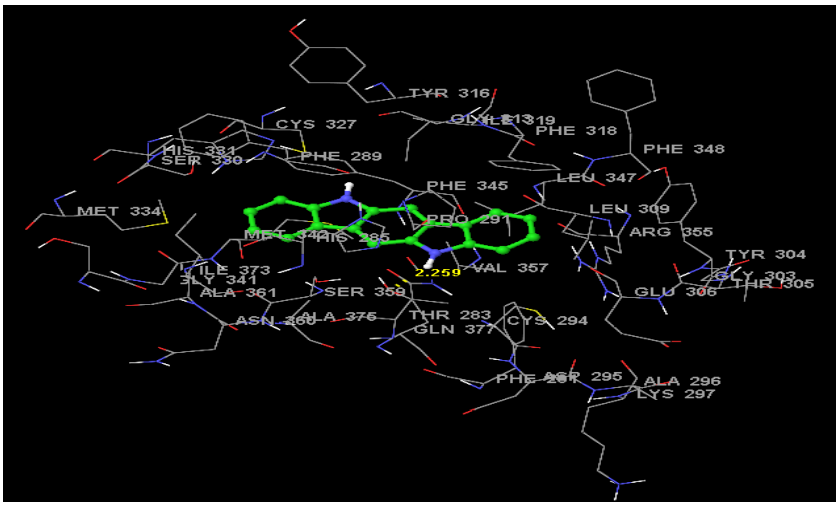

Fig. 5: Docked image of agonist E-80 into the homology model of mAhR. The dotted yellow lines represent the Hydrogen bond.

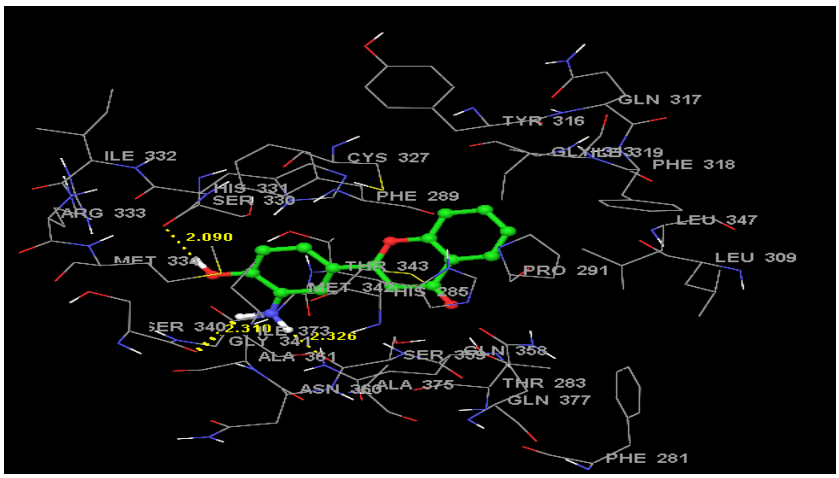

Fig. 6: Docked image of antagonist disubstituted flavones (A-8) into the homology model of mAhR. The dotted yellow lines represent the Hydrogen bonds.

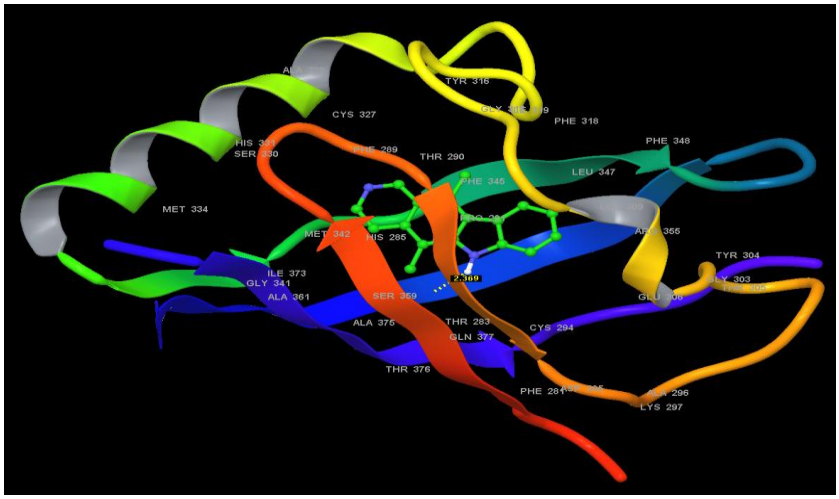

Fig. 7: Docked image of antagonist Ellipticine (19) into the homology model of mAhR. The dotted yellow lines represent the Hydrogen bonds.

