ACS Medicinal Chemistry Letters

Letter

pubs.acs.org/acsmedchemlett

De Novo Prediction of P-Glycoprotein-Mediated Efflux Liability for Druglike Compounds

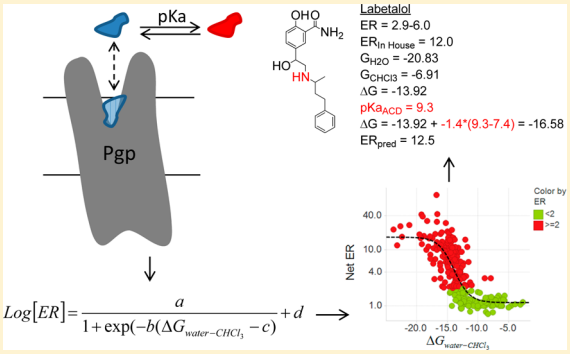
Hakan Gunaydin,*[,†] Matthew M. Weiss,[‡] and Yaxiong Sun[§]

[†]Department of Molecular Structure and [‡]Department of Medicinal Chemistry, Amgen Inc., 360 Binney Street, Cambridge, Massachusetts 02142, United States

[§]Department of Molecular Structure, Amgen Inc., One Amgen Center Drive, Thousand Oaks, California 91320, United States

Supporting Information

ABSTRACT: P-glycoprotein (Pgp) is capable of recognizing and transporting a wide range of chemically diverse compounds in vivo. Overcoming Pgp-mediated efflux can represent a significant challenge when penetration into the central nervous system is required or within the context of developing anticancer therapies. While numerous in silico models have been developed to predict Pgp-mediated efflux, these models rely on training sets and are best suited to make interpolations. Therefore, it is desirable to develop ab initio models that can be used to predict efflux liabilities. Herein, we present a de novo method that can be used to predict Pgp-mediated efflux potential for druglike compounds. A model, which correlates the computed solvation free energy differences obtained in water and chloroform with Pgp-mediated efflux (in logarithmic scale), was successful in predicting Pgp efflux ratios for a wide range of chemically diverse compounds with a R^2 and root-mean-square error of 0.65 and 0.29, respectively.



KEYWORDS: *efflux ratio, P-glycoprotein, prediction*

P-glycoprotein (Pgp) plays an important role in the in vivo disposition of a wide range of clinically relevant drugs.[1] The impact of Pgp on permeability across the blood−brain barrier (BBB) has received considerable attention within the drug discovery community.[2] Active efflux by Pgp at the BBB often represents a significant challenge in the discovery and development of agents that require central nervous system (CNS) penetration for target engagement.[3,4] In addition, overexpression of Pgp in cancer tumor cells has been implicated in the development of drug resistance to anticancer agents.[5]

Pgp is a promiscuous transporter that is capable of recognizing and transporting a wide range of chemically diverse structures.[6] An X-ray crystal structure of mouse Pgp, which shares 87% sequence identity with human Pgp, was recently reported.[1] This structure can provide insight toward understanding the mechanism by which the human Pgp transporter functions. Druglike compounds were hypothesized to be recruited to the Pgp binding pocket while they are either diffusing through the membrane or through the intracellular-facing opening.[1] Upon binding, the intracellular-facing state of Pgp undergoes an ATP-dependent state change that renders it extracellular facing and allows the substrates to be exported into the extracellular domain (Figure 1).[1]

Herein, we report our findings, which correlate the susceptibility of a compound to Pgp-mediated efflux with the free energy required for a compound to transition from a lipophilic to an aqueous environment (Figure 2) since the residues surrounding the Pgp drug binding pocket are mainly hydrophobic and aromatic. We proposed that a significant element contributing to efflux is

