

# UCSF

## UC San Francisco Previously Published Works

### Title

Predicting binding to p-glycoprotein by flexible receptor docking.

### Permalink

<https://escholarship.org/uc/item/9w60j913>

### Journal

PLoS computational biology, 7(6)

### ISSN

1553-734X

### Authors

Dolghih, Elena

Bryant, Clifford

Renslo, Adam R

et al.

### Publication Date

2011-06-01

### DOI

10.1371/journal.pcbi.1002083

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

# Predicting Binding to P-Glycoprotein by Flexible Receptor Docking

Elena Dolgikh<sup>1\*</sup>, Clifford Bryant<sup>2</sup>, Adam R. Renslo<sup>1,2</sup>, Matthew P. Jacobson<sup>1</sup>

**1** Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California, United States of America, **2** Small Molecule Discovery Center, University of California, San Francisco, San Francisco, California, United States of America

## Abstract

P-glycoprotein (P-gp) is an ATP-dependent transport protein that is selectively expressed at entry points of xenobiotics where, acting as an efflux pump, it prevents their entering sensitive organs. The protein also plays a key role in the absorption and blood-brain barrier penetration of many drugs, while its overexpression in cancer cells has been linked to multidrug resistance in tumors. The recent publication of the mouse P-gp crystal structure revealed a large and hydrophobic binding cavity with no clearly defined sub-sites that supports an “induced-fit” ligand binding model. We employed flexible receptor docking to develop a new prediction algorithm for P-gp binding specificity. We tested the ability of this method to differentiate between binders and nonbinders of P-gp using consistently measured experimental data from P-gp efflux and calcein-inhibition assays. We also subjected the model to a blind test on a series of peptidic cysteine protease inhibitors, confirming the ability to predict compounds more likely to be P-gp substrates. Finally, we used the method to predict cellular metabolites that may be P-gp substrates. Overall, our results suggest that many P-gp substrates bind deeper in the cavity than the cyclic peptide in the crystal structure and that specificity in P-gp is better understood in terms of physicochemical properties of the ligands (and the binding site), rather than being defined by specific sub-sites.

**Citation:** Dolgikh E, Bryant C, Renslo AR, Jacobson MP (2011) Predicting Binding to P-Glycoprotein by Flexible Receptor Docking. *PLoS Comput Biol* 7(6): e1002083. doi:10.1371/journal.pcbi.1002083

**Editor:** James M. Briggs, University of Houston, United States of America

**Received:** February 4, 2011; **Accepted:** April 22, 2011; **Published:** June 23, 2011

**Copyright:** © 2011 Dolgikh et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This work was supported by NIH grant AG021601 (to MPJ) and by research support from the Sandler Foundation (to ARR). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** I have read the journal's policy and have the following conflicts: MPJ is a consultant to Schrodinger LLC.

\* E-mail: lena.dolgikh@ucsf.edu

## Introduction

P-glycoprotein (P-gp) is an ATP-dependent transport protein that is selectively expressed at entry points of xenobiotics in tissues such as the intestinal epithelium, capillary brain endothelium, and kidney proximal tubules among others [1]. Acting as an efflux pump, it prevents exogenous substances from entering sensitive organs and, as such, plays a key role in the absorption and blood-brain barrier penetration of many drugs, affecting their distribution and elimination [2,3]. Moreover, overexpression of this protein, also known as MDR1, has been linked to multidrug resistance (MDR) in cancer tumor cells where higher levels of the protein result in increased efflux of chemotherapeutic compounds [4]. Finally, there is also accumulating evidence that P-gp, in addition to its role in drug transport, may transport endogenous molecules such as signaling lipids, and play a role in tumor biology and cancer progression [5].

A major hurdle in the drug discovery process [6,7], P-gp has inspired the development of several assays aimed at identifying its substrates [8,9]. One widely used assay, the monolayer efflux ratio (ER) assay, measures transport rates of molecules in different directions across a single layer of specialized cells. The ratio or difference of the two rates, basal-to-apical and apical-to-basal, is used to identify P-gp substrates. Another commonly used assay, aimed at identifying P-gp inhibitors as well as substrates, is the calcein-AM (CAM) inhibition assay, in which accumulation of the fluorescent calcein molecule inside the cells indicates an interaction between P-gp and the molecule being tested.

Despite being widely used, both assays have limitations [10]. For example, the monolayer efflux assay may fail to identify P-gp substrates with high passive permeability (>300 nm/s) because efflux by P-gp can be masked by the high diffusion rate of the compounds through the membrane. There is also no standard value of the efflux ratio used to distinguish substrates from non-substrates, with cutoff values from 1.5 to 3 being used [11,12,13,14]. Because the CAM assay is based on the competitive inhibition of calcein transport by compounds that interact with P-gp, the assay may not detect P-gp substrates with low passive membrane diffusion rates that reach the P-gp binding site at a much slower rate than the fluorescent compound. Both assays are also expensive and time consuming, and results in different labs can vary significantly. For example, midazolam has been identified as a nonsubstrate [13], an inhibitor [8,11], a substrate [15] and an inducer [16] in different studies. Doxorubicin, resistance to which has been linked to P-gp overexpression, is another example [17]. The drug has been cited repeatedly as a classical P-gp substrate [18,19]; however, it has also been classified as a nonsubstrate by several *in vitro* studies [8,13].

To complement experimental assays, several *in silico* methods have been developed to predict P-gp binding. Among them, pharmacophore models based on anywhere from two [20] to nine [10] features to an ensemble of 100 pharmacophores [21] have been generated. Other approaches have included quantitative structure-activity relationship (QSAR) models and machine-learning algorithms, some of them incorporating up to 70 descriptors [22,23,24]. Even though several of the methods report sensitivity of 80% or

## Author Summary

With many drugs failing in the preclinical stages of drug discovery due to undesirable ADMETox (absorption, distribution, metabolism, excretion and toxicity) properties, improvement of these properties early on in the process, alongside the optimization of the compound activity, is emerging as a new focus in the pharmaceutical field. One of the key players affecting pharmacokinetic profiles of many clinically relevant compounds is an active efflux transporter, P-glycoprotein. Expressed predominantly at various physiological barriers, it can influence drug absorption (intestinal epithelium, colon), drug elimination (kidney proximal tubules) and drug penetration of the blood-brain barrier (endothelial brain cells). Moreover, its increased expression in cancer cells has been linked to resistance to multiple drugs in tumors. In this study we describe a computational approach that allows prediction of which compounds are more likely to interact with P-gp. We have tested the ability of this method to differentiate between binders and nonbinders of P-gp by using consistently measured *in vitro* experimental data. We also implemented a blind test on a series of peptidic cysteine protease inhibitors with encouraging outcome. Overall, our results suggest that this method provides a qualitative, quick, and inexpensive way of evaluating potential drug efflux problem at the early stages of drug development.

higher, the extraordinary chemical diversity of the P-gp substrates, reflected by the large numbers of pharmacophores and descriptors, have frustrated efforts to make sense of the chemical data. A multitude of theories about the number, sizes, and locations of the binding sites has further complicated the issue [25,26,27,28].

The recent publication of the mouse P-gp crystal structure [29] (87% identical amino acid sequence to human P-gp) presents an opportunity to develop new prediction methods that take advantage of the receptor structural information to not only identify molecules that bind to P-gp but also to guide chemical optimization, e.g., to attempt to modify interactions with the protein. Among the three published structures, two were crystallized with stereoisomers of a cyclic inhibitor, QZ59, that defines the drug-binding cavity. All structures are in an inward-facing conformation that is believed to be relevant for initial substrate recognition.

Located in the transmembrane region, the large and hydrophobic drug-binding pocket is lined with various aromatic side chains and has no clearly defined sub-sites. Taking into account the well-known ability of the protein to accommodate ligands of various shapes and sizes, such an arrangement strongly supports the “induced fit” ligand binding model proposed by Loo *et al.* [30]. This, combined with the relatively low resolution of the structure, suggests that it is essential to treat the binding site as flexible while modeling binding site interactions, which we demonstrate explicitly in our results here. In this study, we employ a flexible receptor docking method [31], together with scoring methods that include the Glide XP scoring function [32] and a molecular mechanics scoring function with generalized Born implicit solvent (MM-GB/SA) [33,34], to develop a new prediction algorithm for P-gp binding specificity. We benchmark the method in several ways, including a blind test on a series of peptidic cysteine protease inhibitors, confirming the ability to predict compounds more likely to be P-gp substrates. We also apply this approach to evaluate the ability of P-gp to discriminate endogenous vs. exogenous compounds, and to predict that several endogenous metabolites may be P-gp substrates. Overall, our results suggest that specificity in P-gp is better understood in terms of physicochemical properties of

the ligands (and the binding site), rather than being defined by specific sub-sites. We also suggest that many P-gp substrates bind deeper in the cavity than the cyclic peptide in the crystal structure.

