Computational Toxicology 14 (2020) 100121

Contents lists available at ScienceDirect

Computational Toxicology

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

QSAR modeling of different minimum potency levels for in vitro human CAR activation and inhibition and screening of 80,086 REACH and 54,971 U.S. substances

K.K. Chinen[a], K. Klimenko[b], C. Taxvig[c], N.G. Nikolov[c,1], E.B. Wedebye[c,1,*]

[a] University of California, Los Angeles, CA 90095, USA
[b] Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, 2829 516 Caparica, Portugal
[c] National Food Institute, Technical University of Denmark, 2800 Kgs, Lyngby, Denmark

ARTICLE INFO

Keywords:
QSAR
constitutive androstane receptor (CAR)
Tox21
Adverse Outcome Pathway (AOP)
Screening
REACH

ABSTRACT

The human Constitutive Androstane Receptor (hCAR) is together with the human Pregnane X Receptor (hPXR) a key regulator of the metabolism and excretion of xenobiotics and endogenous compounds. Inhibition or activation of hCAR by xenobiotics can alter protein expression, leading to decreased or enhanced turnover of both xenobiotics and endogenous substances. Impacts from these alternations can potentially disturb physiological homeostasis and cause adverse effects. Tens-of-thousands of manufactured substances of which humans are potentially exposed are not tested for their potential to inhibit or activate hCAR. In this study, the U.S. Toxicology in the 21st Century (Tox21) high-throughput in vitro assay results for hCAR inhibition and activation were used in a comprehensive in-house process to derive training sets for different potency cut-offs, and to develop suites of quantitative structure–activity relationship (QSAR) models with binary outputs. Final, expanded models, which include substances from the external validation sets, were developed for select minimum potency levels. Rigorous cross- and external validations demonstrated good predictive accuracies for the models. The final models were applied to screen 80,086 EU and 54,971 U.S. substances, and the models predicted around 60% of the substances within their respective applicability domains (AD). Finally, statistical comparisons of hCAR predictions and QSAR predictions for a number of other endpoints related to e.g. Pregnane X, aryl hydrocarbon, estrogen and androgen receptors, as well as genotoxicity, cancer, sensitization and teratogenicity from the Danish (Q)SAR database were made to explore possible implications related to hCAR antagonists and agonists. The final models from this study will be made available in the free Danish (Q)SAR Models website. Predictions made with models from this study for 650,000 substances will be made available in the free Danish (Q)SAR Database. Predictions from the models developed in this study can for example contribute to priority setting, read-across cases and weight-of-evidence assessments of chemicals.

Abbreviations: AD, applicability domain; AhR, human Aryl hydrocarbon Receptor; AOP, Adverse Outcome Pathway; AR, Androgen Receptor; BA, balanced accuracy; CAR, Constitutive Androstane Receptor; CoMPARA, Collaborative Modeling Project for Androgen Receptor Activity (SOT); CRS, concentration–response series; CV, cross-validation procedure; CYP, cytochrome P450; DBM, OASIS Database Manager; DTU Food, Technical University of Denmark National Food Institute; EPA, Environmental Protection Agency; ER, Estrogen Receptor; EU, European Union; HTS, high-throughput screening; HepG2, human hepatoma; IATA, Integrated Approaches to Testing and Assessment; LBD, ligand binding domain; LPDM, Leadscope® Predictive Data Miner; MCC, Matthews correlation coefficient; MIE, molecular initiating event; NR, nuclear receptor; NUL, No Upper Limit threshold concentration; OECD, Organisation for Economic Co-operation and Development; PLR, partial logistic regression; PXR, Pregnane X Receptor; qHTS, quantitative high-throughput screening; QMRF, (Q)SAR model reporting format; QPRF, (Q)SAR prediction reporting format; QSAR, quantitative structure–activity relationship; REACH, Registration, Evaluation, Authorisation and Restriction of Chemicals; SD, standard deviation; SDF, structure-data files; TH, thyroid hormone; U.S. EPA, United States Environmental Protection Agency; U.S. Tox21, United States Toxicology Testing in the 21st Century; WoE, weight-of-evidence; VDR, Vitamin D Receptor

* Corresponding author.
E-mail address: ebawe@food.dtu.dk (E.B. Wedebye).
[1] Contributed equally

https://doi.org/10.1016/j.comtox.2020.100121
Received 1 October 2019; Received in revised form 22 January 2020; Accepted 3 February 2020
Available online 04 February 2020

1. Introduction

The constitutive androstane receptor (CAR) belongs to the human nuclear receptor (NR) superfamily, a 48-member group [1,2] of "orphan" and "adopted-orphan" NRs [3–5]. In humans, the CAR protein is encoded by the NR1I3 gene from the NR subfamily 1, group I, member 3. The NR subfamily 1 group I also includes the Vitamin D Receptor (VDR) and the Pregnane X Receptor (PXR) [6–8]. CAR displays so-called constitutive activity, meaning that it is active also in the absence of a ligand [9,10]. Many known CAR agonists are species-specific [11,12]. CAR is expressed mainly in the liver and small intestine [1,3,13] and mediates the induction of metabolizing enzymes, such as cytochrome P450 3A (CYP3A) isoenzymes, conjugation enzymes such as UDP glucuronosyltransferase family 1 member A1, and transporters such as P-glycoprotein [14–17]. Along with the NR PXR, CAR is a principal regulator of the metabolism of xenobiotic compounds [16,18,19]. PXR and CAR cross-regulate their target genes cytochrome P450 (CYP) CYP2B and CYP3A [20]. CAR also plays an important role in the metabolism of a number of endogenous substances such as thyroid and steroid hormones, cholesterol, bile acids, bilirubin, glucose, and lipids [16,19].

CAR activation and inhibition are mechanistic endpoints related to a number of health outcomes. In some cases, the CAR upregulation of xenobiotic metabolism may lead to increased turnover of hormone and other endogenous substances leading to decreased levels in the body [21]. Such interference in the regulation of endogenous hormones may have negative consequences [22], which can be seen in the Adverse Outcome Pathway (AOP): 8 (under development) [23]. According to this AOP, activation of CAR or other NRs like PXR and the aryl hydrocarbon receptor (AhR) as a molecular initiating event (MIE) can cause upregulated thyroid hormone (TH) catabolism, and lead to reduced TH levels, which may result in adverse neurodevelopmental outcomes in mammals [23].

CAR activation can have positive outcomes, and has been found to ameliorate diabetes [24]. However, CAR activation is also the molecular initiating event, according to AOP: 107 [25] (under review), that can lead to hepatocellular adenomas and carcinomas in the mouse and the rat. When mice were exposed to certain xenobiotics, CAR activation was found to be an important factor for tumor development [9,26,27]. In addition, CAR inhibition may decrease the metabolizing potential in the body, which leads to decreased turnover of endogenous hormones as well as decreased detoxification and excretion of xenobiotics [28]. Furthermore, CAR inhibition may lead to hepatic steatosis according to AOP: 58 (under development) [29].

In an effort to reduce animal testing and increase the toxicity-related information level on chemical substances, the Organisation for Economic Co-operation and Development (OECD) and European Union (EU) have paved the way for increasing regulatory use of quantitative structure–activity relationship (QSAR) models e.g. through the EU Registration, Evaluation, Authorization and Restriction of Chemicals (REACH) regulation [30–32]. QSARs are mathematical models that predict properties, such as biological activities, based on chemical structure [33–35]. As QSAR predictions can be rapidly generated for large inventories of substances, their use is very suited for screening and priority setting purposes as exemplified by an earlier large-scale QSAR screening for thyroperoxidase inhibitors by some of the authors of this article [36]. In some cases, QSAR predictions may be used for a 1:1 replacement of experimental tests [37] but especially for higher tier health endpoints, QSAR predictions may rather contribute to Integrated Approaches to Testing and Assessment (IATA) weight-of-evidence (WoE) assessments or to improve read-across cases.

The primary objective of this study was to develop global binary QSAR models that can be used for screening purposes and single-compound identification of possible hCAR antagonists or agonists. A secondary interest in this study was to process the experimental training set data specifically for the development of QSAR models for prediction of minimum potency. We used high-throughput in vitro data sets from the U.S. Tox21 Program's quantitative high-throughput (qHTS) assay for hCAR agonism and for hCAR antagonism [38], and the results were used to train and validate a number of QSAR models for hCAR inhibition and activation. The U.S. Tox21 Program applies qHTS screening with the aim of identifying substances that may adversely affect human health for priority setting purposes. To date, the Tox21 chemical library holds approximately 10,000 diverse chemical substances, such as commercial chemicals, pesticides, food additives/contaminants, and medical compounds [39].

Human CAR agonism data has previously been used as the basis for QSAR modeling [40–44]. In one case, the published models were developed on the qHTS agonism Tox21 data as used in the present study, though using the Tox21 summary calls as is [40,41]. In the other cases, modeling of hCAR agonism was done as a continuous endpoint or as binding preferences between CAR and PXR and using other smaller data sources [42–44]. We have found no published QSARs developed using antagonism datasets.

In this study, we developed a set of criteria, which we applied to process the data for our QSAR model development, including setting a minimum absolute effect which should be seen at a maximum concentration threshold and which should occur at a non-cytotoxic concentration. We also filtered out luciferase inhibitors which were possible false positive agonists or false negative antagonists. The processed data was used to develop initial models for all concentration thresholds. All initial models underwent a Technical University of Denmark, National Food Institute (DTU Food) in-house two times five-fold cross-validation (CV) and external validations with unused actives and inactives. Subsequently, four final expanded models were developed by including external validation data into the training sets of the corresponding initial models. The final models underwent the same CV procedure and external validation for specificity (not for sensitivity as all actives were used in the models). The final models were used to screen 80,086 structurally diverse substances pre-registered and/or registered under the EU REACH chemicals regulation, and the 54,971 unique chemical structures from the U.S. Environmental Protection Agency (U.S. EPA) Collaborative Modeling Project for Androgen Receptor Activity (CoMPARA) inventory [45]. Use of the predictions may include: 1) priority setting; 2) single substance WoE IATA assessments; 3) support for read-across cases e.g. to find suitable source analogs or as part of the basis for making the read-across justification. In addition, hCAR predictions were statistically correlated with predictions from 39 other QSAR models from the free, online Danish (Q)SAR database [64] including PXR binding/activation [36], AhR activation [46], thyroperoxidase (TPO) inhibition [47], estrogen receptor (ER) activation, androgen receptor (AR) antagonism [48], as well as endpoints within genotoxicity, cancer, sensitization and teratogenicity to explore possible roles for the hCAR receptor in relation to other biological activities.

The QSAR predictions from these screenings will be published in the free online Danish (Q)SAR Database. In addition, all final models will be published in the new free, online Danish (Q)SAR Models (accessible from the Danish (Q)SAR Database) for real-time prediction of user-submitted structures and downloads of the detailed results in the (Q)SAR Prediction Reporting Format (QPRF).

2. Materials and methods

2.1. Experimental datasets, definition of endpoints and developed QSAR-targeted data processing

To develop our datasets, we used results from the U.S. Tox21 Program available from the Tox21 Data Browser [28] and structures for the Tox21 substances from PubChem [49]. As part of the U.S. Tox21 Program, the U.S. National Institutes of Health (NIH) screened a total of 9,667 chemical substances for hCAR agonism and antagonism [50,51]

for which substances in the chemical library were not specifically selected to target hCAR agonism and/or hCAR antagonism. Descriptions of the chemical structures, hCAR assays and Tox21 data analysis can be found in more detail in previous publications [52,53].

