JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Endocrine Disruptome—An Open Source Prediction Tool for Assessing Endocrine Disruption Potential through Nuclear Receptor Binding

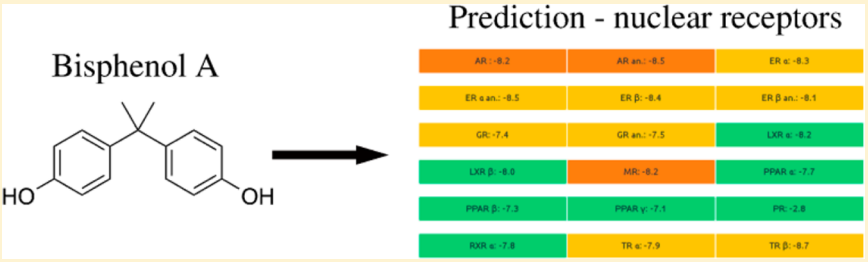
Katra Kolšek,[†,‡] Janez Mavri,[‡,§] Marija Sollner Dolenc,[†] Stanislav Gobec,[†] and Samo Turk*[,†,∥]

[†]Faculty of Pharmacy, University of Ljubljana, Aškerčeva 7, 1000 Ljubljana, Slovenia

[‡]Laboratory for Biocomputing and Bioinformatics, National Institute of Chemistry, Hajdrihova 19, 1000 Ljubljana, Slovenia

[§]EN-FIST Centre of Excellence, Dunajska 156, 1000 Ljubljana, Slovenia

Supporting Information



ABSTRACT: Predicting the endocrine disruption potential of compounds is a daunting but essential task. Here we report a new tool for this purpose that we have termed Endocrine Disruptome. It is a free and simple-to-use Web service that runs on an open source platform called Docking interface for Target Systems (DoTS). The molecular docking is handled via AutoDock Vina. Compounds are docked to 18 integrated and well-validated crystal structures of 14 different human nuclear receptors: androgen receptor; estrogen receptors α and β; glucocorticoid receptor; liver X receptors α and β; mineralocorticoid receptor; peroxisome proliferator activated receptors α, β/δ, and γ; progesterone receptor; retinoid X receptor α; and thyroid receptors α and β. Endocrine Disruptome is free of charge and available at http://endocrinedisruptome.ki.si.

INTRODUCTION

The trend towards replacement, reduction, and refinement (the 3Rs) of animal testing was first described by Russell and Burch in 1959 at the annual meeting of the former American Association of Laboratory Animal Science in Washington D.C.[1] This guiding principle for alternatives to animal testing, now adopted by regulation,[2] led to a flood of new *in vitro*[3] and, more recently, in silico[4,5] approaches for predicting toxicological end points. The latter are especially appealing because of low costs, quick results, and even the ability to operate on purely hypothetical virtual chemicals.[6]

Prediction of endocrine disruption potential of compounds remains a daunting task.[7] Several different methodologies have been developed, but most of them use quantitative structure–activity relationship (QSAR) and/or machine learning models.[8] These methods have exceptional predictivity for structural analogues but are less successful in predicting the endocrine disruption potentials of new, significantly different molecules.[9] From structure-based methods molecular docking and molecular dynamics are used, as in VirtualToxLab.[10,11] The most widely investigated mode of action of endocrine disruptors (EDCs) is direct binding to at least one of the relevant nuclear receptors, so molecular docking can be used as a valuable tool for binding prediction. Most of the available molecular docking software is intended to predict binding affinities of multiple ligands to one target at a time. To screen the endocrine disruption potential of a compound for multiple targets at once, the opposite is required (i.e., binding of one ligand to multiple targets). We therefore undertook the challenge of developing a user-friendly open-source Web-based prediction tool for toxicologists and experts in the field of endocrine disruption. Our aim was to develop a well-validated and easy-to-use Web platform. Rapid prediction of binding probabilities to a variety of targets/nuclear receptors would help experts make informed decisions on further testing. Fourteen human nuclear receptors that regulate reproduction, behavior, development, metabolism, and the immune system[12,13] were chosen: androgen receptor (AR); estrogen receptors α (ER α) and β (ER β); glucocorticoid receptor (GR); liver X receptors α (LXR α) and β (LXR β); mineralocorticoid receptor (MR); peroxisome proliferator activated receptors α (PPAR α), β/δ (PPAR β), and γ (PPAR γ); progesterone receptor (PR); retinoid X receptor α (RXR α); and thyroid receptors α (TR α) and β (TR β). Where available, both agonistic and antagonistic conformations of receptors were evaluated for inclusion in our prediction tool. More than seven different three-dimensional (3D) structures

Received: November 4, 2013
Published: March 14, 2014

dx.doi.org/10.1021/ci400649p | J. Chem. Inf. Model. 2014, 54, 1254–1267

per receptor (103 3D structures overall) were evaluated by docking of different sets of ligands and decoys. Finally, 18 structures with the best predictive power were chosen. In parallel, a Web interface named Docking interface for Target Systems (DoTS) was developed and is available on GitHub (https://github.com/katrakolsek/DoTS). DoTS, together with the 18 validated structures, forms a freely accessible web platform called Endocrine Disruptome (available at http://endocrinedisruptome.ki.si).

MATERIALS AND METHODS

Compounds and Databases. All of the compounds and databases of compounds were carefully prepared and classified in the following way:

Actives: Active compounds downloaded from the DUD-E website.[14] If not available, they were obtained from the ChEMBL database[15] using binding and functional assays with IC_{50} or EC_{50} less than 1 μM.

Agonists: Ligands with IC_{50} or EC_{50} less than 1 μM obtained from functional assays in the ChEMBL database, filtered with word "agonist(ic)" in the assay description.

Antagonists: Ligands with IC_{50} or EC_{50} less than 1 μM obtained from functional assays in the ChEMBL database, filtered with word "antagonist(ic)" in the assay description.

Decoys: Compounds with similar physicochemical properties as ligands but with different chemistry and thus assumed to be inactive. For each ligand, 50 decoys were downloaded or generated from the DUD-E website.

Active database (ACDB): Actives + decoys generated from those compounds.

Agonist database (AGDB): Agonists + decoys generated from those compounds.

Antagonist database (ANDB): Antagonists + decoys generated from those compounds.

Reduced decoys: Decoys clustered to retain only a maximum of 1000 of the most diverse decoys.

Reduced active database (rACDB): Actives + reduced decoys generated from those compounds.

Reduced agonist database (rAGDB): Agonists + reduced decoys generated from those compounds.

Reduced antagonist database (rANDB): Antagonists + reduced decoys generated from those compounds.

Selection and Validation of Receptor Structures. *Preparation of Ligands and Decoys.* Structures of active compounds were obtained from the DUD-E website or, if not available, downloaded from the ChEMBL database. In the latter case, compounds tested in functional and binding assays were used. In addition, structures of agonists and antagonists were downloaded separately from the ChEMBL database, and here, only compounds tested in functional assays were used. Only compounds with IC_{50} or EC_{50} lower than 1 μM and molecular weight below 600 were selected. These criteria are the same as those reported by Mysinger et al.[14] The exceptions were ligands for the TRs, where no limit on molecular weight was applied because of the high molecular weights of the endogenous ligands triiodothyronine (T3) and thyroxine (T4). All duplicate compounds were removed. Protonation was set as it would be at pH 7.4 with the Open Babel toolkit.[16,17] Three types of ligands were finally obtained for most of the receptors: actives, agonists, and antagonists. Because 50 decoys were generated per compound (look below for preparation of decoys), in cases that yielded more than 600 ligands we had to reduce the number of ligands to ultimately get a more manageable number of decoys. In these cases, ligands were clustered according to Murcko scaffolds[18] with RDKit nodes for KNIME, and 600 ligands with the most dissimilar scaffolds were selected.[14] In the opposite case, when fewer than 10 ligands were found, no database was formed (see Table 1).

Table 1. Numbers of Ligands Used in Validation of Nuclear Receptors

	actives	agonists	antagonists
AR	269[a]	290	545
ER α	383[a]	218	122
ER β	367[a]	221	29
GR	258[a]	316	341
LXR α	418	74	*1*
LXR β	451	165	*0*
MR	94[a]	*2*	119
PPAR α	373[a]	300 (804)	*1*
PPAR β	240[a]	492	*2*
PPAR γ	484[a]	411 (985)	*8*
PR	293[a]	379	482
RXR α	131[a]	126	59
TR α	187	*8*	*8*
TR β	103[a]	*9*	*4*

[a]Ligands obtained from the DUD-E website. Others were obtained from the ChEMBL database. Values in *italics* indicate numbers of ligands smaller than 10, which were not used to form databases. Values in parentheses are numbers of ligands before clustering according to Murcko scaffolds.

Decoys for each type of ligand were either downloaded preprepared from the DUD-E website or generated by the DUD-E decoy generator,[14] which generates 50 decoys per ligand. Decoys have similar physicochemical properties as the ligand from which they were generated but have different chemistry (scaffold) and are thus considered inactive. Inactive molecules (i.e., molecules that are known not to bind to receptor in question) from the DUD-E database were docked in our preliminary docking studies (not shown), but we realized that the number of known inactive compounds was too low for proper validation. Because some complete sets of decoys had tens of thousands of compounds, we used subsets of reduced decoys for testing the protocols and the predictive power of different receptor structures. Reduced decoys were prepared by clustering decoys using a Python script utilizing RDKit (see code snippet S1 in the Supporting Information).[19] The script calculates the Tanimoto distance matrix for the whole set and then outputs the 1000 most structurally diverse decoys.

All of the compound structures (actives, agonists, antagonists, and decoys) were preprocessed with a Python script utilizing the Open Babel toolkit.[16,17] Different options for 3D generation were assessed, and we ultimately settled for the following procedure: the 3D conformation was generated by the "make3D" function in 10 steps utilizing MMFF94 followed by 500 steps of local geometry optimization with MMFF94.

