Mitigating the Inhibition of Human Bile Salt Export Pump by Drugs: Opportunities Provided by Physicochemical Property Modulation, In Silico Modeling, and Structural Modification

Daniel J. Warner, Hongming Chen, Louis-David Cantin, J. Gerry Kenna, Simone Stahl, Clare L. Walker, and Tobias Noeske

Department of Medicinal Chemistry, AstraZeneca R&D Montreal, Montreal, Quebec, Canada (D.J.W., L.-D.C.); Computational Sciences, Discovery Sciences (H.C.), and Global Safety Assessment (T.N.), AstraZeneca R&D Mölndal, Mölndal, Sweden; and Molecular Toxicology, Global Safety Assessment, AstraZeneca, Alderley Park, Macclesfield, Cheshire, United Kingdom (J.G.K., S.S., C.L.W.)

ABSTRACT:

The human bile salt export pump (BSEP) is a membrane protein expressed on the canalicular plasma membrane domain of hepatocytes, which mediates active transport of unconjugated and conjugated bile salts from liver cells into bile. BSEP activity therefore plays an important role in bile flow. In humans, genetically inherited defects in BSEP expression or activity cause cholestatic liver injury, and many drugs that cause cholestatic drug-induced liver injury (DILI) in humans have been shown to inhibit BSEP activity in vitro and in vivo. These findings suggest that inhibition of BSEP activity by drugs could be one of the mechanisms that initiate human DILI. To gain insight into the chemical features responsible for BSEP inhibition, we have used a recently described in vitro membrane vesicle BSEP inhibition assay to quantify transporter inhibition for a set of 624 compounds. The relationship between BSEP inhibition and molecular physicochemical properties was investigated, and our results show that lipophilicity and molecular size are significantly correlated with BSEP inhibition. This data set was further used to build predictive BSEP classification models through multiple quantitative structure-activity relationship modeling approaches. The highest level of predictive accuracy was provided by a support vector machine model (accuracy = 0.87, \( \kappa = 0.74 \)). These analyses highlight the potential value that can be gained by combining computational methods with experimental efforts in early stages of drug discovery projects to minimize the propensity of drug candidates to inhibit BSEP.

Introduction

The pharmaceutical industry continues to face high attrition rates throughout the entire drug discovery and development process, with only 8% of all drug candidates that enter phase I studies progressing to the market (U.S. Food and Drug Administration, 2004; Schuster et al., 2005). A major cause of compound attrition in drug development is toxicity (Greaves et al., 2004; Kramer et al., 2007). The financial burden of safety-related dropouts can reach many hundreds of million dollars in late stages of drug development (DiMasi et al., 2003). Therefore, it is important to identify and mitigate potential compound-related safety issues as early as possible during the drug discovery process.

One important cause of drug toxicity in humans is drug-induced liver injury (DILI). This is a major cause for attrition in clinical trials, failed drug registration, and withdrawal of marketed drugs (Abboud and Kaplowitz, 2007). For a few drugs, for example, acetaminophen (Wallace 2004), DILI is a dose-dependent and reproducible process that occurs in humans only after accidental or deliberate overdose and is reproducible in animals (Pirmohamed et al., 1998). However, many drugs cause DILI only infrequently or very rarely in humans, which is not overtly dose-dependent and is not reproducible (Zimmerman 1978). The most common patterns of clinical presentation of DILI in humans are defined as either hepatocellular (i.e., primarily affecting hepatocyte function), cholestatic (primarily affecting the biliary system), or mixed hepatocellular/cholestatic (Mumoli et al., 2006). The underlying mechanisms are complex and include both compound-related properties and factors that are specific to individual susceptible patients (Thompson et al., 2011). A variety of compound-related DILI risk factors have been identified or proposed. These include formation of chemically reactive metabolites, mitochondrial impairment, potent cell cytotoxicity, and inhibition of the human bile salt export pump (BSEP) (Greer et al., 2010; Dawson et al., 2012).

