Toward Predicting Drug-Induced Liver Injury: Parallel Computational Approaches to Identify Multidrug Resistance Protein 4 and Bile Salt Export Pump Inhibitors^S

Matthew A. Welch, Kathleen Köck, Thomas J. Urban, Kim L. R. Brouwer, and Peter W. Swaan

Department of Pharmaceutical Sciences, University of Maryland, Baltimore, Maryland (M.A.W., P.W.S.); Division of Pharmacotherapy and Experimental Therapeutics, UNC Eshelman School of Pharmacy, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina (K.K., T.J.U., K.L.R.B.); Center for Human Genome Variation, Duke University Medical Center, Durham, North Carolina (T.J.U.)

Received February 26, 2015; accepted February 27, 2015

ABSTRACT

Drug-induced liver injury (DILI) is an important cause of drug toxicity. Inhibition of multidrug resistance protein 4 (MRP4), in addition to bile salt export pump (BSEP), might be a risk factor for the development of cholestatic DILI. Recently, we demonstrated that inhibition of MRP4, in addition to BSEP, may be a risk factor for the development of cholestatic DILI. Here, we aimed to develop computational models to delineate molecular features underlying MRP4 and BSEP inhibition. Models were developed using 257 BSEP and 86 MRP4 inhibitors and noninhibitors in the training set. Models were externally validated and used to predict the affinity of compounds toward BSEP and MRP4 in the DrugBank database. Compounds with a score above the median fingerprint threshold were considered to have significant inhibitory effects on MRP4 and BSEP. Common feature pharmacophore models were developed

for MRP4 and BSEP with LigandScout software using a training set of nine well characterized MRP4 inhibitors and nine potent BSEP inhibitors. Bayesian models for BSEP and MRP4 inhibition/noninhibition were developed with cross-validated receiver operator curve values greater than 0.8 for the test sets, indicating robust models with acceptable false positive and false negative prediction rates. Both MRP4 and BSEP inhibitor pharmacophore models were characterized by hydrophobic and hydrogen-bond acceptor features, albeit in distinct spatial arrangements. Similar molecular features between MRP4 and BSEP inhibitors may partially explain why various drugs have affinity for both transporters. The Bayesian (BSEP, MRP4) and pharmacophore (MRP4, BSEP) models demonstrated significant classification accuracy and predictability.

Introduction

Drug-induced liver injury (DILI) is an important cause of drug toxicity and a major reason for withdrawal of drugs from the market (Abboud and Kaplowitz, 2007) or attrition of drug candidates in late development stages, which can be extremely costly. Unfortunately, current in vitro screens or in vivo preclinical studies cannot accurately predict the potential of compounds to cause hepatotoxicity. DILI remains a major concern in drug discovery and clinical development. This obstacle has necessitated a search for alternative technologies, such as computational approaches, to decrease the risk of DILI-associated late-stage failures.

Despite extensive research, the underlying mechanisms of DILI are not well understood. However, it is clear that compound-related properties as well as individual patient characteristics affect the

occurrence of DILI. Formation of reactive metabolites, mitochondrial impairment, and inhibition of canalicular bile acid transport mediated by the bile salt export pump (BSEP) (e.g., troglitazone, bosentan, and erythromycin) (Stieger et al., 2000; Fattinger et al., 2001; Kostrubsky et al., 2003) are known risk factors for the development of DILI in humans. This has been substantiated by large-scale in vitro screening studies revealing that drugs that cause cholestatic DILI have higher potencies as well as frequencies of BSEP inhibition compared with drugs that are not liver toxic or that cause hepatocellular DILI (Morgan et al., 2010; Dawson et al., 2012). BSEP is located at the canalicular membrane of the hepatocyte, where it is involved in the excretion of bile acids into bile under physiologic conditions (Noe et al., 2002). The importance of this protein in bile acid homeostasis is emphasized by the observation that mutations in the BSEP gene *ABCB11* have been associated with progressive familial intrahepatic cholestasis type 2. Although BSEP inhibition may explain bile acid-mediated DILI liability for a large proportion of compounds, a subset of hepatotoxic drugs remains that cannot be explained by BSEP inhibition alone.

In addition to canalicular BSEP, multidrug resistance protein 4 (MRP4) is a bile acid efflux protein localized at the basolateral membrane of hepatocytes. While hepatic expression is low under normal physiologic conditions, MRP4 up-regulation has been demonstrated under cholestatic conditions. MRP4 is hypothesized to

This work was supported by the National Institutes of Health National Institute of General Medical Sciences and National Institutes of Health National Institute of Diabetes and Digestive and Kidney Diseases [Grants R01 GM041935, R01 DK61425, U01DK065201, and a National Institute of General Medical Sciences Collaborative Science Supplement]; and by Deutsche Forschungsgemeinschaft [Grant Ko4186/1-1]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

dx.doi.org/10.1124/dmd.114.062539.

^SThis article has supplemental material available at dmd.aspetjournals.org.

ABBREVIATIONS: AUC, area under the curve; BSEP, bile salt export pump; DHEAS, dehydroepiandrosterone sulfate; DILI, drug-induced liver injury; ECFP, extended-connectivity fingerprints; FCFP, functional-class fingerprints; HBA, hydrogen bond acceptor; MCC, Matthews correlation coefficient; MRP4, multidrug resistance protein 4; PCA, principal component analysis; PI, protease inhibitor; ROC, receiver operator curve.

serve as a back-up system for bile acid efflux from hepatocytes into sinusoidal blood when the normal vectorial transport of bile acids from hepatocytes into bile is compromised (Scheffer et al., 2002; Teng and Piquette-Miller, 2007; Gradhand et al., 2008; Chai et al., 2012). Recently, we screened 88 drugs (BSEP inhibitors and noninhibitors) for inhibition of MRP4-mediated transport of the prototypical substrate [³H]-dehydroepiandrosterone sulfate (DHEAS) and discovered potent MRP4 inhibition among cholestatic BSEP noninhibitors. A statistically significant relationship was observed between the potency of MRP4 inhibition and the probability of cholestatic classification. For each 1% increase in MRP4 inhibition, the probability that a drug was cholestatic increased by 3.1%. Interestingly, many BSEP inhibitors were also MRP4 inhibitors. These data suggested that MRP4 inhibition may serve as a confounding factor in BSEP-mediated DILI or, in some cases, lead to DILI in the absence of BSEP inhibition. Thus, MRP4 inhibition may be an additional risk factor for the development of cholestatic DILI.