Under Tox21 qHTS testing, both assays were applied to screen the approximately 10,000 Tox21 substances for hCAR agonists and antagonists, as well as for cell viability. For the cell culture, Tox21 qHTS testing used human hepatoma (HepG2) cells transfected with a double-stable human CAR and CYP2B6-2.2 kb in both agonist and antagonist mode [54,55]. In addition, both assays screened 16 different concentrations with varying concentration ranges among the different substances [28]. Screening statistics of the agonist assay generated a Z' factor of 0.687 [53,56], and a coefficient of variance close to 6.04% ± 1.56 [53]. Z' factors reflect the assay signal dynamic range and data variation associated with signal measurements. Thus, an indicator of good performance is a Z' factor above 0.5 [56]. These datasets were used as a basis for our study in addition to computer-readable structure-data files (SDF) on the tested chemicals substances structures from PubChem: *a*) AID 1224893 on small molecule antagonists of the hCAR signaling pathway, and *b*) AID 1224892 on small molecule agonists of the hCAR signaling pathway. Assay results were provided by the U.S. Tox21 Program [51,57].

The U.S. Tox21 activity profiling is primarily aimed at identifying potential mechanisms of action to prioritize substances for further in-depth toxicological evaluation. As part of the U.S. Tox21 data processing, concentration–response series (CRS), typically three per substance per assay, are fit to four-parameter Hill equations. Outcomes are then ranked into so-called curve classes specific to the Tox21 program, accounting for efficacy, p-value, asymptotes and inflection [52]. Concentrations of half-maximal (AC50) rather than absolute activity are additionally calculated for activity and cell viability. For instance, if the maximum activity of a substance is 30% inhibition, Tox21's dose–response modeling Hill curve will give an AC50, (i.e. the concentration that causes half-maximal activity) of 15% inhibition. In the end, half-maximal AC50 values for activity and cell viability were applied to make activity outcome summary calls specific to the Tox21 program, which in some cases also integrated results from additional counter screens.

Rather than using the Tox21 summary calls [50,51], in this study we undertook further QSAR-targeted processing of the Tox21 hCAR data by setting criteria for absolute activity for actives, and just as importantly for QSAR models development purposes, by setting criteria to only select the most robust inactives (Fig. 1) . This QSAR-targeted processing was possible because the detailed Tox21 test results for each of the tested concentrations were available in public repositories. For the purpose of defining our endpoints, we did not find any information from regulations, AOPs, scientific literature or other sources that would tell us which potency cut-offs we should apply to identify the most relevant agonists/antagonists from a health impact perspective, or, if certain potency cut-offs formed a better basis for QSAR modeling than others. We therefore decided to apply different potency cut-offs with a 25% absolute effect occurring at or below six different thresholds: 10 μM, 20 μM, 30 μM, 40 μM 50 μM, and No Upper Limit (NUL), constructing a data set for each threshold. A lower cut-off equals a higher minimum potency, which could potentially form the basis for stronger alerts in a model, given that the training set contains a sufficient number of observations. On the other hand, NUL implies that we did not impose a concentration cut-off.

For each substance, our QSAR-targeted data processing led to the assignment of one of the following outcomes: "active", "inactive", or "inconclusive". Only actives and inactives were used for QSAR development and validation. For the data processing, we filtered each test CRS through in-house tools, specifically developed for determining active responses with non-cytotoxic concentrations showing at least 25% effect (in absolute value). To accept only the best Tox21 Hill curve classes for inactives, we required Tox21 Hill curve class 4 (i.e. inactive) and that the substance exhibited no cytotoxicity up to a 10 μM concentration [48]. The complete data processing of each substance for both hCAR agonism and antagonism for each of the selected concentration thresholds (10 μM, 20 μM, 30 μM, 40 μM, 50 μM, and NUL) fell into five main data steps:

1. For agonism activity, only Tox21 curve classes 1.1, 1.2, 2.1, and 2.2, (i.e. according to the Tox21 definition of these curve classes, all complete and incomplete curves with inflection, p-value < 0.05 and efficacy $>$ 3 standard deviations (SD) of control) were accepted.
2. For antagonism activity, only curve classes -1.1, -1.2, -2.1 and -2.2, (i.e. according to the Tox21 definition of these curve classes, all complete and incomplete curves with inflection, p-value < 0.05 and efficacy $>$ 3SD of control) were accepted.
3. The highest concentration with non-cytotoxicity was calculated as the median for all cell viability CRSs for the analyzed substance, using the highest concentration with at least 80% viable cells for each CRS.
4. For agonist "actives", two-thirds of all hCAR activity CRS for the substance were required to fulfill the requirement of step 1 and to have at least 25% effect at a non-cytotoxic concentration (as defined in "3") at or below the selected concentration threshold. For antagonist "actives", two-thirds of all hCAR activity CRS for the substance were required to fulfill the requirement of step 2 and to have at least 25% effect at a non-cytotoxic concentration (as defined in "3") at or below the selected concentration threshold.
5. Substances for which all hCAR activity CRS were curve class 4 with no cytotoxicity up to at least 10 μM were assigned "inactive".

A small number of substances for both agonism and antagonism had at least 50%, but not two-thirds of, all hCAR activity CRS fulfilling the requirement of step 4. We considered these substances to be "active" using expert judgment.

Both hCAR agonism and antagonism assays were luciferase-based. Under certain circumstances, luciferase inhibitors may stabilize the enzyme, giving significant increases in luciferase levels in cells relative to untreated wells within the typical assay incubation time, and potentially leading to increases in luminescent signal in cell-based assays [58]. We therefore removed luciferase inhibitors which were hCAR agonist actives or antagonist inactives. As a cell-based luciferase counter screen specifically for the hCAR agonism and antagonism assays was not available, we applied an *in chemico* luciferase inhibition screen for all Tox21 substances using the Tox21 conclusions as-is [59].

2.2. Structure curation

All data set structures underwent structure curation after the QSAR-targeted data processing (Fig. 2). The curation was performed in OASIS Database Manager (DBM) 1.7.3 [60], including additional algorithms developed in-house. We identified compounds with acceptable structures. Only structures exclusively containing atoms from the following list were kept: H, B, C, N, O, F, Na, Mg, Si, P, S, Cl, K, Ca, Br, and I. Records with structure errors identified by OASIS DBM were removed from the dataset. We then conducted a dissociation simulation by breaking ionic bonds and "neutralizing" the remaining structures. After this step, we removed substances containing two or more organic components, (i.e. "mixtures"), and structures with less than two carbon atoms from the dataset. Furthermore, to assure that every chemical structure was only represented once in the data set, identical structures, (i.e. duplicates), were identified and removed according to the procedure described in Fig. 2.

2.3. Training and external validation sets preparation

For both antagonism and agonism, we split each of the six

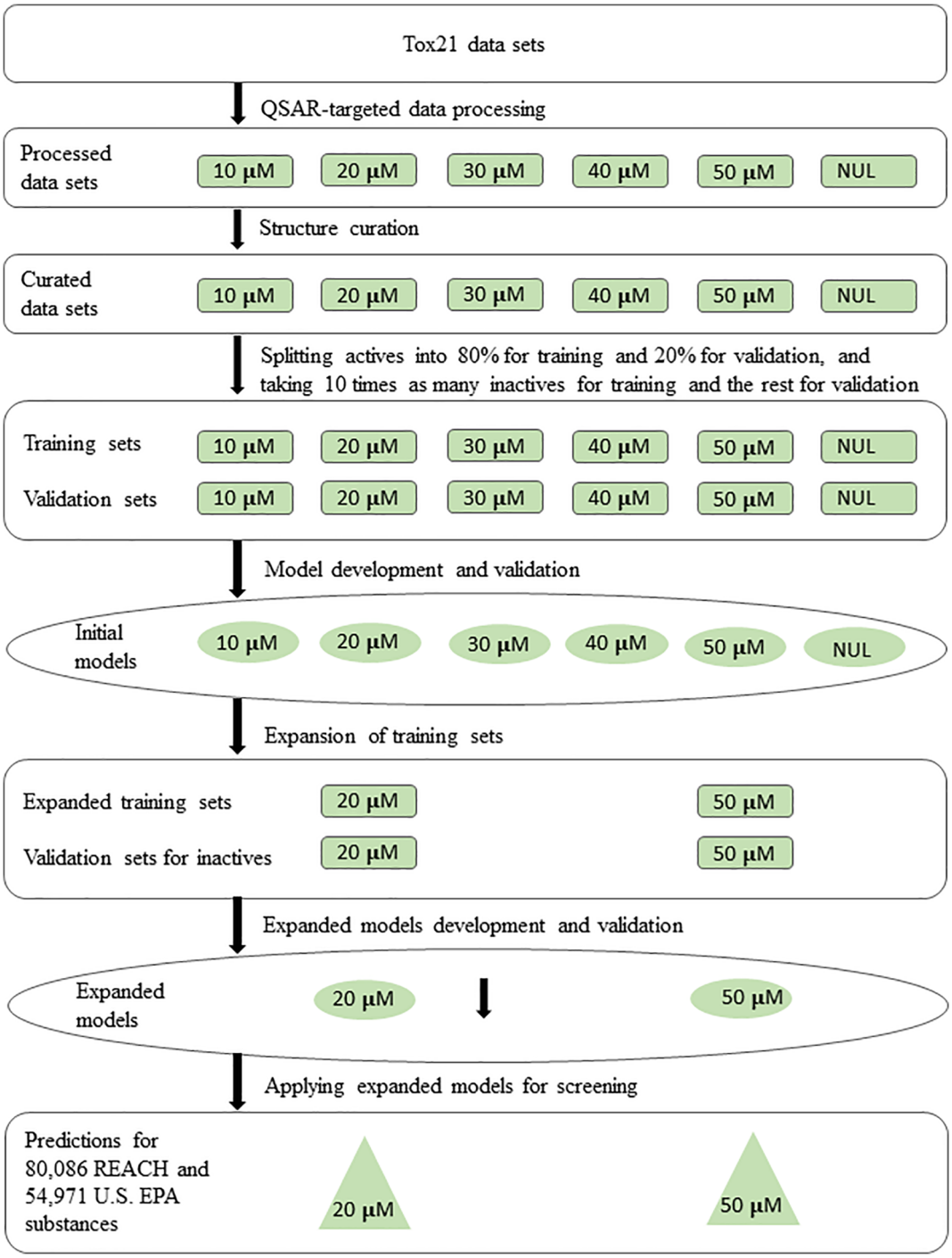

Figure 1. Overview of the process of making training and validation sets, modeling, and predictions for hCAR antagonism and agonism activity. The rectangular boxes are for experimental data sets, the ovals are for models and the triangles are for the screening sets. The process was performed for both the agonism and antagonism sets.

concentration threshold data sets into a training and a validation set (Fig. 1). For each data set, we randomly selected 20% of the active structures and assigned them to the validation set. The remaining 80% active structures were then assigned to the training set. Afterwards, we randomly selected ten times as many training set inactives to create a training set with a 1:10 distribution. This distribution represented the maximum ratio that the applied QSAR modeling software could efficiently model. Any remaining inactives were used for the validation set. This meant that for external validation sets, inactives greatly outnumbered the actives. Only after the models were fully developed, we applied the external validation sets in an independent external validation (i.e. external validation was not used to select models).

Last, we combined the training set and the external validation set for each of a number of selected concentration thresholds to create four final, expanded models, namely for the 20 μM and 50 μM concentration thresholds for both hCAR antagonism and hCAR agonism (Fig. 1). An aim of the expanded models was to possibly improve model accuracy, robustness and/or applicability domain of the models. Due to the 1:10 limitation, some negatives were randomly left out of the expanded models. These negatives were later applied to make independent external validations for specificity for the expanded models.