BSEP (encoded by the \( ABCB11 \) gene) is a liver-specific ABC transporter that is expressed on the canalicular domain of the hepatocyte plasma membrane and that mediates the active secretion of monovalent bile acids or salts into the bile canaliculi (Gerloff et al., 1998; Byrne et al., 2002; Noé et al., 2002). Mutations in the \( ABCB11 \) gene occur via codon deletions, insertions, and various point mutations. These have been shown to cause either progressive familial intrahepatic cholestasis type 2, which is a very rare but severe form of...
cholestatic liver disease that results in fatal liver failure unless treated by liver transplantation or a milder form of liver injury termed benign recurrent intrahepatic cholestasis type 2 (Davit-Spraul et al., 2009). Many drugs that cause cholestatic DILI in humans have been found to inhibit the activity of BSEP and its rat analog Bsep in vitro. In several cases, inhibition of rat Bsep activity in vivo has also been demonstrated (Fattinger et al., 2001; Kostrubsky et al., 2006). Furthermore, it has been shown recently that drugs causing cholestatic DILI in humans exhibit markedly greater potencies and frequencies of BSEP or Bsep inhibition in vitro than drugs that cause hepatocellular DILI or drugs that do not cause DILI (Morgan et al., 2010; Dawson et al., 2012). These findings suggest that mitigation of BSEP inhibition may help reduce the likelihood of DILI for humans.

However, applying high-volume screening for BSEP inhibition in a drug discovery project would be cost-intensive and time-consuming and potentially could delay the rate of progression of the project. A more desirable strategy would be to combine in vitro BSEP screening with in silico BSEP-predicting models (Saito et al., 2009). In the present study, we have investigated the use of various computational algorithms to build a BSEP inhibition model for a set of 624 chemically diverse compounds. The effect of these compounds on BSEP activity was quantified in vitro using a membrane vesicle assay. An accurate classification model for BSEP inhibition has been developed, which can be used for library profiling. The relationship between some molecular physicochemical properties, such as lipophilicity and size, and propensity of BSEP inhibition were also highlighted in this analysis.

Materials and Methods

Experimental Measurement. All standard reagents and chemicals were purchased from Sigma-Aldrich (St. Louis, MO), Merck (Hohenbrunn, Germany), or Wako (Osaka, Japan) and were of the highest purity available. [3H]taurocholate was obtained from American Radiolabeled Chemicals, Inc. (St. Louis, MO). Baculoviruses expressing human BSEP (ABCB11) were provided by Bruno Stieger (Stieger et al., 2000; Noé et al., 2002). Test compounds (>95% purity) were provided by the Compound Management Group, Discovery Sciences, AstraZeneca (Alderley Park, Macclesfield, Cheshire, UK).

BSEP (ABCB11) was expressed in Spodoptera frugiperda S212 insect cells from which inside-out membrane vesicles were prepared as described previously (Dawson et al., 2012) with the modification that at the end of the isolation procedure vesicles were resuspended in buffer containing 50 mM sucrose, 10 mM HEPES/Tris, pH 7.4, and protease inhibitor tablets (Roche, Basel, Switzerland). Inhibition of BSEP-mediated ATP-dependent transport of [3H]taurocholate into vesicles was measured by a rapid filtration method as described previously (Dawson et al., 2012). In brief, 60 µg of BSEP vesicles were incubated with test compound or dimethyl sulfoxide (DMSO) vehicle and 0.5 µM [3H]taurocholate for 5 min at 37°C in buffer containing 5 mM ATP, 10 mM MgCl2, 15.0 mM HEPES/Tris, pH 7.4, 142 mM KNO3, 158 mM sucrose, and 12.5 mM Mg(NO3)2. For each compound, three independent experiments were performed in triplicate for each experiment for each compound test concentration (10, 30, 100, 250, 500, and 1000 µM). The DMSO concentration in all reactions was 2% (v/v). The transport reaction was stopped by addition of vesicles to stop buffer containing 50 mM sucrose, 100 mM KCl, 10 mM HEPES/Tris, pH 7.4, 5 mM EDTA, and 0.1 mM taurocholate. [3H]Taurocholate uptake into vesicles was measured by determining the counts per minute using a TopCount NXT (PerkinElmer Life and Analytical Sciences, Waltham, MA). Counts per minute values determined for reactions containing AMP were subtracted from reactions containing ATP to determine the ATP-dependent uptake activity. The percentage of uptake activity relative to DMSO vehicle control (100%) was determined for each compound test concentration. IC50 values were calculated with nonlinear regression for a sigmoidal dose-response using eq. 1:

$$Y = \text{bottom} + \frac{\text{top} - \text{bottom}}{1 + 10^{\text{logIC}_{50}-X_{\text{half}}}}$$

where X is the common logarithm of the concentration, Y is the response, nH is the variable Hill slope, bottom is 0, and top is 100.