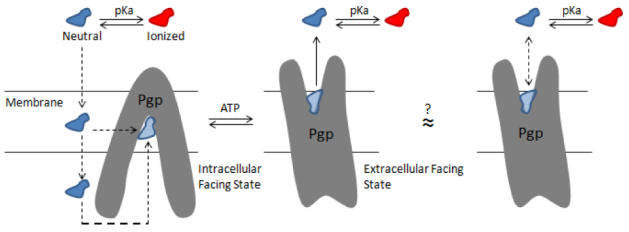


Figure 1. Proposed mechanism for the efflux of compounds by Pgp (adapted from ref 1). Neutral compounds were assumed to bind to the intracellular-facing state. Upon binding, Pgp undergoes an ATP-dependent state change and releases the compound into the extracellular domain. We hypothesized that the probability of transferring compounds from the Pgp binding pocket to the extracellular domain may be used to predict the Pgp-mediated efflux liability.

a compound's preference to be in an aqueous-like environment and (heuristically) hypothesized that Pgp-mediated efflux may be related to the probability by which a compound will be transferred from Pgp binding pocket to the extracellular domain.[7] Accordingly, Pgp transports druglike compounds to the extracellular domain when there is a strong preference for the compound to be in an aqueous-like environment. Alternatively, when the preference to be in the extracellular domain is diminished,

Received: October 5, 2012
Accepted: November 6, 2012
Published: November 6, 2012

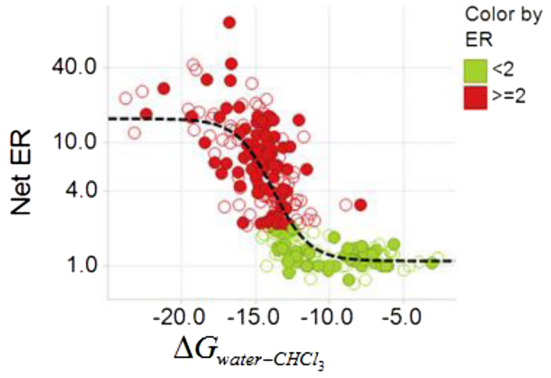


Figure 2. Relationship observed between net ER (logarithmic scale) and computed solvation free energy difference, $\Delta G_{H_2O-CHCl_3}$. The dashed line is the fit obtained for the data with eq 1. Green, compounds with net ER < 2; red, compounds with net ER ≥ 2; open circles, training set; and closed circles, test set.

the ability for Pgp to transport compounds to the extracellular domain is also diminished. This new hypothesis significantly differs from existing hypotheses about how Pgp binds or transports druglike compounds, which rely on *recognition* of polar functional groups such as amides, hydrogen bond donors, or acceptors; we speculated that these metrics may be seen as surrogates for the hydrophilicity of compounds. However, current hypothesis does not invalidate any of the literature reported hypotheses. Because the extracellular-facing state of Pgp is unknown, we proposed that the relative affinities for the compounds to be in the Pgp binding pocket (i.e., hydrophobic environment) or in the extracellular domain (i.e., aqueous-like environment) may be approximated with solvation free energies computed in chloroform (G_{CHCl_3}) and water (G_{H_2O}), respectively (see the Experimental Procedures for details.) We computed G_{CHCl_3} and G_{H_2O} for neutral species. In cases where compounds contain ionizable functional groups, such as basic amines or carboxylic acids, we utilized computed pK_a values to estimate the free energy difference between neutral and ionized forms in an aqueous environment and corrected the G_{H_2O} values with the energy difference obtained with the expression $\Delta G = -1.4(\mathrm{pH} - \mathrm{p}K_a)$. This approximation allowed us to test the hypothesis outlined in Figure 1 by using G_{CHCl_3}, G_{H_2O}, and pK_a values computed for druglike compounds.

