Toxicology and Applied Pharmacology 262 (2012) 301–309

Contents lists available at SciVerse ScienceDirect

Toxicology and Applied Pharmacology

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

QSAR model for human pregnane X receptor (PXR) binding: Screening of environmental chemicals and correlations with genotoxicity, endocrine disruption and teratogenicity

Marianne Dybdahl *, Nikolai G. Nikolov, Eva Bay Wedebye, Svava Ósk Jónsdóttir, Jay R. Niemelä

Department of Toxicology and Risk Assessment, National Food Institute, Technical University of Denmark, Mørkhøj Bygade 19, DK-2860 Søborg, Denmark

ARTICLE INFO

Article history:
Received 4 January 2012
Revised 8 May 2012
Accepted 13 May 2012
Available online 22 May 2012

Keywords:
QSAR
Pregnane X receptor
Teratogenicity
Genotoxicity
Endocrine disruption

ABSTRACT

The pregnane X receptor (PXR) has a key role in regulating the metabolism and transport of structurally diverse endogenous and exogenous compounds. Activation of PXR has the potential to initiate adverse effects, causing drug–drug interactions, and perturbing normal physiological functions. Therefore, identification of PXR ligands would be valuable information for pharmaceutical and toxicological research. In the present study, we developed a quantitative structure–activity relationship (QSAR) model for the identification of PXR ligands using data based on a human PXR binding assay. A total of 631 molecules, representing a variety of chemical structures, constituted the training set of the model. Cross-validation of the model showed a sensitivity of 82%, a specificity of 85%, and a concordance of 84%. The developed model provided knowledge about molecular descriptors that may influence the binding of molecules to PXR. The model was used to screen a large inventory of environmental chemicals, of which 47% was found to be within domain of the model. Approximately 35% of the chemicals within domain were predicted to be PXR ligands. The predicted PXR ligands were found to be overrepresented among chemicals predicted to cause adverse effects, such as genotoxicity, teratogenicity, estrogen receptor activation and androgen receptor antagonism compared to chemicals not causing these effects. The developed model may be useful as a tool for predicting potential PXR ligands and for providing mechanistic information of toxic effects of chemicals.

Introduction

Nuclear receptors constitute a large superfamily of ligand-dependent transcription factors that control gene expression and play a key role in the development, homeostasis and metabolism of living organisms. The pregnane X receptor (PXR) is a member of this superfamily and regulates enzymes and transporters involved in xenobiotic detoxification as well as maintaining homeostatic balance of bile acids, thyroid and steroid hormones (di Masi et al., 2009). PXR is primarily expressed in the liver, intestine and kidney, but its expression is also seen in lung, stomach, the blood–brain barrier, placenta, bone marrow, and specific regions of the brain.

PXR is activated by a structurally diverse array of endogenous and exogenous compounds, including steroids, bile acids, antibiotics, statins, anticancer compounds, PPAR antagonists, and environmental contaminants, such as pesticides and plasticizers (Chen and Nie, 2009; Kretschmer and Baldwin, 2005). Upon interaction with these ligands, PXR forms a heterodimer with the retinoid X receptor in the nucleus and binds to a xenobiotic response element located in the promoter region of the target gene, thereby regulating its transcription. PXR target genes include genes encoding cytochrome P450 (CYP) enzymes, such as CYP3A4 and CYP2B6, conjugation enzymes, such as UDP-glucuronosyltransferases and sulfotransferases, and transporters, such as P-glycoprotein and multidrug resistance-associated proteins (Tolson and Wang, 2010). Because the induced proteins are not only involved in the metabolism and transport of environmental chemicals and drugs, but also of bile acid, thyroid, and steroid hormones, xenobiotics may interfere with normal physiological functions. Moreover, recent studies have provided evidence that PXR activation is involved in lipid metabolism, glucose homeostasis, and inflammation (Moreau et al., 2008; Zhou et al., 2009). The wide array of biological activities of PXR agonists might be of clinical relevance, and their potential in treatment of liver disorders was reviewed recently (Fiorucci et al., 2012).

Several crystal structures of human PXR ligand binding domain (LBD) complexed with xenobiotic ligands reveal that the PXR binding site is unusually large and flexible (Watkins et al., 2001). It can change its shape to accommodate molecules that vary in size from

Abbreviations: AR, androgen receptor; ER, estrogen receptor; EDC, endocrine disrupting chemical; EINECS, European Inventory of Existing Commercial Chemical Substances; HGPRT, hypoxanthine-guanine phosphoribosyl transferase; LBD, ligand binding domain; PLR, partial logistic regression; PLS, partial least squares; SCE, sister chromatid exchange.

* Corresponding author. Fax: +45 3588 7699.
E-mail address: mdyb@food.dtu.dk (M. Dybdahl).

0041-008X/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.taap.2012.05.008

300 Da (such as carbamazepine and clotrimazole) to over 800 Da (rifampicin) in molecular weight. The LBD contains both a ligand binding pocket and a ligand-dependent AF-2 region, which binds to transcriptional cofactors. It has been suggested that different ligands lead to the recruitment of different coactivators and the activation of different genes (Ngan et al., 2009). Up to 28 amino acid residues line the binding site of PXR, although no more than 19 residues are in contact with the ligands in any of the X-ray structures available to date (Ngan et al., 2009). Notably, 20 of the 28 residues are hydrophobic, and the few polar and charged residues, spaced throughout the hydrophobic pocket, may permit a ligand to bind in multiple orientations. Human pharmacophore models have shown that PXR agonists are required to fit to multiple hydrophobic features and at least one hydrogen bond acceptor (Ekins and Erickson, 2002; Ekins et al., 2007; Schuster and Langer, 2005). Very few PXR antagonists have been identified. Most of the known PXR antagonists require a balance between hydrophobic and hydrogen bonding features and are interacting at the AF-2 site, while agonists are bound in the ligand binding pocket (Ekins et al., 2007).

A considerable amount of inter-species variation has been observed in the PXR LBD with human, rabbit and rat sharing only roughly 75% amino acid sequence identity (Jones et al., 2000). There are numerous examples of differences in ligand binding of PXR. For example, SR12813 (a bisphosphonate ester used to lower serum cholesterol levels) is a potent activator of human and rabbit PXR but a weak activator of rat PXR. Thus, it is difficult to extrapolate results obtained with rodent PXR to their human counterparts. In contrast to the significant inter-species variation, very low variation in LBD sequence (lower than the whole genome average for human genes) is seen between human individuals (Zhang et al., 2001). This suggests limited variation in PXR ligands between humans.