The receptor Grid was generated using information reported in the literature[15] about the ligand binding cavity in the Homology model; as well as the site-directed mutagenesis data. The ligand set was then docked onto the mouse AhR LBD using the extra precision scoring mode of Glide[25]. During the docking procedure, ligand was flexible whereas the receptor was held rigid[26]. The best docked pose was saved.

RESULTS AND DISCUSSION

The target sequence of the ligand bind domain of the mouse (268-393) AhR was used as a query to search for homologues protein structure belonging to the category of nuclear receptors that could serve as templates. The x-ray crystalline structure of the high affinity heterodimer of HIF2 alpha and ARNT c-terminal pas domains (resolution: 1.65A^0) with the artificial ligand THS017 (PDB ID 3H7W) showed detectable degree of similarity with the query sequence. The other proteins with a detectable similarity are as shown in **Table 1**.

Table 1: Proteins with detectable similarity

Protein data bank entries	Similarities
3H7W	30%
3H82	24%
2A24	19%

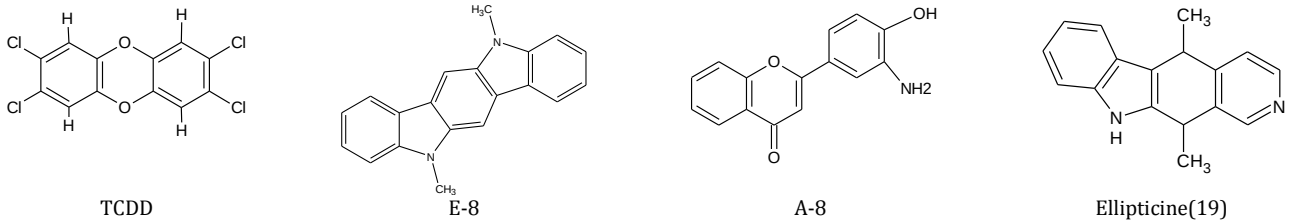

Fig. 8: Structures of Docked compounds: a) TCDD, b) E-80, c) A-8, d) Ellipticine(19) from Table 2, 3 and 4.

Table 2: List of agonists docked into modeled LBD of mouse AhR with pIC50 and G_Score

Serial no.	Compound designation	pIC50	G_Score
1	A-1	9.144	-6.43
2	A-2	8.118	-6.48
3	A-8	6.728	-5.95
4	A-14	4.572	-5.55
5	A-16	10.093	-6.18
6	A-17	10.687	-6.27
7	A-18	9.074	-6.34
8	A-22	9.350	-6.97
9	A-24	8.927	-5.77
10	B-28	3.429	-5.22
11	B-29	6.088	-5.71
12	B-38	7.657	-6.07
13	B-40	8.444	-6.66
14	B-43	5.371	-5.46
15	B-45	8.147	-6.26
16	B-48	7.587	-6.03
17	C-65	6.134	-6.15
18	C-66	5.762	-6.63
19	C-67	6.157	-6.01
20	C-70	5.885	-6.03
21	C-77	7.465	-5.94
22	E-80	8.921	-7.53

Table 3: A list of antagonists belonging to the class of flavones docked into modeled LBD of mouse AhR with pIC50 and G_Score

Serial No.	Compound designation	pIC_{50}	G_Score
1	ANF	-2.3541	-6.97
2	33	-2.9586	-6.1
3	35	-2.9731	-6.67
4	36	-4.0000	-6.56
5	38	-3.6718	-5.97
6	39	-1.0000	-4.56
7	40	-2.6618	-5.87
8	A-1	-2.334	-5.91
9	A-2	-1.996	-5.58
10	A-3	-2.559	-5.87
11	A-4	-0.146	-5.9
12	A-5	-0.176	-6
13	A-7	-3.287	-6.05
14	A-8	-2.868	-7.1
15	A-9	-0.358	-6.51
16	A-10	0.102	-4.92

Table 4: **A list of antagonists belonging to the class of ellipticines docked into modeled LBD of mouse AhR with pIC50 and G_Score**

Serial no.	Compound designation	pIC50	G_Score for mAhR
1	1	-2.4346	-6.38
2	2	-2.9335	-6.58
3	6	-3.0133	-6.84
4	7	-2.4548	-6.35
5	18	-2.1303	-6.9
6	19	-2.5533	-6.3
7	21	-3.0366	-6.72
8	22	-1.9685	-6.44
9	23	-1.6021	-7
10	27	-2.2014	-6.05
11	31	-1.3979	-6.08
12	32	-0.6902	-6.49

The template shows a high degree of structural conservation of typical PAS α and β folds. Phi-psi map, Ramachandran plot chi plot and Distance Matrix Plot of the model were generated as a part of the stereochemical evaluation of the model.