Solvation free energy differences between water and chloroform ($\Delta G_{H_2O-CHCl_3}$) were computed for a diverse set of 282 compounds and were plotted against the logarithm of net efflux ratio (ER) (Log[net ER]) in Figure 2 (see the Experimental Procedures for details.) Net ER was calculated by dividing the ER obtained in Pgp overexpressing cell lines with the ER obtained in parental cell lines. There was a sigmoidal relationship between Log[net ER] and $\Delta G_{H_2O-CHCl_3}$, and this relationship can be described with eq 1, where $a = -1.17$, $b = 0.78$, $c = -13.90$, and $d = 1.22$ (see the Supporting Information for the relationship observed between Log[ER] and $\Delta G_{H_2O-CHCl_3}$). The R^2 and root-mean-square error obtained from such a fit for training (test) set were 0.65 (0.64) and 0.28 (0.29), respectively. Reduced χ^2 for this fit for all of the compounds was 0.9.[8] There were 134 and 148 compounds in the training and test sets, respectively. In this plot, a large negative $\Delta G_{H_2O-CHCl_3}$ suggests a strong preference for the compounds to be in an aqueous environment, while a small negative $\Delta G_{H_2O-CHCl_3}$ suggests a weaker preference for such media.

The relationship observed in Figure 2 allowed us to classify compounds into three categories: (i) compounds that are likely to be Pgp substrates and predicted to be Pgp substrates with high confidence, $\Delta G_{H_2O-CHCl_3} < 16$ kcal/mol (51 compounds); (ii) compounds that are unlikely to be Pgp substrates and predicted not to be Pgp substrates with high confidence, $\Delta G_{H_2O-CHCl_3} > -12$ kcal/mol (75 compounds); and (iii) predictions made with low confidence, $-12 < \Delta G_{H_2O-CHCl_3} < 16$ (156 compounds). The steep relationship observed between Log[net ER] and $\Delta G_{H_2O-CHCl_3}$ around −14 kcal/mol limited our ability to make accurate predictions in this range; the method allowed us to accurately predict the efflux liability for 45% of the compounds. It should be noted that the compounds included in this analysis spanned a wide range of ERs and permeability values. While a strong correlation between $\Delta G_{H_2O-CHCl_3}$ and Pgp-mediated efflux was observed, no such correlation was observed between $\Delta G_{H_2O-CHCl_3}$ and permeability, suggesting that the observed relationship is unique (see the Supporting Information for details).

$$\log[\text{net ER}] = \frac{a}{1 + \exp[-b(\Delta G_{H_2O-CHCl_3} - c)]} + d \quad (1)$$

Equation 1 was next used to predict the net ER for 12 clinically relevant drugs (Table 1).[9−12] The goal of this exercise was not to predict the absolute ER but to test the method's ability to distinguish high efflux compounds from those that would not be expected to succumb to Pgp-mediated efflux.[13] As seen in Table 1, loratadine, tacrine, and midazolam, drugs reported not to be Pgp substates, were predicted to have low ERs (≤1.2). In contrast, compounds that were reported to be substrates for Pgp were predicted to have higher ERs (1.9−16.7).

Predictions made for labetalol, midazolam, and zolmitriptan are exemplified in Figure 3. Computed G_{CHCl_3} and G_{H_2O} for labetalol are −6.91 and −20.83 kcal/mol, respectively; this corresponds to $\Delta G_{H_2O-CHCl_3}$ of −13.92 kcal/mol. Labetalol has a basic amine group, and the ACD predicted pK_a for this amine group is 9.3. Because G_{H_2O} computed for the neutral drug underestimates the polarity of labetalol in water, a pK_a-based correction was applied to account for the free energy difference between the ionized and the netural forms; we obtained a pK_a-corrected solvation free energy difference of $-13.92 + -1.4 \times (9.3 - 7.4) = -16.58$ kcal/mol. Plugging −16.58 for $\Delta G_{H_2O-CHCl_3}$ into eq 1 gave a predicted ER (ER_{pred}) of 12.5. Computed G_{CHCl_3} and G_{H_2O} for midazolam are −6.91 and −20.83 kcal/mol, respectively, and gave a $\Delta G_{H_2O-CHCl_3}$ of −6.98 kcal/mol and an ER_{pred} of 1.2. The method overestimated the ER measured for zolmitriptan; we observed that conformationally flexible molecules with distributed polarity were sometimes predicted incorrectly. The use of conformational ensembles in calculations can potentially alleviate this problem.