## Methods

### Computational Methods

All molecular docking calculations were performed using the mouse P-glycoprotein crystal structure (Protein Data bank [PDB] code 3G60). BLAST [35] sequence alignment with human P-gp revealed 87% overall sequence identity and ~100% identity within the binding cavity with the exception of mSer725/hAla729 directly facing the binding cavity. The docking calculations were performed using Glide (version 5.6) [36] with the OPLS2005 force field [37,38]. The receptor structure was prepared and minimized within the Protein Preparation Wizard.

For rigid docking, a rigid receptor grid defined by a  $10 \times 10 \times 10$  Å inner box was generated. The docking site was either defined by the centroid of the co-crystallized ligand, QZ59-RRR, with a center at (19.1, 52.3, -0.3) Å or defined higher than the original ligand with the center at (19.0, 46.0, -6.0) Å (same as in induced fit docking, see below). All ligands were docked in both standard precision (SP) Glide and extra precision (XP) Glide modes (Figure S1).

Flexible receptor docking was performed using a multi-stage induced fit docking protocol (IFD) [31] as implemented in Schrödinger Suite 2010. Briefly, in the first stage, the van der Waals radii of protein and ligand are scaled by a factor of 0.5 and ligands are docked into the receptor using the default Glide SP mode. Next, Prime is used to predict and optimize selected protein side chains (details below). Finally, the poses are scored and filtered, after which ligands are redocked using Glide XP mode and scored. Final scoring in this work was implemented using the extra precision (XP) Glide [32] scoring function and an MM-GB/SA [33,34] rescoring function.

The specific protocol was developed and refined using well-known P-gp substrates and inhibitors from Table S1. Since no binding modes are known for any compounds (with the exception of QZ59), optimal parameters were selected based on binding scores and the ability to distinguish binders from non-binders (as described below). The parameters we varied included the inner box coordinates, van der Waals radii scaling, the number of poses saved, and the number and identity of ‘trimmed’ (mutated temporarily to Ala) residues in the initial docking stage. Specifically, we chose to delete the side chains of Phe71, Phe332, and Phe728 in the 1<sup>st</sup> docking stage. These three residues are located in the center of the cavity, and were most responsible for preventing potent inhibitors from achieving good scores, by binding deeper in the cavity.

In the primary IFD round, a  $10 \times 10 \times 10$  Å inner box with coordinates (19.0, 47.0, -6.0) Å was used, which is centered deeper in the binding cavity than the cyclic peptide in the crystal structure, and roughly centered on the docked poses of the initial test set. At this stage, all residues lining the cavity were optimized by Prime [39] (Table S2), whereas in the following IFD round, only residues within 5 Å of each ligand were minimized. The number of poses saved during the initial docking was set to 100. For all subsequent docking calculations inner box coordinates (19.0, 46.0, -6.0) Å were used. For each ligand, up to 20 top poses were saved and scored with the Glide XP function and MM-GB/SA.

Ligand coordinates were obtained from the DrugBank [40] and PubChem Compound (<http://pubchem.ncbi.nlm.nih.gov/>) databases and processed using the *Ligprep* 2.4 module. The parameters were assigned based on the OPLS2005 force field. For the QZ59-RRR ligand, selenium atoms were replaced with sulfur atoms. For

molecules with stereocenters, only the known active forms were docked. For drugs used as racemic mixtures, both stereoisomers were investigated. The isomer with the more favorable docking score was used in the data analysis. Ionization states were assigned by *Epkik*, and groups with pKa between 5 and 9 were treated as neutral while those outside the range were treated as charged.

Initial testing of our approach was conducted with two datasets. One of them was comprised of 24 well-known P-gp binders from Table 1 of the review article by Hennessy *et al.* [41] and 102 endogenous molecules selected from the KEGG database [42] to represent different classes of biological compounds. The rationale for this first test was that most endogenous molecules would not be effluxed by P-gp, providing insight into how P-gp discriminates between endogenous and exogenous molecules. The second dataset was based on the Doan *et al.* study of FDA approved drugs [11] that generated consistent experimental data from the monolayer efflux ratio as well as the calcein-inhibition assays. We used the intersection of the results in the two assays to define sets of compounds that were clear P-gp substrates (i.e., positive in both assays) or non-substrates. We did not consider the compounds that were positive in only one of the two assays. The complete list of compounds and their scores are provided in Tables S1 and Tables S3, S4, and S5.

## Synthetic Methods

Reagents and solvents were purchased from Aldrich Chemical, Alfa Aesar, Chem Impex international or TCI America and used as received. Reactions were carried out under an argon atmosphere in oven-dried glassware using anhydrous solvents from commercial suppliers. Air and/or moisture sensitive reagents were transferred via syringe or cannula and were introduced into reaction vessels through rubber septa. Solvent removal was accomplished with a rotary evaporator at ~10–50 Torr. Automated column chromatography was carried out using a Biotage SP1 system and silica gel cartridges from Biotage or Silicycle. Analytical TLC plates from EM Science (Silica Gel 60 F<sub>254</sub>) were employed for TLC analyses. <sup>1</sup>H NMR spectra were recorded on a Varian INOVA-400 400 MHz spectrometer.

Analogs **1** [43,44], **3**, **6**, and **9** were synthesized in one step from commercially available *N*-benzyloxycarbonyl (Cbz) protected amino acids according to the following general procedure. A solution of the *N*-benzyloxycarbonyl protected L-amino acid (0.33 mmol) in 2 mL of DMF was treated with aminoacetonitrile bisulfate (0.37 mmol, 1.1 equiv.), 1-hydroxybenzotriazole (0.33 mmol,

1.0 equiv), *N*-(3-dimethylaminopropyl)-*N'*-ethylcarbodiimide hydrochloride (0.67 mmol, 2.0 equiv.), and *N,N*-diisopropylethylamine (2.0 mmol, 6.0 equiv.). The reaction was stirred at room temperature and monitored until judged complete by TLC or HPLC. The reaction mixture was then poured into ethyl acetate and the resulting organic solution washed in succession with aqueous 1 N HCl (for non-basic analogs only), 50% aqueous NaHCO<sub>3</sub>, saturated aqueous NaCl, and then dried (MgSO<sub>4</sub>), filtered, and concentrated. The crude product thus obtained was purified using automated silica gel flash chromatography (Biotage SP1, ethyl acetate-hexane) to afford the desired products.

Analogs **2**, **4**, **5**, **7**, and **10** were synthesized in two steps from *N*-(benzyloxycarbonyl)-L-serine lactone [45] according to the following procedure. A solution of *N*-(benzyloxycarbonyl)-L-serine lactone (0.45 mmol) in 2 mL of acetonitrile was added dropwise to a solution of the relevant amine or *N*-trimethylsilylamine (1–10 equivalents depending on the amine, see below) in ~3 mL of acetonitrile. The reaction was monitored at room temperature or in some cases heated at 50°C, depending on the reactivity of the amine (see below). When the reaction was judged complete by TLC or HPLC, the reaction mixture was concentrated and the desired amino acid separated from undesired amide side product in one of the following ways. For the amino acid leading to **2**, the crude product was partitioned between ethyl acetate and water and the water phase (containing the desired product) was then lyophilized. For the amino acids leading to **4** and **5**, purification by automated silica gel flash chromatography (Biotage SP1, methanol-dichloromethane) afforded the desired amino acids. For intermediate amino acids leading to **7** and **10**, the crude residue was partitioned between dichloromethane and 1 N aqueous NaOH, followed after phase separation by acidification of the aqueous phase with 1 N HCl to effect precipitation of the amino acid, which was collected on a filter, washed with cold water, and dried. The procedures described above provided the desired amino acids in sufficient purity for use in the subsequent coupling reaction with aminoacetonitrile, which was carried out according to the general coupling protocol described for analogs **1**, **3**, **6**, and **9** above.