The role of hepatic bile acid transport inhibition in the etiology of DILI emphasizes the urgent need to develop screening tools to accurately predict drug–bile acid transporter interactions. While *in vitro* membrane vesicle assays have been developed for BSEP and MRP4 screening, use of these assays early in drug development is time consuming, labor and resource intensive, and requires the physical availability of compounds (including metabolites) for testing. An alternative approach to *in vitro* testing is the use of computational models to predict drug–bile acid transporter interactions and aid in identifying transporter-associated DILI early in the drug discovery process. For example, pharmacophore models have been used in ligand-based drug design to define the key structural characteristics that a molecule must possess to bind to the biologic target (Ekins et al., 2012). Since models for BSEP have been reported recently (Pedersen et al., 2013; Ritschel et al., 2014), the aim of the current study was to develop a comprehensive model for MRP4 inhibition and evaluate its predictive ability. In addition, we developed Bayesian models to delineate molecular features underlying both MRP4 and BSEP inhibition. These *in silico* models were used to identify potential novel MRP4 inhibitors by virtual screening of an existing database and classify drugs as BSEP and MRP4 inhibitors in an effort to correlate these features with DILI incidence.

Materials and Methods

Dataset Composition. A dataset of 86 compounds derived from Köck et al. (2014) was used for MRP4 inhibition modeling, and a dataset of 257 compounds derived from Dawson et al. (2012) and Morgan et al. (2013) was used to develop BSEP inhibition models. The compounds in these datasets were structurally diverse and from various therapeutic classes. They were classified as cholestatic or noncholestatic, according to the DILI type reported in the literature. The compounds were further classified as active for the specified transporter if they had an $IC_{50} \leq 135 \mu M$ for BSEP or a percent inhibition $\geq 21\%$ compared with control at $100 \mu M$ for MRP4; otherwise, they were classified as inactive against that transporter. The MRP4 classifications are based on findings by the Köck et al. that compounds that inhibit by at least 21% have a 50% chance of being cholestatic and the rationale for the BSEP classifications is to identify inhibitor compounds with both potent and moderate cholestatic risk, similar to Morgan et al. These classifications enable the identification of compounds that should be investigated further for their potential to cause cholestasis.

In addition to MRP4 and BSEP datasets, a database of 1510 U.S. Food and Drug Administration–approved drugs was retrieved from DrugBank (<http://www.drugbank.ca>) (Law et al., 2014). The database was modified by removing ionic salts and large polymeric drugs and proteins, resulting in a catalog of 1488 drugs.

Training and Test Set Generation. The MRP4 and BSEP databases were separated into training and test sets by randomly dividing two-thirds of the compounds into the training set and the other one-third into the test set

(Supplemental Table 1). Table 1 enumerates the number of compounds in each set based on the respective transporter as well as the number of compounds that were classified as inhibitors and noninhibitors.

The Bayesian modeling of MRP4 and BSEP used all the compounds in their respective training sets. In contrast, the pharmacophore models were developed using a subset of compounds from the training set. The MRP4 pharmacophore subset was based on clustering of the training set, which produced a subset of nine compounds. Analogously, the BSEP pharmacophore subset contained the strongest inhibitors, also producing a subset of nine compounds. Details of subset generation and composition are explained further within the pharmacophore creation methods section.

The conformational models for pharmacophore creation were produced in LigandScout using the OMEGA conformer generator with the default best quality settings, which produced a maximum of 500 conformations per molecule with an energy window of 10 kcal/mol and root-mean-square threshold difference of 0.4 to identify unique conformers. The common feature pharmacophore was generated using the default settings in LigandScout for ligand-based shared-feature pharmacophore creation with a feature tolerance scale of 1.0.

Principal Component Analysis of Training, Test Set, and DrugBank Molecules. The three-dimensional molecular structures of 86 MRP4 inhibitors and noninhibitors and the 257 BSEP inhibitors were obtained from PubChem (<http://www.ncbi.nlm.nih.gov/pccompound>). Principal component analysis (PCA) plots of each transporter’s training and test sets were produced to ensure that the two sets were representative of each other in terms of molecular descriptors. In addition, the training sets were compared with the modified DrugBank database (see above) to ensure that the training set was representative of currently approved drugs and had predictive power in that chemical space. The PCA plots were generated based on eight molecular descriptors for each drug: ALogP, molecular weight, number of hydrogen bond donors, number of hydrogen bond acceptors, number of rotatable bonds, number of rings, number of aromatic rings, and molecular fractional polar surface area. The molecular descriptors and PCA plots were generated within Discovery Studio 4.0 (DS 4.0) (Accelrys, Inc., San Diego, CA). The two-dimensional plots (Fig. 2) represent only the first two principal components of each comparison for visual clarity.

Common Feature Pharmacophore Generation and Validation. Ligand-based pharmacophores and conformational models were generated using LigandScout (version 3.12, build 20130912; Inte:Ligand, Vienna, Austria) [1,2] with default settings. The pharmacophore models for MRP4 inhibition were generated from a subset of drugs produced by clustering the training sets based on similarity of the pharmacophore radial distribution function. Drugs with similar pharmacophore features were clustered together, and the most potent inhibitors of each cluster were included in the subset to train the pharmacophore model. The rationale of clustering is to generate a pharmacophore from a smaller training set while still maintaining the structural diversity of the original training set. If the common pharmacophore creation failed or produced a pharmacophore with less than three features, the drug that failed to align was removed from the training set. Of the 10 pharmacophores generated per training set, the pharmacophore that aligned with the most compounds in the training and test set and had the highest score was selected for further testing.

The MRP4 common feature pharmacophore was validated within LigandScout through virtually screening the test set for its ability to distinguish actives (i.e., drugs with $\geq 21\%$ MRP4 inhibitory activity) from inactives (i.e., drugs with $< 21\%$ MRP4 inhibitory activity). The conformational models of the test set were generated in an identical manner as the training set. Drugs that aligned with all three pharmacophore features were predicted to be active MRP4 inhibitors.