A significant amount of work has been directed toward understanding the physicochemical properties or molecular attributes that are associated with Pgp-mediated efflux of druglike compounds. Several physicochemical properties have been reported to correlate with Pgp recognition, including cLog *P*, topological polar surface area (TPSA), and the number of hydrogen bond donors (HBD) and acceptors (HBA).[14,15] Net ER data for compounds that are likely Pgp substrates (as defined by net ER ≥ 2) and non-Pgp substrates (as defined by net ER < 2)

dx.doi.org/10.1021/ml300314h | *ACS Med. Chem. Lett.* 2013, 4, 108−112

Table 1. Predicted and Measured ERs for 12 Drug Molecules

drug	net ER[a]	ER_{pred}	G_{H_2O}	G_{CHCl_3}	pK_{aACD}	$\Delta G_{H_2O-CHCl_3}$	ref
loratadine	0.5	1.2	−10.43	−2.68		−7.75	9
tacrine	1.2	1.1	−12.23	−5.81		−6.42	11
midazolam	0.9−1.0	1.2	−8.81	−1.83		−6.98	9, 11
zolmitriptan	2.5	15.4	−26.04	−10.64	9.5	−15.40	11
prednisone	2.8	7.3	−23.51	−8.57		−14.94	9
prazosin	2.9	11.3	−24.36	−8.20		−16.16	9
loperamide	3.8	1.9	−9.44	−0.32	9.5	−9.12	9
methysergide	4.3	9.7	−21.12	−5.47	7.2	−15.65	11
labetalol	2.4−6.5	12.5	−20.83	−6.91	9.3	−13.92	10
quinidine	7.4	5.3	−15.71	−4.52	9.6	−11.19	9
ritonavir	18.4	16.7	−23.68	2.17		−25.85	9
eletriptan	44.7	11.1	−19.09	−7.18	10.4	−11.91	11

[a]Net ER was calculated by dividing the ER obtained in Pgp overexpressing cells with the ER obtained in parental cells.

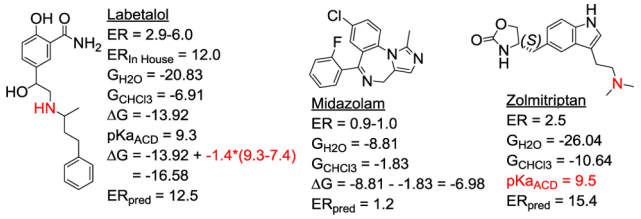


Figure 3. Predicted and measured ERs for labetalol, midazolam, and zolmitriptan. The in-house measured ER for labetalol is 12, and literature-reported values range from 2.9 to 6.0.

were plotted against cLog *P*, TPSA, HBD, and HBA for the same set of 282 Amgen compounds (Figure 4). Histograms given in Figure 4A corroborate the observations reported in the literature[14,15] and can provide empirical guidelines to help minimize Pgp-mediated efflux. For example, the following attributes are commonly used to help design low efflux compounds: the inclusion of <2 HBD, TPSA of <70 Å^2, and cLog *P* of >3. While these guidelines can be useful in the drug design process, it is well-known that compounds with the same physicochemical properties can have substantially different levels of Pgp-mediated efflux. Furthermore, these relationships fail to give clear guidelines to medicinal chemists as there are considerable overlaps between compounds with net ER ≥ 2 and net ER < 2 across each of the previously mentioned properties (Figure 4). Additionally, recognition elements required to interact with certain protein targets often do not align with those that correlate with low efflux, necessitating the need for a more refined approach. In addition to the observed relationships between molecular descriptors and ER, numerous in silico models have been developed to predict the efflux liability for druglike compounds.[16−23] Most of these models rely on existing data sets to predict the ER of new compounds based on their similarity to existing compounds with data. While such models are very useful in late stage projects when interpolations are being made, the predictive power of such models becomes questionable when new chemical space is being explored and there is little or no existing data to "train" these models. Recently, ERs of various compounds have been predicted by using a flexible docking method that makes use of the intracellular-facing state of the mouse Pgp protein.[16] While a correlation between docking score and logarithm of ERs was observed, the relevance between the binding score obtained with the intracellular-facing state and active transport by Pgp remains to be explored.