Analog **8** was prepared in three steps by reaction of indoline with *N*-(benzyloxycarbonyl)-L-serine lactone as described above, using automated silica gel flash chromatography (Biotage SP1, methanol-dichloromethane) to isolate the desired amino acid. The amino acid intermediate was coupled to aminoacetonitrile

**Table 1.** Peptidic cysteine protease inhibitors.

Compound	Mean P <sub>app</sub> (10 <sup>-6</sup> cm/s)		Mean Efflux Ratio	Glide XP (kcal/mol)
	A to B	B to A		
1	24.4	29.6	1.2	-10.9
2	10.6	16.7	1.6	-9.9
3	21.8	30.9	1.4	-12.7
4	19.6	31.8	1.6	-12.9
5	20.3	36.4	1.8	-13.0
6	4.75	42.8	9.0	-14.6
7	10.3	45.5	4.4	-13.2
8	14.5	34.2	2.4	-14.2
9	0.51	2.9	5.8	-12.3
10	8.41	25.8	3.1	-11.0

doi:10.1371/journal.pcbi.1002083.t001

according to the general procedure and finally, the resulting indoline product was oxidized to the desired indole **8** by reaction with 1.05 equivalents of 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ) in dichloromethane for 30 minutes. The final product was purified by automated silica gel flash chromatography (Biotage SP1, ethyl acetate-hexane).

Additional details and NMR data are provided in Supplementary Methods (Text S1).

### MDCK-MDR1 Monolayer Permeability Assay

Permeability measurements were performed by Wuxi Apptec, using the following procedures.

MDCK-MDR1 cells (obtained from Piet Borst at the Netherlands Cancer Institute) were seeded onto polyethylene membranes (PET) in 96-well BD insert systems at  $2 \times 10^5$  cells/cm<sup>2</sup> for 4–6 days to obtain confluent cell monolayer formation. Test compounds were diluted with the transport buffer (HBSS, pH 7.4) from a 10 mM stock solution to a concentration of 2  $\mu$ M and applied to the apical (A) or basolateral (B) side of the cell monolayer. Permeation of the test compounds from the A to B direction or B to A direction was determined in triplicate over a 150-minute incubation at 37°C and 5% CO<sub>2</sub> with a relative humidity of 95%. In addition, the efflux ratio of each compound was also determined. Test and reference compounds were quantified by LC-MS/MS analysis based on the peak area ratio of analyte/IS.

The apparent permeability coefficient  $P_{app}$  (cm/s) was calculated using the equation:

$$P_{app} = (dC_r/dt) \times V_r / (A \times C_0)$$

where  $dC_r/dt$  is the cumulative concentration of compound in the receiver chamber as a function of time ( $\mu$ M/s);  $V_r$  is the solution volume in the receiver chamber (0.075 mL on the apical side, 0.25 mL on the basolateral side);  $A$  is the surface area for the transport, i.e. 0.084 cm<sup>2</sup> for the area of the monolayer;  $C_0$  is the initial concentration in the donor chamber ( $\mu$ M).

The efflux ratio (ER) was calculated using the equation:

$$ER = P_{app}(BA) / P_{app}(AB)$$

Percent recovery was calculated using the equation:

$$\% \text{ Recovery} = 100 \times [(V_r \times C_r) + (V_d \times C_d)] / (V_d \times C_0)$$

$$\% \text{ Total recovery} =$$

$$100 \times [(V_r \times C_r) + (V_d \times C_d) + (V_c \times C_c)] / (V_d \times C_0)$$

Where  $V_d$  is the volume in the donor chambers (0.075 mL on the apical side, 0.25 mL on the basolateral side);  $C_d$  and  $C_r$  are the final concentrations of transport compound in donor and receiver chambers, respectively.  $C_c$  is the compound concentration in the cell lysate solution ( $\mu$ M), and  $V_c$  is the volume of insert well (0.075 mL in this assay).

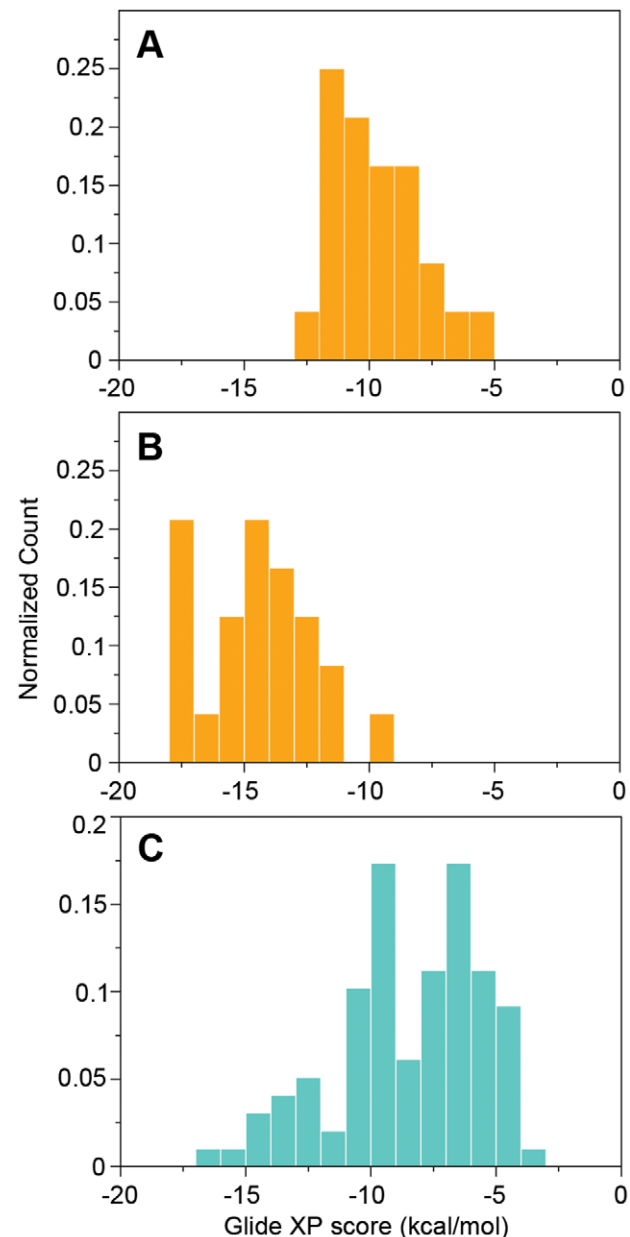
Permeability determinations were performed in triplicate and are reported as mean values. The mean total recovery was greater than 90% for all compounds tested (compounds **1–10**).

## Results

We initially developed the docking strategy using a number of well-known P-gp substrates and inhibitors. Specifically, we used a set of 24 known binders (Hennessy *et al.* [41], Table 1) that

included, among others, HIV protease inhibitors, anthracyclines, vinca alkaloids, and taxanes (Table S1). Initial docking using a rigid receptor and docking box coordinates centered on the co-crystallized ligand generated poses that largely overlapped with the coordinates of the cyclic peptide in the crystal structure, with most of the compounds showing extensive exposure to solvent at the base of the cavity. By contrast, the flexible-receptor docking protocol resulted in ligands receiving much more favorable docking scores (Figure 1) as well as ligand poses in which the ligands bound much ‘deeper’ in the cavity (Figure 2), with little solvent exposure.