Because of its vital role in drug metabolism, it is not surprising that PXR has been found responsible for decreased drug efficacy and increased drug toxicity. The promiscuous nature of PXR ligand binding is a contributing factor in drug–drug interactions due to its capability of binding a variety of structurally diverse molecules that induce CYP3A4. CYP3A4 is considered the main drug-metabolizing CYP in human liver. It accounts for up to 40% of total CYPs in this organ and is responsible for the metabolism of more than 60% of drugs on the market today (di Masi et al., 2009). Bioactivation through CYP3A4 is often suspected as an initiating event in chemical toxicity. The induction of metabolizing enzymes and transporters by PXR has also been regarded as one of the major mechanisms of drug resistance (Chen and Nie, 2009). Activation of PXR may accelerate the metabolism and elimination of chemotherapeutic agents, contributing to resistance to chemotherapy. Given the clinical liabilities associated with developing PXR activating compounds into drugs, it has become one of the primary objectives for drug-discovery programs to attenuate PXR activity (Gao et al., 2007).

Activation of PXR may also lead to disturbance of normal physiological systems, such as the steroid and thyroid hormone systems. Inappropriate activation or antagonism of the estrogen and androgen receptors is the most extensively studied mechanisms for endocrine disruption. However, increasing or decreasing endogenous hormone metabolism could contribute to the detrimental effects of endocrine disrupting chemicals (EDCs), and therefore the potential role of a PXR-mediated mechanism has received increased attention (Kojima et al., 2010; Tabb and Blumberg, 2006). Several EDCs are known to activate PXR, including structurally diverse organochlorine pesticides, polychlorinated biphenyls, plasticizers, fungicides, herbicides, and pharmaceutical compounds, and many of these compounds have previously been associated with developmental toxicity, estrogenic, and/or antiandrogenic effects (Kretschmer and Baldwin, 2005).

Because activation of PXR has the potential to initiate a broad spectrum of adverse effects, identification of PXR agonists would be important information for evaluating health risk of environmental chemicals and drugs. Experimental determination of toxicity profiles requires resources both in terms of cost and time. Therefore reliable in silico alternatives such as quantitative structure–activity relationship (QSAR) models are becoming important tools for rapid and cost-effective prediction of biological activities. Such models may have a great potential for use in the identification of large numbers of potential PXR agonists. To date, the lack of large sets of PXR data has restricted ligand-based computational models to a relatively small universe of molecules (Ekins et al., 2009; Khandelwal et al., 2008; Pan et al., 2011; Ung et al., 2007).

In this study, we developed a QSAR model for the prediction of PXR ligands using data based on a human PXR LBD binding assay (Shukla et al., 2009). The model was built on a training set of 631 molecules (299 positives and 332 negatives). The model was used to screen a library of 51,680 environmental chemicals, for which very limited toxicity testing has been performed. These include chemicals present in our food, environment, and consumer products, e.g. food additives, plasticizers and flame retardants. Such a screening may give an estimate of the number of PXR ligands among environmental chemicals as well as give information about the applicability domain of the PXR model. The PXR predictions were correlated with predictions from other QSAR models including genotoxicity, teratogenicity, estrogen receptor (ER) activation and androgen receptor (AR) antagonism. This approach can be used to elucidate the role of PXR in toxicity and to increase the mechanistic understanding of toxic effects.

Materials and methods

Data set for training and validation of the model. The training set consisted of experimental data for PXR binding kindly provided by Dr. Sunita Shukla (NIH Chemical Genomics Center, National Institutes of Health, Bethesda, Maryland). The data was obtained using a time-resolved fluorescence resonance energy transfer (TR-FRET) assay, which measured human PXR binding. Briefly, the TR-FRET assay reports on the ability of a test ligand to displace a fluorescein-labeled tracer molecule from the nuclear receptor. The assay performance and the data analysis are described in details elsewhere (Shukla et al., 2009). The PXR binding data included compounds from the Sigma-Aldrich LOPAC collection, Tocris/TimTec, bioactive collection, and Biomol.

The classification of active and inactive compounds was as described by Shukla et al. (2009). All compounds defined as “high confidence” actives and an equal number of randomly selected inactive compounds were used for preparation of the training set. The data were checked for structural replicates. In case of duplicate structures, one of the replicates was kept if the compounds had the same activity and both were removed if they had different activity. The data were also searched for salts, and the structures were used in their non-ionized form in the training set. The resulting training set consisted of 631 molecules; 299 actives and 332 inactives.

Two data sets were used for external validation of the predictive model. The first test set (test set 1) consisted of human PXR activation data ($n = 145$) determined by a reporter gene assay (Khandelwal et al., 2008). After removing molecules also present in the training set and two molecules known to be PXR antagonists (fluconazole and ketonazole), the test set consisted of 120 chemicals (68 actives and 52 inactives). The second test set (test set 2) consisted of the remaining 3351 inactive compounds with respect to PXR binding, which were not used in the balanced training set (Shukla et al., 2009).

Modeling methodology. The modeling system Leadscope® Predictive Data Miner (Version 3.04-10, Leadscope, Inc., http://leadscope.com) was used. The molecular structures in the training set were imported into Leadscope and classified by structure into categories using a library of approximately 27,000 structural features (Roberts et al., 2000). The structural features are substructures such as functional

groups, heterocycles and pharmacophores. Eight molecular descriptors, calculated octanol water partition coefficient (logP), hydrogen bond acceptors, hydrogen bond donors, Lipinski score, atom count, molecular weight, polar surface area and rotatable bonds, were calculated for each structure as well. A predictive model was developed based on a sub-set of the structural features and the molecular descriptors, using partial logistic regression (PLR). The sub-set of features was identified with the automatic feature selection option in Leadscope that selects the 30% best correlated features according to a Yates χ^2-test. Another model was built using the scaffold generation function in Leadscope and performances of the two models were evaluated.

Applicability domain. The applicability domain was defined as described in the following, and only predictions within this domain were accepted. Class probabilities (p) for positive predictions were used for defining the domain, requiring $p \geq 0.7$ for actives and $p \leq 0.3$ for inactives. In addition, Leadscope uses structural domain, which requires that a compound has at least 30% Tanimoto structural similarity with a training set compound and contain at least one structural feature from the model, to be considered in domain.