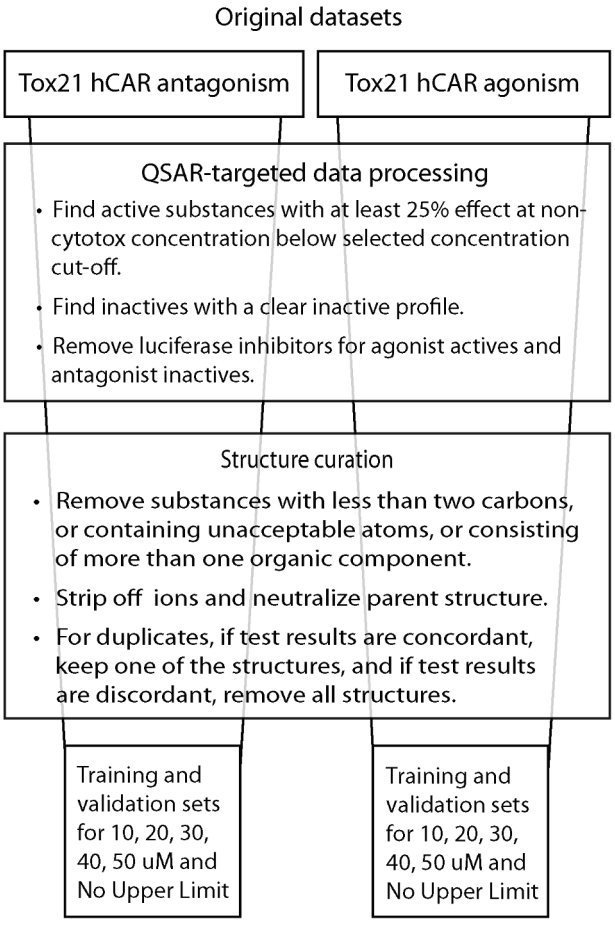

Figure 2. Data processing and structure curation.

2.4. QSAR modeling and selection

To develop all models for both antagonism and agonism, we used the commercial software Leadscope® Predictive Data Miner (LPDM), a component of Leadscope® Enterprise Server version 3.5.3-5 [61]. Structures were first imported into LPDM. Upon import, nine continuous molecular descriptors (AlogP, hydrogen bond acceptors and donors, Lipinski score, molecular weight, parent atom number, parent molecular weight, polar surface area, rotatable bonds) were calculated for each structure. Imported structures also underwent LPDM's systematic substructure analysis for indexing according to 27,000 pre-defined 2D fragment descriptors called "features" [62,63]. LPDM allows the user to generate additional training set-dependent 2D fragment descriptors called "scaffolds," which may or may not coincide with the original feature library. From the entire descriptor set, which includes structural features, scaffolds and molecular descriptors, LPDM automatically selected the top 30% descriptors using the Yates X^2-test. LPDM models binary response variables using partial logistic regression (PLR). A description of the PLR method in LPDM is available in a publication by Valerio et al. [64], which also refers to other publications on the topic and examples of the performance of LPDM PLR models. Briefly, as described in [59],

> "PLR is used for a binary response variable and extracts factors by PLS using the responses as continuous data followed by logistic regression for classifications; this process is repeated until the criteria for optimum number of factors and features are reached. The binary classification model results are given as outcome probabilities from the logistic regression."

Training sets skewed towards a greater number of negatives, however, can lead to models with a higher specificity, (i.e. bigger proportion of the negatives being correctly predicted), at the expense of the sensitivity, (i.e. smaller proportion of the positives being correctly predicted) [64]. LPDM, therefore, offers the option of developing composite models, a method bearing some resemblance to bagging [65] though with full resampling of the smaller class and without replacement of the larger class. With this option, the modeler can set the desired ratio between the two activity classes and include up to ten sub-models with a 1:1 ratio, resampling the smaller class. The sub-models are aggregated to form a composite model. When such a model is used for prediction, the test structure is first predicted by all sub-models individually, and the composite model positive prediction probability is defined as the average positive prediction probability over all sub-models where the test structure is in the defined applicability domain [64]. In our experience, composite models have close-to-equal sensitivity and specificity. In previous research by some of the authors, a "cocktail" model approach was made where the sub-models of composite models are further aggregated with a model on the full skewed training set ("single model") [47]. In earlier work this has been shown to most often increase specificity with only a small penalty on sensitivity compared to the composite models, thus increasing the balanced accuracy compared to either the composite or single models [46].

With the purpose of selecting the specific modeling approach for further hCAR model development, we used the following approaches for all six agonism and six antagonism concentration threshold training sets in the initial model development:

1. single model, i.e. a non-composite model drawing on the full training set
2. composite model, i.e. 10 sub-models
3. composite cocktail model, i.e. single model combined with the 10 composite sub-models

For all models, scaffolds were generated in LPDM from the training set structures and used with the continuous descriptors and features. All models underwent a two-times five-fold CV (i.e. removing 20% and making models on the remaining 80% structures to predict the 20% leave-out) by the LPDM algorithm. Currently, LPDM's variable selection algorithm transfers knowledge of the selected descriptor set from the parent model when developing the CV sub-models. As a result, LPDM's CV may give overly optimistic results. Thus, LPDM's CV was only used to assess the relative performance of the initial models for modeling approach selection. In the end, we selected a number of concentration-thresholds for antagonism and agonism, for which we used the expanded training set made from initial training set and validation set for that concentration threshold to develop new, expanded models (Fig. 1).

2.5. Applicability domain definition

We defined the applicability domain (AD) of our models as a combination of the following three components: 1) model-independent structure requirements; 2) LPDM's definition of a structural domain; 3) DTU Food's in-house definition of class probability refinement on the LPDM's output. We considered a test structure to be in AD if:

1. The test structure exclusively contained H, B, C, N, O, F, Na, Mg, Si, P, S, Cl, K, Ca, Br, and/or I, it was mono-constituent after de-salting, and it contained at least two carbon atoms.
2. The test structure met the following LPDM criteria:
   a) LPDM's algorithms can calculate all molecular descriptors for the structure
   b) the structure of the compound contains at least one structural feature used in the model
   c) the structure of the compound has at least 30% similarity using the Jaccard [66] (also known as Tanimoto) coefficient [64,p. 509,513] with a training set substance (based on Leadscope's

built-in fragment library).

3. The test structure met the following criteria based on the positive prediction probability p between 0 and 1 calculated by LPDM as part of the prediction, with actives having a $p \geq 0.5$ and inactives having a $p < 0.5$ [64,513]: $p \geq 0.7$ is required for an active prediction call and a $p \leq 0.3$ for an inactive prediction call. Predictions closer to the cutoff ($p = 0.5$) are excluded, as they are likely to be less reliable.

2.6. Validation of the models

After using LPDM to guide the selection of the modeling approach, we applied DTU Food's two-times five-fold CV procedure to measure the robustness and performance of the initial antagonism and agonism models (Fig. 1). In this procedure, all DTU Food CV sub-models were developed in isolation from the parent model as completely new models in Leadscope. The DTU Food's CV conditions prevent any transfer of knowledge from the parent model to the CV sub-models. The five-fold approach was chosen as a robust leave-many-out CV approach, and because removing a higher proportion of actives might cause too large of a perturbation in the training set for the relatively small active class.

To execute the DTU Food five-fold CVs, we first randomly divided the training set into five portions, each constituting 20% of the training set structures while preserving the ratio of inactives to actives (10:1). For each of these five portions, the following steps were taken:

1. the 20% portion was removed from the full training set to create a sub-model's training set of 80%.
2. a CV sub-model was built from the reduced training set by applying the same development approach as for the parent model, but without transferring any variable selection information.
3. the 20% left-out portion was predicted by the 80% sub-model.

The whole procedure was performed twice, resulting in 10 CV prediction sets per concentration threshold model. For all in-AD predictions for each concentration threshold model's 10 prediction sets, we also calculated overall sensitivity, specificity and balanced accuracy as well as standard deviations (SD) between prediction results in the ten sub-models [67]. In our study, we used Cooper et al.'s definitions: a) sensitivity is defined as the percentage of experimental actives predicted accurately; b) specificity is the percentage of experimental inactives predicted accurately; c) balanced accuracy (BA) as the average of specificity and sensitivity [67]. We chose these statistical measures of performance as they are particularly appropriate for our cases of models with very imbalanced training sets, as they all three are independent of the balance between the active and inactive class. To determine the percentage of substances with predictions within the AD of the DTU Food CV models, we calculated the total coverage or "the proportion of the full set predicted within the AD of the model" of each threshold concentration model's total 10 CV models [47].

To further evaluate the predictive performance of the initial DTU Food cross-validated threshold-concentration models for antagonism and agonism, we subjected these models to external validations using the set aside validation sets (Fig. 1). Predictions, which were within AD, were then compared with the experimental results. Sensitivity, specificity, balanced accuracy and coverage were calculated for each model. Likewise, we applied the DTU Food CV procedure to the expanded antagonism and agonism models for the selected concentration thresholds (Fig. 1). Since all expanded models contained all of the actives from the QSAR-targeted Tox21 data processing, an external validation could only be performed for specificity with the inactives that were not included in the expanded models. Compliance with the OECD (Q)SAR validation principles for all four expanded models was documented in (Q)SAR Model Reporting Format (QMRF) reports.

2.7. Screening large chemical inventories

To identify possible hCAR activators and inhibitors among current industrial chemicals, we applied the expanded QSAR models to predict two, large regulatory chemical libraries: the REACH pre-registered and/or registered substances compiled for the Danish (Q)SAR Database [68,69], and a U.S. EPA substance list compiled for the U.S. EPA CoMPARA project [70] (Fig. 1). Both the REACH substances and U.S. EPA set already underwent a similar structure preparation as described in 2.2. For our study, 80,086 QSAR-ready REACH structures, and 54,971 QSAR-ready U.S. EPA inventory structures were screened by the expanded QSAR models. As part of the prediction analysis, we calculated the proportion of QSAR-predicted U.S. EPA and REACH-PRS substances within the AD, and of these, how many were predicted as active or inactive.

2.8. Statistical correlations of hCAR predictions with QSAR predictions for other endpoints

To investigate possible statistical (not necessarily causal) associations between hCAR and hPXR, hAhR, hER, hAR, mutagenicity, sensitization, cancer and teratogenicity, we correlated screening results from the REACH set with Leadscope® QSAR predictions on other endpoints contained in the DTU Food-developed Danish (Q)SAR Database. Detailed information for free download on the applied Leadscope® QSAR models can be found in the QSAR Model Reporting Format (QMRF) from the Danish (Q)SAR Database [69]. To gauge the strength of the correlations, specifically for predictions found in the common domain between each of the individual hCAR models and each of the individual other models, we calculated Matthews correlation coefficient (MCC), chi-square (χ^2) test statistic, and statistics for how good the hCAR models are at 'catching' actives from other models as well as how often hCAR models give positive predictions when the other models predict negative, and how good other models are at 'catching' actives from the hCAR models as well as how often the other models give positive predictions when the hCAR models predict negative.

3. Results and discussion

In this study we developed and validated QSAR models for hCAR antagonism and hCAR agonism for a number of different effect concentration thresholds and used final expanded models for 25% absolute effect at maximum 20 μM and 50 μM to screen 80,086 REACH substances and 54,971 U.S. EPA substances for hCAR antagonism and agonism.