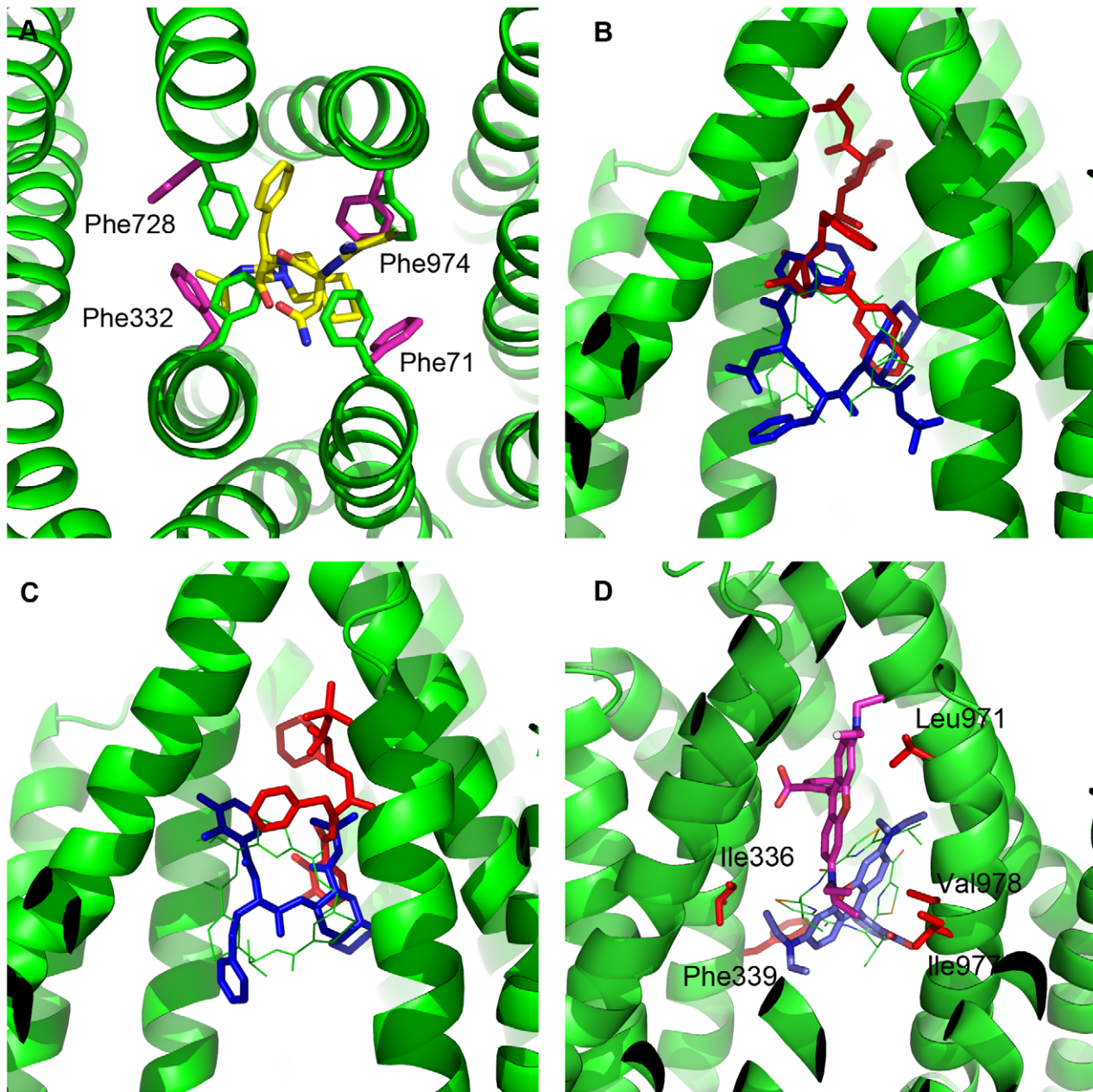
Only side chains in the binding site were treated as flexible (Table S2), and the conformational changes that allowed the ligands to bind more deeply in the cavity were modest. The most



**Figure 1. Distributions of docking scores (Glide XP).** Metabolites/P-gp binders set: **A**) Rigid docking, binders; **B**) Flexible docking, binders; **C**) Flexible docking, metabolites.

doi:10.1371/journal.pcbi.1002083.g001





**Figure 2. Flexible versus rigid docking.** **A**) View of P-gp from above (i.e. viewed from the luminal face, perpendicular to the membrane plane) of the top-scored saquinavir pose (yellow) from the primary round of the flexible receptor docking. Phe71, Phe332, and Phe728 were mutated to Ala in the 1<sup>st</sup> Glide docking stage. The final positions of the three residues and Phe974 are shown in pink, while those of the original crystal structure are in green. **B**) Top-scored poses from second round of flexible receptor docking (red), Glide XP = -17.8 kcal/mol compared to top the pose from rigid docking (blue), Glide XP = -10.0 kcal/mol for saquinavir. **C**) Analogous results for nelfinavir. Flexible docking (red), Glide XP = -15.6 kcal/mol; rigid docking (blue), Glide XP = -8.9 kcal/mol. **D**) Rhodamine B. Flexible receptor docking (pink), Glide XP = -15.3 kcal/mol. Rigid docking (blue), Glide XP = -5.4 kcal/mol. In red, are residues shown experimentally to interact with a cysteine analogue of rhodamine B. QZ59-RRR is shown for reference in light green.  
doi:10.1371/journal.pcbi.1002083.g002

important conformational changes were those of the side chains of Phe71, Phe332, and Phe728, which are located in the center of the cavity. As shown in Figure 2A, the rotamer changes in these side chains result in a more open cavity than that in the initial crystal structure. Representative top poses and Glide XP binding scores of some of the compounds from the final IFD round are shown in Figures 2B and 2C. Additional scores are provided in Table S1, and are represented as a histogram in Figure 1, highlighting the much more favorable docking scores achieved with flexible-receptor

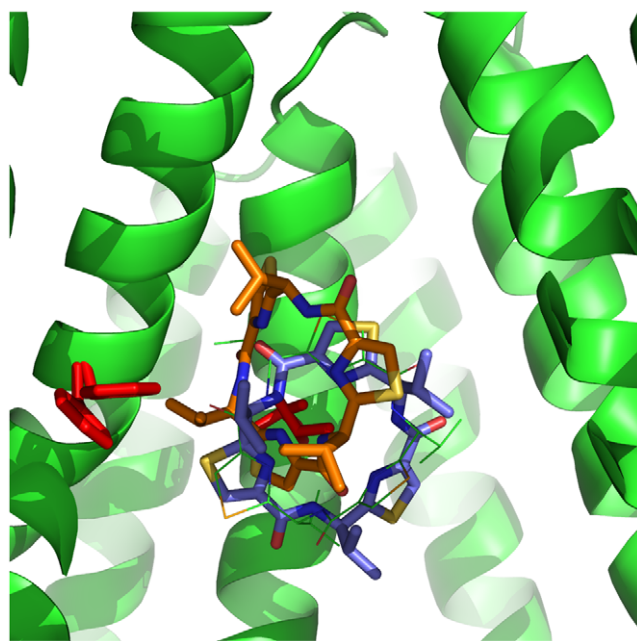
docking. Below, we also demonstrate that the flexible receptor approach greatly improves the ability to discriminate binders from non-binders.

In all of this work, we used two different scoring functions to rank compounds, Glide XP and a molecular mechanics based scoring function (MM-GB/SA), in addition to the default Glide SP scoring function. The results using Glide XP and MM-GB/SA scoring were, on average, remarkably similar, considering the very different functional form and methods of parameterization. Both

scoring functions performed much better than Glide SP in distinguishing binders from non-binders. (On the other hand, the Glide XP and MM-GB/SA scoring functions did not always identify the same poses as top-ranked; see, e.g., doxorubicin in Figure S2. Given the size and flexibility of the binding cavity, it is likely that for some, if not all molecules, several binding modes are possible.) For simplicity, we present mainly the results using Glide XP here, and present the remaining results using MM-GB/SA in supplementary tables and figures, in part because the results with Glide XP are slightly better in some cases. This is not altogether surprising because the molecular mechanics scoring function has been useful primarily in ranking compounds that are chemically similar, e.g., congeneric series, and the series of compounds we use in most of the tests here are quite diverse. However, the similarity of the results using the very different scoring functions is striking, and we use the results with MM-GB/SA scoring to investigate, for example, the role of desolvation in binding by P-gp.

Also shown in Figure 2D are the rigid- and flexible-receptor poses of a well-known (non-drug) P-gp substrate, rhodamine B, the binding mode of which has been partially elucidated by experimental data obtained using its Cys-linked analog. The flexible-receptor pose selected by Glide XP is qualitatively consistent with the experimental data in that the molecule is reasonably close to residues facing the binding cavity and shown to interact with the ligand [26,46]. By comparison, the pose obtained by rigid docking appears less consistent with the experimental data and also has a much less favorable docking score ( $-5.4$  kcal/mol, vs.  $-15.3$  kcal/mol for the flexible receptor pose).