Screening of a large inventory of chemicals for PXR binding. The developed PXR model was used to screen a list of 51,680 chemicals from the EINECS list (European Inventory of Existing Commercial Chemical Substances). The EINECS list contains discrete organic chemicals, inorganic compounds and other substances that were registered for use on the European market between 1971 and 1981. As our modeling systems can only handle discrete organic compounds, a sub-set of molecules with a defined 2D structure, containing at least two carbon atoms and only certain atoms (H, Li, B, C, N, O, F, Na, Mg, Si, P, S, Cl, K, Ca, Br, and I) was selected as prediction set.

Models used for correlation with predicted PXR binders. The resulting screening results were imported into our in-house database tool, built on Oasis Database Manager (Nikolov et al., 2006). The in-house database presently contains predictions from more than 150 QSAR models for up to185,000 organic compounds and can be used to explore correlations between different model predictions. The predicted PXR ligands were correlated with predictions from QSAR models covering various endpoints such as genotoxicity, teratogenicity, ER activation and AR antagonism. Two models for in vitro genotoxicity, the Ames test and the HGPRT mammalian cell gene mutation assay, and one model for in vivo genotoxicity, sister chromatid exchange (SCE) were used. More detailed information on the QSAR models e.g. endpoint description, training set and validation is given in the Supplemental material. The models for the Ames test and the HGPRT assay have very high model performances, matching the best experimental reproducibility among in vitro genotoxicity tests. The other models generally have high specificities, and somewhat lower sensitivities, which make the probability of making false positive predictions lower than making false negative predictions.

Statistical analyses. The performance of the QSAR model was evaluated using Cooper statistics, which describes sensitivity (ability to correctly predict actives), specificity (ability to correctly predict inactives) and concordance (overall accuracy) (Cooper et al., 1979). These measures are defined below, where TP, TN, FP and FN are true positives, true negatives, false positives and false negatives, respectively.

$$Sensitivity = TP/(TP + FN)$$
$$Specificity = TN/(TN + FP)$$
$$Concordance = (TP + TN)/(TP + TN + FP + FN)$$

Matthews's correlation coefficient and chi-squared statistics were calculated to estimate the significance of correlations between different model predictions.

Table 1
Performance of the developed PXR model. The predictive performance was evaluated by 2 × 50% cross-validation and by external validation using test sets 1 and 2.

	In domain (%)	Sensitivity (%)	Specificity (%)	Concordance (%)
Training set (n = 631)	–	82.3	84.6	83.5
Test set 1 (n = 120)	57.5	57.9	83.9	69.6
Test set 2 (n = 3351)	73.9	–	82.1	–

Table 2
The mean values of molecular weight and logP of actives and inactives in the training set. The standard deviations are listed in parentheses.

	MW	Log P
Inactive (n = 332)	308 (163)	1.9 (2.2)
Active (n = 299)	407 (124)	4.4 (1.6)

Results

Validation of the predictive PXR model

The Leadscope model without scaffold generation had the best performance and was used for prediction of PXR binding of the EINECS chemicals. The validation results of the predictive model are shown in Table 1. After leaving 50% of the compounds out by random twice for cross validation, the sensitivity was 82.3%, specificity 84.6%, and concordance 83.5%. Validation with the external test set of 120 compounds from a previous study (Khandelwal et al., 2008) (test set 1) revealed a specificity of 83.9% and a sensitivity of 57.9%. The lower sensitivity most likely reflects the fact that the external test set and the training set are from two different assay types (reporter gene vs. binding), and that the reporter gene assay also identifies compounds acting through other routes than direct ligand binding; i.e. our model which identifies ligand binders will not identify all the compounds which might activate PXR. The large 3351 compounds test set containing only PXR negatives (test set 2) resulted in a specificity of 82.1%.

Model parameters

The Leadscope predictive model was built using 387 structural features and 8 molecular descriptors. The number of partial least squares (PLS) factors was 3. Some of the most important descriptors in the model were logP and molecular weight. From Table 2 it can be seen that PXR binders are in general larger molecules and have higher logP compared to PXR non-binders.

To explore the structural features contributing to PXR activity, the Leadscope model with scaffold generation was used. Cross validation of the model (two times 50%) gave a sensitivity of 77.9%, a specificity of 84.3%, and a concordance of 81.3%. Fig. 1 presents some of the most influential features contributing to positive and negative PXR activity. It can be seen that all positive contributing substructures are essentially hydrophobic, while negatively contributing features appear to be more polar, i.e. contain hydroxyl or other polar groups. It should be noted that the contribution to PXR activity of a given substructure may be influenced by the total molecular structure.

Two examples of positive PXR ligands were investigated to illustrate some of the substructures linked to PXR binding. The ligands chosen were felodipine, a calcium channel blocker, and clotrimazole, an antifungal drug. These molecules are strong PXR binders and have been used in docking models for PXR (Xiao et al., 2011).

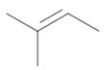
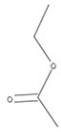
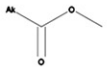
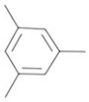
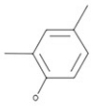
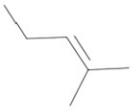
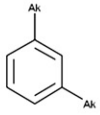
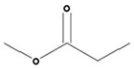
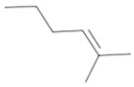
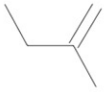
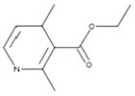
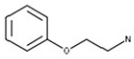
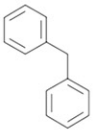
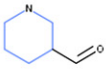
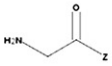
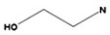
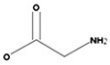
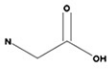
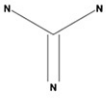
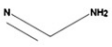
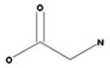
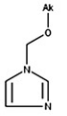
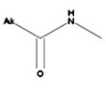
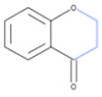
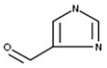
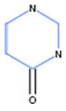
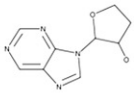
26/4  Scaffold 121	34/9  Scaffold 219	30/9  carboxylate, alkyl, acyc-	23/5  Scaffold 182	12/0  Scaffold 243
12/0  Scaffold 227	14/1  Scaffold 222	21/5  benzene, 1,3-dialkyl-	24/7  Scaffold 195	12/1  Scaffold 233
12/1  Scaffold 221	9/0  Scaffold 434	11/1  Scaffold 188	15/3  Scaffold 100	14/3  pyridine(H),3-carbonyl-
4/30  Alpha-amino acid (free amine)	4/28  Methane, 1-amino(NH2)-,1-carbonyl-	6/29  Ethane, 1-amino,2-hydroxy-	3/21  1-methane-carboxylic acid, 1-amino(NH2)-	8/31  Alpha-amino acid (free acid)
0/12  Guanidine	0/11  Amino(NH2)iminomethyl	0/11  Oxolane, 2-amino-	8/28  1-methane-carboxylic acid, 1-amino-	0/9  Imidazole,1-alkoxymethyl-
3/16  carboxamide(NHR), alkyl, cyc-	3/16  benzopyran, 4-oxo-	0/8  imidazole, 4-carbonyl-	0/8  1,3-diazine(H), 4-oxo-	0/7  Scaffold 117

Fig. 1. The most important features in the PXR model developed in Leadscope with scaffold generation. The features are selected by occurrence and chi-square test. The ratio of ligand/non-ligand for each feature is indicated in upper left corner. Z matches N, S or O.

