New insights in the in vitro characterisation and molecular modelling of the P-glycoprotein inhibitory promiscuity

Giovanni Bocci, Amélie Moreau, Philippe Vayer, Claire Denizot, Olivier Fardel, Yannick Parmentier

New insights in the *in vitro* characterisation and molecular modelling of the P-glycoprotein inhibitory promiscuity

Giovanni Bocci,[a] Amélie Moreau,[b] Philippe Vayer,[b] Claire Denizot,[b] Olivier Fardel,[c] Yannick Parmentier [b]

a) *Laboratory of Chemometrics, Department of Chemistry, Biology and Biotechnology, University of Perugia, Via Elce di Sotto 8, 06123 Perugia, Italy*

b) *Technologie Servier, 25-27 rue Eugène Vignat, BP 11749, 45007 Orléans cedex 1, France*

c) *Univ Rennes, Inserm, EHESP, Irset (Institut de recherche en santé, environnement et travail) - UMR_S 1085, 2 Avenue du Pr Léon Bernard, F-35043 Rennes, France*

ABSTRACT

The presence of several binding sites for both substrates and inhibitors is yet a poorly explored thematic concerning the assessment of the drug-drug interactions risk due to interactions of multiple drugs with the human transport protein P-glycoprotein (P-gp or MDR1, gene ABCB1). In this study we measured the inhibitory behaviour of a set of known drugs towards P-gp by using three different probe substrates (digoxin, Hoechst 33342 and rhodamine 123). A structure-based model was built to unravel the different substrates binding sites and to rationalize the cases where drugs were not inhibiting all the substrates. A separate set of experiments was used to validate the model and confirmed its suitability to either detect the substrate-dependent P-gp inhibition and to anticipate proper substrates for *in vitro* experiments case by case. The modelling strategy described can be applied for either design safer drugs (P-gp as antitarget) or to target specific sub-site inhibitors toward other drugs (P-gp as target).

INTRODUCTION

The assessment of potential drug-drug interactions (DDIs) is a crucial step in the development of any new molecular entity (NME) in current drug discovery (Rodrigues, 2008). Clinically significant DDIs are mainly mediated by pharmacokinetic (PK) mechanisms with alterations in the adsorption, distribution, metabolism and excretion (ADME)

profile of a drug caused by a co-administered drug. DDIs risk is monitored at the very beginning of the NMEs development process because it could reduce both the efficacy and the safety of drugs. In the most unfortunate cases, emerging of adverse drug reactions, due to DDIs, ends up with the NME failure or its withdrawal from the market (Onakpoya et al., 2016; Schuster et al., 2005; Zhang et al., 2009). Thus, considering the cost for a NME just to come up with its failure, early detection of DDIs is a major concern.

The safety impairment of a drug often involves the inhibition of several anti-target proteins such as metabolism enzymes and transporters (Varma et al., 2015). Among these, two of the most studied are the cytochrome P450 3A4 (CYP3A4) and the P-glycoprotein (P-gp). CYP3A4 is the main enzyme involved in the metabolism of drugs and presents several sub-sites for both metabolism and inhibition (Kapelyukh et al., 2008; Kenworthy et al., 1999; Obach et al., 2006). P-gp, also known as ATP-binding cassette sub-family B member 1 (ABCB1) and encoded by the multidrug resistance gene 1 (MDR1), is a membrane efflux transporter belonging to the ATP binding cassette (ABC) transporter subfamily, which is notably expressed at the intestine, liver, kidney and blood-brain barrier level. P-gp drives the elimination of several structurally unrelated xenobiotics from cells through an ATP dependent mechanism and its inhibition is directly linked to DDIs (Lin, 2003). The importance of ABC proteins such as P-gp in regulating drug adsorption, distribution, metabolism, excretion and toxicity (ADMET) has been extensively reviewed (Giacomini and Huang, 2013; Szakacs et al., 2008), as well as the efforts done to predict P-gp-ligand interactions (Chen et al., 2012; Montanari and Ecker, 2015). While P-gp role and function is well known, its behaviour is yet only little understood. P-gp promiscuity in recognizing molecules is due to its large and flexible transport site where its amino acids define a broad and dynamic range of weak interactions. In particular, the high volume of the pocket facilitates the binding of different ligands in different regions (sub-pockets) which may differ for both spatial location and network of interactions. (Martin et al., 2000) An exact mapping of interacting regions is not available because the human P-gp crystallographic structure has not been solved yet. Consequently, elucidating structure-activity relationships and predicting ligands interaction with P-gp are not trivial tasks.

Given the need to anticipate as soon as possible the DDIs risk, *in vitro* P-gp inhibition experiments are routinely carried out for new pharmacologically interesting drugs. However, these tests are extremely sensitive to the experimental conditions (e.g. the nature of test system, consisting in cell line or vesicles, the nature of the probe substrate used, etc.) (https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM581965.pdf; Bentz et al., 2013; Rautio et al., 2006; Zolnerciks et al., 2011). Thus, close experimental protocols may lead to inconsistent results. The importance of the probe used was emphasised by the observation of different inhibitory behaviour emerging from *in vitro* P-gp inhibition experiments when changing the probe substrate used. In a recent work, an alternative assay for the P-gp IC_{50} determination has been proposed (Jouan et al., 2016). The authors evaluated rhodamine 123 (Rh123) to

replace the more problematic digoxin and compared the results obtained with both the probes. For some molecules, while the inhibition was detectable with digoxin, no inhibition was observed with Rh123. These molecules were presumed to not interact with the Rh123 binding site (R-site) of P-gp.

In order to go deeper in understanding this P-gp promiscuity, we tested a third probe substrate: Hoechst 33342 (H33342). This choice was made for two main reasons. First, because it has quite different molecular features such as lipophilicity, size and aromaticity (i.e. different experimental conditions are required for the assay) compared with digoxin and Rh123; hence, it could be used as an alternative probe substrate in P-gp inhibition experiments. Second, because it has been reported to bind on a different site (H-site) (Martin et al., 2000; Shapiro and Ling, 1997). Thus, we suspected it could give rise to a third different inhibitory behaviour respect to Rh123 and digoxin.

Therefore, the safety of these "substrate-dependent" inhibitors may not be fully unravelled even at late stages of drug development. Literature is quite ambiguous when dealing with P-gp substrate-dependent inhibition. Some efforts have been done to get new insights, in particular by agencies such as FDA not to propose standard protocols, but encourage to calibrate our own model (https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM581965.pdf). To face this issue, one can either increase the number of tests including all the probes (exhaustive approach) or use a specific probe, selected from strong evidences and knowledge, to answer a specific question (adaptive knowledge based approach). In this context, molecular modelling tools can be applied and integrated with well-designed experimental data. This could also help to complete missing or ambiguous literature data.