Preparation of Databases. Ligands (actives, agonists, or antagonists) and decoys were combined in databases. The result of database preparation was thus up to three different full databases per receptor: the active database (ACDB), agonist database (AGDB), and antagonist database (ANDB). For testing of the protocols and the predictive power of different receptor structures, up to three different reduced databases per receptor were also created: the reduced active database

(rACDB), reduced agonist database (rAGDB), and reduced antagonist database (rANDB). Not all of the receptors have all three databases if the number of ligands in any case is less than 10.

Preparation of Receptors. At least six 3D structures of the human wild-type isoform for each receptor were selected from the Protein Data Bank (PDB) (see Table 2). Structures with higher resolution and structurally diverse ligands were preferred. Both agonist and antagonist conformers of receptors were considered if available in the PDB. Only chain A for each crystal structure was used for the docking experiments, except in cases where only a small number of crystal structures were available and where the predictive power of those was poor. In these cases, other chains were also used (see Table 2). Receptors were prepared using AutoDock Tools.[20,21] First, only polar hydrogens were added, because AutoDock Vina uses only these for docking. Second, AutoDock 4 (AD4) atom types were assigned, and the size of the grid box was determined to be 22 Å × 22 Å × 22 Å in all cases except for LXRβ, where it was 24 Å × 22 Å × 22 Å. Next, the grid box was positioned around the ligand, and finally, the prepared receptor was saved as a pdbqt file.

Molecular Docking. Molecular docking experiments were performed with Autodock Vina 1.1.2 using default settings.[22] Two sets of docking experiments were performed. Reduced databases were used to select the structure with the best performance. Each reduced database includes one type of ligand together with the corresponding reduced decoy set containing the 1000 most structurally diverse compounds. rACDB and rAGDB were docked into 3D structures crystallized with agonists, and rACDB and rANDB were docked into 3D structures crystallized with antagonists (see Table 2). After the 3D structure with the best validation parameters was selected, the second set of docking experiments was performed with complete databases using the complete set of decoys.

Statistics. 3D structures were evaluated using the area under the curve (AUC) of the receiver operating characteristic (ROC) curve and the ROC-based enrichment factor at 1% (EF1%). EF1% is the percentage of known ligands found in the top 1% of the ranked database and is always compared with the ideal EF1%. The detailed procedure for calculating AUC and EF1% has been described elsewhere.[23,24] Calculations of sensitivity (SE), specificity (SP), ROC curves, AUC, and EF1% were implemented in a Python script (code snippet S2 in the Supporting Information). To create prediction classes, a Python script was written to calculate threshold values, the positive predictive value (PPV), and the negative predictive value (NPV) from a desired SE or SP (code snippet S3 in the Supporting Information).

Web Interface. We call the Web interface the Docking interface for Target Systems (DoTS). The DoTS source code is freely available on GitHub (https://github.com/katrakolsek/DoTS) under a permissive BSD license. On the server side, DoTS requires Python 2.7, Django 1.4 or 1.5, OpenBabel 2.3 with Python bindings, Celery 3.0, and AutoDock Vina. On the user side, it needs a modern Web browser that supports HTML5 canvas element. It should work on most browsers, including Internet Explorer version 9 or above.

Technologies Used in DoTS. DoTS is mostly written in Python and is based on the Django framework.[25] Chemistry handling in the back end is performed with OpenBabel Python bindings.[26] Pan assay interference compounds (PAINS)[27] alerting is implemented as a Python function with embedded previously published SMARTS[28] codes for PAINS.[29] The docking experiments are performed with AutoDock Vina[22] using the same protocols as for validation. We programmed a simple Python interface for AutoDock Vina utilizing the Celery library[30] that enables DoTS to perform asynchronous calculations serving multiple users at the same time. The front end utilizes the latest Web technologies and uses only standard-compliant HTML5. Chemical structure input and depiction is handled using the ChemDoodle JavaScript library.[31] Plots are depicted using the Flot jQuery plugin.[32]

Overview of the Interface. The web interface front end is fairly simple and self-explanatory.

The "Home" page offers a quick overview with the latest news.

The "Prediction" page offers users the possibility to input their structure for the docking experiments. Structures can either be drawn with a structure editor (ChemDoodle Sketcher[31]) or input directly as a SMILES string[33] (Figure S1 in the Supporting Information). Upon submission of the structure, the user is immediately redirected to a new page assigned to their calculation, where the structure of their compound, some basic descriptors, and PAINS[27] alerts are shown. Because the docking itself takes time, these results are not shown right away but appear only when the docking is finished. The page is automatically refreshed every 30 seconds. The docking results are shown as a color-coded table. In order to facilitate interpretation, each cell represents one receptor structure, which is written as its abbreviation and is accompanied by the score of the compound on that receptor structure. The cells are colored from red (representing a high probability of binding) through orange and yellow to green (representing a low probability of binding). The probability thresholds are calculated from the score and the validation experiments.

The "Targets" page offers a quick overview of targets and also a detailed subpage for each target that contains a short target description and results of the validation for the best-performing structure(s) of the target (Figure S2 in the Supporting Information). Results of the validation are shown for agonist and antagonist forms of the target if both structures are available. The ROC curve, with AUC and EF1%, is shown for each structure and accompanied by a table showing score cutoffs for probability thresholds.

The "FAQ" page contains frequently asked questions and answers. It explains what Endocrine Disruptome is, how to use it, and how to interpret the results. It also explains how user data is handled (look below).

The "About" page contains author information and frequently asked questions.

DoTS can be used for any target system and for different kinds of targets, not just nuclear receptors. The target structures have to be validated and all the data input into the system for it to function properly. DoTS needs the following data about the receptor structure:

- raw ROC data with assigned binary classifiers;
- threshold data with score and SE, SP, PPV, and NPV at the given score;
- names of the structure and AutoDock Vina configuration file, both without suffixes.

For more information and examples, consult the source code and supplementary files on the DoTS GitHub page.

Table 2. Validation Parameters for All of the Dockings[a]

Androgen Receptor						
3D Structures Bound with Agonist						
pdb code	2am9	2ax9	2pit	2pnu	3b66	**3l3x**
AUC	0.7898 (rAGDB)	0.7012 (rAGDB)	0.7933 (rAGDB)	0.7785 (rAGDB)	0.5173 (rAGDB)	**0.8163 (rAGDB)**
	0.5429 (rACDB)	0.4632 (rACDB)	0.5589 (rACDB)	0.6116 (rACDB)	0.4368 (rACDB)	**0.5659 (rACDB)**
EF1%	3.4 (ideal 4.4, rAGDB)	3.8 (ideal 4.4, rAGDB)	3.8 (ideal 4.4, rAGDB)	1.4 (ideal 4.4, rAGDB)	1.0 (ideal 4.4, rAGDB)	**4.1 (ideal 4.4, rAGDB)**
	4.4 (ideal 4.7, rACDB)	3.3 (ideal 4.7,rACDB)	4.4 (ideal 4.7, rACDB)	1.8 (ideal 4.7, rACDB)	1.8 (ideal 4.7,rACDB)	**4.4 (ideal 4.7, rACDB)**
3D Structures Bound with Antagonist						
pdb code	**2z4j**	3rlj				
AUC	**0.5824 (rANDB)**	0.4791 (rANDB)				
	0.5552 (rACDB)	0.4237 (rACDB)				
EF1%	**2.8 (ideal 2.8, rANDB)**	1.9 (ideal 2.8, rANDB)				
	4.4 (ideal 4.7, rACDB)	1.8 (ideal 4.7, rACDB)				
Estrogen Receptor α						
3D Structures Bound with Agonist						
pdb code	**1a52**	1gwr	1xpc	2iog	2ouz	2r6w
AUC	**0.7833 (rAGDB)**	0.7360 (rAGDB)	0.7509 (rAGDB)	0.6842 (rAGDB)	0.7381 (rAGDB)	0.7196 (rAGDB)
	0.7388 (rACDB)	0.5304 (rACDB)	0.7344 (rACDB)	0.6722 (rACDB)	0.7573 (rACDB)	0.7782 (rACDB)
EF1%	**4.2 (ideal 5.6, rAGDB)**	4.7 (ideal 5.6, rAGDB)	2.8 (ideal 5.6, rAGDB)	1.9 (ideal 5.6, rAGDB)	1.9 (ideal 5.6, rAGDB)	3.3 (ideal 5.6, rAGDB)
	2.8 (ideal 3.6, rACDB)	3.5 (ideal 3.6, rACDB)	3.6 (ideal 3.6, rACDB)	1.5 (ideal 3.6, rACDB)	3.6 (ideal 3.6, rACDB)	3.6 (ideal 3.6, rACDB)
3D Structures Bound with Antagonist						
pdb code	1err	1gwq	**1sj0**			
AUC	0.7064 (rANDB)	0.4884 (rANDB)	**0.7863 (rANDB)**			
	0.7290 (rACDB)	0.4523 (rACDB)	**0.8044 (rACDB)**			
EF1%	5.9 (ideal 9.2, rANDB)	8.4 (ideal 9.2, rANDB)	**8.4 (ideal 9.2, rANDB)**			
	3.6 (ideal 3.6, rACDB)	3.6 (ideal 3.6, rACDB)	**3.6 (ideal 3.6, rACDB)**			
Estrogen Receptor β						
3D Structures Bound with Agonist						
pdb code	1yye	2giu	2j7x	2nv7	**3ols**	
AUC	0.6755 (rAGDB)	0.8095 (rAGDB)	0.8076 (rAGDB)	0.6763 (rAGDB)	**0.8222 (rAGDB)**	
	0.4816 (rACDB)	0.5406 (rACDB)	0.4981 (rACDB)	0.4868 (rACDB)	**0.5093 (rACDB)**	
EF1%	4.6 (ideal 5.5, rAGDB)	4.1 (ideal 5.5, rAGDB)	5.0 (ideal 5.5, rAGDB)	5.0 (ideal 5.5, rAGDB)	**5.0 (ideal 5.5, rAGDB)**	
	3.7 (ideal 3.7, rACDB)	3.7 (ideal 3.7, rACDB)	3.7 (ideal 3.7, rACDB)	3.7 (ideal 3.7, rACDB)	**3.7 (ideal 3.7, rACDB)**	
3D Structures Bound with Antagonist						
pdb code	1hj1	1l2j	**1qkn**			
AUC	0.6312 (rANDB)	0.4474 (rANDB)	**0.5999 (rANDB)**			
	0.7599 (rACDB)	0.6391 (rACDB)	**0.7842 (rACDB)**			
EF1%	3.5 (ideal 35.5, rANDB)	0.0 (ideal 35.5, rANDB)	**3.5 (ideal 35.5, rANDB)**			
	2.9 (ideal 3.7, rACDB)	1.9 (ideal 3.7, rACDB)	**3.7 (ideal 3.7, rACDB)**			
Glucocorticoid Receptor						
3D Structures Bound with Agonist						
pdb code	1m2z	1p93	**3bqd**	3cld	3e7c	3k22
AUC	0.5313 (rAGDB)	0.5166 (rAGDB)	**0.6196 (rAGDB)**	0.6521 (rAGDB)	0.5338 (rAGDB)	0.4938 (rAGDB)