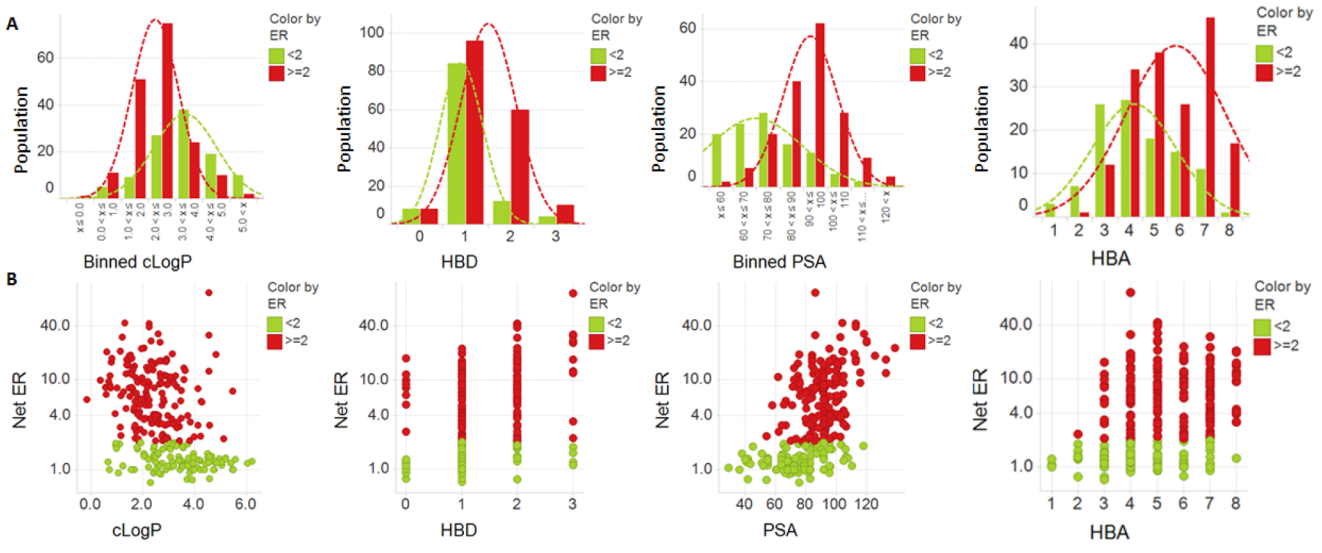


Figure 4. Relationships observed between measured Net ER and cLog *P*, HBD, TPSA, and HBA. Green, compounds with net ER < 2; red, compounds with net ER ≥ 2. Dashed lines in panel A are fits for the observed histograms.

dx.doi.org/10.1021/ml300314h | *ACS Med. Chem. Lett.* 2013, 4, 108−112

Table 2. Fragments That Highlight the Difference between TPSA and Solvation Free Energies

	 1	 2	 3	 4	 5	 6	 7	 8
TPSA	29	29	39	39	42	42	26	26
G_{H2O}	-13.81	-13.17	-7.06	-11.73	-9.61	-13.12	-6.87	-7.67
G_{CHCl3}	-6.48	-6.13	-3.03	-6.25	-4.04	-5.70	-3.27	-3.64
$\Delta G_{H2O\text{-}CHCl3}$	-7.33	-7.04	-4.06	-5.48	-5.57	-7.42	-3.60	-4.03