Starting from our experimental data obtained with digoxin, Rh123 and H33342 as probe substrates, we developed a new structure-based modelling strategy to predict the substrate-dependent inhibition of P-gp. We used the knowledge concerning structural features of human P-gp to develop a structure-based model defined by three transport sub-sites within the main binding pocket. Successively, we rationalized the P-gp ligands behaviour by assuming a competitive inhibition mechanism within the main pocket itself. The pharmacophoric methodology fingerprints for ligands and proteins (FLAP) (Baroni et al., 2007) was employed to conduct all the modelling strategy. The model has been validated with a separate set of experiments and can be further used in routine to improve experimental efficiency by anticipating P-gp inhibition and DDIs depending on the substrate used. We can assert that our study is the first computational approach which, in synergy with experimental testing, gives an exhaustive picture of substrate-dependent P-gp inhibition issue.

MATERIALS AND METHODS

Dataset. The starting set comprises 30 compounds and 69 experimental inhibition records (30 for digoxin, 24 for H33342 and 15 for Rh123). The model validation set comprises 14 compounds of which eight were already tested with two substrates (starting set) and six for which no record was available for any substrate yet. The complete set of molecules, with related SMILES, can be found in Supporting Information 02 (as an excel file). For both starting set and validation set, experimental data about P-gp inhibition with respect to transport of digoxin, H33342 and Rh123 were retrieved from the scientific literature or were generated in the present study using appropriate functional assays (See below) when required, i.e., when no experimental data has been previously reported.

P-gp inhibition assays. Effects of compounds on P-gp-mediated transport was assessed using accumulation (digoxin or Rh123) or efflux (H33342) assays, in P-gp overexpressing MCF7R cells (Jouan et al., 2016). These transport assays were performed in a well-defined transport medium consisting of 136 mM NaCl, 5.3 mM KCl, 1.1 mM KH2PO4, 0.8 mM MgSO4, 1.8 mM CaCl2, 11 mM D-glucose and 10 mM Hepes and adjusted to pH 7.4, as previously described (Chedik et al., 2017). For accumulation assays, MCF7R cells were incubated with 5.25 µM rhodamine 123 (Sigma−Aldrich, Saint-Quentin Fallavier, France) or 12.5 nM [^{3}H]-digoxin (specific activity = 39.8 Ci/mmol) (PerkinElmer, Villebon sur Yvette, France) for 30 min at 37°C, in the absence (control) or presence of chemicals to be tested (a 100 µM concentration was used for each compound, excepted for terfenadine for which 10 µM was retained for solubility consideration) or of 100 µM verapamil, used here as a reference P-gp inhibitor. After washing with cold phosphate-buffered saline (PBS), intracellular accumulations of Rh123 and digoxin were measured by spectrofluorimetry and scintillation counting, respectively. Data were finally expressed as percentages of P-gp substrate accumulation found in control cells. For efflux assays, MCF7R cells were first loaded with 16.2 µM H33342 (Sigma-Aldrich) for 30 min at 37°C. After washing with cold PBS, MCF7R cells were re-incubated in H33342-free medium, in the absence (control) or presence of compounds to be tested or of verapamil for 90 min at 37°C. Intracellular retention of H33342 was finally determined by spectrofluorimetry (excitation and emission wavelengths were 355 and 460 nm, respectively). Data were expressed as percentages of initial loading of the dye. Accumulation and efflux data were statistically analysed through analysis of variance (ANOVA) followed by the Dunnett's post hoc test. The level of significance was p<0.05. Compounds which significantly modulate accumulation (digoxin, Rh123) or retention (H33342) of a P-gp substrate were considered as inhibitors of P-gp-mediated efflux of the substrate.

Modelling. The structure-based modelling strategy applied in this work is based on the pharmacophoric method fingerprints for ligands and proteins (FLAP) (http://www.moldiscovery.com/; Baroni et al., 2007). FLAP is a computational tool, which has been applied in virtual screening (Brincat et al., 2011; Cross et al., 2010; Goracci et al.,

2016), but that can also be used for developing pharmacophore-based models. In this study, FLAP was used to prepare ligands structures, to define the target pocket, to perform docking and to classify ligands after the docking. Fig. 1 shows the flowchart of these main steps. The docking algorithm implemented in FLAP allows to explore the potential binding of ligands with a target protein. This is done by an accurate comparison between molecular interaction fields (MIFs) (Goodford, 1985) of ligands with the MIFs of a pocket previously identified.

Chemical structures, for both P-gp ligands and substrates, were retrieved as SMILES from PubChem (https://pubchem.ncbi.nlm.nih.gov/). The FLAPdb module was used to store molecules into a database by generating their canonical 3D structures. MoKa (Milletti et al., 2007) was used to generate tautomeric and charged forms whose abundance was of at least 20% at pH 7.4. If the stereochemistry was not explicit in the SMILES (e.g. verapamil, disopyramide, etc.), all the possible stereoisomers were generated. Thus, for the same molecule, different structural forms could be generated. In those cases, all of them were considered, because no information was available to distinguish which form or stereoisomer was more likely to interact with P-gp. Hence, they are assumed to have the same chance to do it. On the other hand, not considering all the possible structural forms is a hazard that could significantly decrease the accuracy of pharmacophoric models.

The mouse P-gp crystallographic 3D structure was downloaded from Protein Data Bank (PDB code: 4M1M) (Li et al., 2014) and imported into FLAP. The FLAPsite module was used to compute the main binding pocket by setting the "search for pockets" parameter "sensitivity" equal to 9; all the other parameters were kept by default.

Successively, the FLAPdock module was used to perform the docking. Except for "pocket radius" (which was set to 1.5) and "scale van der Walls external" (which was set to 0.80), the docking parameters were kept by default. In order to identify their binding sub-pockets, probe substrates were docked first. Next, P-gp ligands were docked. Only the best poses were selected (those having the highest S-Score approximated to the first decimal place) for the final classification. If the same molecule was present in more structural forms, best poses were selected for each form. This applies to both substrates and ligands. Concerning substrates, three poses were selected for digoxin, only one for H33342 and four for Rh123. Their location allowed to identify the relative binding sites: D-site, H-site and R-site.

In order to retrieve the classification, the FLAPvs module was used to perform a static cross screening (without alignment) between the selected poses of the three substrates and the selected poses of ligands. This operation was designed to translate the relative substrate-ligand binding position (generated by the docking) into a number that depicts the competition between ligands and substrates for specific sub-pockets. This number equals to the MIFs similarity score Glob-Sum (GS). Hence, for the same ligand, three GS scores are computed: one for digoxin (DGS), one for H33342 (HGS) and one for Rh123 (RGS). Finally, a global GS threshold was set to distinguish substrate-specific inhibitors and non-inhibitors. Ligands having a GS equal to zero are predicted non-inhibitors. Whereas, ligands having a GS major of

5

zero are predicted inhibitors. In most of the cases, more than one ligand pose was selected for the final classification. If that was happening, only the highest GS among the poses was used. It is true that, since all selected poses have the same importance (i.e. the same probability to exist), it would be more correct to average all GS scores. However, in that way we obtained worst results when reproducing the experimental data. The same principle was also used to condense the prediction over probe substrates poses; and it was applied to condense classifications of different structural forms (stereoisomers, tautomers and protomers) for the same ligand. The classification principle just explained was applied to all the P-gp ligands studied in this work.

The pocket dimensions define the applicability domain of our model. Consequently, all the molecules that could fit in the pocket can be predicted. On the other hand, those molecules that are too big to fit in cannot be predicted.

