Science of the Total Environment 576 (2017) 609–616

Contents lists available at ScienceDirect

Science of the Total Environment

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

A high-throughput, computational system to predict if environmental contaminants can bind to human nuclear receptors

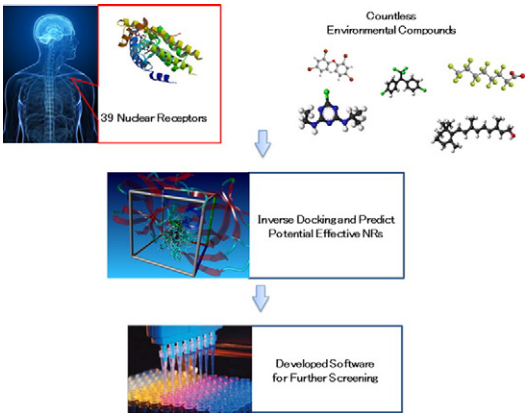
Xiaoxiang Wang [a,f], Xiaowei Zhang [a], Pu Xia [a], Junjiang Zhang [a], Yuting Wang [a], Rui Zhang [a,b], John P. Giesy [a,c,d,e], Wei Shi [a,*], Hongxia Yu [a,*]

[a] State Key Laboratory of Pollution Control and Resource Reuse, School of the Environment, Nanjing University, Nanjing 210046, PR China
[b] School of Resources and Environment, University of Jinan, Jinan, Shandong 250022, PR China
[c] Department of Veterinary Biomedical Sciences and Toxicology Centre, University of Saskatchewan, Saskatoon, Saskatchewan SK S7N5A2, Canada
[d] School of Biological Sciences, University of Hong Kong, Hong Kong, SAR 999077, PR China
[e] Zoology Dept. and Center for Integrative Toxicology, Michigan State University, East Lansing, MI 48824, USA
[f] Multiphase Chemistry Department, Max Planck Institute for Chemistry, 55128 Mainz, Germany

HIGHLIGHTS

- A database covering all existing crystal structures of 39 nuclear receptors was built.
- An inverse docking method was developed to predict highly vulnerable NRs.
- Some rarely reported targets (e.g. LRH-1) are suggested to be vulnerable NRs.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 25 September 2016
Received in revised form 10 October 2016
Accepted 13 October 2016
Available online 27 October 2016

Editor: Jay Gan

Keywords:
Nuclear receptors
Inverse docking
Agonists
Endocrine disrupting activities

ABSTRACT

Some pollutants can bind to nuclear receptors (NRs) and modulate their activities. Predicting interactions of NRs with chemicals is required by various jurisdictions because these molecular initiating events can result in adverse, apical outcomes, such as survival, growth or reproduction. The goal of this study was to develop a high-throughput, computational method to predict potential agonists of NRs, especially for contaminants in the environment or to which people or wildlife are expected to be exposed, including both persistent and pseudo-persistent chemicals. A 3D-structure database containing 39 human NRs was developed. The database was then combined with AutoDock Vina to develop a System for Predicting Potential Effective Nuclear Receptors (SPEN), based on inverse docking of chemicals. The SPEN was further validated and evaluated by experimental results for a subset of 10 chemicals. Finally, to assess the robustness of SPEN, its ability to predict potentials of 40 chemicals to bind to some of the most studied receptors was evaluated. SPEN is rapid, cost effective and powerful for predicting binding of chemicals to NRs. SPEN was determined to be useful for screening chemicals so that

* Corresponding authors.
E-mail addresses: njushiwei@nju.edu.cn (W. Shi), yuhx@nju.edu.cn (H. Yu).

http://dx.doi.org/10.1016/j.scitotenv.2016.10.093
0048-9697/

pollutants in the environment can be prioritized for regulators or when considering alternative compounds to replace known or suspected contaminants with poor environmental profiles.

1. Introduction

The adverse outcome pathway (AOP) is a framework proposed recently for use in toxicology and risk assessment and links exposures to chemicals to series of adverse outcomes (Vinken, 2013). AOPs are initiated by interactions of chemicals with biomolecules (Allen et al., 2014). The AOP framework is being applied to link burgeoning information at the molecular level of organization to adverse outcomes that can be used to make regulatory decisions. As a particular type of interaction, chemicals bind to nuclear receptors (NRs) and by mimicking natural ligands can lead to pleiotropic, adverse effects, such as modulation of the endocrine system (Grun and Blumberg, 2006) and immunodeficiency (Adorini et al., 2006). Contaminants that bind to ligand binding domains of NRs can be either agonists or antagonists. Disruption of endocrine function by contaminants through hormone NRs is one of the most significant issues of concern in environmental toxicology and ecotoxicology (Hopkins and Groom, 2002; Grun and Blumberg, 2006). Various NRs are linked to pathways and outcomes that can be adversely affected by small concentrations of environmental agonists or antagonists. Thus, understanding initiating effects, modulated via NRs, is important for prediction of outcomes of various contaminants. While determining whether a chemical binds to a NR as an agonist or an antagonist was beyond the scope of this study. The assessment of potential effects is a multi-step process, the first step of which is to determine if a chemical has potential to bind with a NR.

