Toxicology Letters 191 (2009) 69–73

Contents lists available at ScienceDirect

Toxicology Letters

journal homepage: www.elsevier.com/locate/toxlet

In silico predication of nuclear hormone receptors for organic pollutants by homology modeling and molecular docking

Bing Wu, Yan Zhang, Jie Kong, Xuxiang Zhang, Shupei Cheng*

State Key Laboratory of Pollution Control & Resource Reuse, School of the Environment, Nanjing University, Nanjing, 210093, China

ARTICLE INFO

Article history:
Received 10 July 2009
Received in revised form 4 August 2009
Accepted 6 August 2009
Available online 14 August 2009

Keywords:
Nuclear hormone receptors
Homology modeling
Molecular docking
Organic pollutants

ABSTRACT

Homology modeling and molecular docking were used to *in silico* predict the rat nuclear hormone receptors of different organic pollutants. Rat aryl hydrocarbon receptor (rAhR), constitutive androstane receptor (rCAR) and pregnane X receptor (rPXR) were chosen as the target nuclear receptors. 3D models of ligand binding domains of rAhR, rCAR and rPXR were constructed by MODELLER 9V6 and assessed by the Procheck and Prosa 2003. Surflex-Dock program was applied to bind the different organic pollutants into the three receptors to predict their affinities. The results of docking experiments demonstrated that three polybrominated dibenzofurans (PBDFs, including TretaBDF, PentaBDF and HexaBDF) and 3,3′,4,4′,5′-pentachlorobiphenyl (PCB126) would be better categorized by rAhR-dependent mechanism, but four polybrominated diphenyl ethers (PBDEs, including BDE47, BDE80, BDE99 and BDE153) and 2,2′,4,4′,5,5′-hexachlorobiphenyl (PCB153) by rCAR and rPXR-dependent mechanism. For benzo(a)pyrene and pyrene, they have high affinities with the three target receptors, which suggests that "crosstalk" among the receptors might occur during the receptor induction. The results of this study are consistent with those of animal experiments reported by previous literatures, which suggest that homology modeling and molecular docking would have the potential to predict the nuclear hormone receptors of environmental pollutants.

1. Introduction

Aryl hydrocarbon receptor (AhR), constitutive androstane receptor (CAR) and pregnane X receptor (PXR) belong to the nuclear hormone receptors which could bind and be activated by a large number of endogenous and xenobiotic ligands (Baroukiabi et al., 2007; Janosek et al., 2006; Kakizaki et al., 2008). The ligand activated nuclear receptors AhR, CAR and PXR could bind to their cognate DNA elements and then activate the transcription of cytochrome P450 1A (CYP1A), CYP2B and CYP3A, respectively (Jacobs et al., 2003). The expression of cytochrome P450 enzymes often acts to detoxify poisonous xenobiotics. However, in some cases, the intermediates in xenobiotic metabolism could themselves be the cause of toxic effects. Thus, it is important to study the interactions of xenobiotics and different nuclear receptors in order to analyze the metabolic process and toxicity of xenobiotics by cytochrome P450.

Many xenobiotics, especially environmental contaminants, have been proved that they can reversibly bind into the special nuclear hormone receptors and activate the function of these receptors in the animal experiments. For example, AhR has high affinities towards the halogenated aromatic hydrocarbons (HAHs) (Hahn, 1998) and polycyclic aromatic hydrocarbons (PAHs) (Savouret et al., 2003). Sanders et al. (2005) found that polybrominated diphenyl ethers (PBDEs) mediated toxicity would be better categorized by CAR and PXR-dependent mechanisms. However, for the increasing number of environmental contaminants, animal experiments are time-consuming and lack hypothesis-driven aim. The approach of homology modeling and molecular docking might be a potential tool to provide the hypothesis-driven aim of animal experiments. Both techniques have been applied to study the interactions between ligands and nuclear receptors (such as AhR, CAR and PXR). However, these studies always focused on the structural and functional characterization of the special receptors (Pandini et al., 2007; Tirona et al., 2004; Windshugel et al., 2007). Few studies were performed to compare the binding affinities of one ligand to the different receptors so as to predict the special receptor of a certain target ligand.