Data Set. The BSEP data set for QSAR modeling comprises 624 internal compounds and external reference compounds (Supplemental Table S4). A BSEP IC50 of 300 µM was used as a threshold value for classifying compounds; i.e., compounds with a geometric mean IC50 value less than 300 µM were regarded as BSEP active compounds (labeled as POS class); otherwise they were regarded as BSEP inactive compounds (labeled as NEG class). An IC50 value of 300 µM determined in the membrane vesicle model was previously found to be a useful operational threshold to identify drugs, which are associated with cholestatic or mixed hepatocellular/cholestatic liver injury in human (Dawson et al., 2012). An in-house Perl script was run to randomly split the BSEP data set into a training set and a test set with the ratio of 7:3, resulting in a training set of 437 compounds (231 POS and 206 NEG) and the test set of 187 compounds (94 POS and 93 NEG).

Physical Property Dependencies. Logistic regression analysis of the relationships between molecular properties and BSEP activity were conducted using JMP8.0 software (SAS Institute Inc., Cary, NC) by estimating the probability of a compound being free from BSEP inhibition as a smooth function of a single parameter [e.g., log10 molecular weight, calculated ClogP, version 4.3; BioByte Corp., Claremont, CA], calculated logD, and acid dissociation constants (Advanced Chemistry Development, Inc., Toronto, ON, Canada). The statistical significance of probabilities based on the relationship compared with the background probabilities were evaluated using a χ² test statistic (G²) evaluated as eq. 2.

$$G^2 = 2 \left( \sum \ln p(\text{background}) - \sum \ln p(\text{relationship}) \right)$$

The observed significance probability (or p value) for the χ² test is the probability of obtaining, by chance alone, a χ² test statistic greater than the one computed. Probabilities of <0.05 were considered significant.

$$R^2 = \frac{\sum \ln p(\text{background}) - \sum \ln p(\text{relationship})}{\sum \ln p(\text{background})}$$

$R^2$ ranges from 0 to 1: a value of 0 would indicate no benefit in the use of a given molecular property to classify BSEP activity, whereas $R^2$ of 1 would indicate perfect classification.

Modeling Methods. Multiple approaches for building a classification model of a compound’s BSEP inhibition activity were investigated. The first and simplest approach used a basic recursive partitioning (RP) algorithm, based solely on the molecular weight and calculated lipophilicity of the structures. The partition scheme was built using JMP software. In contrast to the one-dimensional physical properties modeling above, the scheme was derived from 437 training set compounds and evaluated using the randomly selected 187 test set compounds. Automatic parameter selection was applied at each point in the scheme to maximize the significance of the split. Partitioning was performed until $R^2$ (eq. 3) for the test set ceased to improve, resulting in a total of four partitions, two for each parameter.

To further refine our description of the structure–activity relationships surrounding this transporter, two more elaborate sets of descriptors were calculated for all 624 molecular structures. A set of 196 two-dimensional/three-dimensional descriptors (referred to as the AZdesc set), including descriptors for molecular size, lipophilicity, hydrogen bonding, electrostatics, and topology were calculated with an in-house program. The descriptors themselves are described elsewhere (Katritzky et al., 1998; Bruene 2001; Paine et al., 2010). A second descriptor set (referred to as the AZFP_AZtop14 set) comprises a combination of fingerprint-based descriptors including the in-house developed Ghose-Crippen atom type fingerprints (Ghose and Crippen, 1987) and a set of functional groups fingerprints (Arnold et al., 2004). This collection of structural descriptors was further supplemented with a subset
of the AZdesc set, comprising an additional 14 descriptors to encode molecular bulk properties such as lipophilicity, size, ionic states, and others. Partial least-squares (PLS) and two nonlinear machine learning methods, support vector machine and random forest, were used to build the classification models based on the AZdesc set. For the AZFP_AZtop14 set, which was mostly focused on structural fingerprints, only the two nonlinear techniques (random forest and support vector machine) were used. All three methods are implemented in the in-house machine learning package AZOrange (Stålring et al., 2011), which is an extension of the open source package Orange (http://www.aialab.si/orange/). Classification performance measurements based on the confusion matrix were generated to assess the quality of the classification across all six approaches (Supplemental Table S2).

WizePairZ Methodology. Matched molecular pairs analysis (Griffen et al., 2011) was performed on the full set of 624 compounds using the WizePairZ algorithm (Warner et al., 2010). To deconvolute the intrinsic effects of structural transformations from their associated effect on molecular properties, two variables were considered in the analysis: BSEP activity, expressed as \(-\log_{10}(\text{IC}_{50})\) (or \(p\text{IC}_{50}\)) and ClogP. As in our original study, the mean change of the activity for each molecular pairing was plotted as a function of the change in lipophilicity, allowing those transformations that link the anticipated trend in activity for each molecular pairing to be easily identified.