In addition to the guidelines mentioned above, several strategies have also been used to help modulate Pgp-mediated efflux while maintaining biological activity. For example, in cases where a HBD is required to maintain key interactions with the biological target of interest, it has been found that the inclusion of strategically placed HBAs or steric bulk can serve to mask the HBD and in doing so mitigate Pgp-mediated efflux liability.[24] Although such modifications are made in an attempt to mask polarity, the TPSAs obtained for both analogues may remain the same and provide a disconnect between what is experimentally observed and the probabilistic relationships deduced from plots like those shown in Figure 4A. Table 2 further illustrates this point with several pairs of molecular fragments. Each of these pairs has the same TPSA values but can be differentiated by their calculated solvation free energies. The pyridone (**1**)/6-methylpyridone (**2**) pair represents a case in which the polarity of the fragment is aimed to be manipulated with a steric bulk. The 1,3,5-triazine (**3**)/1,2,3-triazine (**4**) pair represents a case in which the polarity is distributed differently across the fragments. Compound **4** may be perceived to be more polar due to the close proximity of the heteroatoms. Finally, *N*-methylpicolinamide (**5**)/*N*-methylisonicotinamide (**6**) represent a pair of compounds wherein the HBD has the potential to be masked by a hydrogen bond acceptor in **5** but not in **6**.

The pyrazine (**7**)/pyrimidine (**8**) pair is another example that shows that the distribution of polarity can have an impact on ΔG_{H_2O} or $\Delta G_{H_2O-CHCl_3}$. An analysis of our internal data[25] provided 35 pairs of compounds that contain matched molecular **7**/**8** pairs with measured ER. Out of these 35 pairs, 26 pyrimidine analogues have higher net ERs than their matched pyrazine analogues, three pairs have the same net ERs, and five pyrazine analogues have higher net ERs (net $ER_{8<7}$). However, four out of these five pairs with net $ER_{8<7}$ had net ERs that are less than 2 and were considered not to be efflux substrates. We also observed that the pyrimidine analogues, on average, had 2-fold higher net ERs (see the Supporting Information for details).

In summary, we have shown that $\Delta G_{H_2O-CHCl_3}$ can be used as a de novo descriptor to predict the likely Pgp-mediated efflux liability for a diverse set of druglike molecules and supports the hypothesis outlined in this work. The method described herein can be used to rank order target compounds or make quantitative predictions upon fitting the ER data to $\Delta G_{H_2O-CHCl_3}$. Hence, we believe that it is a useful method for predicting efflux liability for druglike compounds when there is little or no data available to guide the medicinal chemistry efforts with in silico models that rely on training sets. The mechanism by which Pgp recognizes or binds druglike molecules is beyond the scope of this work, and the correlation observed between net ER and $\Delta G_{H_2O-CHCl_3}$ was not aimed to address this issue.

EXPERIMENTAL PROCEDURES

Passive permeability was determined in LLCPK-1 cell monolayers, while ER (B → A/A → B) was measured in LLCPK-1 cell monolayers that were transfected with human MDR1. All of the compounds included in this analysis demonstrated passive permeability [(A → B + B → A)/2] of >27 nm/s and efflux of <2.8 in the nontransfected cell line. These measures ensure that the efflux observed in the LLCPK-1 cells transfected with MDR is a robust measurement of efflux mediated by Pgp. Compounds included in this analysis were selected from six different programs and include 10 different subseries. See the Supporting Information for tanimoto similarities among all possible molecular pairs. ERs of the compounds in the data set range from 1 to 94.

G_{H_2O} and G_{CHCl_3} were obtained with single point solvation free energies for the ground state conformations of the compounds obtained from molecular mechanics minimizations with MMFF94x force field with the Born solvation energies. G_{H_2O} and G_{CHCl_3} were computed with the HF/6-31+G(d) method by using Gaussian 03[26] software; this method was shown to produce accurate solvation free energies.[27] The calculation of solvation free energies for one conformation introduces errors. To estimate the magnitude of the errors introduced by using only one molecular mechanics conformation, we calculated the Boltzmann weighted solvation free energies [for a conformational ensemble obtained at the B3LYP/6-31G(d) level] for 10 randomly selected compounds. Even though Boltzmann weighted G_{H_2O} and G_{CHCl_3} were different from those obtained for molecular mechanics obtained ground state conformations, there was a good correlation when Boltzmann weighted $\Delta G_{H_2O-CHCl_3}$ was plotted against $\Delta G_{H_2O-CHCl_3}$ obtained for molecular mechanics ground state conformation (see the Supporting Information for details.) The use of other organic solvents such as benzene and chlorobenzene produced similar results. We chose to use the ΔG_{CHCl_3} due to the close proximity of the chloroform dielectric constant to the dielectric constants of the residues surrounding the Pgp binding pocket. The Supporting Information contains detailed description around the diversity metrics of the compounds used in this work and the derivation of the relationship between ER and the computed solvation free energy differences.