RESULTS

Determination of P-gp-mediated transport inhibition. Inhibitory potential of various compounds towards P-gp mediated transport of digoxin, H33342 and/or Rh123, were collected from the scientific literature when data are available (Supporting information 02). For compounds for which data about P-gp inhibition are lacking, their effects on P-gp-mediated transport of the P-gp substrates were experimentally studied in P-gp-overexpressing human MCF7R cells. As indicated in Fig. 2, prazosin, reserpine and loperamide, like the reference P-gp inhibitor verapamil, significantly enhanced digoxin accumulation in MCF7R cells, thus providing evidence that they inhibit P-gp-mediated transport of digoxin. By contrast, terfenadine did not alter cellular level of digoxin (Fig. 2), thus indicating it has no inhibitory effect towards P-gp-related digoxin efflux. With respect to H33342, carvedilol, mibefradil, loperamide, reserpine, prazosin and terfenadine, unlike itroconazole, enhanced its retention in MCF7R cells (Fig. 2), indicating that they block P-gp-mediated efflux of H33342. For Rh123 accumulation, cisapride zafirkulast, amitriptyline, loperamide, reserpine and terfenadine significantly increased it (Fig. 2), indicating that these compounds inhibit P-gp-mediated transport of Rh123. By contrast, disopyramide, midazolam and prazosin failed to significantly enhance Rh123 accumulation (Fig. 2), thus ruling out any inhibitory effect toward P-gp-related efflux of the dye.

Transport sub-sites detection. In order to identify the location of substrates binding sites, we computed the main binding pocket from the original 3D structure (PDB code: 4M1M) (Li et al., 2014). Successively, we docked the three probe substrates into the main pocket. We used a cut-off on the ligand-pocket similarity ranking score (S-Score) to define which and how many poses to consider. Hence, substrates binding sites were defined on their top-ranked (higher probability) binding poses (see Fig. 3A). Three top-ranked binding poses of digoxin define its binding site (D-site, Fig. 3B), one top-

ranked binding pose of H33342 defines alone its binding site (H-site, Fig. 3D) and four top-ranked binding poses of Rh123 define its binding site (R-site, Fig. 3C).

Docking of P-gp substrate-dependent inhibitors. From starting set of inhibition experiments, some compounds were found to inhibit the transport of some substrates but not of others. Such behaviour was detected for zafirlukast, disopyramide, lansoprazole, itraconazole and sertraline. These compounds were further inspected with docking.

P-gp experimental inhibition by zafirlukast, a compound known to interact with P-gp (Sun et al., 2012), was measured with digoxin and H33342. It was found to inhibit digoxin transport, whereas no substantial inhibition was seen for H33342 up to the highest concentration tested. No information concerning Rh123 inhibition was available at the time in the literature. In Fig. 4A, the top-ranked docking poses for zafirlukast are reported. As shown, since its binding takes place in the digoxin binding region and far from the H33342 binding region, it is predicted to inhibit digoxin and to not inhibit H33342. Similarly, it is also predicted to inhibit Rh123 efflux.

Lansoprazole, another P-gp inhibitor (Pauli-Magnus et al., 2001), was tested *in vitro* with a different combination of substrates: digoxin and Rh123; it was found to inhibit the P-gp mediated transport of the first one and not of the second one (Table 1). In Fig. 4B, the top-ranked poses for lansoprazole are reported. In accordance with experiments, it is predicted to inhibit both digoxin and H33342, whereas no inhibition should be expected with Rh123.

Itraconazole was tested with digoxin and Rh123 and, as for lansoprazole, it resulted to inhibit digoxin, but not Rh123 (Jouan et al., 2016) (Table 1). In Fig. 4C, the top-ranked poses for itraconazole are reported. In this case, there is only a partial agreement between experiments and docking: correctly, itraconazole is predicted to inhibit digoxin transport, but it is also wrongly predicted to inhibit Rh123 transport. Concerning disopyramide and sertraline, we can draw similar conclusions as for zafirlukast and lansoprazole (see Supporting Information 01).

Docking of P-gp substrate-independent inhibitors. From the experiments, several ligands considered in this study were P-gp inhibitors independently from the probe substrate used. In order to confirm their experimental behaviour, but also to better assess the model reliability of predictions, we docked these compounds as well. In Fig. 5, the examples of elacridar and quinidine, two well-known P-gp inhibitors (Rautio et al., 2006; Weiss et al., 2003), are reported. Although with much different orders of potency, both of them in vitro inhibited all the probe substrates (Table 1). The docking results are in agreement with the experiments. The complete list of compounds docked poses is provided in Supporting Information 01; whereas, the complete starting set of experiments and predictions for both substrate-dependent and substrate-independent inhibitors is reported in Table 1.

7

Model validation. To validate the three binding sites model hypothesis, we docked six additional compounds that were known to interact with P-gp: amitriptyline, loperamide, nifedipine, prazosin, reserpine and terfenadine (Doan et al., 2002; Polli et al., 2001). Successively, each of these was tested *in vitro* with the three substrates. Moreover, we also tested *in vitro* eight additional compounds for which the information concerning one substrate was missing (e.g., zafirlukast was tested with Rh123; lansoprazole was tested with H33342; etc.). The results are shown in Table 2. These are in accordance with 88% of model predictions.

Integration of different modelling strategies. It should be noted that our model is a pharmacophoric model, thus based on the modelling of specific interactions, that was not meant to predict globally P-gp inhibition/non-inhibition, but it could be used to refine predictions of other more general strategies. The binding of a molecule to P-gp could not be purely related to specific interactions: in fact, it could be also related to its physiochemical properties which determine if the molecule is reaching or not the binding site. Consequently, we wanted to test the possibility to use our model along with a different methodology to better predict P-gp inhibition. As proposed by Broccatelli and co-workers (Broccatelli et al., 2011), physiochemical and pharmacophoric based models could be used in a composite strategy for the study of P-gp inhibition. Their original model was based on a sequence of different "blocks", composed by the molecular description obtained with the software VolSurf+ and FLAP, as well as PLS and LDA as regression/classification methods. In a recent work, we reviewed the PLS part of the model and we derived a new QSAR pharmacokinetic model for P-gp (Bocci et al., 2017). Here we tried to substitute the LDA part with our model. To test the effect of such substitution we used the dataset proposed by Keogh et al. which is composed of 40 marketed compounds whose inhibitory power toward P-gp was tested *in vitro* with digoxin as substrate (Keogh and Kunta, 2006). Based on this strategy, molecules resulting uncertain from the QSAR model are passed to our structure-based model to retrieve predictions. The result of this operation is shown in Fig. 6. What happens is that our model succeeds in classifying most of the compounds that QSAR was not able to classify.