Results

Measurement of BSEP Inhibition by the Test Compounds. Typical concentration-response curves of test compound effects on BSEP activity are shown in Fig. 1, which includes data obtained with a compound that did not inhibit BSEP (acetaminophen), a compound that exhibited potent BSEP inhibition (bosentan), and two compounds that exhibited less potent inhibition (bezafibrate and labetalol). Experimental results are the means \(\pm\) S.E.M. from three separate test occasions. For each compound, an IC\(_{50}\) value was calculated, and these values are summarized in the supplemental data. As can be seen from Fig. 2, 240 compounds did not exhibit BSEP inhibition. For 269 compounds, BSEP inhibition IC\(_{50}\) values, ranging between 1000 and 10 \(\mu\)M, i.e., within the range of tested compound concentrations, were observed. An additional 115 compounds exhibited very potent BSEP inhibition, which was defined as IC\(_{50}\) < 10 \(\mu\)M.

Molecular Property Dependencies. The influence of several molecular properties on the likelihood of BSEP inhibition with an apparent IC\(_{50}\) value of <300 \(\mu\)M is displayed in Fig. 3. The line of fit (blue) in each plot of Fig. 3 is interpreted as the probability (\(y\)) of a data point appearing below the line [i.e., being free of BSEP inhibition for a given property value (\(x\))]. The scattered dots in the plots are the tested compounds, and their color refers to the class (POS or NEG class) that they belong to. It seems that the likelihood of a compound inhibiting BSEP activity is highly dependent on its molecular weight and calculated lipophilicity. Any compound with a molecular weight greater than 309 is more likely to exhibit a BSEP inhibition than not, and compounds with molecular weights below this threshold become more likely to be free from BSEP inhibition (Fig. 3d). The strong statistical significance of the relationship with molecular weight (\(p < 0.0001\)) and reasonably high \(R^2 (0.22)\) dictates that changes in the likelihood of activity occur fairly sharply. Calculated octanol/water partition coefficients provide an even more powerful means of classifying the compounds into BSEP inhibition active versus BSEP inactive (\(R^2 = 0.29\)). A ClogP of 2.27 marks the point at which the balance shifts from a compound being more likely to be free of BSEP inhibition to becoming a more likely BSEP inhibitor (Fig. 3a).

In an attempt to identify an even more discriminatory lipophilicity metric, we investigated the impact of calculated logD on BSEP inhibition potency below the selected threshold value. Although this approach demonstrated less utility than the analysis based on ClogP described above (data not shown), we noticed an interesting improvement in classification with the use of logD at pH 6.5 over the same calculation based on pH 7.4 (\(R^2 = 0.19\) and 0.17, respectively). This results indicates that whereas BioByte’s logP algorithm provides the more effective classifier in this case, the ionization state of the active species may also play a role, which led us to investigate further the influence of pK\(_a\) on BSEP inhibition (Fig. 3, b and e). The two hypotheses to be tested were that the reduced pH was causing weak acids to appear more hydrophilic to the binding site, thus improving their BSEP inhibition potency classification or that reduced pH caused weak bases to appear more hydrophilic, thus reducing their potency of BSEP inhibition. The data indicate that the latter is true, because although there was no significant relationship with pK\(_a\) (acid 1) (\(p = 0.2751\)), a significant decrease in the likelihood of BSEP inhibition with increasing basicity was observed (\(p = 0.0002\)).

To seek additional support for the argument that the formation of cations disfavors BSEP inhibition potency, we classified all 624 compounds into five ion class categories: neutral compounds, bases, acids, cations, and zwitterions (Fig. 4). Across the full data set, the distribution between BSEP inhibition potency above and below the threshold value was almost completely even, and this was mirrored across the acids, bases, and neutral compounds. However, it is notable that cations formed the least likely class of BSEP inhibitors, with only 2 “active” compounds and 10 “inactive” compounds (17%). The final
paired logistic regression analysis compared the hydrogen bond donor/acceptor characteristics of the data set (Fig. 3, c and f). Here, the only significant relationship was with the hydrogen bond donor count, at which molecules with an increasing number of donors were at reduced risk of BSEP inhibition ($p < 0.0001$). There was no significant relationship between BSEP inhibition category and the acceptor count.