Felodipine

Felodipine and two of the structural features describing the molecule are shown in Fig. 2 (scaffold numbers 47 and 434). It was seen that the training set chemicals containing scaffold number 434 were mainly drugs belonging to the dihydropyridine calcium channel blockers, namely felodipine, nimodipine, nicardipine, nitrendipine, isradipine and cilnidipine. All of the chemicals in the training set containing this feature, nine in total, were positive when tested in the PXR binding assay. Scaffold number 47 was present in 16 positive and 5 negative PXR ligands.

Clotrimazole

Clotrimazole and two of the structural features describing the molecule are shown in Fig. 3 (scaffold numbers 100 and 359). These

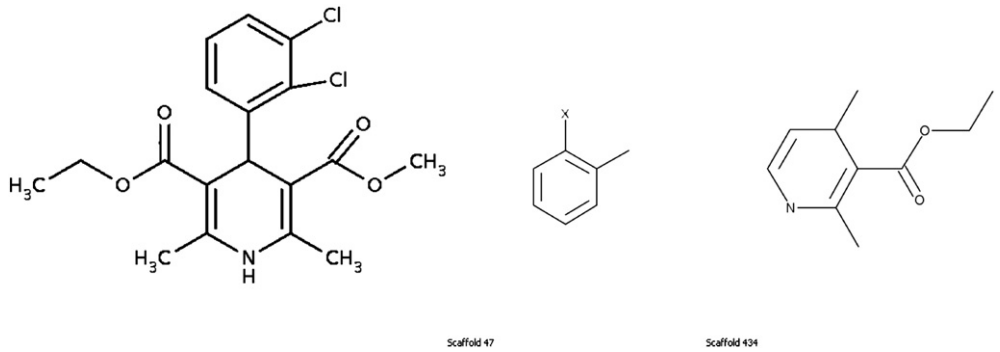


Fig. 2. Felodipine and two describing features.

features illustrate the requirement of PXR agonists to fit to multiple hydrophobic features and at least one hydrogen bond acceptor. The imidazole moiety may form a hydrogen bond with PXR and the aromatic rings can interact with aromatic residues in PXR. Of the chemicals containing scaffold number 100, 15 were positive and 3 were negative when tested in the PXR binding assay, while scaffold number 359 was present in 7 positive and 1 negative chemicals.

Predictions and applicability domain of the EINECS chemicals

A structure set of 51,680 discrete organic chemicals from the EINECS list was screened through the PXR model. As shown in Table 3, 47% of the chemicals were found to be within the domain of the model and of these 35% were predicted as PXR ligands, corresponding to 8516 chemicals.

Comparisons with other endpoints

The QSAR predictions for PXR binding were correlated with corresponding predictions from QSAR models covering various endpoints such as genotoxicity, teratogenicity, ER activation and AR antagonism for the same EINECS structure set. Comparing predicted PXR results with predictions for other endpoints may involve double uncertainty inherent in predictive models but can also be valuable due to the opportunity to draw conclusions from predictions of large chemical libraries. We found statistically significant evidence about correlations between PXR predictions and predictions of other endpoints; the magnitude of correlations between the modeled biological effects is however not estimated in this work. Using QSAR predictions for different properties for a large number of compounds, we were able to identify adverse effects that according to our model predictions were significantly more common among compounds that bind to PXR compared to compounds that do not bind to PXR. This analysis was used to elucidate trends in the predicted data, thereby identifying possible biological processes and pathways affected by PXR ligands. When performing the comparisons, only molecules included in the common domain of the relevant models were used, e.g. molecules with reliable model predictions for both properties being correlated.

The correlations between predictions of different endpoints are summarized in Tables 4 and 5. Matthews's correlation coefficient and chi-squared statistics were calculated to estimate the correlation significance and are summarized in Tables 6 and 7. The values of χ^2 show that all dependencies are statistically significant ($\chi^2 > 3.841$ and $p < 0.05$).

Three genotoxicity models based on in vitro or in vivo data were used for evaluation of a possible association with PXR binding. All genotoxicity models correlated positively with the PXR predictions (Matthew's correlation coefficient between 0.06 and 0.11) indicating that genotoxic effects are more common among compounds that bind to PXR compared to compounds that do not bind to PXR. Fig. 4 illustrates the correlations of predicted PXR binding with predicted genotoxicity.

The correlation of predicted PXR with predicted AR antagonism showed a marked increase in the incidence of predicted PXR positives among AR antagonists compared to non-antagonists, supported also by a high value of the corresponding χ^2 statistics. Predicted ER activators correlated only marginally to predicted PXR positives because of the relatively low number of predicted ER activators available in the common domain of the models, and the rather balanced split of predicted ER activators between the predicted PXR classes.

There was a marked increase in the percentage of predicted PXR ligands among teratogenic compounds compared to non-teratogenic compounds supported by the corresponding statistics. Fig. 5 shows graphically the correlations of predicted PXR binding with predicted endocrine disruption and teratogenicity.