dx.doi.org/10.1021/ci400649p | J. Chem. Inf. Model. 2014, 54, 1254−1267

Table 2. continued

Glucocorticoid Receptor						
3D Structures Bound with Agonist						
	0.4426 (rACDB)	0.4267 (rACDB)	**0.5490 (rACDB)**	0.5174 (rACDB)	0.4526 (rACDB)	0.5064 (rACDB)
EF1%	4.2 (ideal 4.2, rAGDB)	3.2 (ideal 4.2, rAGDB)	**4.2 (ideal 4.2, rAGDB)**	4.2 (ideal 4.2, rAGDB)	1.9 (ideal 4.2, rAGDB)	1.9 (ideal 4.2, rAGDB)
	3.0 (ideal 4.9, rACDB)	1.5 (ideal 4.9, rACDB)	**4.9 (ideal 4.9, rACDB)**	1.9 (ideal 4.9, rACDB)	2.6 (ideal 4.9, rACDB)	2.3 (ideal 4.9, rACDB)
3D Structures Bound with Antagonist						
pdb code	1nhz	3h52 chain A	3h52 chain B	**3h52 chain C**	3h52 chain D	
AUC	0.5445 (rANDB)	0.4658 (rANDB)	0.5134 (rANDB)	**0.6209 (rANDB)**	0.5010 (rANDB)	
	0.6387 (rACDB)	0.5805 (rACDB)	0.5588 (rACDB)	**0.6619 (rACDB)**	0.6263 (rACDB)	
EF1%	0.6 (ideal 3.9, rANDB)	0.9 (ideal 3.9, rANDB)	0.3 (ideal 3.9, rANDB)	**0.0 (ideal 3.9, rANDB)**	1.5 (ideal 3.9, rANDB)	
	1.9 (ideal 4.9, rACDB)	1.1 (ideal 4.9, rACDB)	3.0 (ideal 4.9, rACDB)	**2.6 (ideal 4.9, rACDB)**	2.6 (ideal 4.9, rACDB)	
Liver X Receptor α						
3D Structures Bound with Agonist						
pdb code	1uhl chain B		3ipq	**3ips**	3ipu	
AUC	0.6989 (rAGDB)		0.7700 (rAGDB)	**0.7577 (rAGDB)**	0.7606 (rAGDB)	
	0.7332 (rACDB)		0.7779 (rACDB)	**0.7588 (rACDB)**	0.7334 (rACDB)	
EF1%	0.0 (ideal 14.5, rAGDB)		9.2 (ideal 14.5, rAGDB)	**14.5 (ideal 14.5, rAGDB)**	7.9 (ideal 14.5, rAGDB)	
	1.7 (ideal 3.4, rACDB)		3.2 (ideal 3.4, rACDB)	**2.4 (ideal 3.4, rACDB)**	2.2 (ideal 3.4, rACDB)	
Liver X Receptor β						
3D Structures Bound with Agonist						
pdb code	**1p8d**	1pq6	1pq9	1upv	3loe	
AUC	**0.7412 (rAGDB)**	0.6893 (rAGDB)	0.6628 (rAGDB)	0.4529 (rAGDB)	0.6975 (rAGDB)	
	0.7743 (rACDB)	0.7452 (rACDB)	0.6012 (rACDB)	0.4294 (rACDB)	0.7596 (rACDB)	
EF1%	**5.3 (ideal 7.1, rAGDB)**	5.9 (ideal 7.1, rAGDB)	4.7 (ideal 7.1, rAGDB)	2.4 (ideal 7.1, rAGDB)	5.9 (ideal 7.1, rAGDB)	
	3.2 (ideal 3.2, rACDB)	2.1 (ideal 3.2, rACDB)	3.0 (ideal 3.2, rACDB)	1.7 (ideal 3.2, rACDB)	3.2 (ideal 3.2, rACDB)	
Mineralocorticoid Receptor						
3D Structures Bound with Agonist						
pdb code	2aa2		2aa7		2abi	
AUC	0.4210 (rACDB)		0.3697 (rACDB)		0.4585 (rACDB)	
EF1%	4.2 (ideal 11.6, rACDB)		3.2 (ideal 11.6, rACDB)		4.2 (ideal 11.6, rACDB)	
3D Structures Bound with Antagonist						
pdb code	2aa5		3vhu		**3vhv**	
AUC	0.2848 (rANDB)		0.3646 (rANDB)		**0.4211 (rANDB)**	
	0.4190 (rACDB)		0.5031 (rACDB)		**0.5486 (rACDB)**	
EF1%	2.6 (ideal 9.4, rANDB)		0.9 (ideal 9.4, rANDB)		**6.0 (ideal 9.4, rANDB)**	
	3.2 (ideal 11.6, rACDB)		5.3 (ideal 11.6, rACDB)		**3.2 (ideal 11.6, rACDB)**	
Peroxisome Proliferator-Activated Receptor α						
3D Structures Bound with Agonist						
pdb code	1i7g	2p54	2znn	3g8i	**3kdu**	3vi8
AUC	0.7415 (rAGDB)	0.8146 (rAGDB)	0.7579 (rAGDB)	0.8086 (rAGDB)	**0.8042 (rAGDB)**	0.7939 (rAGDB)
	0.7919 (rACDB)	0.8544 (rACDB)	0.7936 (rACDB)	0.8439 (rACDB)	**0.8530 (rACDB)**	0.7989 (rACDB)
EF1%	1.7 (ideal 4.3, rAGDB)	2.0 (ideal 4.3, rAGDB)	1.7 (ideal 4.3, rAGDB)	2.7 (ideal 4.3, rAGDB)	**3.3 (ideal 4.3, rAGDB)**	2.0 (ideal 4.3, rAGDB)
	2.6 (ideal 3.7, rACDB)	3.2 (ideal 3.7, rACDB)	2.6 (ideal 3.7, rACDB)	2.9 (ideal 3.7, rACDB)	**3.4 (ideal 3.7, rACDB)**	2.4 (ideal 3.7, rACDB)