TABLE 1
Composition of training and test set

Transport Model	MRP4	BSEP
Training set total (inhibitors/noninhibitors)	57 (34/23)	171 (43/128)
Test set total (inhibitors/noninhibitors)	29 (17/12)	86 (22/64)
Pharmacophore training subset ^a	9	9
Pharmacophore test set ^b	77	247

^aSubset of drugs from the training set used to develop the pharmacophore.

^bDrugs not included in the pharmacophore training set were moved to the test set.

Evaluation of Model Performance. The Matthews correlation coefficient (MCC) was used to determine the relative predictive ability of the pharmacophore and Bayesian models. MCC ranges from -1 (no correlation) to +1 (full correlation) and is calculated as follows:

The ROC curve is another method for evaluating models. It is a two-dimensional plot that graphs the sensitivity of a model, its true positive rate, versus the reverse percentage of the specificity of the model, its false positive rate, by the ranked order of the pharmacophore-fit scores. One of the abilities of the ROC curve is the use of the AUC when comparing the ability of different models to correctly classify true positives above false positives. Starting from the bottom left corner, the graph plots the percentage of the actives in the test set properly classified as active, which is defined as the sensitivity or true positive rate, versus the percentage of the inactives improperly classified as active, which is defined as the reverse specificity or false positive rate. In addition to the AUC, the ROC can be used to set a score cutoff that optimizes the tradeoff between sensitivity and specificity.

Results

A

MRP4 inhibitors (45) and BSEP inhibitors (31) overlap by 26 drugs.

B

Cholestatic (14), MRP4 inhibitor (15), and BSEP inhibitor (1) overlap by 1 drug.

Table 2 shows the AUCs of Bayesian models based on the 10-fold cross-validation with training set compounds. AUC values range between 0 and 1, with 0.5 indicating 50% correct prediction and 1 indicating a perfect match between observed and predicted data (Fawcett,

Fig. 2. (A) MRP4: PCA of the training and test set compounds (257 in total) were selected such that they occupy similar areas of the PCA plot. The PCA among the training and test compounds was generated with the following properties: ALogP, molecular weight, molecular fractional polar surface area, number of rings, aromatic rings, rotatable bonds, hydrogen bond acceptors, and hydrogen bond donors. The first principal component explains 0.366 of total variance, and the second principal component explains 0.272 of total variance. When combined, these explain 0.638 of total variance. The principal components are linear combinations of original descriptors. The dominate descriptors in the principal components are determined by the product of the descriptor coefficient while accounting for the magnitude of the descriptor. The first principal component is dominated by molecular weight, number of hydrogen bond acceptors, and number of rotatable bonds. The second principal component is dominated by the molecular fractional polar surface area. Component 1 = $-3.8514 + 0.17609 * [\text{ALogP}] + 0.0029942 * [\text{Molecular_Weight}] + 0.19241 * [\text{Num_H_Donors}] + 0.12966 * [\text{Num_H_Acceptors}] + 0.11058 * [\text{Num_RotatableBonds}] + 0.21601 * [\text{Num_Rings}] + 0.26649 * [\text{Num_AromaticRings}] - 0.91018 * [\text{Molecular_FractionalPolarSurfaceArea}]$. Component 2 = $-0.91763 - 0.22028 * [\text{ALogP}] + 0.00060715 * [\text{Molecular_Weight}] + 0.30038 * [\text{Num_H_Donors}] + 0.1113 * [\text{Num_H_Acceptors}] + 0.0097512 * [\text{Num_RotatableBonds}] - 0.15452 * [\text{Num_Rings}] - 0.34347 * [\text{Num_AromaticRings}] + 4.3042 * [\text{Molecular_FractionalPolarSurfaceArea}]$. (B) BSEP: PCA analysis of the training and test sets. The first and second principal components accounted for 0.391 and 0.344 of total variance, respectively. Together, they explain 0.735 of total variance. The first principal component (x-axis) is governed by the number of hydrogen bond donors/acceptors and number of rings, whereas the second principal component (y-axis) is governed by lipophilicity and the number of aromatic rings. Both principal components are strongly influenced by the fractional polar surface area. Component 1 = $-4.0005 + 0.035399 * [\text{ALogP}] + 0.0031256 * [\text{Molecular_Weight}] + 0.16608 * [\text{Num_H_Donors}] + 0.14138 * [\text{Num_H_Acceptors}] + 0.11881 * [\text{Num_RotatableBonds}] + 0.23456 * [\text{Num_Rings}] + 0.23836 * [\text{Num_AromaticRings}] + 0.48987 * [\text{Molecular_FractionalPolarSurfaceArea}]$. Component 2 = $-0.12988 + 0.2367 * [\text{ALogP}] + 0.00047919 * [\text{Molecular_Weight}] - 0.20734 * [\text{Num_H_Donors}] - 7.2971e-002 * [\text{Num_H_Acceptors}] + 0.019248 * [\text{Num_RotatableBonds}] + 0.16242 * [\text{Num_Rings}] + 0.31811 * [\text{Num_AromaticRings}] - 3.6221 * [\text{Molecular_FractionalPolarSurfaceArea}]$. PCA analysis comparing the training set of MRP4 (C) and BSEP (D) to the DrugBank database of U.S. Food and Drug Administration–approved drugs.

TABLE 2
Characteristics of Bayesian Models for MRP4 and BSEP Inhibition

Bayesian models	MRP4inhib-ECFP_6	MRP4inhib-FCFP_6	BSEPinhib-ECFP_6	BSEPinhib-FCFP_6
Two-dimensional fingerprints	ECFP_6	FCFP_6	ECFP_6	FCFP_6
10-fold XV ROC AUC ^a	0.816	0.793	0.750	0.759
TP/FN/FP/TN ^a	33/1/1/22	33/1/1/22	43/0/3/125	43/0/5/123
External validation ^b	0.819	0.838	0.845	0.871
TP/FN/FP/TN ^b	8/9/1/11	10/7/2/10	18/4/15/49	17/5/10/54
SE (%) ^b	47.1	58.8	81.8	77.3
SP (%) ^b	91.7	83.3	76.7	84.4
Q (%) ^b	65.5	69.0	77.9	82.6
MCC ^b	0.4123	0.4216	0.5238	0.5796

FN, false negative; FP, false positive; Q, overall prediction accuracy; SE, sensitivity; SP, specificity; TN; true negative; TP, true positive.