The image of TCDD docked into the homology model of mouse AhR is shown in **Fig. 4**. The result of docking of TCDD into mAhR showed that the residues lining the ligand binding cavity include Phe289, Met342, His285, Leu347, Tyr316, Ile319, Ala375 and Thr283. This is also in accordance with the literature report [11] thus validating our model further. TCDD when docked into the generated homology model did not show H-bonding interactions. The interactions which TCDD exhibited with the receptor were of hydrophobic involving Ile-319, Ala-375 and π-π stacking interactions involving Phe-289 and Phe-345. **Table 2, 3** and **4** show the sets of agonists[27], antagonists belonging to the class of flavones[28] and antagonists belonging to the class of ellipticines[29] respectively which were docked into the homology model of mouse AhR. The structure of docked compounds shown in **Fig. 8.**

Mutagenesis studies[15] indicate the importance of Thr-283, the mutation of which to either methionine or glutamic acid resulted in complete loss of TCDD and the DNA binding. Since Thr can act as Hydrogen bond Donor, a H-bonding interaction of TCDD with the receptor can also be expected. Although docking of TCDD did not show any H-bonding interaction with the receptor, docking of some other agonists such as E-80[27] showed H-bonding interaction with the receptor as shown in **Fig. 5**, This provided a further correlation between the site directed mutagenesis results and the homology modeling.

Then we decided to dock some of the antagonists. To the best of our knowledge, no report appears in the literature regarding the docking of antagonists into the AhR. The antagonists of AhR can be classified into two types, the compounds belonging to the flavonoid class and the compounds belonging to the nonflavonoid class[29]. Among the nonflavonoid class of AhR antagonists, since a detailed structure-Activity Relationship of ellipticines is reported, we focused our attention on ellipticines. The results of docking indicated that the antagonists, both flavonoids and nonflavonoids form H-bonds with the amino acid residues on the AhR which include His-285, Ser-340, Thr-343 and Thr-283. The flavonoids formed H-bonds with the receptor mainly via groups present on the 3' and 4' position of the B-ring. Highest activity was reported previously in the literature[29] for the compound with 3'-methoxy,4'-nitro substituent. The docking of the compounds into the homology model of AhR explains this on the basis of H-bonding. Groups capable of H-bond formation when present at 3' or 4'-position or both 3'and 4'-positions were shown to give a good dock score when docked into the generated homology model. **Fig. 6** shows the docked image of a 3',4'-disubstituted flavone into the homology model developed. However no special preference was detected for either 3' or 4' position contrary to the literature report[28]. The 4'-halo substituted flavones when docked into the homology model did not generate a favourable dock score, contrary to the literature reports[30] that these compounds showed good binding affinities.

The ellipticines on the other hand, formed H-bond via the central pyrrole ring as shown in **Fig. 7**. Compounds which contained N at 2 or 3 position of the ring did not show involvement of this N in the H-bond formation. The SAR of ellipticines acting as AhR antagonists as reported in the literature[29] also states that ellipticines tolerate only a small substitution, if at all, at positions 1 and 11. An observation of the ellipticine docked into ligand binding cavity of AhR indicates that due to the presence of Phe-289 in the viscinity of the A ring, a compound containing a larger substituent at position 1 will not be accommodated in the ligand binding cavity of the generated model. The H-bonding possibility between 1-aminoellipticines and Ser-340 also explains the SAR observation reported earlier that amino group at position 1 enhances the affinity of these compounds for the receptor.

CONCLUSION

AhR is a nuclear receptor which is activated by various carcinogens acting as agonists at the receptor. In absence of the availability of an

experimental 3D structure of the receptor, we have generated a homology model of the ligand-binding domain of the receptor. The usefulness of the homology models depends upon the ability of these models to explain the differences in binding of various ligands. Results of docking of the agonists have been compared with previous reports. Since much literature is not available on the docking of the antagonists into the homology model of AhR, we have tried to correlate the docking results with the binding affinities of various antagonists and tried to explain some of the SAR observations relating to ellipticines and flavones in terms of binding pattern of these ligands with the homology model of the Ah receptor.

ACKNOWLEDGEMENT

The authors sincerely thank All India Council for Technical Education, India, for sanction of grant (File No: 8023/BOR/RID/RPS-165/2007-08). Support by Schrodinger Inc. is gratefully acknowledged.