dx.doi.org/10.1021/ci400649p | J. Chem. Inf. Model. 2014, 54, 1254−1267

Table 2. continued

Peroxisome Proliferator-Activated Receptor β/δ						
3D Structures Bound with Agonist						
pdb code	2awh	2baw	3d5f	3et2	**3gz9**	3tkm
AUC	0.7795 (rAGDB)	0.7780 (rAGDB)	0.4730 (rAGDB)	0.7607 (rAGDB)	**0.8229 (rAGDB)**	0.8004 (rAGDB)
	0.7871 (rACDB)	0.7819 (rACDB)	0.5351 (rACDB)	0.7867 (rACDB)	**0.8920 (rACDB)**	0.7483 (rACDB)
EF1%	2.4 (ideal 3.0, rAGDB)	2.8 (ideal 3.0, rAGDB)	0.8 (ideal 3.0, rAGDB)	2.0 (ideal 3.0, rAGDB)	**2.4 (ideal 3.0, rAGDB)**	1.8 (ideal 3.0, rAGDB)
	0.4 (ideal 5.2, rACDB)	3.9 (ideal 5.2, rACDB)	0.4 (ideal 5.2, rACDB)	1.7 (ideal 5.2, rACDB)	**3.9 (ideal 5.2, rACDB)**	0.4 (ideal 5.2, rACDB)
Peroxisome Proliferator-Activated Receptor γ						
3D Structures Bound with Agonist						
pdb code	3adx	3b1m	**3et3**	3u9q	3v9t	3v9v
AUC	0.7797 (rAGDB)	0.7391 (rAGDB)	**0.7910 (rAGDB)**	0.7728 (rAGDB)	0.7474 (rAGDB)	0.7666 (rAGDB)
	0.7960 (rACDB)	0.7384 (rACDB)	**0.8111 (rACDB)**	0.7506 (rACDB)	0.7775 (rACDB)	0.7832 (rACDB)
EF1%	2.9 (ideal 3.4, rAGDB)	2.7 (ideal 3.4, rAGDB)	**3.2 (ideal 3.4, rAGDB)**	3.2 (ideal 3.4, rAGDB)	2.9 (ideal 3.4, rAGDB)	2.5 (ideal 3.4, rAGDB)
	2.7 (ideal 3.1, rACDB)	2.7 (ideal 3.1, rACDB)	**2.9 (ideal 3.1, rACDB)**	2.2 (ideal 3.1, rACDB)	2.7 (ideal 3.1, rACDB)	2.7 (ideal 3.1, rACDB)
Progesterone Receptor						
3D Structures Bound with Agonist						
pdb code	1sqn	1zuc	3d90	**3hq5**	3kba	4apu
AUC	0.3756 (rAGDB)	0.3817 (rAGDB)	0.6911 (rAGDB)	**0.7108 (rAGDB)**	0.5631 (rAGDB)	0.4042 (rAGDB)
	0.4826 (rACDB)	0.4774 (rACDB)	0.6165 (rACDB)	**0.6106 (rACDB)**	0.5864 (rACDB)	0.5031 (rACDB)
EF1%	0.8 (ideal 3.6, rAGDB)	2.6 (ideal 3.6, rAGDB)	2.9 (ideal 3.6, rAGDB)	**3.4 (ideal 3.6, rAGDB)**	2.9 (ideal 3.6, rAGDB)	0.8 (ideal 3.6, rAGDB)
	3.1 (ideal 4.4, rACDB)	2.7 (ideal 4.4, rACDB)	3.1 (ideal 4.4, rACDB)	**1.7 (ideal 4.4, rACDB)**	3.7 (ideal 4.4, rACDB)	2.0 (ideal 4.4, rACDB)
3D Structures Bound with Antagonist						
pdb code	2w8y	3zr7	3zra	3zrb		
AUC	0.4667 (rANDB)	0.4628 (rANDB)	0.4788 (rANDB)	0.4838 (rANDB)		
	0.5959 (rACDB)	0.5243 (rACDB)	0.5123 (rACDB)	0.5455 (rACDB)		
EF1%	2.6 (ideal 3.1, rANDB)	0.8 (ideal 3.1, rANDB)	1.6 (ideal 3.1, rANDB)	1.2 (ideal 3.1, rANDB)		
	3.2 (ideal 4.4, rACDB)	2.0 (ideal 4.4, rACDB)	2.4 (ideal 4.4, rACDB)	1.7 (ideal 4.4, rACDB)		
Retinoid X Receptor α						
3D Structures Bound with Agonist						
pdb code	**1mv9**	1mvc	1mzn	2p1t	3e94	3fug
AUC	**0.8616 (rAGDB)**	0.7879 (rAGDB)	0.7683 (rAGDB)	0.7063 (rAGDB)	0.7867 (rAGDB)	0.7050 (rAGDB)
	0.7944 (rACDB)	0.7539 (rACDB)	0.7419 (rACDB)	0.7174 (rACDB)	0.7474 (rACDB)	0.7166 (rACDB)
EF1%	**5.0 (ideal 8.9, rAGDB)**	8.1 (ideal 8.9, rAGDB)	8.1 (ideal 8.9, rAGDB)	7.3 (ideal 8.9, rAGDB)	8.9 (ideal 8.9, rAGDB)	6.5 (ideal 8.9, rAGDB)
	8.6 (ideal 8.6, rACDB)	8.6 (ideal 8.6, rACDB)	8.6 (ideal 8.6, rACDB)	7.8 (ideal 8.6, rACDB)	8.6 (ideal 8.6, rACDB)	8.6 (ideal 8.6, rACDB)
3D Structures Bound with Antagonist						
pdb code	3nsq	3r2a chain A	3r2a chain B	3r2a chain C	3r2a chain D	3r29
AUC	0.4187 (rANDB)	0.3975 (rANDB)	0.2288 (rANDB)	0.4870 (rANDB)	0.3147 (rANDB)	0.3086 (rANDB)
	0.4538 (rACDB)	0.6611 (rACDB)	0.5102 (rACDB)	0.6860 (rACDB)	0.6233 (rACDB)	0.6322 (rACDB)
EF1%	0.0 (ideal 17.9, rANDB)	0.0 (ideal 17.9, rANDB)	0.0 (ideal 17.9, rANDB)	1.6 (ideal 17.9, rANDB)	0.0 (ideal 17.9, rANDB)	0.0 (ideal 17.9, rANDB)
	0.0 (ideal 8.6, rACDB)	3.1 (ideal 8.6, rACDB)	0.8 (ideal 8.6, rACDB)	3.1 (ideal 8.6, rACDB)	3.1 (ideal 8.6, rACDB)	0.8 (ideal 8.6, rACDB)

dx.doi.org/10.1021/ci400649p | J. Chem. Inf. Model. 2014, 54, 1254−1267

Table 2. continued

Thyroid Receptor α						
3D Structures Bound with Agonist						
pdb code	1nav	2h77	2h79	3hzf	**3ilz**	3jzb
AUC	0.7395 (rACDB)	0.6370 (rACDB)	0.6487 (rACDB)	0.8053 (rACDB)	**0.8127 (rACDB)**	0.7288 (rACDB)
EF1%	0.5 (ideal 6.3, rACDB)	0.5 (ideal 6.3, rACDB)	0.0 (ideal 6.3, rACDB)	5.3 (ideal 6.3, rACDB)	**6.3 (ideal 6.3, rACDB)**	2.6 (ideal 6.3, rACDB)
Thyroid Receptor β						
3D Structures Bound with Agonist						
pdb code	1n46	1nax	1xzx	3gws	**3imy**	3jzc
AUC	0.6868 (rACDB)	0.7286 (rACDB)	0.6922 (rACDB)	0.7248 (rACDB)	**0.7895 (rACDB)**	0.7298 (rACDB)
EF1%	9.7 (ideal 10.7, rACDB)	9.7 (ideal 10.7, rACDB)	9.7 (ideal 10.7, rACDB)	5.8 (ideal 10.7, rACDB)	**10.7 (ideal 10.7, rACDB)**	9.7 (ideal 10.7, rACDB)

[a]AUC is the area under the receiver operating characteristic curve. EF1% is the enrichment factor at 1%; the ideal EF1% is the achievable maximum of true positives identified in the top 1% of the ranked results. rACDB is the reduced active database; rANDB is the reduced antagonist database; and rAGDB is the reduced agonist database. Entries in bold indicate the 18 selected structures.

Data Handling. Because of the way that the system is programmed and set up, the data submitted by users remains stored on the Web server but is not freely accessible or browsable by third parties. Each calculation is assigned unique and random URL that is known only to the user who submitted the calculation. Server administrators have full access to all submitted data but it is used only for debugging purposes. Users concerned with data safety can run the tool locally on their computers since all of the source code and models are available on GitHub.

RESULTS AND DISCUSSION

There is a need for a reliable method that can predict the endocrine disruption potential of chemicals, considering several nuclear receptors at once. The method should be fast and easy to use. We were interested in identifying the best-performing structures of nuclear receptors that have good predictive power when used for docking with AutoDock Vina.[22] Moreover, we wanted to make these results available in a freely accessible and easy-to-use Web platform that automatically docks compounds to the selected and validated receptors. For this purpose, we developed the DoTS Web application. Validation results were integrated in DoTS to form the Endocrine Disruptome platform (Figure 1).

Selection and Validation of Receptor Structures. *Preparation of Ligands, Decoys and Databases.* According to the terminology used in this article, ligands are structures that bind to the receptor in question, decoys are structures that theoretically should not bind to the receptor, and databases comprise ligands and their corresponding decoys (see Compounds and Databases in Materials and Methods). The structures of ligands with activities better than 1 μM and molecular weights lower than 600 were obtained from either the DUD-E website[14] or the ChEMBL database.[15] A molecular weight filter was not applied for TRs because of the high molecular weights of endogenous ligands. Three different types of ligands were collected (again see Compounds and Databases): actives, including all compounds with agonistic, antagonistic, and nondefined activity on the target; agonists, including only compounds with agonistic activity; and antagonists, including only compounds with antagonistic activity. The numbers of ligands obtained from the DUD-E website[14] and from the ChEMBL database[15] are shown in Table 1. Agonists for PPAR α and PPAR γ were clustered according to Murcko scaffolds[18] because their large number of ligands would produce too many decoys for a reasonable docking time.[14] Ten groups of ligands contained less than 10 identified ligands (in *italics* in Table 1). In these cases, databases were not created because the validation results would clearly be questionable with such a low number of ligands. Decoys for each type of ligand were downloaded or generated by the DUD-E decoy generator.[14] For each ligand, 50 decoys were created automatically with the DUD-E decoy generator, so the set of decoys consisted of between 10 000 and 30 000 structures. To reduce computational costs of the first phase, the decoys were clustered to retain the 1000 structurally most diverse compounds. These reduced sets of decoys were used along with their corresponding ligands to form the reduced databases, which were used only in the initial experiments of receptor structure selection. After a 3D structure for each receptor was selected, a validation was performed with the databases containing a complete set of decoys.