**BSEP Classification Model.** In an effort to devise a predictive in silico model of BSEP inhibition, we embarked on the exploration of a number of approaches. Because molecular weight (MW) and lipophilicity are useful predictors of BSEP inhibition, as presented in the previous section, we decided to build a classification model based on an RP algorithm regarding these two properties to serve as a baseline model (RP_ClogP_MW) (Fig. 5). At the top of the scheme, the first partition occurs at a ClogP threshold of 1.697, which provides reasonable separation between BSEP active and inactive compounds on the second tier. The second partition takes place to the far left of the scheme (on the more lipophilic compounds), such that those with molecular weight greater than and less than 296.2 are separated. This third tier is where the recursion ends for 204 of these compounds, as those with molecular weight $>296.2$ are deemed almost certainly active with a probability of 91% (red section in the contour plots of Fig. 5b). Classification of the 78 lower weight compounds within this more lipophilic group is more ambiguous but generally favors inactive compounds with a probability of 66% (light blue section in the contour plots of Fig. 5b).

To the right of the scheme, the situation for the less lipophilic compounds is somewhat more complex. In this category, in which ClogP is $<1.697$, compounds are most likely to be inactive, and, overall, this does not change either way with the third tier partitioning at a ClogP of 0.59. Below a ClogP of 0.59, the training set compounds are almost unanimously inactive (95 compounds, 98% inactive).
it had an 85% chance of causing BSEP inhibition with an IC50 of 300 μM. Thus, if a compound was predicted to be positive by the model, the probability of the SVM algorithm and the AZdesc set (SVM_AZdesc) showed the best performance across most of the parameters (Table 1). The probabilities of this model correctly classifying compounds predicted to be positive and negative for BSEP inhibition were 0.85 and 0.90, respectively. Thus, if a compound was predicted to be positive by the model, it had an 85% chance of causing BSEP inhibition with an IC50 of <300 μM. The \( \kappa \) value for this SVM_AZdesc model was the highest (0.74), which provides an indication of the true accuracy of the model because it takes the probability of obtaining such agreement by chance into account. It is also interesting to note that the linear PLS model showed good negative precision and its sensitivity was almost as good as that of the SVM_AZdesc model.

In a comparison of the results from the two descriptor sets, it was clear that the nonlinear models that combined fingerprint and a subset of the bulk property descriptors (AZFP_AZTop14) performed no better than the same methods using the AZdesc set. For this reason, these two models will not be discussed further. In comparison with the RP model, the SVM_AZdesc model has better accuracy in terms of sensitivity and negative precision (0.90 versus 0.84 in both cases). The common denominator in both these metrics and hence the driving force behind this improvement was the false-negative count (compounds predicted to be negative but measured positive), which was reduced from 15 to 9.

To check the robustness of these model generation strategies, a 5-fold cross-validation analysis was performed on all 624 compounds for the nonlinear and PLS modeling methods. The performance metrics for cross-validation analysis are listed in Supplemental Table S2. Although SVM_AZdesc method still performs best in general, the difference between various nonlinear modeling methods is actually very small, and all nonlinear methods perform better than the PLS_AZdesc method.

In situations such as this, in which lipophilicity of compounds has such a strong bearing on their effect on BSEP activity, it is desirable to identify subtle structural changes that reduce activity without compromising other molecular properties. This approach is especially useful in cases in which a chemical series is in late-stage optimization and working with limited opportunities to modify the physicochemical properties. WizePairZ was designed for precisely this purpose, and its application to our BSEP data set identified a total of 103 unique molecular transformations when four bonds of the local environment (counting each transformation and its reverse just once) were included.

Of the four salicylic acid analogs depicted in Fig. 6a, it was only mesalazine (5-amino-2-hydroxybenzoic acid) that exhibited in vitro BSEP inhibition. This compound was neither the most lipophilic in the group (ClogP = 1.06 versus 2.44) nor the one with the highest molecular weight (153 versus 154), although WizePairZ was able to identify a structural motif that correlated with reduced BSEP inhibition activity. The group that is absent from the remaining three structures is the 5-amino group. The presence of the amine marks this compound as the only one theoretically capable of forming zwitterions at physiological pH, which raises the intriguing possibility that this compound could be binding to BSEP in a higher-order state, analogous to that observed in its crystalline form (Banić-Tomišić et al., 1997). To explore this observation further, it would be important to test 5-amino analogs of salicylic amide and thiosalicylic acid. Our analysis of BSEP inhibition activity across different ion classes revealed a low frequency of BSEP inhibitors among the zwitterionic compounds tested, which suggests that this may not be a common feature of compounds that inhibit the transporter.