3.1. The training and validation sets

We started with Tox21 experimental results for 9,667 substances for both hCAR antagonism and hCAR agonism from the Tox21 Data Browser and structures from PubChem. After subjecting the initial data to our QSAR-targeted data processing and structure curation, data sets were reduced to less than half of the original dataset size. Detailed results from the QSAR-targeted data processing, structure curation, and splits into training and validation sets for the individual concentration thresholds can be viewed in Table 1.

3.2. QSAR modeling and selection

QSAR models were developed in LPDM based on all initial training sets using the three different modeling approaches. Furthermore, they were cross-validated by the LPDM two times five-fold CV (Table 2). In all but one case (agonism 10 μM), the cocktail ("3") models had the highest balanced accuracies, due to increased specificities and only slightly decreased sensitivities compared to the composite ("2") models. We, therefore, chose to continue with the cocktail models.

Table 1
Number of processed substances through QSAR-targeted data processing and structure curation, and resulting unique structures in the training and validation sets with the distribution of active and inactive.

	hCAR Antagonism			hCAR Agonism		
	All	Active	Inactive	All	Active	Inactive
Start set	9667	–	–	9667	–	–
10 μM						
After QSAR-targeted data processing	5868	160	5708	6977	108	6869
Acceptable structures	5430	136	5294	6422	106	6316
After luciferase inhibitors removal	5298	136	5162	6356	40	6316
After duplicates removal	4259	107	4152	5098	33	5065
Validation set (20% random for actives)	3313	21	3292	4812	7	4805
Training set (inactives = 10 * actives)	946	86	860	286	26	260
20 μM						
After QSAR-targeted data processing	5897	189	5708	7098	229	6869
Acceptable structures	5459	165	5294	6609	227	6382
After luciferase inhibitors removal	5327	165	5162	6492	110	6382
After duplicate removal	4277	128	4149	5147	84	5063
Validation set (20% random for actives)	3155	26	3129	4410	17	4393
Training Set (inactives = 10 * actives)	1122	102	1020	737	67	670
Expanded training set	1408	128	1280	924	84	840
Reduced validation set (only inactives)			2869			4223
30 μM						
After QSAR-targeted data processing	5920	212	5708	7126	257	6869
Acceptable structures	5481	187	5294	6636	254	6382
After luciferase inhibitors removal	5349	187	5162	6507	125	6382
After duplicate removal	4292	144	4148	5159	96	5063
Validation set (20% random for actives)	3027	29	2998	4312	19	4293
Training Set (inactives = 10 * actives)	1265	115	1150	847	77	770
40 μM						
After QSAR-targeted data processing	5925	217	5708	7129	260	6869
Acceptable structures	5485	191	5294	6639	257	6382
After luciferase inhibitors removal	5353	191	5162	6509	127	6382
After duplicate removal	4292	145	4147	5161	98	5063
Validation set (20% random for actives)	3016	29	2987	4303	20	4283
Training Set (inactives = 10 * actives)	1276	116	1160	858	78	780
50 μM						
After QSAR-targeted data processing	5956	248	5708	7267	398	6869
Acceptable structures	5515	221	5294	6774	392	6382
After luciferase inhibitors removal	5383	221	5162	6592	210	6382
After duplicate removal	4314	170	4144	5234	173	5061
Validation set (20% random for actives)	2818	34	2784	3716	35	3681
Training Set (inactives = 10 * actives)	1496	136	1360	1518	138	1380
Expanded training set	1870	170	1700	1903	173	1730
Reduced validation set (only inactives)			2444			3331
NUL[+]						
After QSAR-targeted data processing	5961	253	5708	7341	472	6869
Acceptable structures	5519	225	5294	6848	466	6382
After luciferase inhibitors removal	5387	225	5162	6641	259	6382
After duplicate removal	4315	172	4143	5271	212	5059
Validation set (20% random for actives)	2797	34	2763	3401	42	3359
Training Set (inactives = 10 * actives)	1518	138	1380	1870	170	1700

[+] For No Upper Limit (NUL) a concentration threshold cut-off was not set.

Our next step was to choose concentration thresholds for which we would expand the training sets with substances from the validation sets. For both the antagonism and the agonism initial models, the differences in the LPDM results for the cocktail models across all concentration thresholds were small and did not point to any concentration threshold standing out as giving better results (Table 2). For both antagonism and agonism, we chose the 20 μM and 50 μM models for further work with expansion and screening. The 20 μM models were chosen to predict higher potency substances and the 50 μM models were chosen to be able to make a wider screening of possible hCAR substances while still being able to interpret the predictions for minimum potency, which would not be possible with the NUL models. More specifically, an active prediction from a 50 μM model means that the chemical is predicted to have minimum 25% effect at maximum 50 μM concentration, i.e. a minimum potency. The 10 μM models were not chosen for higher potency prediction since especially for agonism the model had a quite small number of actives in the training set and was therefore more unstable and with smaller AD than the 20 μM model. The 20 μM and 50 μM expanded models for both hCAR antagonism and hCAR agonism were made using the expanded training sets presented in Table 1 and using the chosen cocktail approach in LPDM.

3.3. Predictive performance of the initial QSAR models

All initial antagonism and agonism QSAR models underwent a two-times five-fold DTU Food in-house CV procedure as well as external validation with the left out 20% actives and all remaining inactives (Table 3). For antagonism, DTU Food CV sensitivities ranged between 54.3% and 74.7%. With rather high SD values (i.e. 11.5% and 16.1%, for Ant-20 μM-QSAR and Ant-30 μM-QSAR, respectively), the CV sensitivities were, in fact, not significantly different from each other. The high standard deviations (SD) of sensitivities for the CV results of both endpoints are likely due to the 20% removal from the relatively small sets of actives, which removes valuable information for chemical classes not highly represented in the sets (Table 3). In contrast, CV specificities ranged between 92.4% and 97.2%. These high specificities had a

Table 2
Results from the LPDM 2 times five-fold CV of all initial models by three approaches: 1) single model, 2) composite model, 3) composite cocktail model to guide selections of modeling approach and concentration threshold.

Initial models	Approach	LPDMs two times five-fold cross validation results		
		Sensitivity (%)	Specificity (%)	Balanced accuracy (%)
Antagonism				
10 μM training set	1	54.1	99.0	76.6
	2	91.2	89.9	90.6
	3	86.6	97.4	92.0
20 μM training set	1	49.3	98.6	74.0
	2	87.8	88.3	88.1
	3	84.5	97.8	91.2
30 μM training set	1	52.9	98.5	75.7
	2	87.2	89.1	88.2
	3	84.9	97.2	91.1
40 μM training set	1	47.3	98.9	73.1
	2	87.8	87.3	87.6
	3	85.9	97.4	91.7
50 μM training set	1	46.0	99.0	72.5
	2	84.1	87.2	85.7
	3	83.5	97.9	90.7
No Upper Limit training set	1	38.5	99.0	68.8
	2	84.6	86.3	85.5
	3	82.4	97.1	89.8
Agonism				
10 μM training set	1	23.5	99.4	61.5
	2	94.7	88.2	91.5
	3	81.3	96.8	89.1
20 μM training set	1	35.8	98.9	67.4
	2	96.6	89.0	92.8
	3	92.5	96.9	94.7
30 μM training set	1	21.1	99.5	60.3
	2	91.4	88.8	90.1
	3	87.7	98.0	92.9
40 μM training set	1	41.8	98.2	70.0
	2	92.2	88.2	90.2
	3	90.2	96.5	93.4
50 μM training set	1	40.4	98.6	69.5
	2	92.0	87.6	89.8
	3	90.3	97.9	94.1
No Upper Limit training set	1	24.5	98.8	61.7
	2	91.0	87.2	89.1
	3	89.1	97.7	93.4

smaller SD range of values, (i.e. 1.2% to 2.9%), which reflected the larger inactive classes in the training sets. Balanced accuracies stayed between 75.7% and 84.0% (Table 3). For the agonism models, CV sensitivity was much lower for the Ag-10 μM-QSAR, (i.e. 28.2%) and ranged between 61.6% and 71.7% for the remaining models (Table 3). Relative to the antagonism values, SD values were in some cases higher (i.e. between 12.2% and 32.3%). Higher SD values reflected the rather small size of the training set active classes. For example, the Ag-10 μM-QSAR initial model had only 26 actives compared to the Ant-10 μM-QSAR initial model with 86 actives. On the other hand, the CV-derived specificities for the initial agonism models ranged between 90.0% and 93.0%. Similar to the antagonism models, these high specificities were attributed to larger inactive classes in the training sets. SD values ranged between 2.2% for the Ag-No-Upper-Limit initial model and

Table 3
Predictive performances for hCAR antagonism and agonism QSAR models.

Model (Cocktail Scaffolds)	DTU Food Cross-Validation† %							External Validation %						
	Sensitivity (SD, %)	Specificity (SD, %)	Balanced accuracy (SD,%)	TP*	FP*	TN*	FN*	Sensitivity	Specificity	Balanced accuracy	Coverage	TP	FP	TN	FN
Ant-10 μM-QSAR (initial)	69.9 (11.0)	93.1 (2.9)	81.5 (5.8)	75	80	1071	32	57.1	94.2	75.7	69.8	8	133	2164	6
Ant-20 μM-QSAR (initial)	54.3 (11.5)	97.2 (1.2)	75.7 (5.7)	55	41	1423	46	83.3	93.5	88.4	62.1	15	126	1814	3
Ant-20 μM-QSAR (expanded)	58.4 (14.2)	97.1 (0.7)	77.8 (7.2)	93	55	1877	65	–	93.4	–	61.3	–	116	1644	–
Ant-30 μM-QSAR (initial)	74.7 (16.1)	93.3 (2.1)	84.0 (8.6)	101	101	1406	35	76.9	92.1	84.5	67.4	10	160	1866	3
Ant-40 μM-QSAR (initial)	74.0 (12.6)	94.1 (2.4)	84.0 (5.9)	90	89	1429	32	56.3	93.5	74.9	66.0	9	129	1845	7
Ant-50 μM-QSAR (initial)	72.4 (14.3)	92.6 (2.4)	82.5 (7.9)	112	129	1599	44	55.0	92.5	73.8	59.9	11	125	1544	9
Ant-50 μM-QSAR (expanded)	72.4 (10.2)	91.6 (1.5)	82.0 (5.0)	130	173	1896	52	–	92.4	–	59.8	–	111	1350	–
Ant-No-Upper-Limit (initial)	70.9 (12.3)	92.4 (2.8)	81.6 (5.5)	111	129	1568	47	58.8	91.7	75.3	57.9	10	133	1469	7
Ag-10 μM-QSAR (initial)	28.2 (32.3)	91.0 (6.3)	55.0 (24.5)	7	25	246	17	100	95.3	97.6	61.7	3	140	2825	0
Ag-20 μM-QSAR (initial)	61.6 (22.3)	93.0 (3.2)	77.3 (12.1)	53	58	784	33	87.5	91.0	89.3	66.2	7	261	2652	1
Ag-20 μM-QSAR (expanded)	72.2 (13.6)	93.5 (1.2)	82.8 (7.0)	72	71	1013	26	–	91.5	–	66.4	–	238	2564	–
Ag-30 μM-QSAR (initial)	71.7 (17.5)	91.0 (4.8)	81.4 (8.8)	57	84	850	25	37.5	92.4	65.0	65.0	3	212	2582	5
Ag-40 μM-QSAR (initial)	67.9 (14.0)	91.6 (3.2)	80.3 (7.1)	62	73	922	30	71.4	92.6	82.0	65.7	10	209	2604	4
Ag-50 μM-QSAR (initial)	70.9 (12.2)	91.5 (3.3)	81.2 (6.6)	118	159	1708	48	77.3	90.9	84.1	67.4	17	226	2258	5
Ag-50 μM-QSAR (expanded)	78.4 (10.6)	91.4 (1.9)	84.9 (5.6)	147	199	2119	42	–	90.6	–	67.7	–	211	2045	–
Ag-No-Upper-Limit (initial)	70.8 (12.5)	90.0 (2.2)	80.4 (6.0)	133	215	1953	53	84.0	89.4	86.7	65.2	21	232	1962	4

† A two times five-fold cross-validation.
* TP: true positives, FP: false positives, TN: true negatives, FN: false negatives. The numbers given are the totals over all ten cross-validation sub-models.