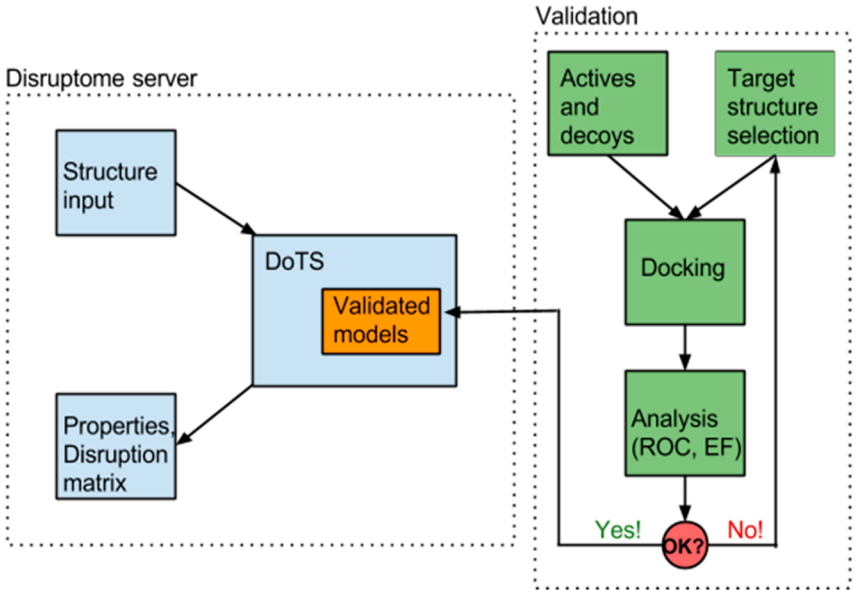


Figure 1. Schematic representation of the Endocrine Disruptome program package. The right side shows the validation steps that were undertaken. The left side shows the Docking interface for Target Systems (DoTS), the software platform on which Endocrine Disruptome runs.

Selection and Validation of Receptor Structures. The first step was to find the best-performing structure for each (agonistic and antagonistic) conformation of each receptor. We are aware that receptor structures are subject to fluctuations, but for practical reasons we considered only one structure per form. This selection docking was carried out on 103 different crystal structures (Table 2) with two reduced databases if both were created. In total, around 200 different docking experiments were performed. Reduced databases were created by combining ligands and 1000 structurally diverse decoys. rACDB and rAGDB were docked into the 3D structures cocrystallized with agonists, and rACDB and rANDB were docked into 3D structures bound to antagonists. Two validation parameters were calculated for each 3D structure. The first of these was the area under the curve (AUC) of the receiver operating characteristic (ROC) curve.[24] The ROC curve is a plot whose ordinate is the sensitivity (SE) and whose abscissa is 1 minus the specificity (1 − SP). SE can be interpreted as a true-positive rate and 1 − SP as a false-positive rate. AUC provides an estimate of the accuracy of the whole docking procedure. For diagnostic tests in medicine, AUC values can be classified as follows: 0.90−1, high accuracy; 0.70−0.90, moderate accuracy; 0.50−0.70, rather low accuracy; 0.50, random choosing of molecules.[34] No such classification was done for the performance of the docking software. For instance, one of the largest docking studies was performed by Mysinger et al.[14] on 102 different targets and they found that 2.0% of targets had AUCs less than 0.50, 19.6% had AUCs between 0.50 and 0.70, 71.6% had AUCs between 0.70 and 0.90, and 6.9% of targets had AUCs greater than 0.90. The second parameter is the ROC-based enrichment factor at 1% (EF1%),[23] which is the percentage of known ligands found in the top 1% of the ranked docking database. For instance, if 10% of the known ligands are found in the top 1% of the ranked database, than EF1% is equal to 10. EF1% depends on the ligand to decoy ratio and is therefore always compared to the ideal EF1%, which is calculated as the number of entries in the whole database divided by the number of known ligands. For instance, rACDB for AR consists of 269 actives (known ligands) and 1000 decoys, so in the entire database there are 1269 molecules. The ideal EF1% would therefore be 1269/269 or 4.7. This parameter indicates the accuracy of the prediction for the top-ranking results. For example, EF1% is important in virtual screenings in medicinal chemistry, where it is expected to find active compounds among the top-ranking ones.[35]

The results for each docking are summarized in Table 2. Analysis of the AUC and EF1% values for the receptors gave useful insights. The predictive power varied significantly between groups of receptors and between the crystal structures of these receptors. For example, the best-performing group of receptors was the group of PPAR receptors. All of the 18 tested crystal structures for the three PPAR receptors had high AUCs and EF1% values. On the other hand, the group of 3-ketosteroid receptors of the estrogen receptor-like superfamily, which includes AR, GR, MR, and PR, had the lowest predictive power and also the most variable validation parameters between crystal structures of each receptor. For example, comparison of the AUCs for PR structures cocrystallized with agonists and docked with the agonist database gave a difference of approximately 0.34 between the best- and worst-performing structures. These results indicate that selection of the appropriate 3D structure is of the utmost importance in the case of the 3-ketosteroid receptors group, while in the case of PPARs, this selection is less important.

One would expect that crystal structures with better resolution would have better performance, but this was not the case. Among the 18 selected structures, only five are the structures with the best resolution of the receptor. These are the AR agonist conformation, the ER α antagonist conformation, the ER β antagonist conformation, MR, and TR β (Table S2 in the Supporting Information).

Nuclear receptors can be divided into classes on the basis of their pattern of dimerization. The majority of nuclear receptor fall into two classes. The first class consists of steroid receptors (ER, AR, GR, MR, PR), which form homodimers in the cytosol upon ligand binding. The second class consists of all ligand-dependent nuclear receptors except steroid receptors. The receptors in the second class heterodimerize with RXR upon

dx.doi.org/10.1021/ci400649p | J. Chem. Inf. Model. 2014, 54, 1254−1267

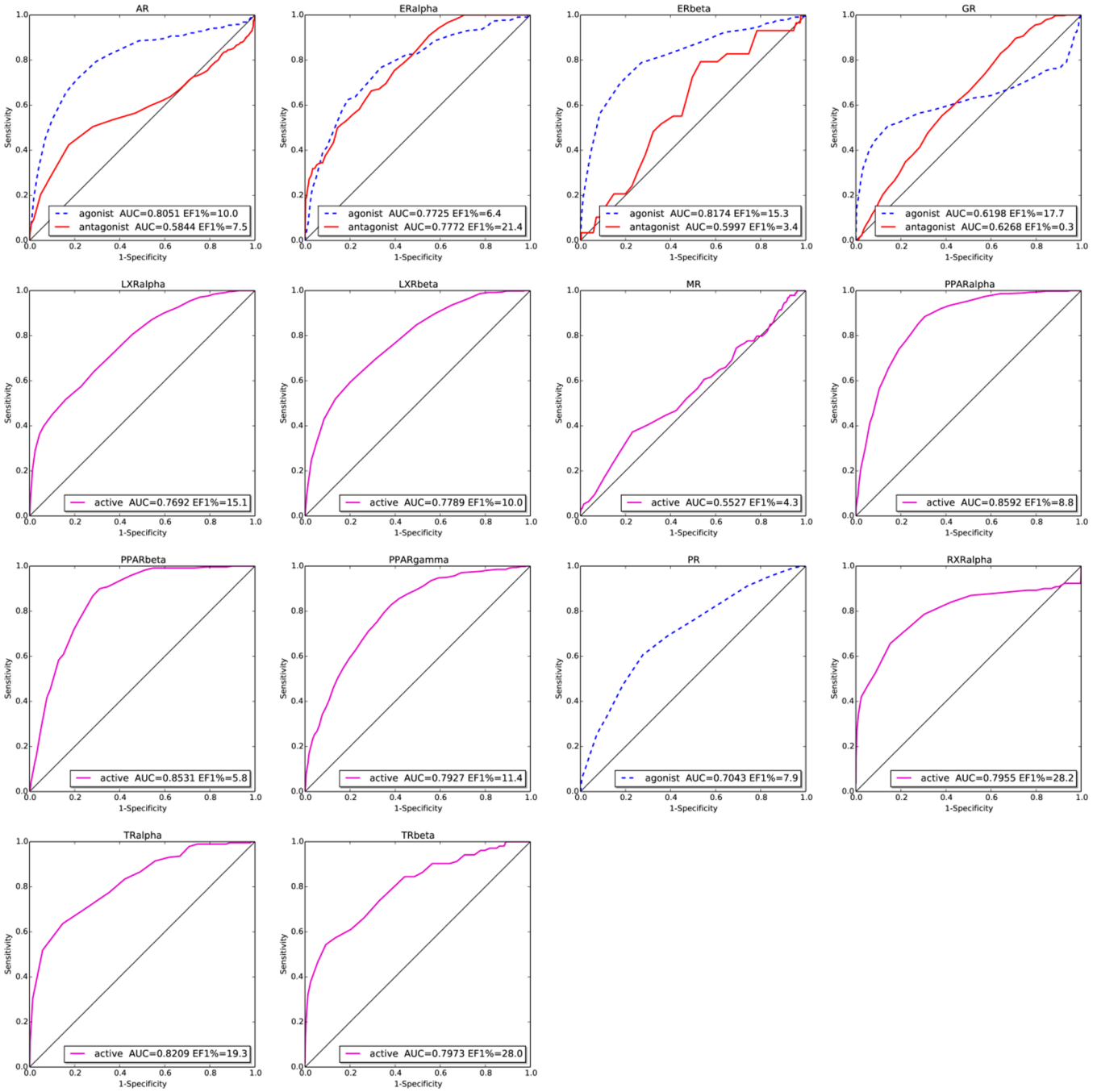


Figure 2. Receiver operating characteristic (ROC) curves with area under the curve (AUC) and enrichment factor at 1% (EF1%; the ideal EF1% is 51) for the 18 selected structures. Blue dotted line, agonist database; red line, antagonist database; magenta line, actives database; black line, AUC = 0.50.

ligand binding. This is a rough classification, and it should be noted that some of these receptors are known also to form heterodimers with others or do not always dimerize but stay as monomers.[36,37] In Table S2 in the Supporting Information we have collected the oligomeric states of all of the crystal structures used in this article. Among all of the tested crystal structures of steroid receptors, 48.98% were crystallized as monomers, 38.78% as dimers, 2.04% as trimers, and 10.20% as tetramers. This ratio had moved in favor of dimers when the selected structures of steroid receptors were analyzed (20.00% monomers, 70.00% dimers, and 10.00% tetramers). All of the other receptors used in this article belong to the second class according to the above definition. For this class, most of the tested structures were crystallized as monomers (39.58%), followed by dimers (37.50%), tetramers (18.75%), hexamers (2.08%), and octamers (2.08%). In this case, the ratio almost stayed the same also in the selected structures. Monomers and dimers were equally distributed, and together they represent 75.00% of the selected crystal structures, with the remaining 25.00% of the structures found to be tetramers.