We determined that the compound to the right of Fig. 6b (proxicromil) was a relatively potent BSEP inhibitor with an IC50 of 30 μM.
Its matched molecular pair in the center (from which only the hydroxyl group has been removed) is remarkable in that its BSEP inhibition activity was reduced by approximately 20-fold. This result suggests that the hydroxyl group plays a role in the interaction of this compound with the transporter. Further support for this interpretation is provided by the BSEP inhibition potency of the structure on the left, in which three aliphatic carbon atoms had been removed and the hydroxyl group had been reintroduced, resulting in an overall modest decrease in ClogP compared with the noninhibitory compound in the center. Previous attempts to model the BSEP structure-activity relationship identified hydroxyl counts as promoting inhibition (Saito et al., 2009). In their case, the group was specifically bonded to aliphatic carbon atoms, whereas for our data the group was linked to an aromatic carbon atom.

Perhaps the striking structure-activity relationship lies among the compounds paired with mianserin, which are depicted in Fig. 6c.
Conversion of mianserin to mepyramine on the left reduces lipophili-
city slightly, whereas antazoline to the right features an increased
ClogP although it is likely that this is offset by the increased basicity
of the dihydroimidazole over the corresponding piperazine. The fairly
subtle differences between the physicochemical properties of the three
compounds can be contrasted with the very large observed differences
in BSEP inhibition activities and suggest that inhibition of BSEP
activity by mianserin is strongly influenced by the conformation of the
molecule.

The final pair of structures provides what is potentially the most
surprising result from the WizePairZ analysis (Fig. 6d). The conver-
sion of the 2-methylindole (mepindolol) to its naphthyl equivalent
(propranolol) induced a near 30-fold drop in BSEP inhibition activity,
despite an increase in ClogP of approximately 0.6. Although this
suggests that the indole in mepindolol makes a significant contribu-
tion to BSEP inhibition, there were no further pairs of the active
species in the data set to corroborate or invalidate this finding.

It should be noted that the absence of multiple observations makes
statistical analysis on the potential to generalize these transformations
impossible. Thus, the examples illustrated are provided on an anec-
dotal basis, intended to fuel hypothesis formulation.

Discussion

Inhibition of BSEP Is Favored by High-Molecular-Weight Hy-
drophobes. How lipophilicity and molecular weight might relate to
toxicity or adverse events has been reported earlier (Leeson and
Springthorpe, 2007), although the precise property ranges over which
this relationship takes hold are surprising. A recent study on 85
marketed drugs demonstrated the correlation between DILI in humans
and in vitro BSEP inhibition, suggesting 300 μM as a useful opera-
tional threshold (Dawson et al., 2012). Using a 300 μM cutoff, we see
a very sharp onset of BSEP inhibition within a lipophilicity and
molecular weight range that is typically considered desirable for many
drug discovery projects (Lipinski et al., 2001; Wenlock et al., 2003).
With a more than 90% chance of BSEP inhibition activity for com-
ounds with ClogP >1.7 and molecular weight >300, we would
consider routine screening for BSEP inhibition of promising chemical
series to be a prudent strategy. Thus, we would suggest a general
guidance when working with compounds on the borderline of BSEP activity to reduce molecular weight and/or lipophilicity. The contour plot (Fig. 5b) indicates how a relatively modest reduction in either of these properties may be enough to tip the balance in the investigator’s favor. Where tolerated by the primary pharmacological target, incorporation of increasingly basic functional groups is likely to be beneficial in reducing the level of BSEP inhibition, as is the systematic replacement of hydrogen bond acceptor features with donors. In particular circumstances, in which compounds exhibit more significant inhibition and the proposed route of administration for the compound allows, introducing cationic groups such as quaternary amines appears to be an almost certain way of eliminating activity.

We have previously constructed a homology model of BSEP (D. J. Warner, unpublished data), taking the X-ray structure of P-glycoprotein as a template (Aller et al., 2009) (Supplemental Figure S3). It was used to align the findings reported here with structural information available on this transporter. Aller et al. described two portals, allowing access to an inner cavity from either side of the transporter from within the inner leaflet. The strong dependence on ClogP observed for our compounds is fully consistent with this hypothesis, because the more lipophilic members of the data set are more likely to become embedded within the vesicle, where they can then gain access to the BSEP cavity. Conversely, the more polar compounds in the set appear to be unable to access this site, regardless of their molecular weight. The correlations with calculated pKa and hydrogen bonding counts both point to the pore being an unfavorable environment for polar hydrogen atoms. This result is probably due to the presence of the histidine and two arginine residues in the transmembrane domain (His72, Arg223, and Arg1034), and the apparent absence of any anionic residues in this region (Supplemental Figure S3a), consistent with an active site environment that has evolved to facilitate the transportation of cholesterol-derived acids from hepatocytes into bile.