As an additional control, we docked QZ59-RRR back into the crystal structure using both rigid and flexible docking protocols. The results are illustrated in Figure 3. Rigid docking reproduced the binding mode of QZ59, but the molecule was ‘flipped’ in comparison to its position in the crystal structure. Induced fit docking produced a similar pose with the molecule slightly shifted



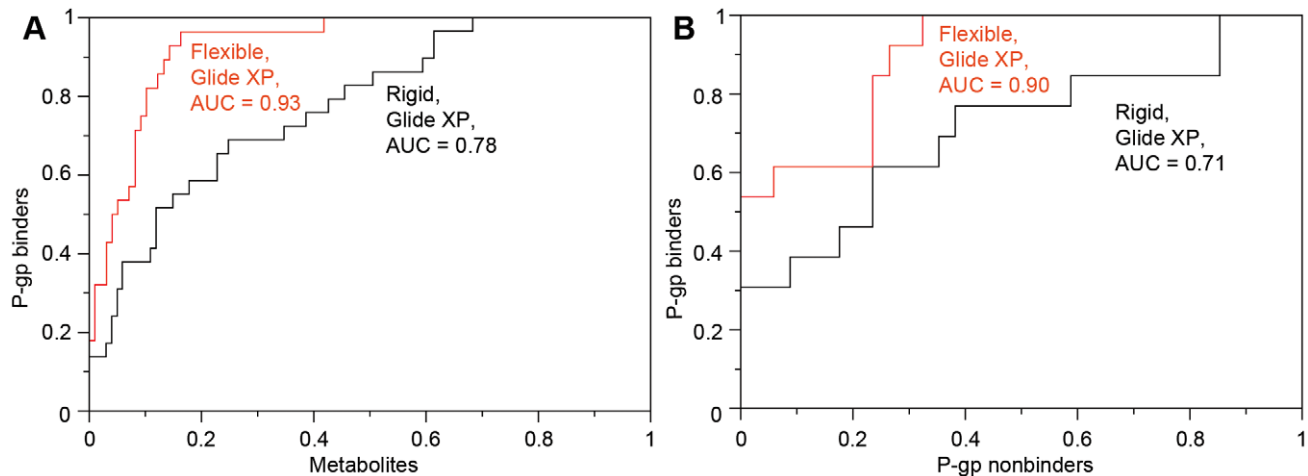
**Figure 3. Flexible and rigid docking results for QZ59-RRR.** Flexible receptor docking (orange), Glide XP =  $-10.2$  kcal/mol. Rigid docking (blue), Glide XP =  $-8.4$  kcal/mol. QZ59-RRR from the crystal structure is shown for reference in light green. In red are two residues believed to be important for ligand binding: Phe724 and Val978. doi:10.1371/journal.pcbi.1002083.g003

upwards from the original position. (The shift was seen regardless of whether the docking box was centered on the original ligand position or shifted deeper into the cavity). The ligand still maintained contact with Phe724 and Val978 deemed to be important for drug binding (as discussed in Aller paper), as well as with the majority of the hydrophobic residues indicated to be within 4–5 Å of the crystal pose (Table S6). The flexible docking score in this case, though not particularly high, is also more favorable than that from rigid docking. Based on the  $IC_{50}$  value reported for QZ59-RRR inhibition of verapamil-stimulated ATPase activity ( $4.8 \pm 2.6$   $\mu$ M), the ligand is a rather weak inhibitor, which could partially explain the weak binding score. In addition, as discussed in Methods, it was not possible to dock the compound containing selenium atoms, and these were replaced with sulfur.

We next docked a set of 102 common metabolites, comprising representatives of four major classes of biological molecules including carbohydrates, amino acids, fatty acids, and nucleic acids (Table S3). We reasoned that most of these metabolites would be non-binders based on the notion that their efflux would be inefficient to cell function (as would inhibition of P-gp by metabolites). As shown in Figure 1, most of the metabolites did in fact have much less favorable docking scores than the drugs discussed above. However, a small fraction of the metabolites received docking scores similar to those of the drugs. For example, of the 26 drugs in Table S1 known to interact with P-gp, 23 had Glide XP scores of  $-12$  kcal/mol or better, and 15 had scores  $< -14$  kcal/mol. By contrast, 18 of the 102 metabolites received scores more favorable than  $-12$  kcal/mol, and only 8 had scores  $< -14$  kcal/mol. Among these metabolites with favorable scores, we subsequently identified four (thyroxin, vitamin D3, progesterone, and cholesterol) that have been reported to interact with P-gp [47,48,49,50] and reassigned them to the binders set (which had little effect on the ROC-type curve). We also searched for literature data on approximately 20 randomly selected metabolites with less favorable docking scores and were unable to find any evidence of these being P-gp substrates. We also could not identify any direct evidence for other top-scoring metabolites, such as riboflavin, retinol, and leukotriene C4, interacting with P-gp, but it is possible that some of these metabolites are currently unrecognized substrates (or inhibitors). In fact, P-gp has been suggested to export naturally derived toxins in healthy cells [51] as well as to play a role in transport of cancer-signaling lipids [5]. Investigation of a more extensive set of biologically relevant molecules is currently under way.

In the absence of any direct evidence for the other metabolites, we consider them non-binders, and we quantify the ability to discriminate the known binders (26 drugs+4 metabolites) and presumed non-binders (98 metabolites) using an ROC-type curve in Figure 4A. Clearly, the flexible-receptor protocol results in much better discrimination between these two sets of compounds (area under the curve, AUC = 0.93) than the rigid receptor docking, either with the docking box centered on the co-crystallized QZ59 ligand (AUC = 0.78, Figure 4A), or with the docking box shifted deeper into the cavity as in the flexible docking results (AUC = 0.83, Figure S1). The results using flexible-receptor docking and the MM-GB/SA scoring function are very similar in this case, AUC = 0.93 (Figure S3).

Next, we tested the ability to qualitatively reproduce results of *in vitro* assays regularly used to evaluate P-gp binding. For that purpose we selected the Doan *et al.* [11] dataset of FDA approved drugs that included results of the monolayer efflux and CAM inhibition assays. Based on the assay results, we defined P-gp binders ( $N = 13$ ) as molecules positive for both assays ( $ER > 1.5$  and  $> 10\%$  CAM inhibition), and nonbinders ( $N = 34$ ) as compounds



**Figure 4. ROC-type curves (Glide XP) for metabolites/P-gp binders set (A) and Doan *et al.* dataset (B).**  
doi:10.1371/journal.pcbi.1002083.g004

negative for both assays ( $ER < 1.5$  and  $< 10\%$  CAM inhibition). The ROC curves obtained with Glide XP and the default treatment of ionization (see Methods) are shown in Figure 4B. The induced fit approach (AUC = 0.90) again outperformed rigid docking (AUC = 0.71), although there is clearly some overlap in the distribution of scores between the binders and non-binders in this set. Some of this overlap is due to the somewhat arbitrary criteria used for distinguishing the sets of compounds, as discussed below.

The results using the MM-GB/SA (AUC = 0.81) scoring function were somewhat worse than using Glide XP (Figure S3), using the default treatment of ionization states. However, when all compounds were docked as neutral species, regardless of their pKa values, the MM-GB/SA scoring function performed much better (AUC = 0.92, Figure S4); the results using the Glide XP scoring function were similar (AUC = 0.92). It is not surprising that the MM-GB/SA scoring function is more sensitive to the treatment of protonation states, since charged compounds have large, unfavorable desolvation energies. It is not completely clear why treating all compounds as neutral results in better discrimination, although we note in this regard that one of the prevalent theories in the field is that P-gp substrates enter the binding cavity from the membrane (where they are assumed to be electrically neutral) rather than directly from the aqueous environment of the cytoplasm [4].