Finally, the 18 structures with the best performance (shown in bold in Table 2) were selected for final evaluation. Four receptors were represented by both agonistic and antagonistic forms.

dx.doi.org/10.1021/ci400649p | J. Chem. Inf. Model. 2014, 54, 1254−1267

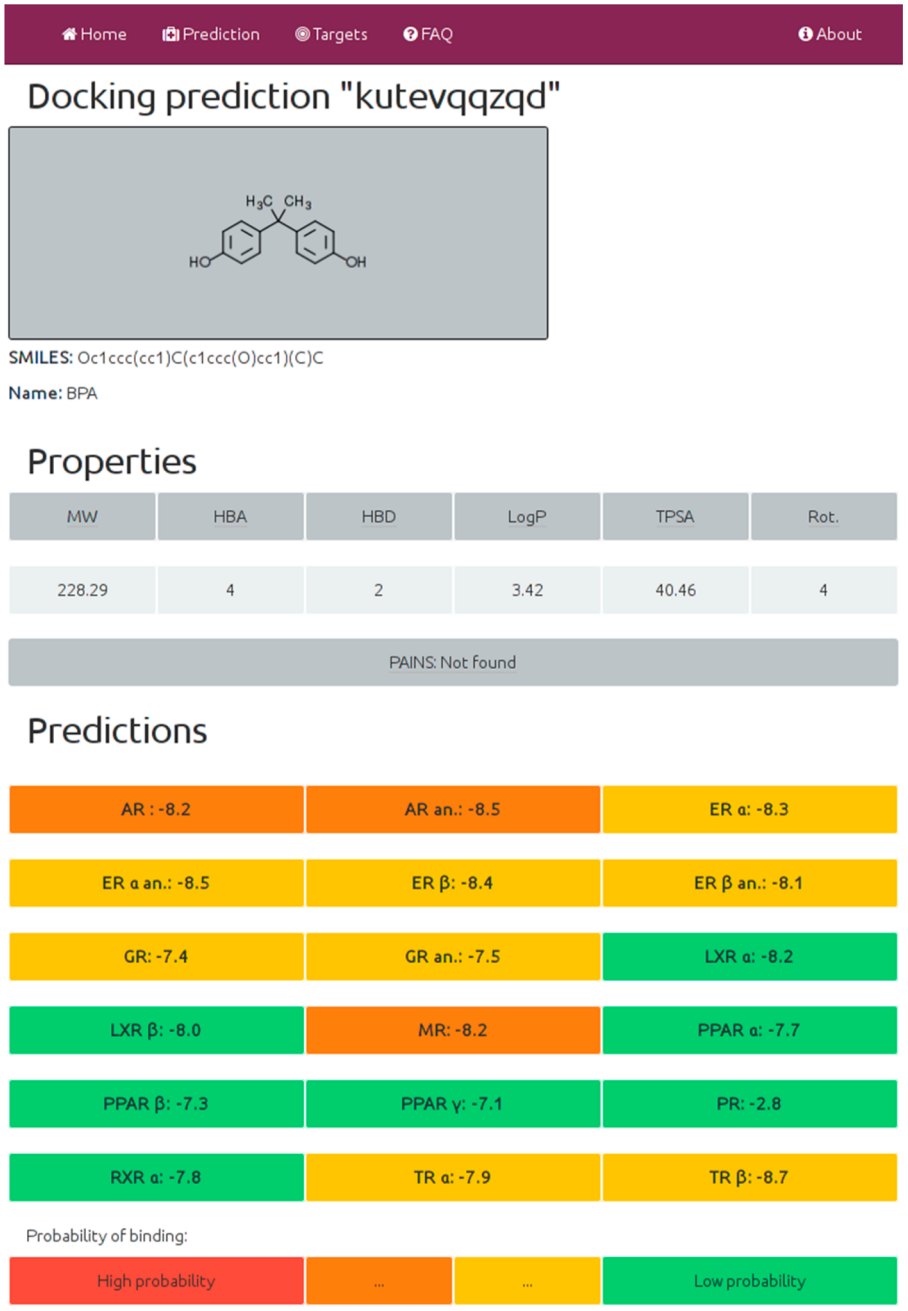


Figure 3. Example of a results page showing calculated physicochemical properties for BPA and its predicted binding to different receptors. The abbreviation "an" stands for the antagonist conformation.

Final Validation of the Chosen Receptors. In the second step, we performed the final validation on the 18 structures selected to have the best validation parameters. This was carried out with the databases containing ligands and full sets of decoys (ACDB, AGDB, and ANDB). The docking results were used to calculate the final performance metrics, including ROC curves, AUC, and EF1% (Figure 2). In four cases, two different structures per receptor were obtained and validated, one with bound agonist and one with bound antagonist. In the cases of MR, PR, and RXR α, only one structure per receptor could be used because of bad performance of the agonist or antagonist form (Table 2). MR 3D structures bound with agonists and PR and RXR α 3D structures bound with antagonists all had too weak predictability. MR and RXR α were therefore only validated with ACDB (Figure 2). Because all of the PR 3D structures performed badly when rACDB was used (Table 2), PR was in the end validated only with AGDB (Figure 2), meaning that the chosen PR structure is much better for prediction of agonists than antagonists. Overall, the best-performing structures were PDB entry 3kdu of PPARα (AUC = 0.8592) and PDB entry 3gz9 of PPAR β (AUC = 0.8531) (Figure 2). Most of the other structures performed well, with the notable exception of MR (AUC = 0.5527), for which the performance was only mediocre. We are not the first group to have problems with the performance of docking software on MR structures.[14] It has to be emphasized therefore that Endocrine Disruptome should be used with special care when interpreting results from MR docking. In conclusion, 13 of the final 18 structures had AUC values better than 0.7.

dx.doi.org/10.1021/ci400649p | J. Chem. Inf. Model. 2014, 54, 1254−1267

Table 3. Additional Examples for Endocrine Disruptome

Chemicals	Triclosan[a]		Gibberellic acid[b]		20-Hydroxyecdysone[c]	
	Prediction	Tests	Prediction	Tests	Prediction	Tests
AR	G: -7.2		Y: -7.8		G: -2.3	
AR an	Y: -7.6	+[52, 53]	R: -9.4		G: -3.1	
ER α	G: -7.6		Y: -8.3	+?[50]	G: -4.9	
ER α an	G: -7.7	+[52, 54]	G: -8.2		G: -6.9	
ER β	G: -7.2		G: -7.9	+?[50]	G: -0.2	
ER β an	G: -7.4	+[52, 54]	G: -6.4		G: -6.0	
GR	Y: -7.6	-[55]	Y: -8.3		Y: -7.9	
GR an	G: -7.1		Y: -7.7		Y: -8.2	
LXR α	G: -8.5		G: -8.2		G: -9.2	
LXR β	G: -8.2		G: -9.6		G: -9.4	
MR	Y: -7.9		Y: -7.0		G: +1.3	
PPAR α	G: -7.4		G: -8.9		G: -4.0	
PPAR β	G: -7.5		G: -5.2		G: -8.6	
PPAR γ	G: -7.3		G: -6.4		G: -6.8	
PR	G: -2.5		Y: -2.9		G: -2.6	
RXR α	G: -7.5	+[56]	G: -8.6		G: -5.7	
TR α	Y: -8.4		G: -4.6		G: -1.2	
TR β	Y: -8.2		G: -5.5		G: -5.7	

Tests are experimental data from not only binding assays but also *in vitro* and *in vivo* assays, meaning they cannot be directly transferred to receptor modulation. To the best of our knowledge, no studies were performed for receptors with blank cells. The colors and letters correspond to the four binding probability classes: G, green; Y, yellow; O, orange (not present here); R, red. Numbers are scores of docking with AutoDock Vina. [a]SMILES: Clc1cc(Cl)ccc1Oc1ccc(Cl)cc1O. [b]SMILES: O=C1O[C@@]23C=C[C@H](O)[C@]1([C@H]2[C@H](C(=O)[O-])[C@@]12[C@H]3CC[C@@](O)(C(=C)C1)C2)C. [c]SMILES: CC(C)(O)CC[C@@H](O)[C@]©(O)[C@H]1CC[C@@]2(O)C3=CC(=O)[C@@H]4C[C@@H](O)[C@@H](O)C[C@]4©[C@H]3CC[C@]12C.

Implementing the Classification of the Docking Results. The validation results were also used to calculate thresholds. Three thresholds were set per structure to enable division into four probability binding classes. For the threshold calculations, SE was used to obtain the corresponding docking scores. It was decided that the true-positive rate or SE is a more important parameter than SP or the true-negative rate for the division of the classes. Thereby we ensured that the smallest possible number of potential endocrine-disrupting chemicals will be lost. Docking scores for each threshold were calculated to correspond to SEs of approximately 0.25, 0.50, and 0.75 (Table S1 in the Supporting Information). The thresholds were translated into four color-coded probability binding classes: class red (SE < 0.25) with a high probability of binding; two intermediate classes, orange (0.25 < SE < 0.50) and yellow (0.50 < SE < 0.75); and class green (SE > 0.75) with a low probability of binding. Because of the poorer performance of the MR structure, instead of SEs we used SPs of approximately 0.95, 0.90, and 0.65 in this case. These numbers correspond approximately to SP values of other nuclear receptors at SEs of 0.25, 0.50 and 0.75, respectively.