ASSOCIATED CONTENT

Supporting Information

Permeability vs ER and cLog *P* relationships, metrics around the diversity of the compounds, derivation of the relationship between ER and $\Delta G_{H_2O-CHCl_3}$ and XYZ coordinates for the drug molecules in Table 1, an example Gaussian input file, and the presentation of the results in the form of confusion plot. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: hakang@amgen.com.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This project was supported by Amgen Innovation Initiative. We thank Dr. Margaret Chu-Moyer, Dr. Terry Rosen, and Dr. Philip Tagari for their support of this research. We thank Dr. Angel Guzman-Perez and Dr. Alan Cheng for useful discussions and valuable suggestions.

■ ABBREVIATIONS

Pgp, P-glycoprotein; BBB, blood−brain barrier; CNS, central nervous system; ER, efflux ratio; G_{H_2O}, solvation free energy in water; G_{CHCl_3}, solvation free energy in chloroform; $\Delta G_{H_2O-CHCl_3}$, solvation free energy difference between water and chloroform; HBD, hydrogen bond donor; HBA, hydrogen bond acceptor; TPSA, topological polar surface area

■ REFERENCES

(1) Aller, S. G.; Yu, J.; Ward, A.; Weng, Y.; Chittaboina, S.; Zhuo, R.; Harrell, P. M.; Trinh, Y. T.; Zhang, Q.; Urbatsch, I. L.; Chang, G. Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science* **2009**, *323*, 1718−1722.

(2) Abbott, N. J.; Romero, I. A. Transporting therapeutics across the blood-brain barrier. *Mol. Med. Today* **1996**, *2*, 106−113.

(3) Kusuhara, H.; Sugiyama, Y. Efflux transport systems for drugs at the blood-brain barrier and blood-cerebrospinal fluid barrier (Part 2). *Drug Discovery Today* **2001**, *6*, 206−212.

(4) Kusuhara, H.; Sugiyama, Y. Efflux transport systems for drugs at the blood-brain barrier and blood-cerebrospinal fluid barrier (Part 1). *Drug Discovery Today* **2001**, *6*, 150−156.

(5) Ambudkar, S. V.; Dey, S.; Hrycyna, C. A.; Ramachandra, M.; Pastan, I.; Gottesman, M. M. Biochemical, cellular, and pharmacological aspects of the multidrug transporter. *Annu. Rev. Pharmacol. Toxicol.* **1999**, *39*, 361−398.

(6) Gottesman, M. M.; Pastan, I. Biochemistry of multidrug resistance mediated by the multidrug transporter. *Annu. Rev. Biochem.* **1993**, *62*, 385−427.

(7) See the Supporting Information for the derivation and assumptions made to arrive to this conclusion.

(8) We assumed that the error in measured efflux data is within 2-fold.

(9) Feng, B.; Mills, J. B.; Davidson, R. E.; Mireles, R. J.; Janiszewski, J. S.; Troutman, M. D.; de Morais, S. M. In vitro P-glycoprotein assays to predict the in vivo interactions of P-glycoprotein with drugs in the central nervous system. *Drug Metab. Dispos.* **2008**, *36*, 268−275.

(10) Hellinger, E.; Bakk, M. L.; Pocza, P.; Tihanyi, K.; Vastag, M. Drug penetration model of vinblastine-treated Caco-2 cultures. *Eur. J. Pharm. Sci.* **2010**, *41*, 96−106.

(11) Mahar Doan, K. M.; Humphreys, J. E.; Webster, L. O.; Wring, S. A.; Shampine, L. J.; Serabjit-Singh, C. J.; Adkison, K. K.; Polli, J. W. Passive permeability and P-glycoprotein-mediated efflux differentiate central nervous system (CNS) and non-CNS marketed drugs. *J. Pharmacol. Exp. Ther.* **2002**, *303*, 1029−1037.