Integrating Computational and Experimental Approaches. Our proposal for the integration of computational and experimental approaches in BSEP screening is based on our experience with the QSAR model developed above (SVM_AZdesc). If we were to prioritize only those compounds with high molecular weight and lipophilicity for experimental testing, there is a danger that active inhibitors will fail to be identified at an early stage (affecting 16% of active compounds). Our results show how the additional information encoded by this model facilitates the correct identification of positive compounds from the low-molecular-weight category.

Here we attempt to provide deeper insight into the origin of this improvement, giving drug designers the confidence to use the model in prioritizing compounds for experimental testing. In Fig. 7, we show 10 structures intended to illustrate from where the crucial improvements originate. None of these compounds were involved in training the model. The first two compounds in this figure, benzylpenicillin and buphenin, contradict our hypothesis by exposing flaws in the model. Although they would both fall into the “high” ClogP/molecular weight category, they are incorrectly classified as inactive by the model. However, these deficiencies are outweighed by the compounds in Fig. 7b for which exactly the reverse is true. All 8 of these compounds would be incorrectly classified according to their physical properties alone, because according to the scheme no compound with a molecular weight less than 296.2 can ever be classified as positive, regardless of its lipophilicity. Here we see how the additional information in the AZdesc set enables the model to correctly identify these compounds as active inhibitors of BSEP. Regrettably, the nature of some of the descriptors and the modeling method used confine this computational approach to a black box with limited interpretability. Future work should be focused on further development of transparent and predictive BSEP models, which are beyond the lipophilicity and size factors.

Lipophilicity Dependence and Inhibition of Multiple Transporters. A preference for large lipophiles is an emerging theme from published analyses of cation transporter inhibition (Kido et al., 2011; Karlgren et al., 2012), with previous studies also demonstrating reasonable summaries of experimental data sets using only a handful of simple descriptors (Ahlin et al., 2008). In contrast, it was reported that lipophilicity has very little effect on P-glycoprotein efflux (Hitchcock, 2012). In this study, only 17% of compounds remain misclassified after the simple partitioning scheme, yet they are the ones potentially holding the most information when it comes to understanding the specific recognition features that are important for BSEP. According to the published trends, we would not expect small, polar compounds to inhibit other transporters, making the BSEP active compounds in this category ideal tool compounds for further study. Ultimately, the identification (and possibly even design) of compounds that are able to discriminate between proteins in this class will be important to fully understand the consequences of specific transporter inhibition in vivo.

Limitations of the Current Analyses. At present, the relationship between in vitro potency of BSEP inhibition in membrane vesicle assays and compound concentrations within human hepatocytes in vivo, which cause functional BSEP inhibition remains undefined. Our in vitro BSEP assay conditions were similar to those used in a study reported by Dawson et al. (2012), who evaluated BSEP inhibition by 85 drugs and found that a BSEP IC50 threshold value of 300 µM provided useful discrimination between the majority of the tested drugs that caused cholestatic or mixed hepatocellular/cholestatic DILI and drugs that caused hepatocellular DILI or did not cause DILI. An IC50 cutoff value of 300 µM was therefore used in the current analysis. Of note, Morgan et al. (2010) found that numerous drugs associated with DILI in humans exhibit in vitro BSEP IC50 values less than 25 µM. Two of our potent BSEP inhibitors, verapamil and tamoxifen, have been reported by Morgan et al. not to exhibit significant BSEP inhibition. The differences in BSEP inhibition potency could be related to the different assay conditions (temperature, incubation time, and substrate concentration) used in the two studies.