^aXV ROC AUC based on training set compounds (green shaded region).

^bPredictive performance validation by test set compounds (blue shaded region) (Ung et al., 2007; Khandelwal et al., 2008). SE = TP/(TP + FN); SP = TN/(TN + FP); Q = (TP + TN)/(TP + TN + FP + FN); MCC = [(TP * TN) - (FN * FP)]/[(TP + FP)(TP + FN)(TN + FN)(TN + FP)]^{1/2}.

2006). The AUC values associated with the four individual models indicated good internal consistency and prediction accuracy. While both the MRP4 inhibitor and the BSEP inhibitor Bayesian classification had similar ROC AUC scores for both the internal leave-one-out cross-validation and the external test set validation, the sensitivity and specificity varied more between the models. The MRP4 inhibitor models had a significantly lower sensitivity, especially the MRP4 ECFP model, compared with the BSEP inhibitor models, but the trade-off was a higher specificity, minimizing false positives. Bayesian classification modeling of BSEP inhibitors resulted in more predictive models, as demonstrated by their relatively higher Matthews correlation coefficient

compared with MRP4 inhibitor models, which could be due to the larger size of the training set (171 BSEP compounds versus 59 MRP4 compounds). In addition to the external validation performed here, the BSEP Bayesian FCFP model was used to predict the classification of five strong inhibitors and five noninhibitors from a previous screen for BSEP inhibitors (Pedersen et al., 2013). The model was able to correctly classify 9 of the 10 compounds, only incorrectly classifying MK571 as a noninhibitor.

Fingerprints can be defined as molecular fragments that characterize the structural features of drug molecules. Figures 3 and 4 display the five most predictive structural fragments for both favorable and

Category BayesianModel-MRP4_inhib_FCFP_6: good features from FCFP_6

G1: 260714409 6 out of 6 good Bayesian Score: 0.383	G2: -367494947 6 out of 6 good Bayesian Score: 0.383	G3: -415216134 5 out of 5 good Bayesian Score: 0.370	G4: 713358128 5 out of 5 good Bayesian Score: 0.370	G5: -55265897 5 out of 5 good Bayesian Score: 0.370

Category BayesianModel-MRP4_inhib_FCFP_6: bad features from FCFP_6

B1: 1069584379 0 out of 4 good Bayesian Score: -1.257	B2: -1099193755 0 out of 4 good Bayesian Score: -1.257	B3: 692823813 0 out of 3 good Bayesian Score: -1.060	B4: 1995860413 0 out of 3 good Bayesian Score: -1.060	B5: 545355516 0 out of 3 good Bayesian Score: -1.060

Fig. 3. Favorable and unfavorable molecular features for interactions with MRP4. Each feature is a fragment-like fingerprint, up to six bond lengths in diameter, which occurs within the larger parent molecule. The squiggle and asterisks indicate that the bond extends further but does not specify the atom type. The favorable features or good features are labeled G1–G5, and the unfavorable features or bad features are labeled B1–B5. A feature is considered good if it frequently occurs within compounds that were classified as inhibitors and bad if it frequently occurs in compounds that are noninhibitors. The large integer after the colon is the unique hash identifier for the shown fingerprint. The Bayesian score is the normalized probability assigned to that feature.

Category BayesianModel-BSEPinhib-FCFP₆: good features from FCFP₆

G1: 1862216547 5 out of 5 good Bayesian Score: 0.816	G2: -1096767684 4 out of 4 good Bayesian Score: 0.767	G3: 346066116 4 out of 4 good Bayesian Score: 0.767	G4: 391786003 4 out of 4 good Bayesian Score: 0.767	G5: -1576063084 3 out of 3 good Bayesian Score: 0.697
---	--	--	---	--

Category BayesianModel-BSEPinhib-FCFP₆: bad features from FCFP₆

B1: 946589555 0 out of 12 good Bayesian Score: -1.603	B2: -1943140669 0 out of 10 good Bayesian Score: -1.460	B3: -14048077 0 out of 8 good Bayesian Score: -1.293	B4: 309602933 2 out of 26 good Bayesian Score: -1.163	B5: -451251206 0 out of 6 good Bayesian Score: -1.093
--	--	---	---	--

Fig. 4. Favorable and unfavorable molecular features for interactions with BSEP.

unfavorable inhibitory activity against MRP4 and BSEP using FCFP₆ fingerprints. Supplemental Figure 2 contains an expanded figure of structural fragments favorable and unfavorable for inhibition of MRP4 and BSEP using both FCFP₆ and ECFP₆ fingerprints. Structural elements depicted in Figs. 3 and 4 were identified in inhibitors and noninhibitors among the training set compounds, respectively. Oxygen atoms tended to be predictive of favorable inhibitory activity for both MRP4 and BSEP; however, negatively ionized oxygen atoms frequently occurred in the MRP4 model but not in the BSEP model even though both are considered anion transporters. This is in agreement with the study by Pedersen et al. (2013) who reported that BSEP substrates tend to be anionic but inhibitors were more likely to be neutral at a physiologic pH. Accordingly, positively charged secondary and tertiary amines frequently occur among the MRP4 fingerprints associated with noninhibition. Thus, the identified fingerprints could be helpful in distinguishing inhibitors and noninhibitors of MRP4 among novel compounds.

MRP4 Pharmacophore Development. The MRP4 training set of 57 drugs was imported into LigandScout 3.12 and clustered according to pharmacophore radial distribution function-code similarity, with the maximum conformations set to three and the cluster distance set to 0.5. This algorithm clusters compounds that have similar individual three-dimensional pharmacophore characteristics.

The following nine drugs that represent the strongest inhibitors in their respective cluster were used to generate MRP4 inhibition pharmacophores: nitrenedipine, sulindac, sorafenib, clobetasol propionate, benz-bromarone, glafenine, furosemide, finasteride, and simvastatin. The remaining 77 compounds not selected for the training set were moved to the test set for pharmacophore validation. The ligand-based common feature pharmacophore produced from the nine compounds had two hydrophobic features and a hydrogen bond acceptor feature (Fig. 5A).