6.3% for the Ag-10 μM-QSAR initial model. Balanced accuracies had a slightly lower range between 55.0% and 81.4%. From the results in Table 3 for the initial models for the different concentration thresholds there was no clear trend in the accuracy and robustness of the models, except that the two smallest agonism models, i.e. the Ag-10 μM-QSAR and Ag-20 μM-QSAR, were less robust than the other agonism models, likely due to the small active class. In fact, the SDs were overlapping between the different antagonism models and between the different agonism models. This lack of clear trends in the accuracies of the models for different concentration thresholds could potentially be due to: 1) the substances tested positive at higher concentrations could be just as distinct members of the active structure class as the substances tested positive at lower concentrations, 2) the uncertainties due to the modeling procedure could be bigger than the model accuracy difference between thresholds, 3) the experimental test results could have bigger uncertainties than the model accuracy differences between thresholds (despite a good Z' factor), or 4) a combination of some or all of the above. For the four expanded models the accuracies are not significantly higher than for the corresponding initial models, however except for the Ant-20 μM-QSAR the SD decreases in the expanded models, indicating higher robustness of the expanded models compared to the corresponding initial models.

Unlike the CV sensitivities, external validation with the 20% left-out actives and the remaining inactives resulted in a wider range of sensitivities for both antagonism and agonism models, though specificities remained relatively high and within a narrow range. For antagonism, the sensitivities ranged between 55.0% for the Ant-50 μM-QSAR initial model and 83.3% for the Ant-20 μM-QSAR initial model (Table 3). Specificities, however, were higher, i.e. between 91.7% for the Ant-No-Upper-Limit initial model and 94.2% for the Ant-10 μM-QSAR initial model. The range for balanced accuracy was also slightly wider (i.e. 73.8% to 88.4%) when compared to the CV balanced accuracies (i.e. 75.7% to 84.0%).

External validations for agonism gave sensitivities between 37.5% for the Ag-30 μM-QSAR initial model and 100.0% for the Ag-10 μM-QSAR initial model (Table 3). Specificities ranged between 89.4% for the Ag-No-Upper-Limit initial model and 95.3% for the Ag-10 μM-QSAR initial model. Balanced accuracies ranged between 65.0% and 97.6% for the Ag-30 μM-QSAR and for the Ag-10 μM-QSAR initial models, respectively.

For both antagonism and agonism, sensitivities lacked a clear trend, and varied much more than specificities. Variance was most likely due to the small number of actives in AD in the validation sets, as reflected in the true positive (TP) and false negative (FN) numbers (Table 3). On the other hand, specificities were much more stable due to the high number of inactives in the validation sets. High specificities are desirable because it means that the models do not generate a high percentage of false positives, which is important especially for endpoints with low prevalence, like hCAR antagonism and agonism. A possible trend for both antagonism and agonism models could be the slight decrease in specificity with increasing concentration thresholds. This slight trend could possibly indicate that including actives with lower potency in the training set leads to decreasing quality of the positive alerts in the models. Because of the uncertainties of the sensitivities, comparisons between CV results and external validation results are most relevant for specificity. As shown in Table 3, the specificities from the external validations are close to the specificities from the CVs, with the latter being a few percent higher (SD taken into consideration) in some cases.

Documentation of the four expanded models in QMRF reports is available from the Danish (Q)SAR Database [69].

3.4. Predictive performance of the 20 μM and 50 μM expanded QSAR models

All expanded QSAR models underwent a two-times five-fold DTU Food CV procedure as well as external validation for specificity using the inactives not included in the expanded training sets (Table 3). According to the DTU Food CV results, the expanded antagonism 20 μM and 50 μM models had sensitivities of 58.4% and 72.4%, specificities of 97.1% and 91.6%, and BAs of 77.8% and 82.0%. For the expanded agonism 20 μM and 50 μM models, sensitivities were 72.2% and 78.4%, specificities were 93.5% and 91.4%, and BAs were 82.8% and 84.9%. In all cases, the results for the expanded antagonism and agonism models were not dissimilar from the CV results of the corresponding initial models, taking SD into account. With the exception of sensitivity for the 20 μM antagonism model, for both antagonism and agonism, the SDs for sensitivity and specificity were, in all cases, lower for the expanded models than the corresponding initial models, thus, indicating more stability.

When looking at the external validation results, the expanded antagonism 20 μM and 50 μM models had specificities of 93.4% and 92.4%. For the expanded agonism 20 μM and 50 μM models, external validation specificities were 91.5% and 90.6%. Notably, the specificity results for all antagonism and agonism expanded models had a difference of < 1% from the external validation specificity results for the corresponding initial models. Although smaller than the 50 μM models, the 20 μM antagonism and agonism expanded models had slightly higher external validation specificities on these relatively large inactive validation sets. We, therefore, speculated that some less discriminating positive alerts entered into the 50 μM model compared to the 20 μM models, especially for antagonism, when weaker actives were included in the training set.

3.5. Screening results

We screened the U.S. EPA and REACH inventories using the expanded models for hCAR 20 μM and 50 μM antagonism and agonism (Table 4). The 20 μM and 50 μM antagonism models had coverages between 54.3% and 63.1% (with the 50 μM model having the lower coverages in accordance with external validation results from Table 3),

Table 4
The coverage (AD) and the number of active/inactive predictions of the U.S. EPA and REACH inventories in the expanded 20 μM and 50 μM hCAR antagonism models and the expanded 20 μM and 50 μM hCAR agonism models.

Datasets	No. substances	Ant-20 μM-QSAR (expanded)			Ant-50 μM-QSAR (expanded)		
		In AD	Active (%)	Inactive (%)	In AD	Active (%)	Inactive (%)
REACH[†]	80,086	63.1%	8,058 (16.0)	42,441 (84.0)	57.0%	7,680 (16.8)	37,931 (83.2)
U.S. EPA[††]	54,971	59.6%	5,175 (15.8)	27,577 (84.2)	54.3%	5,062 (17.0)	24,767 (83.0)
Datasets	No. substances	Ag-20 μM-QSAR (expanded)			Ag-50 μM-QSAR (expanded)		
		In AD	Active (%)	Inactive (%)	In AD	Active (%)	Inactive (%)
REACH[†]	80,086	61.1%	8,265 (16.9)	40,631 (83.1)	63.7%	10,289 (20.2)	40,711 (79.8)
U.S. EPA[††]	54,971	59.8%	5,731 (17.4)	27,121 (82.6)	63.0%	7,254 (20.9)	27,389 (79.1)

[†] REACH pre-registered and/or registered substances.
[††] U.S. EPA inventory of man-made chemical structures to which humans are potentially exposed.

and predicted between 15.8% and 17.0% substances as being active within their respective ADs from the two inventories. The 20 μM and 50 μM agonism models, on the other hand, had coverages between 59.8% and 63.7%, and predicted between 16.9% and 20.9% of the substances in AD from the two inventories to be active. The prevalences of actives among the tested Tox21 substances were between 2.5% and 4.0%, depending on the concentration threshold, after we did the QSAR-targeted data processing (prevalences can be derived from Table 1). While the experimentally tested Tox21 library of substances was not selected based on suspicion of hCAR antagonism or agonism, it is not known how well the library reflects the true prevalence of hCAR antagonists and agonists of the U.S. EPA and REACH inventories. As shown by the rather large external validations, both the applied antagonism and the agonism models showed high specificities (90.7% − 93.2%), indicating that they do not give many false positive predictions. Nevertheless, as can be derived from the external validation specificity results the models may still have false positive rates of around 6.8% − 9.3%. These false positive rates may explain some though not all of the high percents, (i.e. 15.8% − 20.9%), of active predictions for hCAR antagonism and agonism in the two large inventories, thereby indicating a possible high number of hCAR antagonists and agonists in the U.S. EPA and REACH inventories.

3.6. Statistical correlations of hCAR predictions with QSAR predictions for other endpoints

To explore possible biological pathways and toxicity properties affected by hCAR antagonists and agonists, REACH screening results for both the expanded hCAR 50 μM antagonism and agonism models were statistically correlated with QSAR predictions within the same REACH set using DTU Food-developed and commercial LPDM QSAR models for endpoints related e.g. to PXR, AhR, ER, AR, TPO, mutagenicity, cancer, sensitization and teratogenicity endpoints (Table 5).

In Table 5 A and B for correlations with hCAR antagonism and agonism predictions, respectively, the rows are sorted by column “hCAR Pos of M1 Pos (%)” in descending order. At the top of both lists, there are endpoints related to AhR, AR, ER, PXR/CYP3A4, cancer in rats and TPO. More specifically, hCAR positive predictions are found at a much higher percentage among positive rather than among negative predictions from models for these endpoints. To investigate the risk of chance correlations from the results given in Table 5, we calculated Pearson $\chi 2$ test (one degree of freedom) probabilities for being chance correlations and almost all of them are nearly 0 ($<$ 1E-10, results not shown). In addition we calculated false discovery rate by Benjamini-Yekutieli procedure with alpha equal to 1E-18 (i.e. extremely low risk of chance correlation). With the exception of a few (which had absolute MCC values $<$ 0.05 in Table 5 A and B), all correlations rejected the null hypothesis, i.e. were found not to be chance correlations by this test.

These statistical correlations of QSAR predictions between CAR and another endpoint indicate overlap in the chemical classes predicted positive for both endpoints as well as negative for both endpoints. To identify classes of substances that induce hCAR agonism, Lynch et al. grouped the Tox21 primary screen of 10,000 substances based on their structural similarities into 1,014 clusters, of which 62 clusters were significantly enriched with hCAR activators [53]. These clusters included the known CAR activators flavones and prazoles as well as a potential novel cluster containing nitazoxanide and tenonitrozole. It is known that CAR shares response elements with PXR [71] and that the two receptors have a significant overlap in their ligand specificities [53]. It is likewise known that CAR and PXR coordinate the metabolism and excretion of xenobiotic substances together with AhR [72]. As shown in Table 5, both hCAR antagonism and agonism correlated strongly with PXR and even more strongly with AhR. We therefore wanted to explore further how well QSAR predictions from the hCAR agonism model correlated with QSAR predictions from our earlier models on PXR and AhR agonism [36,46] in a trilateral analysis.

The tabulated counts in Table 6 show that 23,004 substances out of 80,086 REACH substances in the Danish (Q)SAR Database are in the common domain of the three models. Of these 23,004 substances, the vast majority were predicted negative by all three models, (i.e. 20,164 substances or 88%). hCAR had 561 positive predictions, for which the other two models predicted negative (corresponding to 2% of the common domain), hPXR had 738 (3%) positive predictions, for which the two other models predicted negative, and hAhR had 21 (0.1%) positive predictions, for which the two other models predicted negative. However, all three models gave positive predictions for 243 substances. hCAR and hPXR shared the biggest number of common positive predictions, namely the 243 substances which were positive in all three models, plus 1,233 substances only positive in the hCAR and hPXR models, in total 1,476 (6%). In addition, the AhR model rarely gave a positive prediction when the hCAR model gave a negative prediction, (i.e. 21 plus 3 cases).