Efforts have been made to detect interactions between emerging pollutants and some NRs including androgen receptors (AR) or estrogen receptors (ER) (Fang et al., 2003; Blair et al., 2000), while it is far from enough to assess the effects of countless compounds on human through variety types of NRs. Thus, high throughput methods to determine the binding potential of NRs with chemicals of concern to humans or wildlife are needed for timely assessment. Since the initial event determining such interactions is binding to NRs, if binding affinities can be predicted it would allow an initial prioritization of which chemicals are likely to cause adverse effects via NR-mediated pathways. And it will also give some insight into what in vitro transactivation assays would be appropriate or what endpoints would be appropriate to monitor during in vivo exposures. This information would also be useful to determine potential cross-talk between or among pathways and to determine consistent effects among chemicals, thus informing categorization during assessments of hazard or risk. Finally, this information would be useful to determine which chemicals would be most likely to cause the same or similar effects such that they should be considered together during assessments of hazard or risk. Integrative experimental approaches, such as *High Information Content Toxicity Screening* (USEPA) and *Molecular Screening and Toxicogenomics* (Toxicogenomic), have been proposed to solve this problem. However, all of these strategies are time-consuming and expensive and some require use of live animals.

Developments in computational chemistry have demonstrated potential to supplement experimental testing for chemical hazard assessment. Several computational tools have been developed to predict potential targets of chemicals based on inverse docking (Chen and Zhi, 2001; Kumar et al., 2014), whereby a small molecule is docked into a panel containing multiple receptors (Fig. 1). Recently, an online tool was developed based on this method and used to investigate cosmetic ingredients (Kolšek et al., 2014; Plošnik et al., 2015). In this study, a database containing 39 human NRs was constructed and a software program, based on inverse docking, was developed to predict effective NRs for several chemicals of emerging concern. The System for Predicting Potential Effective Nuclear Receptors (SPEN) was used to predict the most probable effective NRs of 40 emerging environmental contaminants (Fig. 2).

2. Methods

2.1. Target database and SPEN based on AutoDock Vina

The target database for inverse docking contained 39 NRs (Table 1). There are 48 types of NRs in humans, but only 39 of these NRs have known structures (Zhao et al., 2015). The 3D crystal structures of LBDs (agonist conformation) of NRs can be obtained from the Protein Data Bank (PBD; http://www.rcsb.org/pdb/). After obtaining most of the necessary 3D structures from the PDB, additional residues were added with Swiss-PdbViewer 4.0 (Guex and Peitsch, 1997). Further information which included removing water molecules and buffers because the existing water in crystal structure could affect the binding, assigning charges and adding polar hydrogens, was obtained by use of AutoDock

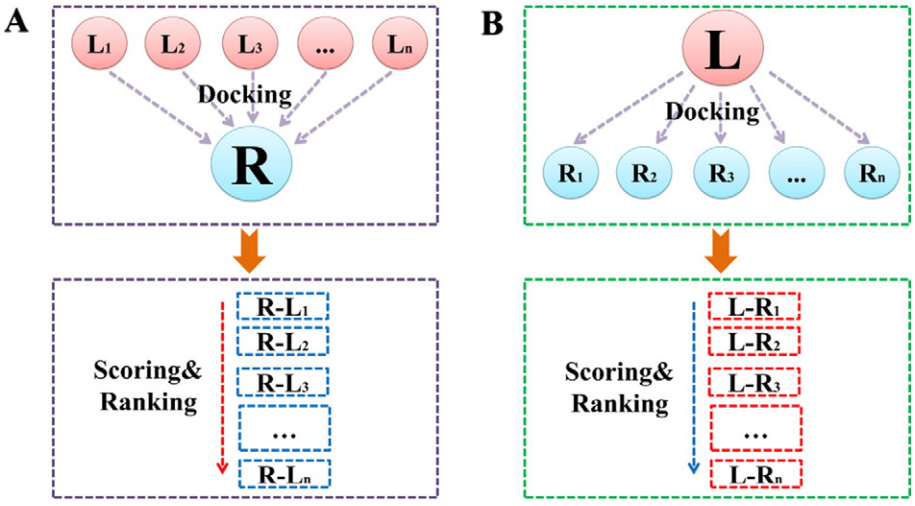


Fig. 1. Schematic representations of docking (A) and inverse docking (B). The term "L" and "R" represent ligand and receptor, respectively.