It was also investigated as to what extent the predicted teratogenicity might be linked to AR antagonism and ER activation (Tables 5 and 7). Both ER activation and AR antagonism correlated positively with the predictions for teratogenicity. However, the comparison between predicted teratogenicity and AR antagonism resulted in only

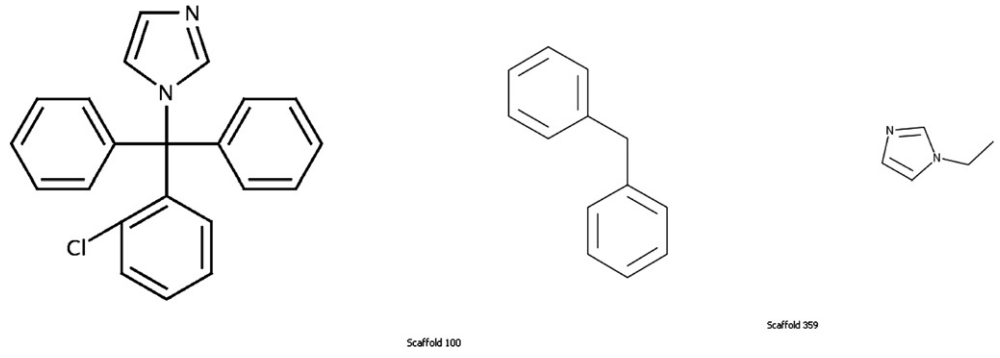


Fig. 3. Clotrimazole and two describing features.

Table 3
Domain of the PXR model within 51,680 EINECS chemicals and prediction for PXR binding.

		In domain	In domain and predicted PXR positive
EINECS chemicals	n	24,524	8516
	%	47	35

2% difference in the incidence of positives, and χ^2 of 9.1, so this dependence can be regarded as marginal.

The correlations of PXR binding with other endpoints in the in-house QSAR database were also performed using PXR predictions from a QSAR model built in the SciQSAR modeling system. The SciQSAR model was built on the same training set as the Leadscope model. Predictions from the two modeling systems resulted in the same trends, which support the validity of the correlation results (data not shown).

Examples of molecules which were positive for PXR binding and teratogenicity

Screening of the EINECS list identified 592 chemicals, which were positive for both PXR binding and teratogenicity. Examples of the chemical classes, which were identified by manual inspection of the chemicals, included steroids, genotoxic agents, carcinogens, estrogen receptor agonists, androgen receptor antagonists, tetracyclins, thyroid modulators, anticoagulants and retinoids.

Discussion

In the present study, we have developed a global QSAR model for the identification of PXR ligands. The Leadscope modeling system was used to construct the model, and data from an in vitro assay measuring human PXR binding were used to train the model. The training set included 631 molecules representing a variety of chemical compounds (Shukla et al., 2009). The model was used to screen a large inventory of environmental chemicals, and the predictions were correlated with predictions from a number of other QSAR model endpoints.

Model validation

The first part of this study was focused on the development and validation of a QSAR model for PXR binding including identification of descriptors important for binding to the receptor. Cross-validation of the model showed that a robust model was developed (sensitivity 82%, specificity 85%, and concordance 84%). The external validation set predominantly consisted of drugs, imidazole derivates similar to clotrimazole, steroids, molecules with different heterocyclic ring systems, as well as many other diverse molecules (Khandelwal et al., 2008). The PXR activity in the external test set was determined by a human cell-based reporter gene assay. The PXR binding assay identifies PXR activation by direct binding to the ligand binding domain, whereas the cell-based assay can identify compounds which activate either through direct ligand binding or through modulation of PXR activation by signaling pathways. It is important to note that comparisons between different assay formats may produce discrepant results and need careful interpretation. For example, active compounds from the PXR binding assay which do not reproduce in a cell-based assay could act as antagonists, may not enter the cell or may be degraded in the cellular environment. On the other hand, compounds which are active in the cell-based assay and negative in the binding assay may not have been able to displace the tracer molecule. In a recent study, compounds were profiled in a PXR binding assay and a cell-based PXR assay and the concordance rate (compounds with activity in both assays relative to compounds with activity in either assay) was found to be 71% (Shukla et al., 2011). The external validation of our model resulted in a sensitivity of 58%, a specificity of 84%, and a concordance of 70%. The lower sensitivity indicates that a number of compounds identified as positives in the cell-based PXR assay were not picked up in the binding assay. This may be expected assuming that the cell-based assay can identify compounds acting through other routes than direct ligand binding. In order to provide more detailed information on the mechanism of PXR activation, future studies would benefit from comparing the results from human cell-based PXR reporter systems to the results using the PXR binding assay.

Table 4
Correlations of predicted PXR ligands with predicted data for endpoints for genotoxic effects, endocrine disruption and teratogenicity. Only predictions in domain of both the PXR model and the respective toxicity model were included.

	Total (n)	PXR pos (n)	PXR pos (%)
Genotoxicity			
Ames pos	2686	1147	43
Ames neg	15,194	5273	35
HGPRT pos	2298	980	43
HGPRT neg	7065	2292	32
SCE pos	4305	1662	39
SCE neg	6558	1846	28
Endocrine disruption			
AR antagonism pos	1509	802	53
AR antagonism neg	14,098	4512	32
ER activation pos	400	166	42
ER activation neg	9464	3160	33
Reproductive toxicity			
Teratogenicity pos	1127	592	53
Teratogenicity neg	10,899	3366	31

Table 5
Correlations of predicted AR and ER activation with predicted teratogenicity. Only predictions in domain of both the teratogenicity model and the AR antagonism or ER activation model were included.

	Total (n)	AR pos (n)	AR pos (%)	Total (n)	ER pos (n)	ER pos (%)
Teratogenicity pos	1749	203	11.6	1099	96	8.7
Teratogenicity neg	16,050	1503	9.4	12,887	379	2.9

Descriptors

The large and promiscuous ligand binding pocket of PXR accepts molecules of widely varying sizes, and is likely to be capable of binding small molecules in multiple orientations. Human PXR agonist pharmacophore models have shown that agonists are required to fit to multiple hydrophobic features and at least one hydrogen bond acceptor (Ekins and Erickson, 2002; Ekins et al., 2007; Schuster and Langer, 2005). The Leadscope modeling system enables the identification of structural features and molecular descriptors of importance for PXR binding. Some of the most important descriptors in our PXR model were molecular weight and logP. The importance of molecular weight may indicate that there is a lower limit to the contraction of the PXR LBD that can occur to accommodate small ligands. In a recent

Table 6
Correlation coefficients and χ^2 statistics for PXR predictions compared to predictions of other endpoints.

	MCC[a]	χ^2
Ames	0.06	63.5
HGPRT	0.09	79.4
SCE	0.11	130.0
AR antagonism	0.13	271.4
ER activation	0.03	11.3
Teratogenicity	0.13	216.7

[a] MCC: Matthew's correlation coefficient.

Table 7
Correlation coefficients and χ^2 statistics for teratogenicity predictions compared to predictions of AR antagonism and ER activation.