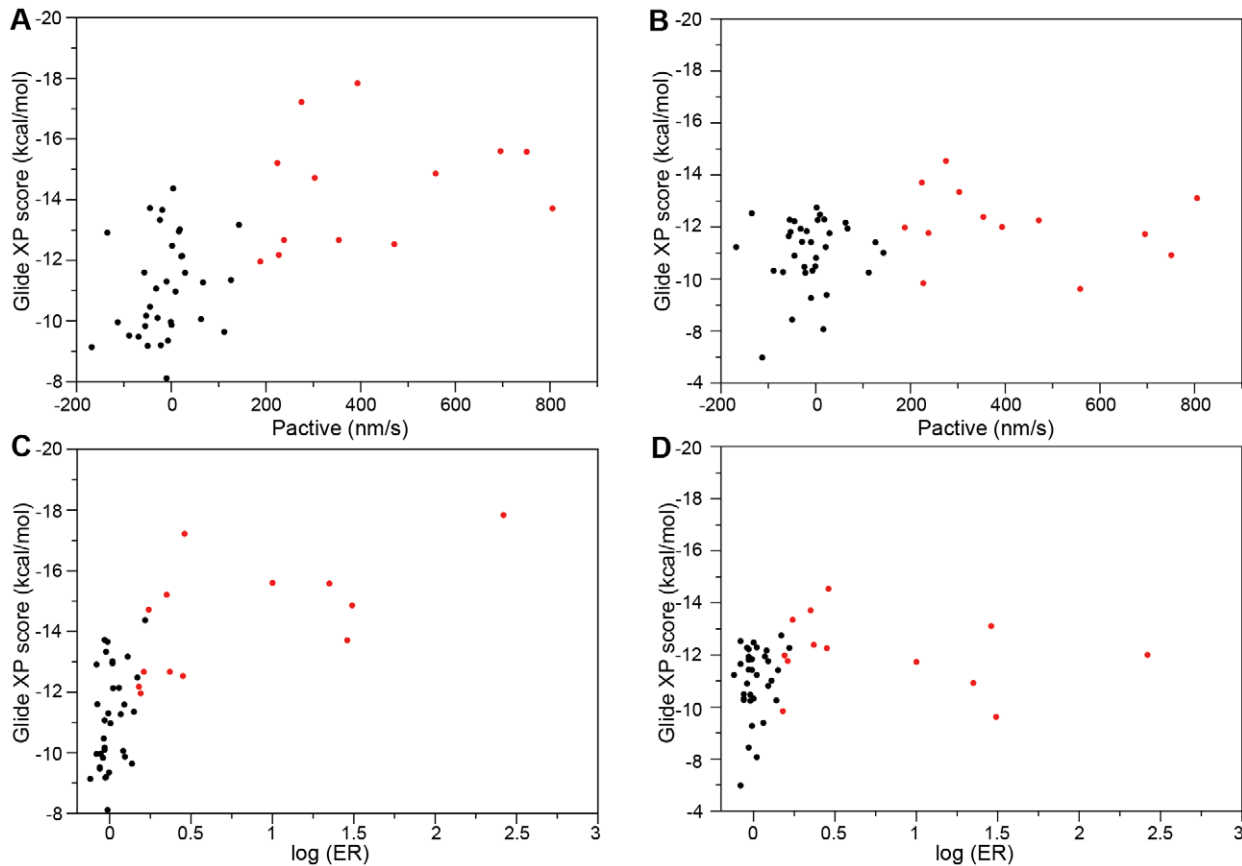
It is currently challenging to measure binding affinities to P-gp, and the results of standard assays are generally interpreted qualitatively (i.e., is it a substrate or not), as we have done here. However, the ratio or difference of the rates of permeability in the two directions (basal-to-apical,  $P_{BA}$ , and apical-to-basal,  $P_{AB}$ ) provides a quantitative measure of how ‘strong’ a substrate a given compound is. Although there is no reason to expect the computed docking scores to necessarily correlate well with these metrics, there is, in fact, a reasonable correlation with both the difference in the rates ( $P_{active} = P_{BA} - P_{AB}$ ) and the more commonly used log of the efflux ratio (Figure 5A). (Results using the MM-GB/SA scoring function are qualitatively similar, and again the results using this scoring function are better when all molecules are treated as neutral (Figures S5 and S6). One advantage of representing the data this way is that it avoids somewhat arbitrary criteria used for classifying the compounds as substrates and non-substrates. The plots illustrate the two different, if somewhat overlapping ranges of the binding scores for P-gp binders and nonbinders. When the same data was plotted versus rigid docking

scores, the two classes were undistinguishable (Figure 5B). When interpreting these plots, one must keep in mind that efflux ratio values are a result of a complex interplay between the binding affinities of the compounds and kinetic aspects of the P-gp efflux, and may be influenced by rate of passive membrane permeability, and we do not claim that it should be possible to quantitatively predict the efflux ratio based on docking calculations alone.

The P-gp binding site is highly hydrophobic and the ligand desolvation energy may have a significant effect on the ligand binding. To investigate this point, we computed free energy of transferring compounds from water to a low dielectric solvent (chloroform) using an approach described previously for predicting passive membrane permeability [52,53]. The plot of the P-gp binding scores vs. free energy of desolvation showed no correlation between the two (Figure S7), indicating that binding scores reflect specific interactions with P-gp and are not dominated solely by the polarity/hydrophobicity (quantified using the solvation free energy) of the compounds. Some of these specific interactions are illustrated in Figure S2 and include pi-stacking, cation-pi interactions and hydrogen bonding. These results also suggest that the physicochemical properties of the ligands that define their passive membrane permeability are different from the physicochemical determinants that define their interactions with P-gp. This finding in turn suggests that it might be possible to reduce a ligand’s P-gp binding using chemical modifications without dramatically reducing its membrane permeability. We note that the computations we perform here attempt to predict the (path-independent) thermodynamics of transferring a ligand from water to the P-gp binding site, and thus our computations do not provide any direct information about whether the ligand enters the binding site through the membrane, or directly from the cytoplasm.

Finally, we have performed a first ‘blind’ test of the method, using a series of peptidic cysteine protease inhibitors bearing natural and unnatural amino acid residues. These compounds were originally designed to test hypotheses concerning passive membrane permeability, and those results will be reported elsewhere. However, we also tested the compounds in a cell-monolayer assay (performed by WuXi AppTec), specifically using P-gp transfected MDCK cells. The results are summarized in Table 1 and Figure 6, where we again plot the predicted docking scores for binding to P-gp versus measures of the asymmetry of the permeability across the monolayer. The compound that showed the strongest evidence for P-gp mediated efflux, compound **6**, had





**Figure 5. Flexible (A, C) and rigid (B, D) docking binding scores (Glide XP) versus  $P_{\text{active}}$  and log of efflux ratio.** Compounds from Doan *et al.* dataset [11].  
doi:10.1371/journal.pcbi.1002083.g005

an efflux ratio of 9 with moderate passive membrane permeability. Encouragingly, this compound had the most favorable Glide XP score ( $-14.6$ ) among the series, comparable to many of the P-gp substrates in the benchmarking results discussed above. Similarly, the compound with the least favorable docking score ( $-9.9$ , compound **2**) had a much lower efflux ratio (1.6), and the compound with the lowest measured efflux ratio (1.2, compound **1**) had the second lowest docking score ( $-10.9$ ).

The results are not perfect. Compound **9** has relatively polarized efflux ( $ER = 5.8$ ), but has a docking score that would classify it only as a ‘possible’ binder ( $-12.3$ ). However, we note that this compound has  $10\times$  lower passive membrane permeability than any other member of the series,  $P_{AB} = 5 \times 10^{-7}$  cm/s, making the determination of P-gp mediated efflux more uncertain. In general, the correlation between the docking scores and the experimental results is ‘noisier’ than in the benchmarking study using the Doan *et al.* data set. However, we note that the range of experimental efflux ratios, as well as the range of docking scores, is narrower in this series of peptidic compounds, which is expected due to the compounds being much more chemically similar. As such, it is gratifying to be able to confirm an ability to predict compounds more/less likely to be P-gp substrates, even in the more challenging case of a more chemically homogeneous series, albeit not with ‘quantitative’ accuracy.

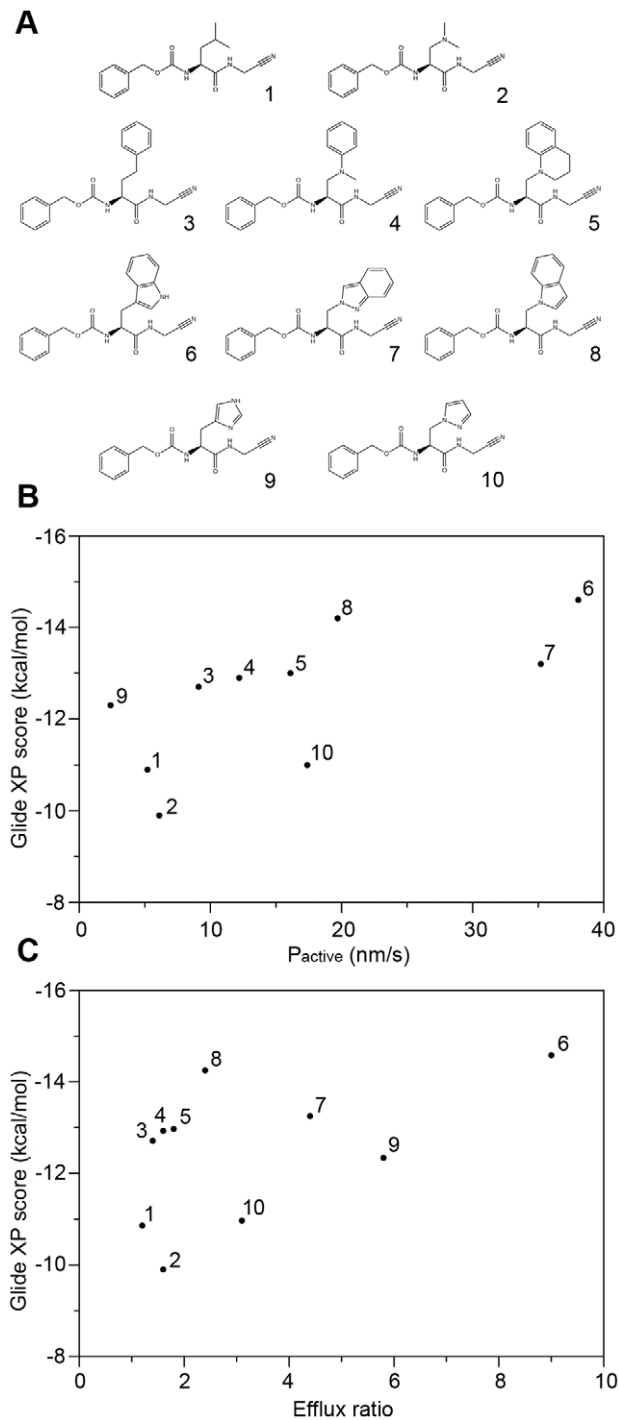
In this regard, we again emphasize that, even with perfect ability to predict binding affinities to P-gp (which we certainly do not claim), there is no reason to expect any simple relationship between the docking scores and experimental measures of P-gp

mediated efflux. Thus the docking results, as well as the experimental results, should be interpreted in qualitative terms. As general guidelines based on the work presented here, we view a Glide XP docking score of approximately  $-14$  or lower, using the flexible docking protocol, to be a predictor of P-gp interaction, while scores of  $-12$  or greater predict non-interaction. Intermediate scores of approximately  $-12$  to  $-14$  are less conclusive, but many compounds in this range show evidence of being relatively ‘weak’ substrates.