In addition to calculating SE and SP for each threshold, we also calculated the PPV and NPV[38,39] for each threshold (Table S1 in the Supporting Information). PPV is the proportion of positive test results that are true positives, and NPV is the proportion of correctly predicted negative results.[38] Because Endocrine Disruptome is intended to be used for screening, NPV is more important, that is, for the toxicologist's screening purposes it is more important that inactive molecules are predicted correctly. Endocrine Disruptome has very high NPVs for all of the thresholds and is therefore very good at excluding the nonbinding molecules. All of these validation data are available and integrated in Endocrine Disruptome, making it easier for users to understand the performance of a given model and thus to interpret the results.

Web Interface: Features and Overview. *DoTS Features.* The Docking interface for Target Systems (DoTS) is the underlying software on which Endocrine Disruptome runs. It relies heavily on established open-source technologies and modern Web standards. It enables administrators to set up an easily accessible Web platform for simultaneous docking to many different targets. It needs data from ROC curves as well as SE and SP and thus forces users to perform a validation to make sure that only validated models are included. Once set up and with targets added, DoTS enables users with a basic knowledge of computational chemistry to make docking predictions on several targets at the same time. The results are presented in a convenient color-coded table, making it easy for users to understand and make an informed decision regarding their own compound.

Interpretation of Results. Endocrine Disruptome can be a powerful tool when used with caution and intelligence. It should be noted that Endocrine Disruptome does not consider any of the pharmacokinetic parameters, including bioaccumulation and metabolism; it makes a prediction only for the structure of interest. It should also be noted that validation was

performed with fairly potent ligands. All of our ligands had IC_{50} or EC_{50} values less than 1 μM. The prediction results should therefore be interpreted with this in mind. Weak ligands with much higher binding constants would not be classified in the red class but can still be problematic, especially in chronic exposure. In addition to the docking itself, DoTS calculates basic physicochemical properties of the input compounds and attempts to alert users if PAINS[27] are detected. In recent years, great progress in the identification of PAINS and other types of unwanted compounds has been made.[27,29,40] This type of compound is particularly problematic in target- and cell-based assays, since they can give false-positive results. Lessons from medicinal chemistry can also be applied in the field of toxicology. This is why we implemented PAINS filters in DoTS to alert the user to the potentially problematic character of a compound. In this way toxicologists are stimulated to be especially careful when evaluating compounds with the PAINS scaffold.

Examples. Four structures were submitted for prediction to Endocrine Disruptome: the two industrial chemicals bisphenol A (BPA) and triclosan, the plant hormone gibberellic acid, and the insect hormone 20-hydroxyecdysone. BPA is a well-known endocrine disruptor (Figure 3). Its calculated physicochemical properties are not unusual and are well within druglike[41] chemical space. PAINS[27] groups were also not detected. As for docking predictions, the platform assigns "orange" for the binding probability to AR, AR an ("an" stands for the antagonist form), and MR. Class "yellow" is assigned for ER α, ER α an, ER β, ER β an, GR, GR an, TR α, and TR β. The rest are classified as "green". These results accord well with what is known about BPA. In *in vitro* assays, BPA exhibits antiandrogenic,[42] antimineralocorticoid[43] and antithyroid activities.[42] However, it is an agonist of GR[44,45] and a relatively weak estrogen agonist, with higher affinity for ER β than for ER α.[42,46] In some cell lines it exhibits mixed antagonist and agonist activities on ER α.[46] BPA alters the expression of PR,[47] PPAR α,[48] PPAR β,[48] and RXR α.[49] No binding or functional assay results could be found for class "green" receptors.

The results of dockings for the other three compounds are summarized in Table 3. The docking predictions for triclosan are less in concord with experimental data as they are in the case of BPA. Endocrine Disruptome correctly predicts binding to antagonist form of AR but it does not predict binding to antagonist forms of ER α and β, and RXR α. It also incorrectly predicts binding to GR. For other receptors and their forms experimental data is not available. In the case of gibberellic acid, Endocrine Disruptome predicts binding to agonist form of ER α but it does not predict binding to ER β. Test in which ER activity was determined does not distinguish if effects are result of binding on ER α and β.[50] For other receptors and their forms experimental data is not available. It would be interesting to see *in vitro* results on AR especially since *in vivo* studies suggest androgenic activity.[51] 20-Hydroxyecdysone seems not to have any significant endocrine disrupting activity according to our docking studies, except for weak GR modulation. Unfortunately, no experimental results for this compound have been published, and the correctness of this prediction remains to be seen.

CONCLUSION

EDCs interfere with the endocrine system in humans and other animals. The focus of current study is human nuclear receptors. From the environmental perspective, receptors from other vulnerable species such as reptiles or amphibians would be also interesting and will be evaluated for the next version. Over 200 docking experiments on 103 crystal structures were performed to identify the best-performing structures of 14 human nuclear receptors, and 18 structures with the best validation results were identified and selected. These structures are integrated into Endocrine Disruptome (http://endocrinedisruptome.ki.si), a user-friendly open-source Web-based prediction tool. Hopefully toxicologists and other experts will find it useful to help them make informed decisions for further *in vitro* and *in vivo* testing. In addition, we have programmed the Docking interface for Target Systems (DoTS) open-source software, whose source code can be found on GitHub (https://github.com/katrakolsek/DoTS). This software can be used with other validated crystal structures to predict binding to other systems and has potential in drug design. We anticipate that Endocrine Disruptome and DoTS will prove increasingly useful, including in their potential applications in drug design and molecular pharmacology.

ASSOCIATED CONTENT

Supporting Information

Structure input page (Figure S1); receptor overview page (Figure S2); clustering with RDKit (code snippet S1); functions for calculating ROC, EF1%, and AUC (code snippet S2); script for calculating score, PPV, and NPV at a given Se or Sp (code snippet S3); statistical parameters for each threshold (Table S1); and resolution and ligand data for all of the crystal structures (Table S2). This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: turk@bio.mx.

Present Address

‖S.T.: BioMed X Innovation Center, Im Neuenheimer Feld 583, 69120 Heidelberg, Germany. Telephone: +49 6221 65 456 14.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This study was supported by a Young Researcher Grant awarded to K.K. by the Slovenian Research Agency. Financial support from the Slovenian Research Agency is gratefully acknowledged.

ABBREVIATIONS

DUD-E, directory of useful decoys, enhanced; ROC, receiver operator characteristic curve; AUC, area under the curve; EF1%, enrichment factor at 1%; SE, sensitivity; SP, specificity; PPV, positive predictive value; NPV, negative predictive value; QSAR, quantitative structure−activity relationship; AR, androgen receptor; ER α, estrogen receptor α; ER β, estrogen receptor β; GR, glucocorticoid receptor; LXR α, liver X receptor α; LXR β, liver X receptor β; MR, mineralocorticoid receptor; PPAR α, peroxisome proliferator activated receptor α; PPAR β, peroxisome proliferator activated receptor β/δ; PPAR γ, peroxisome proliferator activated receptor γ; PR, progesterone receptor; RXR α, retinoid X receptor α; TR α, thyroid receptor α; TR β, thyroid receptor β; an, antagonist; ACDB,

active database; AGDB, agonist database; ANDB, antagonist database; r, reduced.

REFERENCES

(1) Burch, R. L.; Russell, W. M. S. The Principles of Humane Experimental Technique. http://altweb.jhsph.edu/pubs/books/humane_exp/het-toc (accessed Feb 12, 2014).

(2) REACH—Registration, Evaluation, Authorisation and Restriction of Chemicals. http://ec.europa.eu/enterprise/sectors/chemicals/reach/index_en.htm (accessed June 13, 2013).

(3) Eisenbrand, G.; Pool-Zobel, B.; Baker, V.; Balls, M.; Blaauboer, B. J.; Boobis, A.; Carere, A.; Kevekordes, S.; Lhuguenot, J. C.; Pieters, R.; Kleiner, J. Methods of in vitro toxicology. *Food. Chem. Toxicol.* **2002**, *40*, 193−236.

(4) Raunio, H. In silico toxicology—Non-testing methods. *Front. Pharmacol.* **2011**, *2*, No. 33.

(5) Vedani, A.; Smiesko, M. In silico toxicology in drug discovery—Concepts based on three-dimensional models. *ATLA, Altern. Lab. Anim.* **2009**, *37*, 477−96.

(6) Muster, W.; Breidenbach, A.; Fischer, H.; Kirchner, S.; Muller, L.; Pahler, A. Computational toxicology in drug development. *Drug Discovery Today* **2008**, *13*, 303−10.

(7) Endocrine Disruptor Screening Program for the 21st Century. http://www.epa.gov/endo/pubs/edsp21_work_plan_summary%20_overview_final.pdf (accessed Feb 4, 2014).

(8) Devillers, J.; Marchand-Geneste, N.; Carpy, A.; Porcher, J. M. SAR and QSAR modeling of endocrine disruptors. *SAR QSAR Environ. Res.* **2006**, *17*, 393−412.

(9) Cherkasov, A.; Muratov, E. N.; Fourches, D.; Varnek, A.; Baskin, II; Cronin, M.; Dearden, J.; Gramatica, P.; Martin, Y. C.; Todeschini, R.; Consonni, V.; Kuz'min, V. E.; Cramer, R.; Benigni, R.; Yang, C.; Rathman, J.; Terfloth, L.; Gasteiger, J.; Richard, A.; Tropsha, A. QSAR Modeling: Where Have You Been? Where Are You Going To? *J. Med. Chem.* **2014**, DOI: 10.1021/jm4004285.

(10) Vedani, A.; Dobler, M.; Smiesko, M. VirtualToxLab—A platform for estimating the toxic potential of drugs, chemicals and natural products. *Toxicol. Appl. Pharmacol.* **2012**, *261*, 142−53.