DISCUSSION

P-gp plays a major role in ADMET processes and its implication in determining DDIs still is a main concern. Consequently, researchers are encouraged to study more effective *in vitro* protocols and to develop more reliable *in silico* models. Owing to the lack of standard protocols for *in vitro* P-gp inhibition experiments, determining high intra/inter-laboratory variability of results still remains a problem (Hayeshi et al., 2008), and a lot of efforts have been done to optimize these assays (Polli et al., 2001; Szakacs et al., 2008). In particular, the choice of the proper probe substrate is a drastic parameter (Rautio et al., 2006). Indeed, not only their molecular features, such as permeability, may have a deep impact on the results, but also the binding of the test molecule and the substrate to different sites may produce misleading

results. Assuming only a competitive inhibition mechanism, a molecule binding to a different region respect to the probe substrate do not block the efflux of the latter, but interacts with the protein anyway. Hence, since several substrates binding sites are present for P-gp, using only one probe substrate it could be possible to fully reveal the potential inhibition only if it interacts strongly the same with all the different sub-sites. If a probe is not or less efficient for the inhibition of at least one sub-site, P-gp related DDI would be under estimated.

To illustrate this point, the experimental data reported in the present study (Fig. 2) indicate that terfenadine was able to inhibit P-gp-mediated transport of H33342 and Rh123, but not that of digoxin. In the same way, prazosin significantly blocked P-gp-related efflux of digoxin and H33342, but not that of Rh123. By contrast, other compounds such as reserpine and loperamide counteracted transport of the three P-gp substrates. Overall, the nature of the substrate is therefore crucial for assessing P-gp inhibition potential, as previously suggested (Rautio et al., 2006).

To understand at a molecular level the possible mechanisms beyond these observations, we developed a structure-based model based on the experimental inhibition data obtained. In terms of *in silico* modelling, the high flexibility and the large size of its active sites make the study of this protein not trivial at all. Furthermore, the absence of a human P-gp structure crystallized, even if it would not solve all the issues, is a big limitation in understanding the behaviour of this protein. Nevertheless, the use of the recent mouse P-gp crystal structure could provide reliable information (Li et al., 2014). Hence, we decided to use it and try to give an explanation to the substrate-dependent inhibition detected.

First, we identified three well-distinguished binding regions for digoxin, Rh123 and H33342 and we named them respectively D, R and H sites. In terms of respective location, our results are in accordance with some previous findings. R and H sites are located in opposite trans membrane alpha-helices domains (Lugo and Sharom, 2005; Pajeva et al., 2013; Shapiro and Ling, 1997). Furthermore, the large D-site partially overlaps with the smaller R-site as reported by Rautio et al. (Rautio et al., 2006) However, if we compare the R and H sites with those recently characterized by Ferreira and co-workers (Ferreira et al., 2013), substantial differences are emerging. In fact, our H and R sites are located much deeper in the main pocket and their respective positions are inverted from those described by Ferreira et al. These findings may suggest a wrong assignment for R and H sites in the previous study. Nevertheless, the authors postulated that these sites might interchange in one another during the efflux cycle and we could speculate that the crystallographic "snapshot" has registered this event. Furthermore, Li et al. highlighted the residues involved in what they call "drug translocation pathway" deep within the pocket which corroborates our finding (Li et al., 2014). Anyway, since none of the tests we did substantially changed the R and H sites locations, we decided to proceed without modifying our pocket assignment. The model validation proved that our decision was correct. Concerning the D-site, although its existence has been explored previously (Ledwitch et al., 2016), no molecular description has been provided so far. Our study provides a first

9

characterization of the digoxin binding site in terms of location respect to the other sites and affinities with several P-gp ligands.

In a second step the docking of P-gp inhibitors was performed. The whole binding pocket previously computed to dock them was used. As previously done for substrates, the same cut-off principle on the S-Score was used to select the best poses for each inhibitor. Successively, by inspecting their location respect to the three substrates binding sites, we expected to unravel their eventual substrate-dependent inhibition. Assuming our model was purely competitive, it did not consider other types of inhibition mechanism such as the non-competitive or acting outside of the active site. In addition, the binding affinity was not taken into account because we saw that none of the docking scores reproduced or had a trend similar to the experimental findings. Hence, the model predictions were based only on the respective location of substrates and inhibitors: if the ligand is located on the same region of the substrate, the inhibition takes place, whereas, if it is located away from the substrate, the inhibition does not occur. In this way, we were able to explain why zafirlukast, lansoprazole, sertraline and disopyramide were not blocking the efflux of all the probe substrates. From the experiments, zafirlukast and disopyramide were inhibiting digoxin efflux but not H33342 efflux. This suggested that the two compounds could compete with digoxin on the D-site, but not with H33342 on the H-site. The docking agreed with this hypothesis (see Fig. 4A). This happens when all the top-ranked binding poses, for both the compounds, are located on the D-site and not on the H-site. Concerning lansoprazole and sertraline, these compounds were inhibiting digoxin but not Rh123 (sertraline inhibited H33342 as well) suggesting a similar interaction mechanism to the previous two compounds. Hence, lansoprazole and sertraline competitively inhibit digoxin, whereas Rh123 inhibition does not occur because none of the compounds binds to the R-site (see Fig. 4B). Fig. 4C depicts the case of itraconazole, which is one of the few cases where the model disagrees with the experiments. Indeed, while the docking was able to reproduce the digoxin inhibition, it fails to reproduce the Rh123 non-inhibition by locating itraconazole poses also on the R-site. It is difficult to conclude if this was an isolate case or a more spread limitation of the model. Probably a different protein conformation is required to locate itraconazole correctly respect to all the substrates binding sites. The model may wrongly predict the binding site of big and quite rigid molecules (also having elongated shape), such as itraconazole.

To confirm the model robustness, we docked several compounds for which no substrate-dependent inhibition was observed. We expected that, for these compounds, the docking would be less selective; hence, it is likely to locate them on more substrates binding sites. In Fig. 5, the docking result for elacridar and quinidine are reported. These compounds differ of several orders of potency for inhibiting P-gp and this could be due to a different binding mechanism. Indeed, there is a substantial difference in terms of binding. We can see that elacridar poses are filling all the bottom of the active sites, impairing in one shot all the substrates binding sites. On the other hand, quinidine could bind with equal

probability to all sites. Hence, the results are the same, but substantial differences in the magnitude of inhibition could emerge.

These results are in agreement with the 4-sites model proposed by Martin et al. where three specific sites and one nonspecific site were postulated for P-gp ligands (Martin et al., 2000). Our docking model illustrates at the molecular level the proposal of Martin et al based on experimental inhibition patterns. For both elacridar (Fig. 5A) and nifedipine (Supporting Information 01 - Fig. S14) the assignment to the nonspecific site seems to be confirmed by our findings. This is the site where "pure" P-gp inhibitors (i.e. not transported by P-gp) are supposed to bind (Ferreira et al., 2013; Martin et al., 2000). However, since the aim of our study is not to differentiate transported inhibitors from non-transported inhibitors, we did not needed to define a separate fourth site for P-gp.

Recent site-directed mutagenesis experiments on P-gp showed that when residues of the primary drug-binding site are mutated, substrates and inhibitors bind to secondary sites on the transporter (Chufan et al., 2013). In addition, the transport of several substrates by multiple-residues mutant P-gp is also not significantly altered, showing that substrates bound at secondary sites are still transported. The effect of mutagenesis on P-gp binding was not studied here as our goal was not to characterise precisely the binding sub-sites in the P-gp structure. Our goal was to point first the existence at a molecular level of P-gp sub-sites according to Martin and co-workers and, second, the ability for ligand to be classified according to a molecular based classification of inhibitors.