## Discussion

In summary, we have developed an *in silico* method suitable for predicting compounds that are more likely or less likely to interact with P-gp. We tested the ability of this method to differentiate between binders and nonbinders of P-gp using consistently measured experimental data from P-gp efflux and calcein-inhibition assays. Treating the P-gp binding cavity as flexible is critical for obtaining good results. We suspect that this observation reflects intrinsic flexibility of the binding site, but may also be related to the relatively low resolution of the crystal structures; most of the side chains orientations are not well defined by the electron density. Encouragingly, a first ‘blind’ test of the flexible-receptor approach on a series of peptidic protease inhibitors provided additional evidence for the predictive ability of the method.

Overall, our results suggest that specificity in P-gp is better understood in terms of physicochemical properties of the ligands (and the binding site), rather than being defined by specific



**Figure 6. Series of peptidic cysteine protease inhibitors.** Compound structures (A). Plot of binding scores (Glide XP) vs.  $P_{active}$  (B) and experimentally measured efflux ratio values (C). doi:10.1371/journal.pcbi.1002083.g006

sub-sites. We also suggest that many P-gp substrates bind deeper in the cavity than the cyclic peptide in the crystal structure. Finally, we also make testable predictions concerning metabolites that may be P-gp substrates.

We have not explored whether we can distinguish substrates and inhibitors of P-gp. Experimentally distinguishing these two classes is not simple, and available evidence suggests that substrates

can competitively inhibit efflux of other compounds to varying extents [19]. From the standpoint of the computations, it is clear that both substrates and inhibitors must bind to P-gp, presumably in the ‘inward’ configuration represented in the crystal structures. We speculate that the mode of binding (i.e., where in the binding site the compounds bind) might relate to these complicated considerations. However, the biophysical basis of coupling between ligand binding and ATP hydrolysis (enhancement and inhibition) remain poorly characterized, making further progress difficult at this time.

## Supporting Information

**Figure S1** ROC-type curves from rigid docking of metabolites/P-gp binders data set. Glide SP and XP with inner docking box coordinates (19.1, 52.3, -0.3) centered on the original ligand are in black and green, respectively. Glide SP and XP results with the inner box coordinates (19.0, 46.0, -6.0), located deeper in the cavity, are in red and blue, respectively. (TIF)

**Figure S2** Top-ranked poses from flexible docking. Both Glide XP and MM-GB/SA identified the similar poses as top ranked for digoxin (A) and loperamide (B) and the same pose for etoposide (C). (only Glide XP pose is shown for clarity). The different top-scored poses identified by the two methods for doxorubicin are shown in D. QZ59 is shown for reference in light green. For etoposide, two residues, Tyr303 and Ser725, forming hydrogen-bonding interactions with the ligand are shown. Residues Phe71, Phe953, and Phe974, positioned for cation-pi interaction, are shown for loperamide. (TIF)

**Figure S3** ROC-type curves (MM-GB/SA scoring) for metabolites/P-gp binders set (A) and Doan *et al.* dataset (B). Default treatment of protonation states (see Methods). (TIF)

**Figure S4** ROC-type curves (MM-GB/SA scoring) for flexible docking of Doan *et al.* dataset using the default treatment of protonation states (labeled as “pH = 7”) and when treating all compounds as neutral. (TIF)

**Figure S5** Flexible docking binding scores (MM-GB/SA) plotted versus  $P_{active}$  (A) and efflux ratio (B) for the Doan *et al.* dataset (pH 7). (TIF)

**Figure S6** Flexible docking binding scores (MM-GB/SA) plotted versus  $P_{active}$  (A) and efflux ratio (B) for the Doan *et al.* dataset (neutral). (TIF)

**Figure S7** Free energy of desolvation versus MM-GB/SA binding scores. (TIF)

**Table S1** Docking scores of P-gp binders selected from Hennessy *et al.*, 2007, Table 1. (DOCX)

**Table S2** Mouse P-gp binding cavity residues that were optimized with Prime in the first IFD round. (DOCX)

**Table S3** Docking scores of metabolites selected from KEGG database. (DOCX)

**Table S4** Docking scores for compounds from the Doan *et al.* dataset (at pH 7). (DOCX)

**Table S5** Docking scores for compounds (treated as neutral) from the Doan *et al.* dataset. (DOCX)

**Table S6** Binding cavity residues within 5 Å of QZ59 ligand in the original crystal structure and in the top flexible receptor docking pose. (DOCX)