	MCC[a]	χ^2
AR antagonism	0.02	9.1
ER activation	0.09	103.6

[a] MCC: Matthew's correlation coefficient.

study, compounds with molecular weights less than 300 Da appeared to be too small to be able to interact with PXR (Xiao et al., 2011). Besides, the three-dimensional shapes of the compounds also appeared to be rather crucial for fitting into the PXR LBD. Several larger molecules with molecular weights greater than 300 (e.g., dexamethazone, doxorubicin, sulfazalazine and pioglitazone) were also shown to be incapable of interacting with PXR, probably because the molecular shapes of these compounds are linear and/or elongated in one dimension, resulting in mismatch with the PXR LBD. Leadscope also identified structural features that contributed positively or negatively to PXR binding. It appeared that all positive contributing substructures are essentially hydrophobic, while negatively contributing features possess hydroxyl or other substitutions which are likely not optimally placed to facilitate interactions with hydrogen bonding features in PXR. This is in agreement with other studies, which have used docking or ligand-based modeling to study the structure–activity relationship for PXR activators (Ekins et al., 2009; Gao et al., 2007). Two molecules from the training set, felodipine and clotrimazole, were used to illustrate the features in the model. Results from a recent docking study showed that both clotrimazole and felodipine formed mostly hydrophobic interactions with PXR, and the imidazole moiety of clotrimazole and the nitrogen atom in the pyridine core of felodipine formed hydrogen bonds with PXR (Xiao et al., 2011). The great interest in avoiding PXR activation in drug development has led to suggestions such as introducing polar groups to the end of an activator, adding larger more rigid groups as well as removing central H-bond acceptors (Gao et al., 2007; Zimmermann et al., 2010).

Screening of a large inventory of chemicals for PXR binding

We found that 35% of the EINECS chemicals in domain of the model were predicted to be PXR ligands. To our knowledge, this is the first study which examines a large set of environmental chemicals in a QSAR screening to identify PXR ligands. It was recently reported that 11% of more than 2800 clinically utilized drugs had human PXR activity

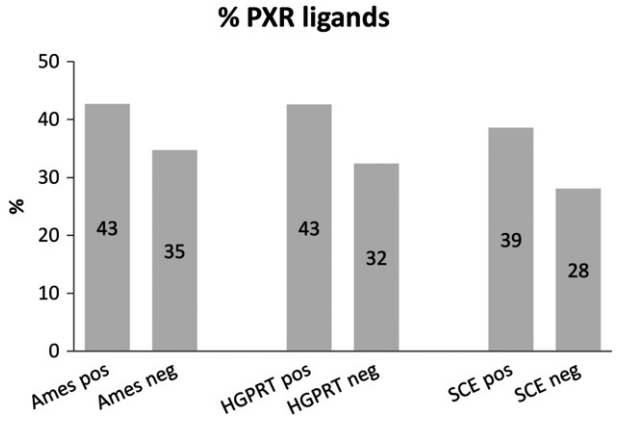


Fig. 4. The percentage of predicted PXR ligands among compounds predicted to be positive or negative in models for genotoxicity. Two models for in vitro genotoxicity, the Ames test and the HGPRT mammalian cell gene mutation assay, and one model for in vivo genotoxicity, sister chromatid exchange (SCE) were used. The percentages of predicted PXR ligands are indicated on the bars.

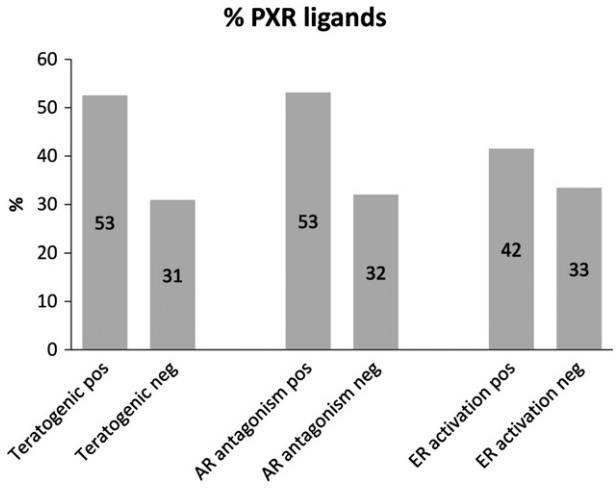


Fig. 5. The percentage of predicted PXR ligands among compounds predicted to be positive or negative in models for teratogenicity, AR antagonism and ER activation. The percentages of predicted PXR ligands are indicated on the bars.

using a reporter gene assay and 7% were positive in a PXR binding assay (Shukla et al., 2011). In another recent study, human PXR activity of 200 pesticides (e.g. organochlorines, diphenyl ethers, organophosphorus, pyrethroids, and carbamates) was characterized using a reporter gene assay (Kojima et al., 2011). Of the 200 pesticides tested, 106 activated PXR, and the PXR activators included representatives from all groups studied. Together, these findings indicate that a large number of chemicals and drugs possess human PXR activity.

Correlations with other endpoints

The aim of the second part of this study was to investigate the role of PXR in toxicity and to elucidate potential mechanisms and modes of action underlying toxic effects. Given the role of PXR as a master regulator of the metabolism of steroids, bile acids and xenobiotics, PXR activators could potentially perturb many different pathways of key biological functions. A large set of environmental chemicals was used for exploring the correlations of PXR with various other endpoints.

Toxic effects of xenobiotics are often ascribed to the chemical reactivity of metabolites generated through biotransformation, leading to covalent binding to DNA, proteins, and lipids (Walsh and Miwa, 2011). The CYP enzyme system is considered to be one of the most important groups of enzymes involved in biotransformation, one important example being CYP3A4. Many xenobiotics are metabolized by more than one CYP enzyme, and CYP3A4 may represent only one pathway. When chemicals are metabolized by CYP3A4, a main target gene of PXR, they may undergo deactivation or be activated into potentially reactive metabolites, which may damage DNA and lead to mutations. In the present study, three different models for genotoxicity were used to evaluate the correlations between PXR binding and genotoxicity. We found a significant higher percentage of predicted PXR ligands among chemicals predicted to be genotoxic compared to non-genotoxic chemicals, indicating that genotoxic effects are more common among compounds that bind to PXR compared to compounds that do not bind to PXR.