Our model performance was verified in predicting the substrate-dependent inhibition, by using a different set of experiments. A first group of experiments was meant to fill in the blanks for all the probes. Thus, four experiments were targeted with H33342 (carvedilol, itraconazole, lansoprazole and mibefradil) and four with Rh123 (cisapride, disopyramide, midazolam and zafirlukast). A second group of experiments included six additional molecules (for which interaction with P-gp was already detected in the literature or was functionally *in vitro* measured in the present study) which were tested with all probes. These compounds were amitriptyline, loperamide, nifedipine, prazosin, reserpine and terfenadine. First, compounds were docked to predict their inhibitory behaviour and, successively, data from docking were confronted to experimental P-gp inhibition data.

The results are reported in Table 2. The model correctly anticipates most of the experimental results. There is a complete agreement between predictions and experiments for the probe H33342. The agreement for Rh123 is 9 over 10 correctly predicted and for digoxin is 4 over 6 correctly predicted. The model performances in prediction are reported in Table 3. As already mentioned, there can be several explanations on the origin of wrong predictions and these are related to some interaction aspects not considered by the model such as the non-competitive mechanism, the binding affinity and the protein flexibility. A future aim could be the incorporation of such information in the model through mechanism data and molecular dynamics studies.

The case of terfenadine is quite odd because it is the only compound, in our set, that inhibited H33342 and Rh123, but not digoxin. In the future, it could be quite interesting to further inspect its mechanism of interaction and enrich the model with other compounds bearing similar information.

Both our experimental and computational findings suggest that the type of inhibition (whether it is dependent or independent from the substrate used) would depend only on the number of transport sub-sites that can be blocked by the test compound. Consequently, in presence of a molecule binding to only one of them, the other sites may continue to work independently from each other.

We used our model to improve the predictability of the ligand-based P-gp inhibition model developed by Broccatelli et al. (Broccatelli et al., 2011) and recently reviewed by us (Bocci et al., 2017). Here we substitute the LDA part with our model resulting in a more efficient combined model. Both physiochemical and pharmacophoric (ligand-based or structure-based) models are proposed to predict P-gp inhibition. The first considers the holistic physiochemical molecular features, such as lipophilicity and molecular size, that are expected to be the main determinants of the interaction (or not) of a molecule with a protein. Differently, pharmacophoric models dig more into the reasons of interaction by inspecting in detail H-bonding interactions, hydrophobic interactions and size/shape fit between protein and ligand. Hence, pharmacophoric models should be a successive step to a first physiochemical passage. Indeed, the physiochemical model provides a filter for molecules which reach and interact with the P-gp active site. From this prospective, the model presented in this study is not intended as an alternative way to separate P-gp inhibitors from P-gp non-inhibitors, but it can provide a supplement were other strategies do not return predictions. By using the dataset proposed by Keogh and co-workers (Keogh and Kunta, 2006), we predicted the inhibition with the physiochemical model. Then, we docked those molecules resulting not predicted by the previous model. As it can be seen in Fig. 6, the coverage of correct predictions increases consistently, although the model accuracy does not change much (data not shown).

The strategy applied in this study takes advantages in the complementarity between modelling and experimental teams, which is the key for successful models and can be used to improve experimental efficiency. Indeed, only molecules interacting with P-gp (QSAR predictions) are proposed for further *in vitro* testing with the most appropriate probe depending on the site identified by the structure-based model. This results in most efficient DDI experiment *in vitro* focused on the most appropriate probe.

Previous P-gp models were mainly meant to discriminate binders from non-binders or to differentiate ligands by their mechanism of interaction (e.g. inhibitors/non-substrates, substrate-inhibitors, etc) (Ferreira et al., 2013; Zeino et al., 2014). In contrast our approach was most focus on the need to classify compounds according to sub-site classification to provide a real help in the study of DDIs.

The model presented provides an improvement in the classification of DDIs for P-gp because it is not based on the chemical similarity comparison between substrates and inhibitors structures, but on similarity scoring between ligands and sites of interaction. Our sub-site model is now available and validated, some advantage from its use could be in drug design, in the interpretation and analysing observations, in drug development for more efficient assays or in drug delivery. In medicinal drug design our model can help to select the appropriate probe to be used in experimental phase. In the case of drugs potentially interacting with P-gp but with specific sub-sites interaction unknown, the model could access the molecular interpretation for DDI by demonstrating the interaction in a specific sub-site or when all sub-sites are involved. It can also be used in virtual optimisation to get proposed optimized structures. The model can also be used to give the opportunity to visualise, at a molecular level the results obtained. The scientist can better understand the mechanism behind odd inhibition patterns. Moreover, it can help to design structures that shall avoid interacting with sub-sites used by another drug for a specific treatment; thus, reducing the global DDIs risk and improve access to the target organ. This could be used in cancer therapy for drug delivery using drug cocktail where drugs should not interact on the same P-gp sub-site but both inhibiting P-gp resulting in an improved absorption of drug in tumours.

CONCLUSIONS

Here we described a new computational strategy to predict the P-gp substrate-dependent inhibition in synergy with *in vitro* experiments. We used the P-gp crystal structure to detect the D-site, H-site and R-site, which are the hypothetical binding sites of P-gp substrates digoxin, H33342 and Rh123 respectively. Successively, the docking of P-gp experimental substrate-dependent inhibitors has unravelled the reasons for their behaviour.

The model is based on the simple competition of the ligand towards each sub-site detected. In this way, we were able to rationalize most of our stating experimental data (see Table 3). The model is consistent also for those compounds showing substrate-independent inhibition. The performances observed from the initial set of experiments are supported by a separate set of experiments whose results confirm the majority of the model predictions (Table 3). To our knowledge, this is the only P-gp structure-based model supported by experimental validation.

The modelling work and the results obtained provide a deeper knowledge of the complex P-gp functionalities and may substantially simplify the design of new drugs. The model described can be used prior *in vitro* testing to target the proper substrate in inhibition experiments for a more efficient detection of drug-drug interactions. In combination with QSAR physiochemical-based models, it can speed up experimental studies by focusing only on compounds interacting with P-gp. It could elucidate the sometimes fuzzy literature data concerning P-gp inhibition where the impact of the probe used could not always be well documented. If addressing P-gp as target (e.g. in cancer therapies), it could be used to design sub-site specific competitors to enhance the permeability of other co-administered drugs.

In the future, the model may significantly benefit of additional data experiments (especially non-competitive) and molecular dynamics studies with better P-gp structures focused on a small subset of differently behaving P-gp ligands. Furthermore, this approach could be extended in the near future to other P-gp substrates and to other membrane transport proteins, once their structures will become available.

ACKNOWLEDGMENTS

REFERENCES

FLAP version 2.2.0 is distributed by Molecular Discovery Ltd. http://www.moldiscovery.com/.

U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), In Vitro Metabolism and Transporter Mediated Drug-Drug Interaction Studies, Guidance for Industry. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM581965.pdf.