**Text S1** Supplementary methods. (RTF)

## References

- Thiebaut F, Tsuruo T, Hamada H, Gottesman MM, Pastan I, et al. (1987) Cellular localization of the multidrug-resistance gene product P-glycoprotein in normal human tissues. *Proc Natl Acad Sci U S A* 84: 7735–7738.
- Szakacs G, Varadi A, Ozvegy-Laczka C, Sarkadi B (2008) The role of ABC transporters in drug absorption, distribution, metabolism, excretion and toxicity (ADME-Tox). *Drug Discov Today* 13: 379–393.
- Schinkel AH, Wagenaar E, Mol CA, van Deemter L (1996) P-glycoprotein in the blood-brain barrier of mice influences the brain penetration and pharmacological activity of many drugs. *J Clin Invest* 97: 2517–2524.
- Ambudkar SV, Dey S, Hrycyna CA, Ramachandra M, Pastan I, et al. (1999) Biochemical, cellular, and pharmacological aspects of the multidrug transporter. *Annu Rev Pharmacol Toxicol* 39: 361–398.
- Fletcher JL, Haber M, Henderson MJ, Norris MD (2010) ABC transporters in cancer: more than just drug efflux pumps. *Nat Rev Cancer* 10: 147–156.
- Ayrton A, Morgan P (2008) Role of transport proteins in drug discovery and development: a pharmaceutical perspective. *Xenobiotica* 38: 676–708.
- Hochman JH, Yamazaki M, Ohe T, Lin JH (2002) Evaluation of drug interactions with P-glycoprotein in drug discovery: in vitro assessment of the potential for drug-drug interactions with P-glycoprotein. *Curr Drug Metab* 3: 257–273.
- Polli JW, Wring SA, Humphreys JE, Huang L, Morgan JB, et al. (2001) Rational use of in vitro P-glycoprotein assays in drug discovery. *J Pharmacol Exp Ther* 299: 620–628.
- Feng B, Mills JB, Davidson RE, Mireles RJ, Janiszewski JS, et al. (2008) In vitro P-glycoprotein assays to predict the in vivo interactions of P-glycoprotein with drugs in the central nervous system. *Drug Metab Dispos* 36: 268–275.
- Li WX, Li L, Eksterowicz J, Ling XB, Cardozo M (2007) Significance analysis and multiple pharmacophore models for differentiating P-glycoprotein substrates. *J Chem Inf Model* 47: 2429–2438.
- Mahar Doan KM, Humphreys JE, Webster LO, Wring SA, Shampine IJ, et al. (2002) Passive permeability and P-glycoprotein-mediated efflux differentiate central nervous system (CNS) and non-CNS marketed drugs. *J Pharmacol Exp Ther* 303: 1029–1037.
- Wang Q, Rager JD, Weinstein K, Kardos PS, Dobson GL, et al. (2005) Evaluation of the MDR-MDCK cell line as a permeability screen for the blood-brain barrier. *Int J Pharm* 288: 349–359.
- Gombar VK, Polli JW, Humphreys JE, Wring SA, Serabjit-Singh CS (2004) Predicting P-glycoprotein substrates by a quantitative structure-activity relationship model. *J Pharm Sci* 93: 957–968.
- Crivori P, Reinach B, Pezzetta D, Poggesi I (2006) Computational models for identifying potential P-glycoprotein substrates and inhibitors. *Mol Pharm* 3: 33–44.
- Summerfield SG, Read K, Begley DJ, Obradovic T, Hidalgo IJ, et al. (2007) Central nervous system drug disposition: the relationship between in situ brain permeability and brain free fraction. *J Pharmacol Exp Ther* 322: 205–213.
- Seelig A (1998) A general pattern for substrate recognition by P-glycoprotein. *Eur J Biochem* 251: 252–261.
- Ueda K, Cardarelli C, Gottesman MM, Pastan I (1987) Expression of a full-length cDNA for the human “MDR1” gene confers resistance to colchicine, doxorubicin, and vinblastine. *Proc Natl Acad Sci U S A* 84: 3004–3008.
- Loo TW, Clarke DM (1994) Mutations to amino acids located in predicted transmembrane segment 6 (TM6) modulate the activity and substrate specificity of human P-glycoprotein. *Biochem* 33: 14049–14057.
- Shapiro AB, Ling V (1997) Positively cooperative sites for drug transport by P-glycoprotein with distinct drug specificities. *Eur J Biochem* 250: 130–137.
- Garrigues A, Loiseau N, Delaforge M, Ferte J, Garrigues M, et al. (2002) Characterization of two pharmacophores on the multidrug transporter P-glycoprotein. *Mol Pharmacol* 62: 1288–1298.
- Penzotti JE, Lamb ML, Evensen E, Grootenhuis PD (2002) A computational ensemble pharmacophore model for identifying substrates of P-glycoprotein. *J Med Chem* 45: 1737–1740.
- de Cerqueira Lima P, Golbraikh A, Oloff S, Xiao Y, Tropsha A (2006) Combinatorial QSAR modeling of P-glycoprotein substrates. *J Chem Inf Model* 46: 1245–1254.
- Huang J, Ma G, Muhammad I, Cheng Y (2007) Identifying P-glycoprotein substrates using a support vector machine optimized by a particle swarm. *J Chem Inf Model* 47: 1638–1647.
- Cabrera MA, Gonzalez I, Fernandez C, Navarro C, Bermejo M (2006) A topological substructural approach for the prediction of P-glycoprotein substrates. *J Pharm Sci* 95: 589–606.
- Dey S, Ramachandra M, Pastan I, Gottesman MM, Ambudkar SV (1997) Evidence for two nonidentical drug-interaction sites in the human P-glycoprotein. *Proc Natl Acad Sci U S A* 94: 10594–10599.
- Loo TW, Bartlett MC, Clarke DM (2003) Methanethiosulfonate derivatives of rhodamine and verapamil activate human P-glycoprotein at different sites. *J Biol Chem* 278: 50136–50141.
- Martin C, Berridge G, Higgins CF, Mistry P, Charlton P, et al. (2000) Communication between multiple drug binding sites on P-glycoprotein. *Mol Pharmacol* 58: 624–632.
- Ayesh S, Shao YM, Stein WD (1996) Co-operative, competitive and non-competitive interactions between modulators of P-glycoprotein. *Biochim Biophys Acta* 1316: 8–18.
- Aller SG, Yu J, Ward A, Weng Y, Chittaboina S, et al. (2009) Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science* 323: 1718–1722.
- Loo TW, Bartlett MC, Clarke DM (2003) Simultaneous binding of two different drugs in the binding pocket of the human multidrug resistance P-glycoprotein. *J Biol Chem* 278: 39706–39710.
- Sherman W, Day T, Jacobson MP, Friesner RA, Farid R (2006) Novel procedure for modeling ligand/receptor induced fit effects. *J Med Chem* 49: 534–553.
- Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, et al. (2006) Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. *J Med Chem* 49: 6177–6196.
- Ghosh A, Rapp CS, Friesner RA (1998) Generalized born model based on a surface integral formulation. *J Phys Chem B* 102: 10983–10990.
- Huang N, Kalyanaraman C, Irwin JJ, Jacobson MP (2006) Physics-based scoring of protein-ligand complexes: enrichment of known inhibitors in large-scale virtual screening. *J Chem Inf Model* 46: 243–253.
- Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–3402.
- Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, et al. (2004) Glide: A new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *J Med Chem* 47: 1739–1749.
- Kaminski GA, Friesner RA, Tirado-Rives J, Jorgensen WL (2001) Evaluation and reparametrization of the OPLS-AA force field for proteins via comparison with accurate quantum chemical calculations on peptides. *J Phys Chem B* 105: 6474–6487.
- Jorgensen WL, Tiradorives J (1988) The Opls Potential Functions for Proteins - Energy Minimizations for Crystals of Cyclic-Peptides and Crambin. *J Am Chem Soc* 110: 1657–1666.
- Jacobson MP, Friesner RA, Xiang Z, Honig B (2002) On the role of the crystal environment in determining protein side-chain conformations. *J Mol Biol* 320: 597–608.
- Wishart DS, Knox C, Guo AC, Cheng D, Shrivastava S, et al. (2008) DrugBank: a knowledgebase for drugs, drug actions and drug targets. *Nucleic Acids Res* 36: D901–D906.
- Hennessy M, Spiers JP (2007) A primer on the mechanics of P-glycoprotein the multidrug transporter. *Pharmacol Res* 55: 1–15.
- Kanehisa M, Goto S, Furumichi M, Tanabe M, Hirakawa M (2010) KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Res* 38: D355–360.
- Palmer JT, Bryant C, Wang DX, Davis DE, Setti EL, et al. (2005) Design and synthesis of tri-ring P3 benzamide-containing aminonitriles as potent, selective, orally effective inhibitors of cathepsin K. *J Med Chem* 48: 7520–7534.
- Pansare SV, Huyer G, Arnold LD, Vederas JC (1998) Synthesis of N-protected alpha-Amino acids from N-(benzyloxycarbonyl)-L-serine via its beta-lactone:

## Acknowledgments

We thank Dr. Siegfried Leung for helpful conversations.

## Author Contributions

Conceived and designed the experiments: ED CB ARR MPJ. Performed the experiments: ED CB. Analyzed the data: ED CB ARR MPJ. Contributed reagents/materials/analysis tools: ARR MPJ. Wrote the paper: ED CB ARR MPJ.

- N-(benzyloxycarbonyl)-beta-(pyrazol-1-yl)-L-alanine. *Organic Syntheses* Vol. 9: 58.
45. Arnold LD, Kalantar TH, Vederas JC (1985) Conversion of Serine to Stereochemically Pure Beta-Substituted Alpha-Amino-Acids Via Beta-Lactones. *J Am Chem Soc* 107: 7105–7109.
  46. Loo TW, Clarke DM (2002) Location of the rhodamine-binding site in the human multidrug resistance P-glycoprotein. *J Biol Chem* 277: 44332–44338.
  47. Mitchell AM, Tom M, Mortimer RH (2005) Thyroid hormone export from cells: contribution of P-glycoprotein. *J Endocrinol* 185: 93–98.
  48. Fan J, Liu S, Du Y, Morrison J, Shipman R, et al. (2009) Up-regulation of transporters and enzymes by the vitamin D receptor ligands, 1alpha,25-dihydroxyvitamin D3 and vitamin D analogs, in the Caco-2 cell monolayer. *J Pharmacol Exp Ther* 330: 389–402.
  49. Yang CP, DePinho SG, Greenberger LM, Arceci RJ, Horwitz SB (1989) Progesterone interacts with P-glycoprotein in multidrug-resistant cells and in the endometrium of gravid uterus. *J Biol Chem* 264: 782–788.
  50. Wang E, Casciano CN, Clement RP, Johnson WW (2000) Cholesterol interaction with the daunorubicin binding site of P-glycoprotein. *Biochem Biophys Res Commun* 276: 909–916.
  51. Lee JS, Paull K, Alvarez M, Hose C, Monks A, et al. (1994) Rhodamine efflux patterns predict P-glycoprotein substrates in the National Cancer Institute drug screen. *Mol Pharmacol* 46: 627–638.
  52. Rezai T, Bock JE, Zhou MV, Kalyanaraman C, Lokey RS, et al. (2006) Conformational flexibility, internal hydrogen bonding, and passive membrane permeability: successful in silico prediction of the relative permeabilities of cyclic peptides. *J Am Chem Soc* 128: 14073–14080.
  53. Kalyanaraman C, Jacobson MP (2007) An atomistic model of passive membrane permeability: application to a series of FDA approved drugs. *J Comput Aided Mol Des* 21: 675–679.