The two hydrophobic features were 5.01 Å apart, while the hydrogen bond acceptor was 4.81 Å from the neighboring hydrophobic feature and 8.86 Å from the distal hydrophobic group. All nine drugs in the training set aligned with all three pharmacophore features. Two representative compounds were aligned to the pharmacophore to illustrate scale and similarity in how the molecules align with the respective molecular features comprising the pharmacophore (Fig. 5B). These compounds were chosen because their steroid backbone renders them particularly rigid, increasing the likelihood that the representative conformer is close to its bioactive conformation. Additionally, they contain few atoms that can engage in intermolecular interactions, which further confirms that these are the requisite features for MRP4 recognition. Exclusion volumes, i.e., spheres that cannot be occupied and represent steric hindrance, were generated initially using the noninhibitors in the 57 compound training set; however, consideration of exclusion volumes rendered the models more likely to incorrectly classify MRP4 inhibitors as noninhibitors during external validation and were subsequently omitted during database screening.

Quantitative Validation of the MRP4 Pharmacophore Model. The MRP4 pharmacophore model was able to correctly classify 30 of the 42 actives in the test set and 22 of the 35 inactives, featuring a model sensitivity of 71.4% and specificity of 62.8%. The area under the receiver operating characteristic curve was 0.70, which is considered a fair quality model (Fig. 6).

Based on the virtual screening results from the external test set validation, the model has its highest positive predictive value, the number of true positives over the sum of true and false positives, at a pharmacophore-fit score cutoff of 37.75. The positive predictive value of the model at this cutoff is 0.826, selecting 19 true positives, 45.2% of total actives in the set, but only four false positives, 11.4% of total inactives in the set. The pharmacophore-fit score cutoffs allow for

Fig. 5. Pharmacophore model of inhibitors of MRP4-mediated transport of DHEAS. (A) The pharmacophore model with the measured distances between the three features. (B) The pharmacophore model aligned with chemical groups of two drugs from the training set: clobetasol propionate (orange) and finasteride (lavender). Yellow spheres represent hydrophobic features, and the red sphere represents a hydrogen bond acceptor. On the stick model, red represents oxygen atoms, blue represents nitrogen atoms, green represents halogen atoms, and the rest are carbons. Both hydrophobic features align with methyl groups, and the hydrogen bond acceptor aligns with a ketone group. Hydrogen atoms are not displayed for clarity.

selecting drugs with a higher likelihood of being classified correctly beyond those that align to the pharmacophore within the tolerance of the features.

The inactive compounds that were incorrectly classified included dexamethasone, naloxone, clopamide, vinblastine, tolbutamide, probenecid, indinavir, flupirtine, chlorpropamide, alprenolol, chlorpheniramine, fluorescein, and timolol. Interestingly, the false positive with the highest pharmacophore-fit score was dexamethasone, a glucocorticoid that had no significant inhibitory activity ($5 \pm 34\%$). This compound was aligned with clobetasol propionate, another glucocorticoid included in the nine training set compounds, which was a strong inhibitor ($101 \pm 23\%$) (Fig. 7).

As can be seen in Fig. 7, clobetasol propionate (orange) and dexamethasone (gray) have a high degree of structural similarity. From this observation, molecular properties that could be mediating the significant difference in inhibitory activity were investigated. The molecular property that exhibited the most significant difference was calculated log P, which is 4.18 for clobetasol propionate and 1.68 for dexamethasone, rendering clobetasol propionate a more hydrophobic compound. The difference in calculated log P values was evaluated for all 87 compounds tested by Köck et al. Actives trended toward higher log P values than inactives. A Pearson correlation coefficient for the calculated log P and a compound considered active ($\geq 21\%$ MRP4 inhibition) was 0.634, and the correlation coefficient for the percent MRP4 inhibition and calculated log P was 0.508. Fig. 8 represents a box

Fig. 6. ROC curve of pharmacophore model of MRP4 inhibitors from virtually screening the test set ($N = 77$ compounds).

plot of the calculated log P values for the compounds classified as inactive and active. The mean and median of the inactive compounds' calculated log P values were 0.38 and 0.69, respectively, and the active compounds' calculated log P values were 3.64 and 3.84, respectively. It is interesting to note that numerous sulfonamides or sulfamides, such as clopamide, tolbutamide, probenecid, and chlorpropamide, were classified as false positives. These molecules may either be a poor match to the models or their features are incorrectly parameterized within the Bayesian and pharmacophore modeling algorithms.

The actives that were not properly classified as active by the model included 19-norethindrone, clozapine, desipramine, diphenhydramine, etoposide, maprotiline, nitrofurantoin, nortriptyline, oxybutynin, praziquantel, promethazine, and ticlodipine. Eight of these 12 drugs have similar structures containing an amine group, which is predicted to be positively ionized at a physiologic pH. In addition, six of these drugs contain two aromatic rings whose distance is comparable to the distance observed between the two hydrophobic features in the MRP4 model. Compounds that fall into this category are clozapine, desipramine, diphenhydramine, maprotiline, nortriptyline, and promethazine. Oxybutynin and ticlodipine contain an amine predicted to be positively charged, but they have only one aromatic group. Nitrofurantoin is a compound that continually failed to match any structural similarity search to other known inhibitors; therefore, we speculate that it is binding in a different manner than any of the other inhibitors, perhaps at an allosteric site of the transporter.

Qualitative Validation of the MRP4 Inhibitor Model. In addition to the quantitative validation from virtually screening the test set, two compounds that qualitatively strengthen confidence in the model are DHEAS, the substrate used to generate the data, and felbinac, a potent MRP4 inhibitor from a separate screening of MRP4 inhibitors (Morgan et al., 2013). DHEAS was not included in either the training set or the test set, but the MRP4 pharmacophore model would be expected to align to the substrate that was used experimentally to generate the inhibition data. Figure 9A depicts how the two methyl groups on DHEAS align to the hydrophobic features in the pharmacophore, and one of the oxygen atoms from the sulfate group aligns with the hydrogen bond acceptor feature. The alignment of the pharmacophore model to DHEAS is of particular interest because of its structural rigidity due to the steroid backbone structure. The only significant intramolecular motion that DHEAS can undergo is the rotation of the sulfate group. In addition to its rigidity, DHEAS contains few atoms that can participate in intermolecular interactions.