Verapamil did not show significant inhibition of human BSEP-mediated [3H]taurocholate transport in Sf9 insect vesicles up to a 135 µM test concentration in the study by Morgan et al., whereas inhibition of [3H]taurocholate uptake into canalicular membrane vesicles isolated from rat liver has been demonstrated for verapamil with an IC50 value of 92.5 µM (Horikawa et al., 2003), although a species difference cannot be excluded in this case. Likewise, tamoxifen did not affect human BSEP-mediated [3H]taurocholate transport in High Five insect cell-derived membrane vesicles up to a concentration of 30 µM (Byrne et al., 2002), which is consistent with the findings by Morgan et al. (2010). However, in a study undertaken using SK-E2 cells expressing human BSEP, both verapamil and tamoxifen inhibited transporter activity, which was determined using the fluorescent substrates BODIPY or dihydrofluorescein (Wang et al., 2003).

The Sf21 insect cell expression system used in our BSEP inhibition studies has a low content of cholesterol (Paulusma et al., 2009). An increase in the cholesterol content of vesicles has been shown to increase taurocholate uptake activity, with moderate effects in insect cell-derived membrane vesicles but stronger effects in vesicles derived from liver canalicular membranes. This typically results in an increase in the transport velocity Vmax with having no significant effect on the affinity of taurocholate for the BSEP transporter (Km value) (Kis et al., 2009; Paulusma et al., 2009). Of importance, for a limited number of drugs, it has been demonstrated in the Sf21 insect cell model that BSEP inhibition potencies are not affected by the
membrane cholesterol content (Kis et al., 2009). We therefore consider that the BSEP inhibition potencies for the compounds evaluated in this study using the insect vesicles are likely to be indicative of values that would be observed in the presence of high concentrations of cholesterol.

The kinetics of BSEP inhibition may differ among drugs. Saito et al. (2009) described a BSEP inhibition QSAR model, which did not provide good prediction of inhibition by compounds, which were noncompetitive inhibitors of BSEP activity. Further studies are needed to explore the effect of the mechanism of BSEP inhibition on the performance of QSAR models.

In conclusion, we have presented a thorough investigation of the physicochemical properties of compounds in relation to their inhibition of the BSEP transporter. We conclude that, for compounds for which molecular weight is greater than 309, the most important parameter influencing BSEP inhibition potency is lipophilicity as calculated by ClogP. We also identified a statistically significant relationship with the acid dissociation constants of basic compounds, indicating that more basic compounds and those with more hydrogen bond donors tend to be less potent inhibitors. The two properties with the most marked effect on activity (molecular weight and ClogP) were used to construct a simple recursive partitioning scheme that allowed improved classification \( R^2 = 0.5 \) in training and 0.36 in tests) over any single property and served as a benchmark against which we compared more elaborate modeling strategies.

Our evaluation of a collection of molecular descriptor sets and modeling algorithms yielded a further improvement in classification of a collection of 187 compounds to which the modeling algorithm had not previously been exposed. The source of this improvement appeared to originate from the ability of the QSAR model to identify low-molecular-weight active compounds for which the partitioning scheme failed. These improved classifications appear to be the result

---

**Fig. 7.** Improvement in false-negative classification through use of the SVM_AZdesc model versus simple property-based recursive partitioning. a, two active compounds that are incorrectly flagged as negative by SVM_AZdesc but not by the RP method. b, eight active compounds that are incorrectly flagged as negative by the RP method but not by SVM_AZdesc. The molecular weight and calculated logP of each compound are expressed below the BSEP activities in the format molecular weight/ClogP.
of better molecular descriptions of the molecules as encoded by the AZdesc set rather than the nature of the classification scheme used (i.e., linear versus nonlinear methods). We have provided the full data set with this publication to allow others to emulate (or possibly surpass) this result.

Finally, we present a collection of interesting structural modifications causing a reduction in BSEP activity. Although there is no statistical evidence to support the generalizability of these modifications, with further data they could serve as an interesting starting point for those wishing to further examine the structure-activity relationship surrounding this transporter.

Acknowledgments

We are grateful to our AstraZeneca colleagues Sarah Dawson, John Cuff, Dearg Brown, Mike Rolf, and the team from Reagents and Assay Development for excellent scientific and technical assistance. We also thank Prof. Dr. Bruno Stieger (University of Zürich, Zürich, Switzerland) for the provision of the BSEP baculovirus stocks. The data set used in the analysis was generated under contract by Ricerca Biosciences LLC (Bothell, WA), and we thank Gonzalo Castillo for his excellent management of this activity.

Authorship Contributions

Participated in research design: Warner, Chen, Cantin, Kenna, Stahl, and Noeske.

Conducted experiments: Walker.

Performed data analysis: Warner and Chen

Wrote or contributed to the writing of the manuscript: Warner, Chen, Cantin, Kenna, Stahl, Walker, and Noeske.

References