Modulation of steroid metabolism through PXR has been suggested as a mechanism for endocrine disruption. We found an overrepresentation of predicted PXR ligands among predicted AR antagonists compared to AR non-antagonists and also among predicted ER activators compared to ER non-activators. These findings suggest that some estrogenic as well as anti-androgenic mechanisms may be mediated through PXR. There are a number of routes through which chemicals can alter steroid receptor activity without directly binding to steroid receptors, e.g. activation into endocrine disrupting

metabolites or metabolic deactivation. For example, methoxychlor and benzophenone are not estrogenic, but they are rapidly converted to metabolites with estrogenic activity. It has been shown that methoxychlor and benzophenone can activate PXR and induce some of the enzymes responsible for producing the metabolites that display estrogenic activity (Mikamo et al., 2003). Another example is sulfonation, which plays an important role in steroid hormone deactivation, because sulfonated hormones often fail to bind to and activate their cognate receptors and therefore lose their hormonal activities. Recently, a novel PXR-mediated and metabolism-based mechanism to reduce androgen activity was reported (Zhang et al., 2010). The study showed that activation of PXR lowered androgen activity by inducing the expression of CYP3As and sulfotransferase 2A1, which are enzymes important for the metabolic deactivation of androgens.

An interesting observation in this study was the significant correlation of PXR binding with teratogenic risk. There was an overrepresentation of predicted PXR ligands among the compounds with predicted teratogenic effects compared to non-teratogenic compounds. By inspection of the compounds predicted positive for both teratogenicity and PXR binding, there were several chemical groups known to be responsible for teratogenicity, such as steroids, genotoxic and carcinogenic chemicals. Human teratogenicity is a complex endpoint with many mechanisms of action (van Gelder et al., 2010). Because PXR is a key regulator of many metabolic enzymes and transporters, activation of PXR may lead to perturbations of many different pathways. Thus, PXR could potentially be an important factor in many different teratogenic mechanisms. The impact of AR antagonism and ER activation on teratogenic risk was also evaluated. Compared to PXR binding, these mechanisms seem to explain only a small fraction of the teratogenic effect of the EINECS chemicals. It was recently reported that disruption of steroid hormone homeostasis as a result of PXR activation contributed to the reproductive toxicity of triazole antifungals (Goetz and Dix, 2009). Perturbation of other hormone systems through PXR, such as the thyroid hormone system, may be critical during development. Thus, up-regulation of UDP-glucuronosyltransferase via activation of PXR leads to increased elimination of thyroid hormones and hypothyroxinemia, which is known to cause impaired brain development in humans (Miller et al., 2009).

Conclusions

A global QSAR model for human PXR binding was developed. To our knowledge, this is the most comprehensive QSAR model for PXR activity developed so far. The model was used to screen 51,680 environmental chemicals, and approximately 35% of the chemicals within domain of the model were predicted to be PXR ligands.

The predicted PXR ligands were found to be overrepresented among compounds predicted to cause adverse effects, such as genotoxicity, teratogenicity, estrogen receptor activation and androgen receptor antagonism compared to compounds not causing these effects. The obtained correlations were biologically plausible supporting the validity of using a QSAR based approach for the study of biological associations.

The large number of predicted PXR ligands may show dose-additive effects and cause adverse effects even when each chemical is present at a low concentration. The developed model may become a valuable tool for screening large numbers of chemicals for PXR binding and for providing mechanistic information of toxic effects of chemicals.

Conflict of interest statement

The authors declare that they have no conflicts of interest.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.taap.2012.05.008.

Acknowledgments

We would like to thank Dr. Sunita Shukla for providing the PXR binding data.