(11) Jacobs, M. N. In silico tools to aid risk assessment of endocrine disrupting chemicals. *Toxicology* **2004**, *205*, 43−53.

(12) Grün, F.; Blumberg, B. Environmental obesogens: Organotins and endocrine disruption via nuclear receptor signaling. *Endocrinology* **2006**, *147*, s50−5.

(13) Diamanti-Kandarakis, E.; Bourguignon, J. P.; Giudice, L. C.; Hauser, R.; Prins, G. S.; Soto, A. M.; Zoeller, R. T.; Gore, A. C. Endocrine-disrupting chemicals: An Endocrine Society scientific statement. *Endocr. Rev.* **2009**, *30*, 293−342.

(14) Mysinger, M. M.; Carchia, M.; Irwin, J. J.; Shoichet, B. K. Directory of useful decoys, enhanced (DUD-E): Better ligands and decoys for better benchmarking. *J. Med. Chem.* **2012**, *55*, 6582−94.

(15) Gaulton, A.; Bellis, L. J.; Bento, A. P.; Chambers, J.; Davies, M.; Hersey, A.; Light, Y.; McGlinchey, S.; Michalovich, D.; Al-Lazikani, B.; Overington, J. P. ChEMBL: A large-scale bioactivity database for drug discovery. *Nucleic Acids Res.* **2012**, *40*, D1100−7.

(16) The Open Babel Package, version 2.3.2. http://openbabel.org (accessed April 2, 2013).

(17) O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R. Open Babel: An open chemical toolbox. *J. Cheminform.* **2011**, *3*, No. 33.

(18) Bemis, G. W.; Murcko, M. A. The properties of known drugs. 1. Molecular frameworks. *J. Med. Chem.* **1996**, *39*, 2887−93.

(19) RDKit: Open-source cheminformatics. http://www.rdkit.org (accessed April 10, 2013).

(20) Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J. AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.* **2009**, *30*, 2785−91.

(21) Sanner, M. F. Python: A programming language for software integration and development. *J. Mol. Graphics Modell.* **1999**, *17*, 57−61.

(22) Trott, O.; Olson, A. J. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455−61.

(23) Li, H.; Zhang, H.; Zheng, M.; Luo, J.; Kang, L.; Liu, X.; Wang, X.; Jiang, H. An effective docking strategy for virtual screening based on multi-objective optimization algorithm. *BMC Bioinf.* **2009**, *10*, No. 58.

(24) Triballeau, N.; Acher, F.; Brabet, I.; Pin, J. P.; Bertrand, H. O. Virtual screening workflow development guided by the "receiver operating characteristic" curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4. *J. Med. Chem.* **2005**, *48*, 2534−47.

(25) Django: The Web framework for perfectionists with deadlines. https://www.djangoproject.com/ (accessed Jan 10, 2013).

(26) O'Boyle, N. M.; Morley, C.; Hutchison, G. R. Pybel: A Python wrapper for the OpenBabel cheminformatics toolkit. *Chem. Cent. J.* **2008**, *2*, No. 5.

(27) Baell, J. B.; Holloway, G. A. New substructure filters for removal of pan assay interference compounds (PAINS) from screening libraries and for their exclusion in bioassays. *J. Med. Chem.* **2010**, *53*, 2719−40.

(28) *SMARTS Theory Manual*; Daylight Chemical Information Systems: Santa Fe, NM; http://www.daylight.com/dayhtml/doc/theory/theory.smarts.html (accessed Feb 12, 2014).

(29) Saubern, S.; Guha, R.; Baell, J. B. KNIME Workflow to Assess PAINS Filters in SMARTS Format. Comparison of RDKit and Indigo Cheminformatics Libraries. *Mol. Inf.* **2011**, *30*, 847−50.

(30) Celery: Distributed Task Queue. http://www.celeryproject.org/ (accessed June 20, 2013).

(31) ChemDoodle Web Components. http://web.chemdoodle.com/ (accessed June 24, 2013).

(32) Flot: Attractive JavaScript plotting for jQuery. http://www.flotcharts.org/ (accessed May 21, 2013).

(33) *SMILES Theory Manual*; Daylight Chemical Information Systems: Santa Fe, NM; http://www.daylight.com/dayhtml/doc/theory/theory.smiles.html (accessed Feb 12, 2014).

(34) Swets, J. A. Measuring the accuracy of diagnostic systems. *Science* **1988**, *240*, 1285−93.

(35) Klebe, G. Virtual ligand screening: Strategies, perspectives and limitations. *Drug Discovery Today* **2006**, *11*, 580−94.

(36) Germain, P.; Staels, B.; Dacquet, C.; Spedding, M.; Laudet, V. Overview of nomenclature of nuclear receptors. *Pharmacol. Rev.* **2006**, *58*, 685−704.

(37) Mangelsdorf, D. J.; Thummel, C.; Beato, M.; Herrlich, P.; Schutz, G.; Umesono, K.; Blumberg, B.; Kastner, P.; Mark, M.; Chambon, P.; Evans, R. M. The nuclear receptor superfamily: The second decade. *Cell* **1995**, *83*, 835−9.

(38) Altman, D. G.; Bland, J. M. Diagnostic tests 2: Predictive values. *BMJ* **1994**, *309*, 102.

(39) Parikh, R.; Mathai, A.; Parikh, S.; Chandra Sekhar, G.; Thomas, R. Understanding and using sensitivity, specificity and predictive values. *Indian J. Ophthalmol.* **2008**, *56*, 45−50.

(40) Bruns, R. F.; Watson, I. A. Rules for identifying potentially reactive or promiscuous compounds. *J. Med. Chem.* **2012**, *55*, 9763−72.

(41) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Delivery Rev.* **2001**, *46*, 3−26.

(42) Rubin, B. S. Bisphenol A: An endocrine disruptor with widespread exposure and multiple effects. *J. Steroid Biochem.* **2011**, *127*, 27−34.

(43) Odermatt, A.; Gumy, C. Disruption of glucocorticoid and mineralocorticoid receptor-mediated responses by environmental chemicals. *Chimia* **2008**, *62*, 335−9.

(44) Prasanth, G. K.; Divya, L. M.; Sadasivan, C. Bisphenol-A can bind to human glucocorticoid receptor as an agonist: An in silico study. *J. Appl. Toxicol.* **2010**, *30*, 769−74.

(45) Sargis, R. M.; Johnson, D. N.; Choudhury, R. A.; Brady, M. J. Environmental endocrine disruptors promote adipogenesis in the 3T3-

L1 cell line through glucocorticoid receptor activation. *Obesity* **2010**, *18*, 1283−8.
(46) Vandenberg, L. N.; Maffini, M. V.; Sonnenschein, C.; Rubin, B. S.; Soto, A. M. Bisphenol-A and the great divide: A review of controversies in the field of endocrine disruption. *Endocr. Rev.* **2009**, *30*, 75−95.
(47) Lee, H. R.; Jeung, E. B.; Cho, M. H.; Kim, T. H.; Leung, P. C.; Choi, K. C. Molecular mechanism(s) of endocrine-disrupting chemicals and their potent oestrogenicity in diverse cells and tissues that express oestrogen receptors. *J. Cell. Mol. Med.* **2013**, *17*, 1−11.
(48) Grasselli, E.; Cortese, K.; Voci, A.; Vergani, L.; Fabbri, R.; Barmo, C.; Gallo, G.; Canesi, L. Direct effects of Bisphenol A on lipid homeostasis in rat hepatoma cells. *Chemosphere* **2013**, *91*, 1123−9.
(49) Nishizawa, H.; Manabe, N.; Morita, M.; Sugimoto, M.; Imanishi, S.; Miyamoto, H. Effects of in utero exposure to bisphenol A on expression of RARα and RXRα mRNAs in murine embryos. *J. Reprod. Dev.* **2003**, *49*, 539−45.
(50) Gawienowski, A. M.; Stadnicki, S. S.; Stacewicz-Sapuntzakis, M. Synergistic uterotrophic effect of gibberellic acid and estradiol in the immature mouse. *Life Sci.* **1977**, *20*, 785−8.
(51) Gawienowski, A. M.; Stadnicki, S. S.; Stacewicz-Sapuntzakis, M. Androgenic properties of gibberellic acid in the chick comb bioassay. *Experientia* **1977**, *33*, 1544−5.
(52) Ahn, K. C.; Zhao, B.; Chen, J.; Cherednichenko, G.; Sanmarti, E.; Denison, M. S.; Lasley, B.; Pessah, I. N.; Kultz, D.; Chang, D. P. Y.; Gee, S. J.; Hammock, B. D. In vitro biologic activities of the antimicrobials triclocarban, its analogs, and triclosan in bioassay screens: Receptor-based bioassay screens. *Environ. Health Persp.* **2008**, *116*, 1203−10.
(53) Chen, J.; Ahn, K. C.; Gee, N. A.; Gee, S. J.; Hammock, B. D.; Lasley, B. L. Antiandrogenic properties of parabens and other phenolic containing small molecules in personal care products. *Toxicol. Appl. Pharmacol.* **2007**, *221*, 278−84.
(54) Gee, R. H.; Charles, A.; Taylor, N.; Darbre, P. D. Oestrogenic and androgenic activity of triclosan in breast cancer cells. *J. Appl. Toxicol.* **2008**, *28*, 78−91.
(55) Christen, V.; Crettaz, P.; Oberli-Schrammli, A.; Fent, K. Some flame retardants and the antimicrobials triclosan and triclocarban enhance the androgenic activity in vitro. *Chemosphere* **2010**, *81*, 1245−52.
(56) PUBCHEM_BIOASSAY: qHTS assay for small molecule antagonists of retinoid X receptor alpha signaling. https://www.ebi.ac.uk/chembl/assay/inspect/ChEMBL1794471 (accessed Jan 29, 2014).