Fig. 7. Structural alignment of glucocorticoids clobetasol propionate (orange) and dexamethasone (gray). Clobetasol propionate, a potent MRP4 inhibitor, inhibits MRP4-mediated transport of DHEAS by $101 \pm 23\%$. In contrast, dexamethasone exhibits no significant inhibitory effect ($5 \pm 34\%$ inhibition). The orange circles indicate identical chemical groups in proximity with each other. On the stick model, red represents oxygen atoms, green represents halogen atoms, and the rest represents carbon atoms.

From the DHEAS pharmacophore (Fig. 9B), it appears that the two methyl groups can participate in hydrophobic interactions, while the ketone can be a hydrogen bond acceptor, and all the oxygen atoms in the sulfate group, which is negatively ionized at physiologic pH, can act as hydrogen bond acceptors.

Felbinac is of interest for the same reasons as DHEAS, namely, its structural rigidity and minimal possible intermolecular interactions. In addition, felbinac is a potent inhibitor of MRP4-mediated transport of β -estradiol 17-(β -D-glucuronide), with an IC_{50} of 8.2 μ M (Morgan et al., 2013). Felbinac aligns well with the MRP4 pharmacophore model and, as shown in Fig. 9D, engages in only a limited number of intermolecular interactions. Both hydrophobic and aromatic interactions with the phenyl groups in the biphenyl compound and the oxygen atoms of the negatively ionized carboxylate group are able to act as hydrogen bond acceptors. The two phenyl groups are locked rigidly on perpendicular planes and, therefore, only the carboxylate group is able to rotate.

Log P Filtering Improves Model Specificity. The MRP4 pharmacophore model's specificity can be significantly improved if the

test set was to be filtered after screening with a calculated log P cutoff of 2.92, which corresponds to the start of the lower quartile of the actives. The sensitivity would decrease to 52.38% (22 true actives of 42 total actives), but the specificity would increase to 91.43% (3 false actives of 35 total inactives). This marked improvement in specificity demonstrates the significant influence log P plays in MRP4 inhibition.

BSEP Inhibitor Pharmacophore Development and Validation.

The BSEP inhibitor pharmacophore was produced with a subset of drugs from the training set that represented the nine strongest BSEP inhibitors, according to Morgan et al., which were also previously tested for MRP4 inhibition by our group. This subset included nitrendipine, fenofibrate, ritonavir, pioglitazone, rosiglitazone, valinomycin, simvastatin, benzbromarone, and lopinavir. The common feature pharmacophore produced with this set of nine compounds had three features: two hydrophobic features and one hydrogen bond acceptor (Supplemental Fig. 1). This is similar to the MRP4 pharmacophore, which may explain the high degree of inhibition overlap; however, the distances between the features were different. The distance between the two hydrophobic features was 6.29 Å, the distance between the hydrogen bond acceptor and proximal hydrophobic feature was 4.67 Å, and the distance between the hydrogen bond acceptor and distal hydrophobic feature was 6.69 Å. All nine drugs that were used to train the pharmacophore hit all three features.

The BSEP inhibitor pharmacophore was validated by virtual screening of the test set, and the model correctly classified 46 of the 56 inhibitors and incorrectly classified 120 of the 191 noninhibitors. The model selectivity was 82.7%, but the specificity was 37.2%. The poor model specificity is partly due to the higher proportion of noninhibitors to inhibitors in the test set (191 versus 56), but is also indicative of the difficulty of modeling BSEP through a pharmacophore approach.

Discussion

A ligand-based pharmacophore and Bayesian modeling approach is presented here, describing the molecular properties and chemical features necessary for human MRP4 and BSEP interaction. Since these transport proteins have been associated with DILI, these models may be useful in predicting DILI liability of novel compounds. The models were developed from our laboratories' previous work and data from other groups. An advantage of Bayesian classification modeling is the ability to easily interpret how the model weighs the various molecular properties, and which molecular properties are most predictive for classification. The Bayesian classification model was developed by creating up to 11 bins for each molecular property. For discrete properties, such as the number of rings or the hydrogen-bond acceptor atoms, all the compounds with the same value were put into the same bin. For continuous properties, such as molecular weight or ALogP, a bin was assigned a value range and all the compounds that fell within that range were put into that bin. The ranges for binning continuous values were created such that the number of compounds in each bin was evenly distributed. A normalized probability was then calculated for each bin according to the fraction of compounds in the bin that were active, i.e., if all the compounds in a bin were active, that bin was assigned a higher probability. The probability was normalized by adjusting for bins with few compounds. A Bayesian score was then calculated by summing the normalized probabilities of all the bins to which a compound was assigned.

By inspecting the normalized probabilities of individual bins, the molecular properties that contain bins with high probabilities indicate molecular properties that are well correlated with either inhibition or noninhibition. For example, with the MRP4 model, a trend was observed with a higher ALogP correlating to MRP4 inhibition (14 of 15

Fig. 8. MRP4: comparison of calculated Log P of compounds classified as inactive ($< 21\%$ MRP4 inhibitory activity; $n = 37$) compared with those classified as active ($\geq 21\%$ MRP4 inhibitory activity; $n = 50$). The mean and median log P values of the inactives are 0.38 and 0.69, respectively, and 3.64 and 3.84, respectively, for the actives.

Fig. 9. DHEAS, an MRP4 substrate, and felbinac, an MRP4 inhibitor, aligned with the MRP4 inhibitor pharmacophore. Both compounds are also depicted with their individual pharmacophore, which shows all possible intermolecular interactions. (A) DHEAS aligned to the MRP4 pharmacophore. (B) DHEAS pharmacophore showing all possible intermolecular interactions. (C) Felbinac aligned to the MRP4 pharmacophore. (D) Felbinac pharmacophore showing all possible interactions. Yellow spheres represent hydrophobic features, red spheres represent hydrogen bond acceptor features, the red star represents a negatively ionizable feature, and the purple torus represents an aromatic ring feature. On the stick models, red represents oxygen atoms, yellow represents phosphorus atoms, and the rest represents carbon atoms. Hydrogen atoms are not displayed for clarity.

drugs with an ALogP of 3.8 or higher were inhibitors, while only 2 of 12 drugs with a ALogP of 0.94 or lower were inhibitors). A trend was also observed with large molecular weight drugs more likely to inhibit MRP4 compared with smaller drugs (23 of 25 drugs with molecular weight ≥ 356 Da). It should again be noted that the value range of a bin for molecular properties that are continuous is a result of evenly dividing the ordered drugs into 11 bins. In addition to those continuous properties, 23 of 25 drugs with three or four rings were inhibitors compared with 1 of 9 compounds with zero rings or one ring.