References

Chen, Y., Nie, D., 2009. Pregnane X receptor and its potential role in drug resistance in cancer treatment. Recent Pat. Anticancer Drug Discov. 4, 19–27.
Cooper, J.A., Saracci, R., Cole, P., 1979. Describing the validity of carcinogen screening-tests. Br. J. Cancer 39, 87–89.
di Masi, A., De Marinis, E., Ascenzi, P., Marino, M., 2009. Nuclear receptors CAR and PXR: molecular, functional, and biomedical aspects. Mol. Aspects Med. 30, 297–343.
Ekins, S., Erickson, J., 2002. A pharmacophore for human pregnane X receptor ligands. Drug Metab. Dispos. 30, 96–99.
Ekins, S., Chang, C., Mani, S., Krasowski, M.D., Reschly, E.J., Iyer, M., Kholodovych, V., Ai, N., Welsh, W.J., Sinz, M., Swaan, P.W., Patel, R., Bachmann, K., 2007. Human pregnane X receptor antagonists and agonists define molecular requirements for different binding sites. Mol. Pharmacol. 72, 592–603.
Ekins, S., Kortagere, S., Iyer, M., Reschly, E.J., Lill, M.A., Redinbo, M.R., Krasowski, M.D., 2009. Challenges predicting ligand–receptor interactions of promiscuous proteins: the nuclear receptor PXR. PLoS Comput. Biol. 5, e1000594.
Fiorucci, S., Zampella, A., Distrutti, E., 2012. Development of FXR, PXR and CAR agonists and antagonists for treatment of liver disorders. Curr. Top. Med. Chem. 12, 605–624.
Gao, Y.D., Olson, S.H., Balkovec, J.M., Zhu, Y., Royo, I., Yabut, J., Evers, R., Tan, E.Y., Tang, W., Hartley, D.P., Mosley, R.T., 2007. Attenuating pregnane X receptor (PXR) activation: a molecular modelling approach. Xenobiotica 37, 124–138.
Goetz, A.K., Dix, D.J., 2009. Mode of action for reproductive and hepatic toxicity inferred from a genomic study of triazole antifungals. Toxicol. Sci. 110, 449–462.
Jones, S.A., Moore, L.B., Shenk, J.L., Wisely, G.B., Hamilton, G.A., McKee, D.D., Tomkinson, N.C.O., LeCluyse, E.L., Lambert, M.H., Willson, T.M., Kliewer, S.A., Moore, J.T., 2000. The pregnane X receptor: a promiscuous xenobiotic receptor that has diverged during evolution. Mol. Endocrinol. 14, 27–39.
Khandelwal, A., Krasowski, M.D., Reschly, E.J., Sinz, M.W., Swaan, P.W., Ekinst, S., 2008. Machine learning methods and docking for predicting human pregnane X receptor activation. Chem. Res. Toxicol. 21, 1457–1467.
Kojima, H., Takeuchi, S., Nagai, T., 2010. Endocrine-disrupting potential of pesticides via nuclear receptors and aryl hydrocarbon receptor. Afr. J. Health Sci. 56, 374–386.
Kojima, H., Sata, F., Takeuchi, S., Sueyoshi, T., Nagai, T., 2011. Comparative study of human and mouse pregnane X receptor agonistic activity in 200 pesticides using in vitro reporter gene assays. Toxicology 280, 77–87.
Kretschmer, X., Baldwin, W., 2005. CAR and PXR: xenosensors of endocrine disrupters? Chem. Biol. Interact. 155, 111–128.
Mikamo, E., Harada, S., Nishikawa, J., Nishihara, T., 2003. Endocrine disruptors induce cytochrome P450 by affecting transcriptional regulation via pregnane X receptor. Toxicol. Appl. Pharmacol. 193, 66–72.
Miller, M.D., Crofton, K.M., Rice, D.C., Zoeller, R.T., 2009. Thyroid-disrupting chemicals: interpreting upstream biomarkers of adverse outcomes. Environ. Health Perspect. 117, 1033–1041.
Moreau, A., Vilarem, M.J., Maurel, P., Pascussi, J.M., 2008. Xenoreceptors CAR and PXR activation and consequences on lipid metabolism, glucose homeostasis, and inflammatory response. Mol. Pharm. 5, 35–41.
Ngan, C., Beglov, D., Rudnitskaya, A.N., Kozakov, D., Waxman, D.J., Vajda, S., 2009. The structural basis of pregnane X receptor binding promiscuity. Biochemistry 48, 11572–11581.
Nikolov, N., Grancharov, V., Stoyanova, G., Pavlov, T., Mekenyan, O., 2006. Representation of chemical information in OASIS centralized 3D database for existing chemicals. J. Chem. Inf. Model. 46, 2537–2551.
Pan, Y., Li, L., Kim, G., Ekins, S., Wang, H., Swaan, P.W., 2011. Identification and validation of novel human pregnane X receptor activators among prescribed drugs via ligand-based virtual screening. Drug Metab. Dispos. 39, 337–344.
Roberts, G., Myatt, G., Johnson, W., Cross, K., Blower, P., 2000. LeadScope: software for exploring large sets of screening data. J. Chem. Inf. Comput. Sci. 40, 1302–1314.
Schuster, D., Langer, T., 2005. The identification of ligand features essential for PXR activation by pharmacophore modeling. J. Chem. Inf. Model. 45, 431–439.
Shukla, S.J., Nguyen, D., MacArthur, R., Simeonov, A., Frazee, W.J., Hallis, T.M., Marks, B.D., Singh, U., Eliason, H.C., Printen, J., Austin, C.P., Inglese, J., Auld, D.S., 2009. Identification of pregnane X receptor ligands using time-resolved fluorescence resonance energy transfer and quantitative high-throughput screening. Assay Drug Dev. Technol. 7, 143–169.
Shukla, S.J., Sakamuru, S., Huang, R., Moeller, T.A., Shinn, P., VanLeer, D., Auld, D.S., Austin, C.P., Xia, M., 2011. Identification of clinically used drugs that activate pregnane X receptors. Drug Metab. Dispos. 39, 151–159.
Tabb, M., Blumberg, B., 2006. New modes of action for endocrine-disrupting chemicals. Mol. Endocrinol. 20, 475–482.
Tolson, A.H., Wang, H., 2010. Regulation of drug-metabolizing enzymes by xenobiotic receptors: PXR and CAR. Adv. Drug Deliv. Rev. 62, 1238–1249.
Ung, C.Y., Li, H., Yap, C.W., Chen, Y.Z., 2007. In silico prediction of pregnane X receptor activators by machine learning approaches RID B-2587-2010. Mol. Pharmacol. 71, 158–168.
van Gelder, M.M., van Rooij, I.A., Miller, R.K., Zielhuis, G.A., de Jong-van den Berg, L.T., Roeleveld, N., 2010. Teratogenic mechanisms of medical drugs. Hum. Reprod. Update 16, 378–394.

Walsh, J.S., Miwa, G.T., 2011. Bioactivation of drugs: risk and drug design. Annu. Rev. Pharmacol. Toxicol. 51, 145–167 (2011 51).
Watkins, R., Wisely, G., Moore, L., Collins, J., Lambert, M., Williams, S., Willson, T., Kliewer, S., Redinbo, M., 2001. The human nuclear xenobiotic receptor PXR: structural determinants of directed promiscuity. Science 292, 2329–2333.
Xiao, L., Nickbarg, E., Wang, W., Thomas, A., Ziebell, M., Prosise, W.W., Lesburg, C.A., Taremi, S.S., Gerlach, V.L., Le, H.V., Cheng, K., 2011. Evaluation of in vitro PXR-based assays and in silico modeling approaches for understanding the binding of a structurally diverse set of drugs to PXR. Biochem. Pharmacol. 81, 669–679.
Zhang, J., Kuehl, P., Green, E.D., Touchman, J.W., Watkins, P.B., Daly, A., Hall, S.D., Maurel, P., Relling, M., Brimer, C., Yasuda, K., Wrighton, S.A., Hancock, M., Kim, R.B., Strom, S., Thummel, K., Russell, C.G., Hudson, J.R., Schuetz, E.G., Boguski, M.S., 2001. The human pregnane X receptor: genomic structure and identification and functional characterization of natural allelic variants. Pharmacogenetics 11, 555–572.
Zhang, B., Cheng, Q., Ou, Z., Lee, J.H., Xu, M., Kochhar, U., Ren, S., Huang, M., Pflug, B.R., Xie, W., 2010. Pregnane X receptor as a therapeutic target to inhibit androgen activity. Endocrinology 151, 5721–5729.
Zhou, C., Verma, S., Blumberg, B., 2009. The steroid and xenobiotic receptor (SXR), beyond xenobiotic metabolism. Nucl. Recept. Signal. 7, e001.
Zimmermann, K., Wittman, M.D., Saulnier, M.G., Velaparthi, U., Sang, X., Frennesson, D.B., Struzynski, C., Seitz, S.P., He, L., Carboni, J.M., Li, A., Greer, A.F., Gottardis, M., Attar, R.M., Yang, Z., Balimane, P., Discenza, L.N., Lee, F.Y., Sinz, M., Kim, S., Vyas, D., 2010. SAR of PXR transactivation in benzimidazole-based IGF-1R kinase inhibitors. Bioorg. Med. Chem. Lett. 20, 1744–1748.