4. Conclusions

Our study presents two main results. First, we developed an in-house QSAR-targeted data processing approach to extract Tox21 experimental results for QSAR development of different absolute minimum potency classes. Second, we developed, validated, and applied global, binary QSAR models for hCAR antagonism and agonism in vitro. All initial models were based on 10 μM, 20 μM, 30 μM, 40 μM, 50 μM and No Upper Limits threshold concentrations for both agonism and antagonism. All initial models underwent DTU Food CV and external validation and showed high specificities of around 90–95% and good balanced accuracies, (i.e. 75.7–84.0% for CV and 65.0–97.6% for external validation). For both antagonism and agonism, new, expanded models were developed for the 20 μM and 50 μM thresholds by incorporating the external validation set actives and ten times as many inactives into the training sets. These expanded models also underwent DTU Food CV and external validation for specificity, only as there were no additional actives for a full external validation. In all but one case, the results for the expanded antagonism and agonism models were slightly better compared to the corresponding initial models. However, taking SD into account, the results were not dissimilar from the DTU Food CV results related to their corresponding initial models. External validations of specificity showed similar performance between initial and the corresponding expanded models.

Our four expanded models were used to screen two, large chemical inventories from the U.S. and EU. Of the substances predicted within the ADs of the expanded models, the 20 μM agonism model predicted 8,265 (16.9%) REACH substances and 5,731 (17.4%) U.S. EPA substances to be positive; the 20 μM antagonism model predicted 8,058 (16.0%) REACH substances and 5,175 (15.8%) U.S. EPA substances to be positive. For antagonism, the 50 μM expanded model predicted a slightly lower number of positives than the 20 μM expanded model due to fewer substances in the AD. For agonism, the 50 μM expanded model predicted a slightly higher number of positives than the 20 μM expanded model. Finally, we explored if a number of biological pathways and toxicity properties correlated statistically, (i.e. not causally), with predicted hCAR antagonists and agonists. This was done by correlating QSAR predictions from the expanded hCAR 50 μM antagonism and agonism models with QSAR predictions for endpoints related to PXR, AhR, ER, AR, mutagenicity, sensitization, cancer and teratogenicity endpoints for 80,086 REACH substances contained in the Danish (Q)SAR Database. A number of strong statistical correlations were found to e.g. other receptor endpoints, genotoxicity endpoints and rodent carcinogenicity.

Predictions from the developed hCAR antagonism and agonism QSAR models may be utilized in the future to support e.g. screening and priority setting, read-across and IATA WoE assessments. Predictions for 650,000 substances from the expanded antagonism and agonism 20 μM

Table 5
A and B Correlations between prediction for REACH substances from the (A) hCAR Ant-50 μM (expanded) model and (B) hCAR Ag-50 μM (expanded) model and predictions from other DTU Food-developed and commercial LPDM QSAR models from the Danish (Q)SAR Database (only predictions in the domains of the pairwise correlated models were counted). “Pos” stands for positive predictions, ”Neg” stands for negative predictions. The lists in both A and B are sorted by the deccending order of the percent of “hCAR Pos of M1 Pos”

A hCAR Ant-50 μM (expanded) Model 1 (M1)	M1 Pos, hCAR Neg (#)	M1 Neg, hCAR Pos (#)	M1 Pos, hCAR Pos (#)	M1 Neg, hCAR Neg (#)	hCAR Pos of M1 Pos (%)	hCAR Pos of M1 Neg (%)	M1 Pos of hCAR Pos (%)	M1 Pos of hCAR Neg (%)	χ2	MCC
AhR agonism (rational) human in vitro	84	3435	408	27,383	83	11	11	0	2317	0,27
AhR agonism (random) human in vitro	340	2587	1064	24,347	76	10	29	1	5207	0,43
ER alpha Binding (human in vitro) - Balanced	812	783	2540	16,457	76	5	76	5	10,522	0,71
AR antagonism human in vitro	275	4006	836	29,277	75	12	17	1	3551	0,32
ER alpha agonism human in vitro	344	2080	731	14,135	68	13	26	2	2254	0,36
CYP3A4 induction human in vitro	623	2414	1247	29,655	67	8	34	2	6426	0,44
Frameshift Ames mutagens	1813	1815	3041	20,073	63	8	63	8	7899	0,54
Maximum Recommended Daily Dose (MRDD) in humans in vivo	1765	1294	2680	15,851	60	8	67	10	6539	0,55
PXR agonism rat in vitro	314	4736	459	32,621	59	13	9	1	1404	0,19
FDA cancer rat in vivo	2207	2013	2903	15,506	57	11	59	12	4778	0,46
PXR agonism human in vitro	1114	2063	1391	22,372	56	8	40	5	4507	0,41
PXR binding human in vitro (new model)	2791	1676	3376	22,222	55	7	67	11	7988	0,52
TPO inhibition rat in vitro	1990	1670	2354	24,713	54	6	58	7	7507	0,49
Bacterial reverse mutation test (Ames) in vitro	1468	4125	1646	29,112	53	12	29	5	3488	0,31
FDA cancer rodent in vivo	2113	1736	2288	15,879	52	10	57	12	4185	0,44
Liver Specific Cancer in Rat or Mouse in vivo	2300	1462	2480	13,366	52	10	63	15	3974	0,45
Syrian Hamster Embryo Cell Transformation in vitro	2857	840	3018	15,766	51	5	78	15	6547	0,54
Mutations in HGPRT Locus in Chinese Hamster Ovary Cells	995	1136	992	18,447	50	6	47	5	3950	0,43
Micronucleus Test in Mouse Erythrocytes in vivo	2601	1306	2274	17,284	47	7	64	13	4689	0,45
Teratogenic potential in humans in vivo	1624	1984	1042	13,607	39	13	34	11	1144	0,25
Respiratory sensitisation in humans in vivo	4127	290	2607	9966	39	3	90	29	3701	0,47
FDA cancer mouse in vivo	1536	4668	904	21,357	37	18	16	7	518	0,13
Sister Chromatid Exchange in Mouse Bone Marrow Cells in vivo	6216	696	3131	12,197	33	5	82	34	3003	0,37
Chromosome Aberrations in Chinese Hamster Lung Cells in vitro	5331	2596	2492	17,234	32	13	49	24	1315	0,22
Ashby structural alerts	4761	3618	2000	21,600	30	14	36	18	854	0,16
Mutations in Thymidine Kinase Locus in Mouse Lymphoma Cells in vitro	6985	2441	2508	12,298	26	17	51	36	345	0,12
Unscheduled DNA Synthesis in Rat Hepatocytes in vitro	4710	3367	1560	14,462	25	19	32	25	103	0,07
Potent Ames Mutagens (Reversions ≥ 10 Times Controls)	11,724	516	3428	4519	23	10	87	72	368	0,14
Allergic contact dermatitis in vivo	8589	759	2402	17,099	22	4	76	33	2161	0,27
Sex-Linked Recessive Lethal Test in Drosophila m. in vivo	2403	3856	608	16,339	20	19	14	13	2	0,01
CYP2C9 substrates in humans in vivo	636	3428	158	16,488	20	17	4	4	4	0,01
CYP2D6 substrates in humans in vivo	791	3516	137	18,539	15	16	4	4	1	−0,01
Direct Acting Ames Mutagens (without S9)	10,646	2810	1777	6896	14	29	39	61	711	−0,18
Chromosome Aberrations in Chinese Hamster Ovary Cells in vitro	4503	4204	700	16,183	13	21	14	22	137	−0,07
Dominant Lethal Mutations in Rodents in vivo	8033	3028	1146	5115	12	37	27	61	1439	−0,29
Comet Assay in Mouse in vivo	7093	3178	921	12,610	11	20	22	36	278	−0,11
Severe skin irritation in rabbit in vivo	7288	3263	923	14,722	11	18	22	33	200	−0,09
Base-Pair Ames Mutagens	11,482	2115	1207	7219	10	23	36	61	726	−0,18

(*continued on next page*)

Table 5 (*continued*)

A hCAR Ant-50 μM (expanded) Model 1 (M1)	M1 Pos, hCAR Neg (#)	M1 Neg, hCAR Pos (#)	M1 Pos, hCAR Pos (#)	M1 Neg, hCAR Neg (#)	hCAR Pos of M1 Pos (%)	hCAR Pos of M1 Neg (%)	M1 Pos of hCAR Pos (%)	M1 Pos of hCAR Neg (%)	χ2	MCC
B hCAR Ag-50 μM (expanded) Model 1 (M1)	M1 Pos, hCAR Neg (#)	M1 Neg, hCAR Pos (#)	M1 Pos, hCAR Pos (#)	M1 Neg, hCAR Neg (#)	hCAR Pos of M1 Pos (%)	hCAR Pos of M1 Neg (%)	M1 Pos of hCAR Pos (%)	M1 Pos of hCAR Neg (%)	χ2	MCC
AhR agonism (rational) human in vitro	64	4568	671	29,793	91	13	13	0	3447	0,31
AhR agonism (random) human in vitro	298	3458	1487	26,897	83	11	30	1	6697	0,46
AR antagonism human in vitro	532	5019	1375	30,648	72	14	22	2	4317	0,34
PXR agonism human in vitro	987	1390	2279	26,525	70	5	62	4	11,826	0,62
CYP3A4 induction human in vitro	852	2793	1783	32,380	68	8	39	3	8220	0,47
ER alpha Binding (human in vitro) - Balanced	1384	2435	2857	15,924	67	13	54	8	5623	0,50
FDA cancer rat in vivo	1728	2551	3361	18,773	66	12	57	8	6917	0,51
Frameshift Ames mutagens	1963	2663	3712	20,073	65	12	58	9	7524	0,51
Bacterial reverse mutation test (Ames) in vitro	1180	4928	2227	32,148	65	13	31	4	5815	0,38
ER alpha agonism human in vitro	678	4213	1150	13,035	63	24	21	5	1211	0,25
FDA cancer rodent in vivo	1695	2487	2770	17,741	62	12	53	9	5401	0,47
PXR binding human in vitro (new model)	2168	2109	3477	25,485	62	8	62	8	9756	0,54
TPO inhibition rat in vitro	2396	2858	3040	25,365	56	10	52	9	6615	0,44
Syrian Hamster Embryo Cell Transformation in vitro	3084	1071	3755	16,612	55	6	78	16	7445	0,55
Liver Specific Cancer in Rat or Mouse in vivo	2825	1759	3175	13,700	53	11	64	17	4212	0,44
Micronucleus Test in Mouse Erythrocytes in vivo	3093	1266	2717	18,457	47	6	68	14	5549	0,47
Mutations in Thymidine Kinase Locus in Mouse Lymphoma Cells in vitro	5397	1633	4569	14,851	46	10	74	27	4469	0,41
Mutations in HGPRT Locus in Chinese Hamster Ovary Cells	1351	2406	1133	18,285	46	12	32	7	1980	0,29
Maximum Recommended Daily Dose (MRDD) in humans in vivo	3126	1372	2528	18,472	45	7	65	14	4852	0,44
Ashby structural alerts	4603	4359	2983	23,536	39	16	41	16	2041	0,24
Sister Chromatid Exchange in Mouse Bone Marrow Cells in vivo	6384	256	3823	13,041	37	2	94	33	5082	0,47
Chromosome Aberrations in Chinese Hamster Ovary Cells in vitro	2798	4720	1620	18,039	37	21	26	13	525	0,14
Chromosome Aberrations in Chinese Hamster Lung Cells in vitro	5674	2871	3063	18,254	35	14	52	24	1789	0,24
PXR agonism rat in vitro	905	6256	466	34,692	34	15	7	3	348	0,09
Potent Ames Mutagens (Reversions ≥ 10 Times Controls)	11,826	499	5434	4451	31	10	92	73	900	0,20
FDA cancer mouse in vivo	2043	5901	796	22,677	28	21	12	8	84	0,05
Respiratory sensitisation in humans in vivo	5459	887	2040	8470	27	9	70	39	911	0,23
Teratogenic potential in humans in vivo	2608	2064	972	14,787	27	12	32	15	518	0,16
Comet Assay in Mouse in vivo	5754	2611	2101	14,872	27	15	45	28	500	0,14
Allergic contact dermatitis in vivo	10,610	1690	3253	15,867	23	10	66	40	1119	0,19
Unscheduled DNA Synthesis in Rat Hepatocytes in vitro	6143	4843	1654	13,350	21	27	25	32	85	−0,06
Severe skin irritation in rabbit in vivo	8712	4994	1329	12,972	13	28	21	40	781	−0,17
Sex-Linked Recessive Lethal Test in Drosophila m. in vivo	3035	5448	458	17,459	13	24	8	15	199	−0,09
Base-Pair Ames Mutagens	11,053	3794	1660	7588	13	33	30	59	1410	−0,24
CYP2D6 substrates in humans in vivo	1064	4330	152	20,729	13	17	3	5	19	−0,03
CYP2C9 substrates in humans in vivo	775	3984	108	18,128	12	18	3	4	19	−0,03
Dominant Lethal Mutations in Rodents in vivo	8371	3543	1148	5856	12	38	24	59	1667	−0,30
Direct Acting Ames Mutagens (without S9)	12,948	4996	1200	5546	8	47	19	70	4865	−0,44