Compared with the MRP4 Bayesian model, the BSEP Bayesian model was more predictive of negative properties, thus predicting more noninhibitors than inhibitors. This is likely due to the fact that the training set contained more noninhibitors relative to inhibitors compared with the MRP4 training set. For the BSEP model, only 7 of 90 drugs with a molecular weight less than 337 Da were inhibitors. Correspondingly, only 11 of 90 compounds with four or fewer rotatable bonds were inhibitors and, similar to the MRP4 model, only 3 of 72 drugs with an ALogP of 2.03 or lower were inhibitors.

Potential biologic implications of the association between high calculated log P values and MRP4 inhibition were considered. These data lead to speculation that molecules must partition first into the bilayer to be an inhibitor for MRP4 or that higher log P values correspond with increased hydrophobic interactions within the protein environment, thus rendering these molecules stronger competitive inhibitors.

It is worth mentioning that these computational models are based on data collected from membrane vesicle assays, and this *in vitro* system could have an influence on experimental transport inhibition results. The methods generally involve short incubation periods with the test compounds in which the degree of partitioning of the test compound into the membrane of the vesicles could be in flux. If the test compound exerts its inhibition while imbedded in the membrane, this could cause a skew in the data, in which compounds with higher log P values have a higher rate of partitioning into the membrane than those with lower log P values (Nagar and Korzekwa, 2012).

Comparison with Previous Models. Previous studies identified important molecular features for BSEP inhibition (Pedersen et al., 2013), and pharmacophore models have been proposed for BSEP (Ritschel et al., 2014) and MRP4 (Fukuda et al., 2013). While there was good corroboration between the important molecular features, there were several notable differences between the pharmacophores previously reported and those developed here, which can be ascribed primarily to differences in the training sets used to generate the pharmacophores.

Molecular properties that were reported previously to have a statistically significant difference between strong BSEP inhibition ($> 50\%$ inhibition) and noninhibitors were log $D_{7.4}$, molecular weight, saturated nonpolar surface area, log P, number of rotatable bonds, unsaturated nonpolar surface area, number of hydrogen-bond acceptors, and net charge at pH 7.4. The molecular properties that remain statistically significant for weak inhibitors (27–50% inhibition) and noninhibitors were only molecular weight and saturated nonpolar surface area. In comparison, our best Bayesian BSEP model (FCFP_6) ranked the top four feature bins based on their normalized probability in descending order: molecular weight, FCFP_6 fingerprint, number of rotatable bonds, and ALogP. However, since a compound can have only one value for a molecular property but multiple fingerprints, the fingerprints tend to be the predominate factor influencing the final Bayesian score for a compound. The pharmacophore model for BSEP is also in agreement with these previously reported molecular properties since they consisted of two hydrophobic features and a hydrogen-bond accepting feature. The hydrophobic features were associated with the nonpolar surface area and high log P, while the hydrogen-bond accepting feature was associated with the number of

hydrogen bond acceptors correlated with inhibition. The presence of only one hydrogen bond accepting features in the pharmacophore model, however, suggests that a large number is not essential for inhibition but could provide more opportunities for hydrogen bonding in the correct spatial arrangement.

Two common feature MRP4 pharmacophores were reported previously in the same paper: one based on five protease inhibitors (PIs) and the other based on a more diverse set of 10 drugs that were inhibitors based on literature reports. The pharmacophore based on five PIs resulted in a pharmacophore with four hydrogen bond acceptors (HBAs), one hydrogen bond donor, and three hydrophobic features, and the pharmacophore based on 10 drugs resulted in two HBAs and a hydrophobic feature. The PI-based pharmacophore featured a large number of features due to the small number of compounds in the training set and their high degree of structural similarity, while the pharmacophore based on 10 diverse compounds was based on the findings from multiple laboratories. The MRP4 pharmacophore we developed is more appropriate for predicting potential MRP4 inhibitors to identify the cholestatic potential of compounds because it was based on data from a single laboratory utilizing one assay, was developed based on a diverse set of compounds, and contained a large number of inhibitors.

The other recently reported BSEP pharmacophore (Ritschel et al., 2014) was trained using five compounds of limited structural diversity and resulted in a pharmacophore that had eight features: four hydrophobic features and two HBAs that had an associated vector feature. The authors found the pharmacophore too stringent so they modified it by making only four hydrophobic features in the core of the pharmacophore essential. The advantage of the BSEP pharmacophore presented in this paper is that it is derived from pharmaceuticals instead of a chemical library and was developed with more diverse compounds, which results in a pharmacophore with fewer features but one more equipped to deal with a larger chemical space. The BSEP pharmacophore reported in this paper, however, is able to align with the hydrophobic and hydrogen bond accepting features of the previously reported BSEP pharmacophore. This suggests that the pharmacophore reported in this paper may be convergent with previously reported pharmacophore, although it is less stringent.

In conclusion, ongoing studies will use these models in an ensemble fashion against drugs that are predicted to be MRP4 or BSEP inhibitors to further validate the models. These models, when used in combination, may aid in the a priori identification of potential cholestasis-inducing compounds during the early stages of drug development.

Authorship Contributions

Participated in research design: Brouwer, Köck, Swaan, Welch.

Conducted experiments: Welch.

Performed data analysis: Swaan, Welch.

Wrote or contributed to the writing of the manuscript: Welch, Köck, Urban, Brouwer, Swaan.