(12) Polli, J. W.; Wring, S. A.; Humphreys, J. E.; Huang, L.; Morgan, J. B.; Webster, L. O.; Serabjit-Singh, C. S. Rational use of in vitro P-glycoprotein assays in drug discovery. *J. Pharmacol. Exp. Ther.* **2001**, *299*, 620−628.

(13) While the net ER of the drugs listed in Table 1 spans the range 0.5−44.7, eq 1 bottoms out at 1.1 and tops out at 16.7. Hence, predicted values span the range from 1.1 to 16.7. In addition, literature-reported net ER values in Table 1 were obtained from different cell lines with different compound concentrations and Pgp expression levels.

(14) Wager, T. T.; Liras, J. L.; Mente, S.; Trapa, P. Strategies to minimize CNS toxicity: In vitro high-throughput assays and computational modeling. *Expert Opin. Drug Metab. Toxicol.* **2012**, *8*, 531−542.

(15) Gatlik-Landwojtowicz, E.; Aanismaa, P.; Seelig, A. Quantification and characterization of P-glycoprotein-substrate interactions. *Biochemistry* **2006**, *45*, 3020−3032.

(16) Dolghih, E.; Bryant, C.; Renslo, A. R.; Jacobson, M. P. Predicting binding to p-glycoprotein by flexible receptor docking. *PLoS Comput. Biol.* **2011**, *7*, e1002083.

(17) Li, W. X.; Li, L.; Eksterowicz, J.; Ling, X. B.; Cardozo, M. Significance analysis and multiple pharmacophore models for differentiating P-glycoprotein substrates. *J. Chem. Inf. Model.* **2007**, *47*, 2429−2438.

(18) Garrigues, A.; Loiseau, N.; Delaforge, M.; Ferte, J.; Garrigos, M.; Andre, F.; Orlowski, S. Characterization of two pharmacophores on the multidrug transporter P-glycoprotein. *Mol. Pharmacol.* **2002**, *62*, 1288−1298.

(19) Penzotti, J. E.; Lamb, M. L.; Evensen, E.; Grootenhuis, P. D. A computational ensemble pharmacophore model for identifying substrates of P-glycoprotein. *J. Med. Chem.* **2002**, *45*, 1737−1740.

(20) de Cerqueira Lima, P.; Golbraikh, A.; Oloff, S.; Xiao, Y.; Tropsha, A. Combinatorial QSAR modeling of P-glycoprotein substrates. *J. Chem. Inf. Model.* **2006**, *46*, 1245−1254.

(21) Huang, J.; Ma, G.; Muhammad, I.; Cheng, Y. Identifying P-glycoprotein substrates using a support vector machine optimized by a particle swarm. *J. Chem. Inf. Model.* **2007**, *47*, 1638−1647.

(22) Garg, P.; Verma, J. In silico prediction of blood brain barrier permeability: An Artificial Neural Network model. *J. Chem. Inf. Model.* **2006**, *46*, 289−297.

(23) Keefer, C. E.; Chang, G.; Kauffman, G. W. Extraction of tacit knowledge from large ADME data sets via pairwise analysis. *Bioorg. Med. Chem.* **2011**, *19*, 3739−3749.

(24) Hitchcock, S. A. Structural modifications that alter the P-glycoprotein efflux properties of compounds. *J. Med. Chem.* **2012**, *55*, 4877−4895.

(25) This data set is different than the 282 compound data set used in Figures 2 and 4. Similar compounds were removed form the 282 compound data set to ensure diversity.

(26) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision D.01; Gaussian, Inc.: Wallingford, CT, 2003.

(27) Takano, Y.; Houk, K. N. Benchmarking the Conductor-like Polarizable Continuum Model (CPCM) for Aqueous Solvation Free Energies of Neutral and Ionic Organic Molecules. *J. Chem. Theory Comput.* **2005**, *1*, 70−77.