Table 6
Correlations between prediction for REACH substances from the hCAR Ant-50 μM (expanded) model and hPXR agonism and hAhR agonism (rational model) QSAR predictions from the Danish (Q)SAR Database (only predictions in the common domain of all three correlated models were counted).

hCAR agonism	hPXR agonism	hAhR agonism (rational)	Count
Negative	Negative	Negative	20,164 (88%)
Negative	Negative	Positive	21 (0.1%)
Negative	Positive	Negative	738 (3.2%)
Negative	Positive	Positive	3 (0.01%)
Positive	Negative	Negative	561 (2%)
Positive	Negative	Positive	41 (0.1%)
Positive	Positive	Negative	1,233 (5%)
Positive	Positive	Positive	243 (1%)

and 50 μM models along with their QMRFs, as well as training and validation sets for all initial and expanded models, will be published in the free online Danish (Q)SAR Database, and the four expanded models will be made available via the free online Danish (Q)SAR Models website [69].

CRediT authorship contribution statement

K.K. Chinen: Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing - original draft, Writing - review & editing, Funding acquisition. **K. Klimenko:** Methodology, Data curation, Visualization, Writing - review & editing. **C. Taxvig:** Methodology. **N.G. Nikolov:** Methodology, Conceptualization, Supervision, Software, Data curation, Investigation, Formal analysis, Validation, Visualization, Writing - original draft, Writing - review & editing. **E.B. Wedebye:** Methodology, Conceptualization, Project administration, Resources, Funding acquisition, Methodology, Validation, Visualization, Data curation, Supervision, Writing - original draft, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We would like to thank Informatics leader, Senior scientist Ruili Huang from the U.S. National Institute of Health for support on the Tox21 data interpretation. This work was financially supported by the Danish Environmental Protection Agency as well as by the University of California, Los Angeles under the Dissertation Year Fellowship.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.comtox.2020.100121.

References

[1] P. Honkakoski, T. Sueyoshi, M. Negishi, Drug-activated nuclear receptors CAR and PXR, Ann. Med. 35 (2003) 172–182, https://doi.org/10.1080/07853890310008224.
[2] J.M. Maglich, A. Sluder, X. Guan, Y. Shi, D.D. McKee, K. Carrick, K. Kamdar, T.M. Willson, J.T. Moore, Comparison of complete nuclear receptor sets from the human, Caenorhabditis elegans and Drosophila genomes, Genom. Biol. 2 (2001), https://doi.org/10.1186/gb-2001-2-8-research0029 Research0029-0021.
[3] A. di Masi, E. De Marinis, P. Ascenzi, M. Marino, Nuclear receptors CAR and PXR: molecular, functional, and biomedical aspects, Mol. Aspects Med. 30 (2009) 297–343, https://doi.org/10.1016/j.mam.2009.04.002.
[4] E. Kachaylo, V. Pustylnyak, V. Lyakhovich, L. Gulyaeva, Constitutive androstane receptor (CAR) is a xenosensor and target for therapy, Biochem. (Moscow) 76 (2011) 1087, https://doi.org/10.1134/S0006297911100026.
[5] J. Sonoda, L. Pei, R.M. Evans, Nuclear receptors: decoding metabolic disease, FEBS Lett. 582 (2008) 2–9, https://doi.org/10.1016/j.febslet.2007.11.016.
[6] S.P. Alexander, J.A. Cidlowski, E. Kelly, N. Marrion, J.A. Peters, H.E. Benson, E. Faccenda, A.J. Pawson, J.L. Sharman, C. Southan, The concise guide to PHARMACOLOGY 2015/16: nuclear hormone receptors, Br. J. Pharmacol. 172 (2015) 5956, https://doi.org/10.1111/bph.1335.
[7] F. Molnár, J. Küblbeck, J. Jyrkkärinne, V. Prantner, P. Honkakoski, An update on the constitutive androstane receptor (CAR), Drug Metab. Drug Interact. 28 (2013) 79–93, https://doi.org/10.1515/dmdi-2013-0009.
[8] P. Wang, X. Xiao, K.-C. Chou, NR-2L: a two-level predictor for identifying nuclear receptor subfamilies based on sequence-derived features, PLoS One 6 (2011) e23505, https://doi.org/10.1371/journal.pone.0023505.
[9] X.C. Kretschmer, W.S. Baldwin, CAR and PXR: xenosensors of endocrine disrupters? Chem. Biol. Interact. 155 (2005) 111–128.
[10] J.T. Moore, L.B. Moore, J.M. Maglich, S.A. Kliewer, Functional and structural comparison of PXR and CAR. Biochim. Biophys. Acta (BBA)-General Sub. 1619 (2003) 235–238, https://doi.org/10.1016/S0304-4165(02)00481-6.
[11] M. Qatanani, D. Moore, CAR, the continuously advancing receptor, in drug metabolism and disease, Curr. Drug Metab. 6 (2005) 329–339.
[12] H. Gong, W. Xie, Chapter 7: Animal models of xenobiotic nuclear receptors and their utility in drug development, Wiley Online Library, 2008, pp. 185–210.
[13] P. Lu, W. Xie, Xenobiotic receptors in the crosstalk between drug metabolism and energy metabolism, Drug Metab. Dis.: Elsevier (2017) 257–278, https://doi.org/10.1016/B978-0-12-802949-7.00011-0.
[14] J. Yan, W. Xie, A brief history of the discovery of PXR and CAR as xenobiotic receptors, Acta Pharmaceut. Sin. B 6 (2016) 450–452, https://doi.org/10.1016/j.apsb.2016.06.011.
[15] J.G. DeKeyser, C.J. Omiecinski, Constitutive Androstane Receptor, Elsevier, The Boulevard, Langford Lane, Kidlington OX5 1GB, United Kingdom, 2010, pp. 169–181.
[16] M.M. Tabb, B. Blumberg, New modes of action for endocrine-disrupting chemicals, Mol. Endocrinol. 20 (2006) 475–482, https://doi.org/10.1210/me.2004-0513.
[17] C. Xu, C.Y.-T. Li, A.-N.T. Kong, Induction of phase I, II and III drug metabolism/transport by xenobiotics, Arch. Pharmacal. Res. 28 (2005) 249, https://doi.org/10.1007/BF02977789.
[18] A. Poso, P. Honkakoski, Ligand recognition by drug-activated nuclear receptors PXR and CAR: structural, site-directed mutagenesis and molecular modeling studies, Mini Rev. Med. Chem. 6 (2006) 937–943, https://doi.org/10.2174/138955706777935008.
[19] J. Hakkola, C. Bernasconi, S. Coecke, L. Richert, T. Andersson, O. Pelkonen, Cytochrome P450 induction and xeno-sensing receptors pregnane X receptor, constitutive androstane receptor, aryl hydrocarbon receptor and peroxisome proliferator-activated receptor α at the crossroads of toxicokinetics and toxicodynamics, Basic Clin. Pharmacol. Toxicol. 123 (2018) 42–50, https://doi.org/10.1111/bcpt.13004.
[20] G.A. Francis, E. Fayard, F. Picard, J. Auwerx, Nuclear receptors and the control of metabolism, Annu. Rev. Physiol. 65 (2003) 261–311, https://doi.org/10.1146/annurev.physiol.65.092101.142528.
[21] M. Qatanani, J. Zhang, D.D. Moore, Role of the constitutive androstane receptor in xenobiotic-induced thyroid hormone metabolism, Endocrinology 146 (2005) 995–1002.
[22] M.D. Miller, K.M. Crofton, D.C. Rice, R.T. Zoeller, Thyroid-disrupting chemicals: interpreting upstream biomarkers of adverse outcomes, Environ. Health Perspect. 117 (2009) 1033–1041, https://doi.org/10.1289/ehp.0800247.
[23] K.P. Friedman, M.E. Gilbert, K.M. Crofton, AOP: 8: Upregulation of Thyroid Hormone Catabolism via Activation of Hepatic Nuclear Receptors, and Subsequent Adverse Neurodevelopmental Outcomes in Mammals, 2016. Available: https://aopwiki.org/aops/8 [accessed September 12 2019].
[24] B. Dong, P.K. Saha, W. Huang, W. Chen, L.A. Abu-Elheiga, S.J. Wakil, R.D. Stevens, O. Ilkayeva, C.B. Newgard, L. Chan, Activation of nuclear receptor CAR ameliorates diabetes and fatty liver disease, Proc. Nat. Acad. Sci. 106 (2009) 18831–18836, https://doi.org/10.1073/pnas.0909731106.
[25] R.C. Peffer, K.A. Bailey, L.-K. Lake, Kristin , B.G. Lake, AOP: 107: Constitutive androstane receptor activation leading to hepatocellular adenomas and carcinomas in the mouse and the rat, 2018. Available: https://aopwiki.org/aops/107 [accessed September 12 2019].
[26] W. Huang, J. Zhang, M. Washington, J. Liu, J.M. Parant, G. Lozano, D.D. Moore, Xenobiotic stress induces hepatomegaly and liver tumors via the nuclear receptor constitutive androstane receptor, Mol. Endocrinol. 19 (2005) 1646–1653, https://doi.org/10.1210/me.2004-0520.
[27] Y. Yamamoto, R. Moore, T.L. Goldsworthy, M. Negishi, R.R. Maronpot, The orphan nuclear receptor constitutive active/androstane receptor is essential for liver tumor promotion by phenobarbital in mice, Cancer Res. 64 (2004) 7197–7200, https://doi.org/10.1158/0008-5472.CAN-04-1459.
[28] U.S. Tox21 Program, Tox21 Data Browser: Assays, 2019. Available: https://tripod.nih.gov/tox21/assays/ [accessed September 19 2019].
[29] M. Angrish, B. Chorley, AOP: 58: NR1I3 (CAR) suppression leading to hepatic steatosis, 2018. Available: https://aopwiki.org/aops/58 [accessed September 12 2019].
[30] OECD, OECD Quantitative Structure-Activity Relationships Project [(Q)SARs], 2018. Available: http://www.oecd.org/chemicalsafety/risk-assessment/oecdquantitativestructure-activityrelationshipsprojectqsars.htm [accessed December 3, 2018 2018].
[31] ECHA, QSAR models, 2018. Available: https://echa.europa.eu/support/registration/how-to-avoid-unnecessary-testing-on-animals/qsar-models [accessed