In this study, homology modeling was used to construct the 3D model of ligand binding domains (LBDs) of rat AhR, CAR and PXR. Then, the predicted models were applied to bind the different organic pollutants by molecular docking. Free energy of binding was considered as the criteria to identify the binding affinities to analyze the specialization of interaction between receptor and ligand. The data from animal experiments was used to validate the results of docking experiments. This study is a useful attempt to

* Corresponding author. Tel.: +86 25 83595995; fax: +86 25 83595995.
E-mail address: chengsp@nju.edu.cn (S. Cheng).

0378-4274/$ – see front matter © 2009 Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.toxlet.2009.08.005

in silico predict the special nuclear hormone receptors for different organic pollutants and might provide a potential tool to search hypothesis-driven aim of animal experiments.

2. Computational methods

2.1. Sequence alignment and homology modeling

The primary sequences of rAhR, rCAR and rPXR were obtained from the Swiss-Prot database (http://www.expasy.ch/sprot/). The LBDs of three proteins were chosen as the target sequences (Table 1). BLAST algorithm against Protein Data Bank (PDB) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to carry out the sequence homology searches. The sequence and crystal structure of each template protein were extracted from Swiss-Prot and PDB databases (Table 1). Multiple sequence alignments among the target and template sequences were performed by ClustalW 2.0.10 program with default parameters (http://www.ebi.ac.uk/Tools/clustalw2/index.html).

MODELLER 9v6 program (Sali and Blundell, 1993) was used to construct initial 3D structural model of rAhR, rCAR and rPXR LBDs. MODELLER can implement comparative protein structure modeling by satisfying spatial restraints in terms of probability density functions. In this study, 50 runs of modeler were carried out using standard parameters and the outcomes were ranked on the basis of the internal scoring function of the program. The model with the highest score was chosen as the target model. Then, energy minimizations of chosen models were performed using GROMACS 3.3 according to the software protocol (Pandini et al., 2007; Van der Spoel et al., 2005).

2.2. Model evaluation

Model evaluation involved analysis of geometry, stereochemistry, and energy distribution of the predicted models. Firstly, 3D visualization programs Swiss-PdbViewer 4.01 (Guex and Peitsch, 1997) and Rasmol 2.7.4 (Goodsell, 2005) were carried out to peruse the reliability of the alignment and modeling of variable surface loops of predicated models (Lutfullah et al., 2008). Then, stereochemical quality of the homology models was checked by Procheck program (Laskowski et al., 1993). The energetic architecture of model folds was determined by Prosa 2003 program (Van Brussel et al., 1998).

2.3. Target pollutants

According to the previous literatures, 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (TCDD), 16,17-androstene-3-ol (ATE) and 16a-carbonitrile (PCN) have been proved to be the known ligands of rAhR, rCAR and rPXR, respectively (Hahn, 1998; Shan et al., 2004; Tirona et al., 2004). Thus, they were chosen as the reference ligands of docking experiment. A total of 11 environmental pollutants, whose data of animal experiment were available, were selected as target ligands to *in silico* predict their special nuclear hormone receptors, including 3,4-benzo(a)pyrene (BaP), pyrene, 3,3′,4,4′,5′-pentachlorobiphenyl (PCB126), 2,2′,4,4′,5,5′-hexachlorobiphenyl (PCB153), 2,3,7,8-tetrabromodibenzofuran (TetraBDF), 1,2,3,7,8-pentabromodibenzofuran (PentaBDF), 1,2,3,4,7,8-hexabromodibenzofuran (HexaBDF), 2,2′,4,4′-tetrabromodiphenyl ether (BDE47), 3,3,5,5-tetrabromodiphenyl ether (BDE80), 2,2′,4,4′,5′-pentabromodiphenyl ether (BDE99) and 2,2′,4,4′,5,5′-hexabromodiphenyl ether (BDE153).

Initial conformations of compounds were obtained from the Chemical Book Database (http://www.chemicalbook.com/ProductIndex.aspx). Compounds not included in the Database were constructed from the structures of similar compounds (Yang et al., 2009). The geometries of these compounds were subsequently optimized in Sybyl 7.3 (Tripos Inc., St. Louis, MO). Relevant energy minimization of target compounds was conducted using Tripos Force Field (distance-dependent dielectric) with atom charge calculated by Gasteiger–Hückel method to reach a final energy convergence gradient value of 0.001 kcal/mol. The optimized structures offered reasonable starting conformations for further molecular docking.