PubChem. https://pubchem.ncbi.nlm.nih.gov/ (accessed 1/7/2016).

Baroni, M., Cruciani, G., Sciabola, S., Perruccio, F., Mason, J.S., 2007. A Common Reference Framework for Analyzing/Comparing Proteins and Ligands. Fingerprints for Ligands And Proteins (FLAP): Theory and Application. J. Chem. Inf. Model. 47, 279-294.

Bentz, J.E., O'Connor, M., Bednarczyk, D., Coleman, J., Lee, C.A., Palm, J.E., Pak, A., Perloff, E.S., Reyner, E.L., Balimane, P., 2013. Variability in P-glycoprotein inhibitory potency (IC50) using various in vitro experimental systems: implications for universal digoxin DDI risk assessment decision criteria. Drug Metab. Dispos., dmd. 112.050500.

Bocci, G., Carosati, E., Vayer, P., Arrault, A., Lozano, S., Cruciani, G., 2017. ADME-Space: a new tool for medicinal chemists to explore ADME properties. Sci. Rep. 7, 6359.

Brincat, J.P., Carosati, E., Sabatini, S., Manfroni, G., Fravolini, A., Raygada, J.L., Patel, D., Kaatz, G.W., Cruciani, G., 2011. Discovery of Novel Inhibitors of the NorA Multidrug Transporter of Staphylococcus aureus. J. Med. Chem. 54, 354-365.

Broccatelli, F., Carosati, E., Neri, A., Frosini, M., Goracci, L., Oprea, T.I., Cruciani, G., 2011. A novel approach for predicting P-glycoprotein (ABCB1) inhibition using molecular interaction fields. J. Med. Chem. 54, 1740-1751.

Chedik, L., Bruyere, A., Le Vee, M., Stieger, B., Denizot, C., Parmentier, Y., Potin, S., Fardel, O., 2017. Inhibition of human drug transporter activities by the pyrethroid pesticides allethrin and tetramethrin. PLoS One 12, e0169480.

Chen, L., Li, Y., Yu, H., Zhang, L., Hou, T., 2012. Computational models for predicting substrates or inhibitors of P-glycoprotein. Drug discovery today 17, 343-351.

Chufan, E.E., Kapoor, K., Sim, H.-M., Singh, S., Talele, T.T., Durell, S.R., Ambudkar, S.V., 2013. Multiple transport-active binding sites are available for a single substrate on human P-glycoprotein (ABCB1). PLoS One 8, e82463.

Cross, S., Baroni, M., Carosati, E., Benedetti, P., Clementi, S., 2010. FLAP: GRID Molecular Interaction Fields in Virtual Screening. Validation using the DUD Data Set. J. Chem. Inf. Model. 50, 1442-1450.

Doan, K.M.M., Humphreys, J.E., Webster, L.O., Wring, S.A., Shampine, L.J., Serabjit-Singh, C.J., Adkison, K.K., Polli, J.W., 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.

Ferreira, R.J., Ferreira, M.-J.U., dos Santos, D.J., 2013. Molecular docking characterizes substrate-binding sites and efflux modulation mechanisms within P-glycoprotein. J. Chem. Inf. Model. 53, 1747-1760.

Giacomini, K., Huang, S.M., 2013. Transporters in drug development and clinical pharmacology. Clinical Pharmacology & Therapeutics 94, 3-9.

Goodford, P.J., 1985. A computational procedure for determining energetically favorable binding sites on biologically important macromolecules. J. Med. Chem. 28, 849-857.

Goracci, L., Deschamps, N., Randazzo, G.M., Petit, C., Dos Santos Passos, C., Carrupt, P.-A., Simões-Pires, C., Nurisso, A., 2016. A Rational Approach for the Identification of Non-Hydroxamate HDAC6-Selective Inhibitors. Sci. Rep. 6, 29086.

Hayeshi, R., Hilgendorf, C., Artursson, P., Augustijns, P., Brodin, B., Dehertogh, P., Fisher, K., Fossati, L., Hovenkamp, E., Korjamo, T., 2008. Comparison of drug transporter gene expression and functionality in Caco-2 cells from 10 different laboratories. Eur. J. Pharm. Sci. 35, 383-396.

Jouan, E., Le Vée, M., Mayati, A., Denizot, C., Parmentier, Y., Fardel, O., 2016. Evaluation of P-glycoprotein inhibitory potential using a rhodamine 123 accumulation assay. Pharmaceutics 8, 12.

Kapelyukh, Y., Paine, M.J., Maréchal, J.-D., Sutcliffe, M.J., Wolf, C.R., Roberts, G.C., 2008. Multiple substrate binding by cytochrome P450 3A4: estimation of the number of bound substrate molecules. Drug Metab. Dispos. 36, 2136-2144.

Kenworthy, K., Bloomer, J., Clarke, S., Houston, J., 1999. CYP3A4 drug interactions: correlation of 10 in vitro probe substrates. Br. J. Clin. Pharmacol. 48, 716.

Keogh, J.P., Kunta, J.R., 2006. Development, validation and utility of an in vitro technique for assessment of potential clinical drug–drug interactions involving P-glycoprotein. Eur. J. Pharm. Sci. 27, 543-554.

Ledwitch, K.V., Barnes, R.W., Roberts, A.G., 2016. Unravelling the complex drug–drug interactions of the cardiovascular drugs, verapamil and digoxin, with P-glycoprotein. Biosci. Rep. 36, e00309.

Li, J., Jaimes, K.F., Aller, S.G., 2014. Refined structures of mouse P-glycoprotein. Protein Sci. 23, 34-46.

Lin, J.H., 2003. Drug–drug interaction mediated by inhibition and induction of P-glycoprotein. Adv. Drug Delivery Rev. 55, 53-81.

Lugo, M.R., Sharom, F.J., 2005. Interaction of LDS-751 with P-glycoprotein and mapping of the location of the R drug binding site. Biochemistry 44, 643-655.

Martin, C., Berridge, G., Higgins, C.F., Mistry, P., Charlton, P., Callaghan, R., 2000. Communication between multiple drug binding sites on P-glycoprotein. Mol. Pharmacol. 58, 624-632.

Milletti, F., Storchi, L., Sforna, G., Cruciani, G., 2007. New and Original pKa Prediction Method Using Grid Molecular Interaction Fields. J. Chem. Inf. Model. 47, 2172-2181.

Montanari, F., Ecker, G.F., 2015. Prediction of drug–ABC-transporter interaction—Recent advances and future challenges. Adv. Drug Delivery Rev. 86, 17-26.

Obach, R.S., Walsky, R.L., Venkatakrishnan, K., Gaman, E.A., Houston, J.B., Tremaine, L.M., 2006. The utility of in vitro cytochrome P450 inhibition data in the prediction of drug-drug interactions. J. Pharmacol. Exp. Ther. 316, 336-348.

Onakpoya, I.J., Heneghan, C.J., Aronson, J.K., 2016. Post-marketing withdrawal of 462 medicinal products because of adverse drug reactions: a systematic review of the world literature. BMC medicine 14, 10.

Pajeva, I.K., Hanl, M., Wiese, M., 2013. Protein Contacts and Ligand Binding in the Inward- Facing Model of Human P- Glycoprotein. ChemMedChem 8, 748-762.