REFERENCES:

1. Schmidt JV, Bradfield CA. Ah receptor signaling pathways. Annu Rev Cell Dev Bio 1996;12(1):55-89.
2. Burbach KM, Poland A, Bradfield CA. Cloning of the Ah-receptor cDNA reveals a distinctive ligand-activated transcription factor. P Natl Acad Sci USA 1992;89(17):8185-9.
3. Kewley RJ, Whitelaw ML, Chapman-Smith A. The mammalian basic helix-loop-helix/PAS family of transcriptional regulators. Int J Biochem Cell B 2004;36(2):189-204.
4. Whitlock Jr JP. Mechanistic aspects of dioxin action. Chem Res Toxicol 1993;6(6):754-63.
5. Huang ZJ, Edery I, Rosbash M. PAS is a dimerization domain common to Drosophila period and several transcription factors. Nature. 1993;364:259-62.
6. Chen HS, Perdew GH. Subunit composition of the heteromeric cytosolic aryl hydrocarbon receptor complex. J Biol Chem 1994;269(44):27554-8.
7. Perdew GH. Association of the Ah receptor with the 90-kDa heat shock protein. J Biol Chem 1988;263(27):13802-5.
8. Reyes H, Reisz-Porszasz S, Hankinson O. Identification of the Ah receptor nuclear translocator protein (Arnt) as a component of the DNA binding form of the Ah receptor. Science. 1992;256(5060):1193-5.
9. Soshilov A, Denison MS. Role of the Per/Arnt/Sim domains in ligand-dependent transformation of the aryl hydrocarbon receptor. J Biol Chem 2008;283(47):32995-3005.
10. Elferink CJ, Gasiewicz TA, Whitlock JP. Protein-DNA interactions at a dioxin-responsive enhancer. Evidence that the transformed Ah receptor is heteromeric. J Biol Chem 1990;265(33):20708-12.
11. Safe SH. Modulation of gene expression and endocrine response pathways by 2, 3, 7, 8-tetrachlorodibenzo-p-dioxin and related compounds. Pharmacol Therapeut 1995;67(2):247-81.
12. McGuire J, Okamoto K, Whitelaw ML, Tanaka H, Poellinger L. Definition of a dioxin receptor mutant that is a constitutive activator of transcription. J Biol Chem 2001;276(45):41841-9.
13. Coumailleau P, Poellinger L, Gustafsson JÅ, 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(42):25291-300.
14. Perdew GH, Bradfield CA. Mapping the 90 kDa heat shock protein binding region of the Ah receptor. IUBMB Life 1996;39(3):589-93.
15. 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(3):696-708.
16. Erbel PJA, 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. P Natl Acad Sci USA 2003;100(26):15504-9.
17. Card PB, Erbel PJA, 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(3):664-77.
18. 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(1):13-8.
19. Jacobs MN, Dickins M, Lewis DFV. Homology Modeling of the Nuclear Receptor Ligand Binding Domains from the Human Estrogen Receptor α (hERα) Crystal Structure. J Steroid Biochem Mol Biol 2003;84:117-32.
20. 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(25):5972-83.
21. Taft CA, da Silva VB, da Silva CHT. Current topics in computer aided drug design. J Pharm Sci 2008;97(3):1089-98.
22. Altschul SF, Madden TL, Schäffer AA, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997;25(17):3389-402.
23. Key J, Scheuermann TH, Anderson PC, Daggett V, Gardner KH. Principles of ligand binding within a completely buried cavity in HIF2 PAS-B. J Am Chem Soc 2009;131(48):17647-54.
24. Jorgensen WL, Maxwell DS, Tirado-Rives J. Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. J Am Chem Soc 1996;118(45):11225-36.
25. D'Mello P, Gadhwal MK, Joshi U, Shetgiri P. Modeling of COX-2 Inhibotory Activity of Flavonoids. Int J Pharm Pharm Sci. 2011;3(4):33-40.
26. Subramaniana AK, Cardin CJ. Molecular modelling studies of binding of DACA derivatives into g-quadruplex DNA: comparison of force field and quantum polarized ligand docking methods. Int J Pharm Pharm Sci. 2012;4(1):509-14.
27. Piparo EL, Koehler K, Chana A, Benfenati E. Virtual screening for aryl hydrocarbon receptor binding prediction. J Med Chem 2006;49(19):5702-9.
28. Lu YF, Santostefano M, Cunningham BDM, Threadgill MD, Safe S. Substituted flavones as aryl hydrocarbon (Ah) receptor agonists and antagonists. Biochem Pharmacol 1996;51(8):1077-87.
29. Gasiewicz TA, Kende AS, Rucci G, Whitney B, Jeff Willey J. Analysis of structural requirements for ah receptor antagonist activity: Ellipticines, flavones, and related compounds. Biochem Pharmacol 1996;52(11):1787-803.
30. Lu YF, Santostefano M, Cunningham BDM, Threadgill MD, Safe S. Identification of 3 -methoxy-4 -nitroflavone as a pure aryl hydrocarbon (Ah) receptor antagonist and evidence for more than one form of the nuclear Ah receptor in MCF-7 human breast cancer cells. Arch Biochem Biophys 1995;316(1):470-7.