References

- Abboud G and Kaplowitz N (2007) Drug-induced liver injury. *Drug Saf* **30**:277–294.
- Chai J, He Y, Cai SY, Jiang Z, Wang H, Li Q, Chen L, Peng Z, He X, and Wu X, et al. (2012) Elevated hepatic multidrug resistance-associated protein 3/ATP-binding cassette subfamily C 3 expression in human obstructive cholestasis is mediated through tumor necrosis factor alpha

- and c-Jun NH2-terminal kinase/stress-activated protein kinase-signaling pathway. *Hepatology* **55**:1485–1494.
- Dawson S, Stahl S, Paul N, Barber J, and Kenna JG (2012) In vitro inhibition of the bile salt export pump correlates with risk of cholestatic drug-induced liver injury in humans. *Drug Metab Dispos* **40**:130–138.
- Ekins S, Polli JE, Swaan PW, and Wright SH (2012) Computational modeling to accelerate the identification of substrates and inhibitors for transporters that affect drug disposition. *Clin Pharmacol Ther* **92**:661–665.
- Fattinger K, Funk C, Pantze M, Weber C, Reichen J, Stieger B, and Meier PJ (2001) The endothelin antagonist bosentan inhibits the canalicular bile salt export pump: a potential mechanism for hepatic adverse reactions. *Clin Pharmacol Ther* **69**:223–231.
- Fawcett T (2006) An introduction to ROC analysis. *Pattern Recognit Lett* **27**:861–874.
- Fukuda Y, Takenaka K, Sparreboom A, Cheepala SB, Wu C-P, Ekins S, Ambudkar SV, and Schuetz JD (2013) Human immunodeficiency virus protease inhibitors interact with ATP binding cassette transporter 4/multidrug resistance protein 4: a basis for unanticipated enhanced cytotoxicity. *Mol Pharmacol* **84**:361–371.
- Gradhand U, Lang T, Schaeffeler E, Glaeser H, Tegude H, Klein K, Fritz P, Jedlitschky G, Kroemer HK, and Bachmakov I, et al. (2008) Variability in human hepatic MRP4 expression: influence of cholestasis and genotype. *Pharmacogenomics J* **8**:42–52.
- Khandelwal A, Bahadduri PM, Chang C, Polli JE, Swaan PW, and Ekins S (2007) Computational models to assign biopharmaceutics drug disposition classification from molecular structure. *Pharm Res* **24**:2249–2262.
- Khandelwal A, Krasowski MD, Reschly EJ, Sinz MW, Swaan PW, and Ekins S (2008) Machine learning methods and docking for predicting human pregnane X receptor activation. *Chem Res Toxicol* **21**:1457–1467.
- Köck K, Ferslew BC, Netterberg I, Yang K, Urban TJ, Swaan PW, Stewart PW, and Brouwer KLR (2014) Risk factors for development of cholestatic drug-induced liver injury: inhibition of hepatic basolateral bile acid transporters multidrug resistance-associated proteins 3 and 4. *Drug Metab Dispos* **42**:665–674.
- Kostrubsky VE, Strom SC, Hanson J, Urda E, Rose K, Burliegh J, Zocharski P, Cai H, Sinclair JF, and Sahi J (2003) Evaluation of hepatotoxic potential of drugs by inhibition of bile-acid transport in cultured primary human hepatocytes and intact rats. *Toxicol Sci* **76**:220–228.
- Law V, Knox C, Djoumbou Y, Jewison T, Guo AC, Liu Y, Maciejewski A, Arndt D, Wilson M, and Neveu V, et al. (2014) DrugBank 4.0: shedding new light on drug metabolism. *Nucleic Acids Res* **42**:D1091–D1097.
- Morgan RE, Trauner M, van Staden CJ, Lee PH, Ramachandran B, Eschenberg M, Afshari CA, Qualls CW Jr, Lightfoot-Dunn R, and Hamadeh HK (2010) Interference with bile salt export pump function is a susceptibility factor for human liver injury in drug development. *Toxicol Sci* **118**:485–500.
- Morgan RE, van Staden CJ, Chen Y, Kalyanaraman N, Kalanzi J, Dunn RT, 2nd, Afshari CA, and Hamadeh HK (2013) A multifactorial approach to hepatobiliary transporter assessment enables improved therapeutic compound development. *Toxicol Sci* **136**:216–241.
- Nagar S and Korzekwa K (2012) Commentary: nonspecific protein binding versus membrane partitioning: it is not just semantics. *Drug Metab Dispos* **40**:1649–1652.
- Noé J, Stieger B, and Meier PJ (2002) Functional expression of the canalicular bile salt export pump of human liver. *Gastroenterology* **123**:1659–1666.
- Obuchowski NA and Lieber ML (1998) Confidence intervals for the receiver operating characteristic area in studies with small samples. *Acad Radiol* **5**:561–571.
- Pedersen JM, Matsson P, Bergström CA, Hoogstraate J, Norén A, LeCluyse EL, and Artursson P (2013) Early identification of clinically relevant drug interactions with the human bile salt export pump (BSEP/ABCB11). *Toxicol Sci* **136**:328–343.
- Ritschel T, Hermans SM, Schreurs M, van den Heuvel JJ, Koenderink JB, Greupink R, and Russel FG (2014) In silico identification and in vitro validation of potential cholestatic compounds through 3D ligand-based pharmacophore modeling of BSEP inhibitors. *Chem Res Toxicol* **27**:873–881.
- Rogers D, Brown RD, and Hahn M (2005) Using extended-connectivity fingerprints with Laplacian-modified Bayesian analysis in high-throughput screening follow-up. *J Biomol Screen* **10**:682–686.
- Scheffler GL, Kool M, de Haas M, de Vree JM, Pijnenborg AC, Bosman DK, Elferink RP, van der Valk P, Borst P, and Schepers RJ (2002) Tissue distribution and induction of human multidrug resistant protein 3. *Lab Invest* **82**:193–201.
- Stieger B, Fattinger K, Madon J, Kullak-Ublick GA, and Meier PJ (2000) Drug- and estrogen-induced cholestasis through inhibition of the hepatocellular bile salt export pump (Bsep) of rat liver. *Gastroenterology* **118**:422–430.
- Teng S and Piquette-Miller M (2007) Hepatoprotective role of PXR activation and MRP3 in cholic acid-induced cholestasis. *Br J Pharmacol* **151**:367–376.
- Ung CY, Li H, Yap CW, and Chen YZ (2007) In silico prediction of pregnane X receptor activators by machine learning approaches. *Mol Pharmacol* **71**:158–168.
- Xia X, Maliski EG, Gallant P, and Rogers D (2004) Classification of kinase inhibitors using a Bayesian model. *J Med Chem* **47**:4463–4470.
- Zweig MH and Campbell G (1993) Receiver-operating characteristic (ROC) plots: a fundamental evaluation tool in clinical medicine. *Clin Chem* **39**:561–577.

Address correspondence to: Peter W. Swaan, Ph.D., Department of Pharmaceutical Sciences, University of Maryland, Baltimore MD 21201. E-mail pswaan@rx.umaryland.edu