Pauli-Magnus, C., Rekersbrink, S., Klotz, U., Fromm, M.F., 2001. Interaction of omeprazole, lansoprazole and pantoprazole with P-glycoprotein. Naunyn-Schmiedeberg's Arch. Pharmacol. 364, 551-557.

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

Rautio, J., Humphreys, J.E., Webster, L.O., Balakrishnan, A., Keogh, J.P., Kunta, J.R., Serabjit-Singh, C.J., Polli, J.W., 2006. In vitro p-glycoprotein inhibition assays for assessment of clinical drug interaction potential of new drug candidates: a recommendation for probe substrates. Drug Metab. Dispos. 34, 786-792.

Rodrigues, A.D., 2008. Drug-drug interactions. CRC Press.

Schuster, D., Laggner, C., Langer, T., 2005. Why drugs fail-a study on side effects in new chemical entities. Curr. Pharm. Des. 11, 3545-3559.

Shapiro, A.B., Ling, V., 1997. Positively cooperative sites for drug transport by P- glycoprotein with distinct drug specificities. Eur. J. Biochem. 250, 130-137.

Sun, Y.-L., Kathawala, R.J., Singh, S., Zheng, K., Talele, T.T., Jiang, W.-Q., Chen, Z.-S., 2012. Zafirlukast antagonizes ATP-binding cassette subfamily G member 2-mediated multidrug resistance. Anti-Cancer Drugs 23, 865-873.

Szakacs, G., Varadi, A., Özvegy-Laczka, C., Sarkadi, B., 2008. The role of ABC transporters in drug absorption, distribution, metabolism, excretion and toxicity (ADME–Tox). Drug discovery today 13, 379-393.

Varma, M.V., Pang, K.S., Isoherranen, N., Zhao, P., 2015. Dealing with the complex drug–drug interactions: towards mechanistic models. Biopharm. Drug Dispos. 36, 71-92.

Weiss, J., Dormann, S.-M.G., Martin-Facklam, M., Kerpen, C.J., Ketabi-Kiyanvash, N., Haefeli, W.E., 2003. Inhibition of P-glycoprotein by newer antidepressants. J. Pharmacol. Exp. Ther. 305, 197-204.

Zeino, M., Saeed, M.E., Kadioglu, O., Efferth, T., 2014. The ability of molecular docking to unravel the controversy and challenges related to P-glycoprotein—a well-known, yet poorly understood drug transporter. Invest. New Drugs 32, 618-625.

Zhang, L., Zhang, Y.D., Zhao, P., Huang, S.-M., 2009. Predicting drug–drug interactions: an FDA perspective. The AAPS journal 11, 300-306.

Zolnerciks, J.K., Booth- Genthe, C.L., Gupta, A., Harris, J., Unadkat, J.D., 2011. Substrate- and species- dependent inhibition of p- glycoprotein- mediated transport: Implications for predicting in vivo drug interactions. J. Pharm. Sci. 100, 3055-3061.

MANUSCRIPT TABLES & FIG.S

Compound	digoxin	H33342	Rh123
amiodarone	INHIBITOR	INHIBITOR	INHIBITOR
diltiazem	INHIBITOR	INHIBITOR	INHIBITOR
elacridar	INHIBITOR	INHIBITOR	INHIBITOR
felodipine	INHIBITOR	INHIBITOR	INHIBITOR
nicardipine	INHIBITOR	INHIBITOR	INHIBITOR
nitrendipine	INHIBITOR	INHIBITOR	INHIBITOR
quinidine	INHIBITOR	INHIBITOR	INHIBITOR
sertraline	INHIBITOR	INHIBITOR	NONINHIBITOR
verapamil	INHIBITOR	INHIBITOR	INHIBITOR
zosuquidar	INHIBITOR	INHIBITOR	INHIBITOR
alfentanil	INHIBITOR	INHIBITOR	
amlodipine	INHIBITOR	INHIBITOR	
astemizole	INHIBITOR	INHIBITOR	
cisapride	INHIBITOR	INHIBITOR	
disopyramide	INHIBITOR	NONINHIBITOR	
ketoconazole	INHIBITOR	INHIBITOR	
loratadine	INHIBITOR	INHIBITOR	
lovastatin	INHIBITOR	INHIBITOR	
midazolam	INHIBITOR	INHIBITOR	

nefazodone	INHIBITOR	INHIBITOR	
pimozide	INHIBITOR	INHIBITOR	
quinine	INHIBITOR	INHIBITOR	
sildenafil	INHIBITOR	INHIBITOR	
zafirlukast	INHIBITOR	NONINHIBITOR	
carvedilol	INHIBITOR		INHIBITOR
isradipine	INHIBITOR		INHIBITOR
itraconazole	INHIBITOR		NONINHIBITOR
lansoprazole	INHIBITOR		NONINHIBITOR
mibefradil	INHIBITOR		INHIBITOR
troglitazone	INHIBITOR		INHIBITOR

Table 1. Starting set experiments and predictions. The P-gp inhibitor or non-inhibitor nature of the compounds was determined from experimental data collected from the literature or generated in the study (see Supporting information 02). Green cells are those for which the docking agrees with the experiments. Red cells are those for which the experiment resulted opposite to the docking prediction. White cells correspond to experiments not performed or for which no experimental data was available at the beginning of the study.

Compound	digoxin	H33342	Rh123
cisapride			INHIBITOR
disopyramide			NONINHIBITOR
midazolam			NONINHIBITOR
zafirlukast			INHIBITOR
carvedilol		INHIBITOR	
itraconazole		NONINHIBITOR	
lansoprazole		INHIBITOR	
mibefradil		INHIBITOR	
amitriptyline	INHIBITOR	INHIBITOR	INHIBITOR
loperamide	INHIBITOR	INHIBITOR	INHIBITOR
nifedipine	INHIBITOR	INHIBITOR	INHIBITOR
prazosin	INHIBITOR	INHIBITOR	NONINHIBITOR
reserpine	INHIBITOR	INHIBITOR	INHIBITOR
terfenadine	NONINHIBITOR	INHIBITOR	INHIBITOR

Table 2. Experimental results for the model validation. Green cells are those for which the docking agrees with the experiments. Red cells are those for which the experiment resulted opposite to the docking prediction. White cells correspond to experiments already performed (see Table 1).