2.4. Flexible molecular docking

The Surflex-Dock program of Sybyl 7.3 was employed to dock the target pollutants into the rAhR, rCAR and rPXR LBDs, respectively. Surflex-Dock could automatically dock ligands into a receptor's ligand binding site using a protomol based approach and assess the affinity by an empirically derived scoring function. The method has been proved to be one of the most effective docking techniques (Kellenberger et al., 2004). In this study, prior to docking, the hydrogen atoms were added in predicted models using the Biopolymer modulators of Sybyl 7.3. The Kollman-all atom charges were assigned to protein atoms.

Protomol for Surflex-Dock was generated according to the software protocol. Two important factors, "proto_bloat" and "proto_thresh", can significantly affect the size and extent of the protomol. "Proto_thresh" determines how far the protomol extents into the concavity of the target site, while "proto_bloat" impacts how far the protomol extents outside of the concavity (Holt et al., 2008). Considering the purposes of this study, "proto_thresh" was set to 0.5 and "proto_bloat" was set to 1 for all protomols generated. Other parameters were employed with default setting in all runs. Protomols were visualized with Sybyl 7.3 to ensure proper coverage of the desired target area.

Surflex-Dock's scoring function, which contains hydrophobic, polar, repulsive, entropic, and salvation terms, was trained to estimate the dissociation constant (K_d) expressed in $-\log(K_d)$ unit (Jain, 2007). After running Surflex-Dock, the scores of docked conformers could be ranked in a molecular spread sheet. The best score conformer would be selected as the docking results. In this study, the scores of binding were converted to the free energy of binding (kcal/mol) in order to better compare the binding affinities between ligand and three target receptors, and to predict the preference receptor. The free energy of binding was calculated as following equation, where $RT = 0.59$ kcal/mol (Holt et al., 2008):

$$\text{free energy of binding} = RT \ln(10^{-\text{pKd}}).$$

3. Results and discussion

3.1. Construction of receptor model

Crystal structures, as potential templates of target proteins, were obtained from the BLAST search for rAhR, rCAR and rPXR, respectively. Template selection was performed on the basis of sequence similarity, residues completeness, crystal resolution and functional similarity. Table 1 shows the basic information on the selected templates used in this study.

For rAhR, among the available candidate templates, the sequence identities between target and templates were low (≤30%). Pandini et al. (2007) found that, despite low level of sequence similarity, mouse AhR could be well constructed using the crystal structures of HIf-2a and ARNT by homology modeling. HIf-2a and ARNT, like the AhR, belong to the basic helix-loop-helix (bHLH)/Per-Arnt-Sim- (PAS) family of transcriptional factors that are key regulators of gene expression networks underlying many essential biological processes. They have similar functions. Thus, three crystal structures of HIf-2a and ARNT LBD (PDB ID: 1P97, 2A24 and 3F1N) were selected as multi-templates to construct the 3D model of rAhR LBD. The sequence identities between rAhR LBD and its templates were 30% (Table 1). In general, sequence identities of 30% are enough to construct the 3D model of target proteins through the homology modeling. Fig. 1 shows the ribbon schematic representation of the final modeled structure of the rAhR LBD.

Table 1
Basic information on the target and template proteins.

Target	Swiss-Prot ID	Residues of LBD[a]	Template		
			PDB ID	Protein	Sequence identity
rAhR	P41738	273-384	1P97	Human hypoxia-inducible factor, HIF-2α	30%
			2A24	HIF-2a/ARNT PAS-B Heterodimer	30%
			3F1N	Heterodimer of HIF2 alpha and ARNT C-terminal PAS domains	30%
rCAR	Q9QUS1	114-345	1XNX	Mouse constitutive androstane receptor, mCAR	89%
rPXR	Q9R1A7	202-431	3CTB	Human tethered PXR-LBD/SRC-1p apoprotein	76%

[a] the residues of chosen ligand binding domains.