December 3 2018].
[32] E. Benfenati, A. Lombardo, A. Roncaglioni, Chapter 9: Computational toxicology and REACH, Computational Toxicology: Risk Assessment for Chemicals, Hoboken, NJ, 2018, pp. 245–268, https://doi.org/10.1002/9781119282594.ch9.
[33] E. Benfenati, Theory, guidance and applications on QSAR and REACH, 2012. E-book available at: http://wwworchestra-qsar eu/documents/333.
[34] ECHA, Guidance on information requirements and chemical safety assessment. Chapter r.6: Qsars and grouping of chemicals. Helsinki, Finland, 2008. Available: https://echa.europa.eu/documents/10162/13632/information_requirements_r6_en.pdf [accessed September 15, 2019].
[35] OECD, Guidance Document on the Validation of (Quantitative) Structure-Activity Relationships [(Q)SAR] Models. ENV/JM/MONO(2007)2. OECD, Paris, France, 2007. Available: http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono(2007)2&doclanguage=en [accessed September 15, 2019].
[36] S.A. Rosenberg, M. Xia, R. Huang, N.G. Nikolov, E.B. Wedebye, M. Dybdahl, QSAR development and profiling of 72,524 REACH substances for PXR activation and CYP3A4 induction, Comput. Toxicol. 1 (2017) 39–48, https://doi.org/10.1016/j.comtox.2017.01.001.
[37] U.S. FDA, M7(R1) Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals To Limit Potential Carcinogenic Risk: ICH Guidance for Industry. U.S. FDA (U.S. Food & Drug Administration), Silver Spring, MD, 2018. Available: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/m7r1-assessment-and-control-dna-reactive-mutagenic-impurities-pharmaceuticals-limit-potential [accessed January 19, 2020].
[38] NTP, Tox21: Toxicology in the 21st Century (Tox21), 2019. Available: https://ntp.niehs.nih.gov/results/tox21/index.html [accessed November 26 2019].
[39] NTP, Tox21: Toxicology in the 21st Century (Tox21). Available: https://ntp.niehs.nih.gov/results/tox21/index.html [accessed February 23 2019].
[40] Y. Matsuzaka, Y. Uesawa, Optimization of a deep-learning method based on the classification of images generated by parameterized deep snap a novel molecular-image-input technique for quantitative structure-activity relationship (QSAR) analysis, Front. Bioeng. Biotechnol. 7 (2019) 65, https://doi.org/10.3389/fbioe.2019.00065.
[41] Y. Matsuzaka, Y. Uesawa, Prediction model with high-performance constitutive androstane receptor (CAR) using deepsnap-deep learning approach from the Tox21 10K compound library, Int. J. Mol. Sci. 20 (2019) 4855, https://doi.org/10.3390/ijms20194855.
[42] J. Jyrkkärinne, B.r. Windshügel, T. Rönkkö, A.J. Tervo, J. Küblbeck, M. Lahtela-Kakkonen, W. Sippl, A. Poso, P. Honkakoski, Insights into ligand-elicited activation of human constitutive androstane receptor based on novel agonists and three-dimensional quantitative structure− activity relationship, J. Med. Chem. 51 (2008) 7181–7192, https://doi.org/10.1021/jm800731b.
[43] A.M. Dring, L.E. Anderson, S. Qamar, M.A. Stoner, Rational quantitative structure–activity relationship (RQSAR) screen for PXR and CAR isoform-specific nuclear receptor ligands, Chem. Biol. Interact. 188 (2010) 512–525, https://doi.org/10.1016/j.cbi.2010.09.018.
[44] H. Kato, N. Yamaotsu, N. Iwazaki, S. Okamura, T. Kume, S. Hirono, Precise prediction of activators for the human constitutive androstane receptor using structure-based three-dimensional quantitative structure–activity relationship methods, Drug Metab. Pharmacokinet. 32 (2017) 179–188, https://doi.org/10.1016/j.dmpk.2017.02.001.
[45] U.S. EPA, CoMPARA: Collaborative Modeling Project for Androgen Receptor Activity (SOT), 2019. Available: https://cfpub.epa.gov/si/si_public_record_report.cfm?Lab=NCCT&dirEntryId=339684 [accessed September 12 2019].
[46] K.O. Klimenko, S.A. Rosenberg, M. Dybdahl, E.B. Wedebye, N.G. Nikolov, QSAR modelling of a large imbalanced aryl hydrocarbon activation dataset by rational and random sampling and screening of 80,086 REACH pre-registered and/or registered substances, PLoS One 14 (2019), https://doi.org/10.1371/journal.pone.0213848.
[47] S.A. Rosenberg, E.D. Watt, R.S. Judson, S. Simmons, K.P. Friedman, M. Dybdahl, N.G. Nikolov, E.B. Wedebye, QSAR models for thyroperoxidase inhibition and screening of US and EU chemical inventories, Comput. Toxicol. 4 (2017) 11–21, https://doi.org/10.1016/j.comtox.2017.07.006.
[48] A.M. Vinggaard, J. Niemelä, E.B. Wedebye, G.E. Jensen, Screening of 397 chemicals and development of a quantitative structure− activity relationship model for androgen receptor antagonism, Chem. Res. Toxicol. 21 (2008) 813–823, https://doi.org/10.1021/tx7002382.
[49] NIH, PubChem: Bioassay, 2018. Available: https://pubchem.ncbi.nlm.nih.gov/# [accessed December 8, 2018 2018].
[50] NIH, PubChem AID 1224892: qHTS assay to identify small molecule agonists of the constitutive androstane receptor (CAR) signaling pathway: Summary, 2019. Available: https://pubchem.ncbi.nlm.nih.gov/bioassay/1224892 [accessed September 20 2019].
[51] NIH, PubChem AID 1224893: qHTS assay to identify small molecule antagonists of the constitutive androstane receptor (CAR) signaling pathway: Summary, 2019. Available: https://pubchem.ncbi.nlm.nih.gov/bioassay/1224893 [accessed September 12 2019].
[52] R. Huang, A quantitative high-throughput screening data analysis pipeline for activity profiling, Human Press, New York, NY, 2016, pp. 111–122, https://doi.org/10.1007/978-1-4939-6346-1_12.
[53] C. Lynch, B. Mackowiak, R. Huang, L. Li, S. Heyward, S. Sakamuru, H. Wang, M. Xia, Identification of modulators that activate the constitutive androstane receptor from the Tox21 10K compound library, Toxicol. Sci. (2018).
[54] NIH, PubChem AID 1224838: qHTS assay to identify small molecule antagonists of the constitutive androstane receptor (CAR) signaling pathway, 2019. Available: https://pubchem.ncbi.nlm.nih.gov/bioassay/1224838 [accessed September 12 2019].
[55] NIH, Pubchem AID 1224839: qHTS assay to identify small molecule agonists of the constitutive androstane receptor (CAR) signaling pathway, 2019. Available: https://pubchem.ncbi.nlm.nih.gov/bioassay/1224839 [accessed January 22 2020].
[56] J.-H. Zhang, T.D. Chung, K.R. Oldenburg, A simple statistical parameter for use in evaluation and validation of high throughput screening assays, J. Biomol. Screen. 4 (1999) 67–73, https://doi.org/10.1177/108705719900400206.
[57] NIH, Pubchem AID 1224839: qHTS assay to identify small molecule agonists of the constitutive androstane receptor (CAR) signaling pathway, 2019. Available: https://pubchem.ncbi.nlm.nih.gov/bioassay/1224839 [accessed December 8 2020].
[58] D.S. Auld, J. Inglese, Interferences with luciferase reporter enzymes, in: G.S. Sittampalam, A. Grossman, K. Brimacombe et al., (Eds.), Assay Guidance Manual [Internet], Eli Lilly & Company and the National Center for Advancing Translational Sciences, Bethesda (MD), 2016 [Updated 2018]. https://www.ncbi.nlm.nih.gov/books/.
[59] NIH, PubChem AID 1224835: qHTS assay to identify small molecule inhibitors of firefly luciferase, 2019. Available: https://pubchem.ncbi.nlm.nih.gov/bioassay/1224835 [accessed September 20 2019].
[60] N. Nikolov, V. Grancharov, G. Stoyanova, T. Pavlov, O. Mekenyan, Representation of chemical information in OASIS centralized 3D database for existing chemicals, J. Chem. Inf. Model. 46 (2006) 2537–2551.
[61] Leadscope, Leadscope, 2018. Available: http://www.leadscope.com [accessed September 15 2019].
[62] G. Roberts, G.J. Myatt, W.P. Johnson, K.P. Cross, P.E. Blower, LeadScope: software for exploring large sets of screening data, J. Chem. Inf. Comput. Sci. 40 (2000) 1302–1314, https://doi.org/10.1021/ci0000631.
[63] Leadscope, Leadscope: Enterprise User Manual Version 2.2, Leadscope, Inc., Columbus, OH, 2002. Available: www.leadscope.com.
[64] L.G. Valerio Jr, C. Yang, K.B. Arvidson, N.L. Kruhlak, A structural feature-based computational approach for toxicology predictions, Expert Opin. Drug Metab. Toxicol. 6 (2010) 505–518, https://doi.org/10.1517/17425250903499286.
[65] L. Breiman, Bagging predictors, Mach. Learning 24 (1996) 123–140, https://doi.org/10.1007/BF00058655.
[66] P. Jaccard, Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines, Bull. Soc. Vaudoise Sci. Nat. 37 (1901) 241–272.
[67] J. Cooper II, R. Saracci, P. Cole, Describing the validity of carcinogen screening tests, Br. J. Cancer 39 (1979) 87.
[68] DTU Food: National Food Institute, User Manual for the Danish (Q)SAR Database, 2018. Available: http://qsardb.food.dtu.dk/Danish_QSAR_Database_Draft_User_manual.pdf [accessed December 10 2018].
[69] DTU Food: National Food Institute, Danish (Q)SAR Database, 2019. Available: http://qsar.food.dtu.dk [accessed September 13 2019].
[70] K. Mansouri, N. Kleinstreuer, E. Watt, J. Harris, CoMPARA: Collaborative Modeling Project for Androgen Receptor Activity, 2017. Available: https://www.researchgate.net/publication/316606155_CoMPARA_Collaborative_Modeling_Project_for_Androgen_Receptor_Activity.
[71] M.T. Cherian, S.C. Chai, T. Chen, Small-molecule modulators of the constitutive androstane receptor, Expert Opin. Drug Metab. Toxicol. 11 (2015) 1099–1114.
[72] L. Li, J.D. Stanton, A.H. Tolson, Y. Luo, H. Wang, Bioactive terpenoids and flavonoids from Ginkgo biloba extract induce the expression of hepatic drug-metabolizing enzymes through pregnane X receptor, constitutive androstane receptor, and aryl hydrocarbon receptor-mediated pathways, Pharm. Res. 26 (2009) 872.