	digoxin	H33342	Rh123
Starting set	97%	92%	88%

18

Validation set	67%	100%	90%

Table 3. Model performances. Percentages of correctly predicted molecules for each probe substrate in both sets of experiments.
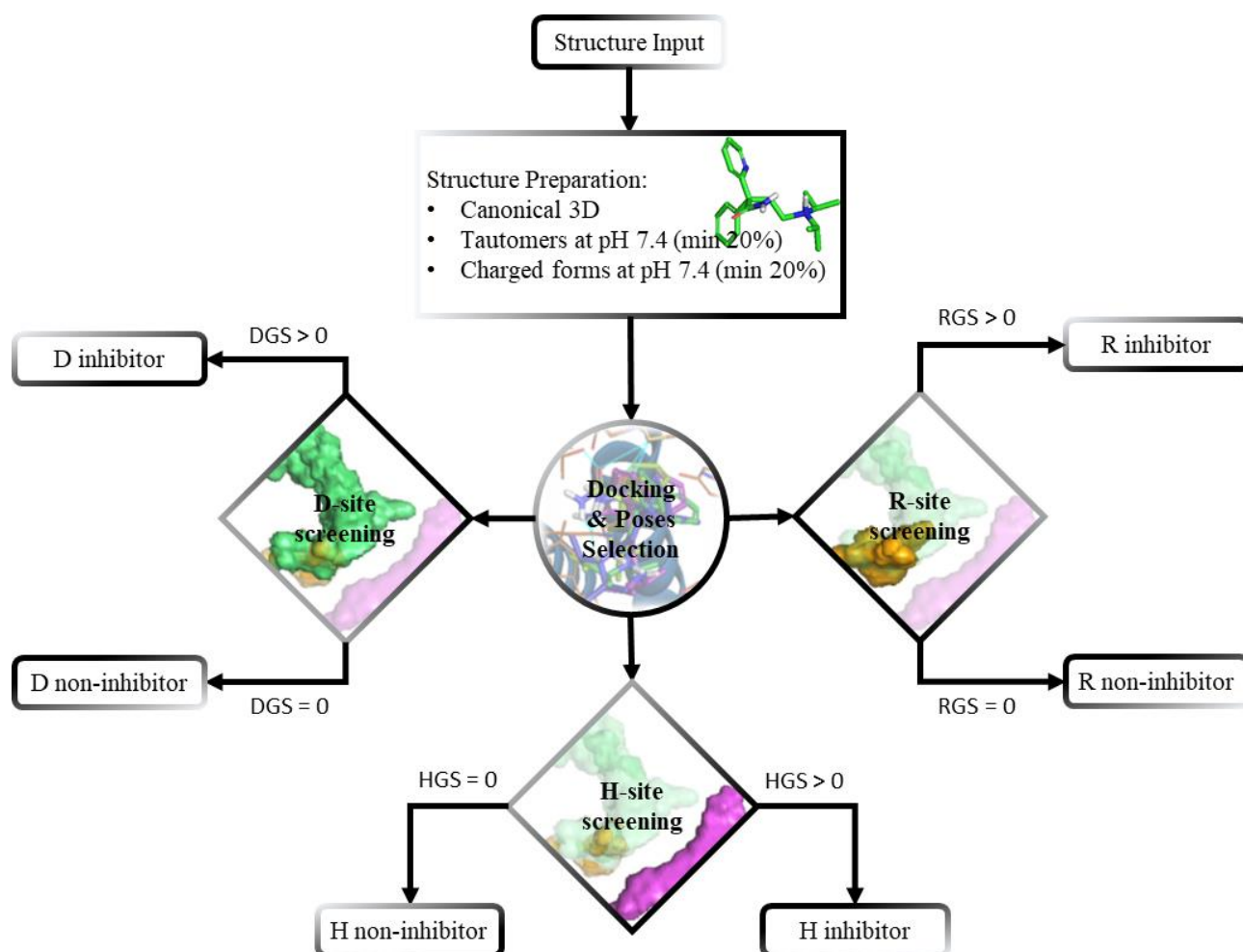

Fig. 1. Flowchart of the model steps to generate predictions for new ligands. DGS (digoxin Glob-Sum), HGS (H33342 Glob-Sum), RGS (Rh123 Glob-Sum).
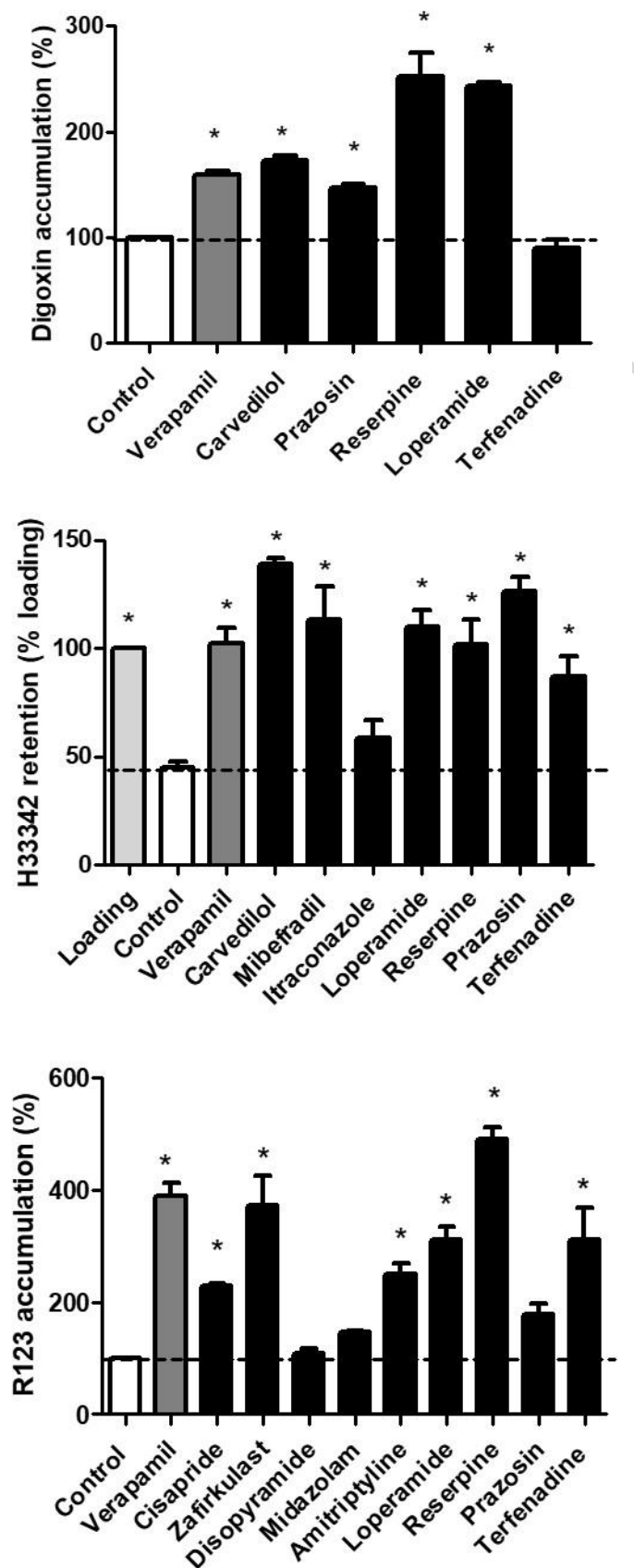
Fig. 2. Inhibitory effects of compounds on P-gp-mediated transport of digoxin, H33342 and Rh123. Accumulation of

digoxin and Rh123 and retention of H33342 in P-gp-expressing MCF7R cells were determined in the absence (control)

or presence of various compounds (each at 100 µM, excepted terfenadine for which a 10 µM concentration was used), as indicated in Experimental Section. Data are the means ± SEM of at least three independent assays. *, p<0.05 when compared to control.
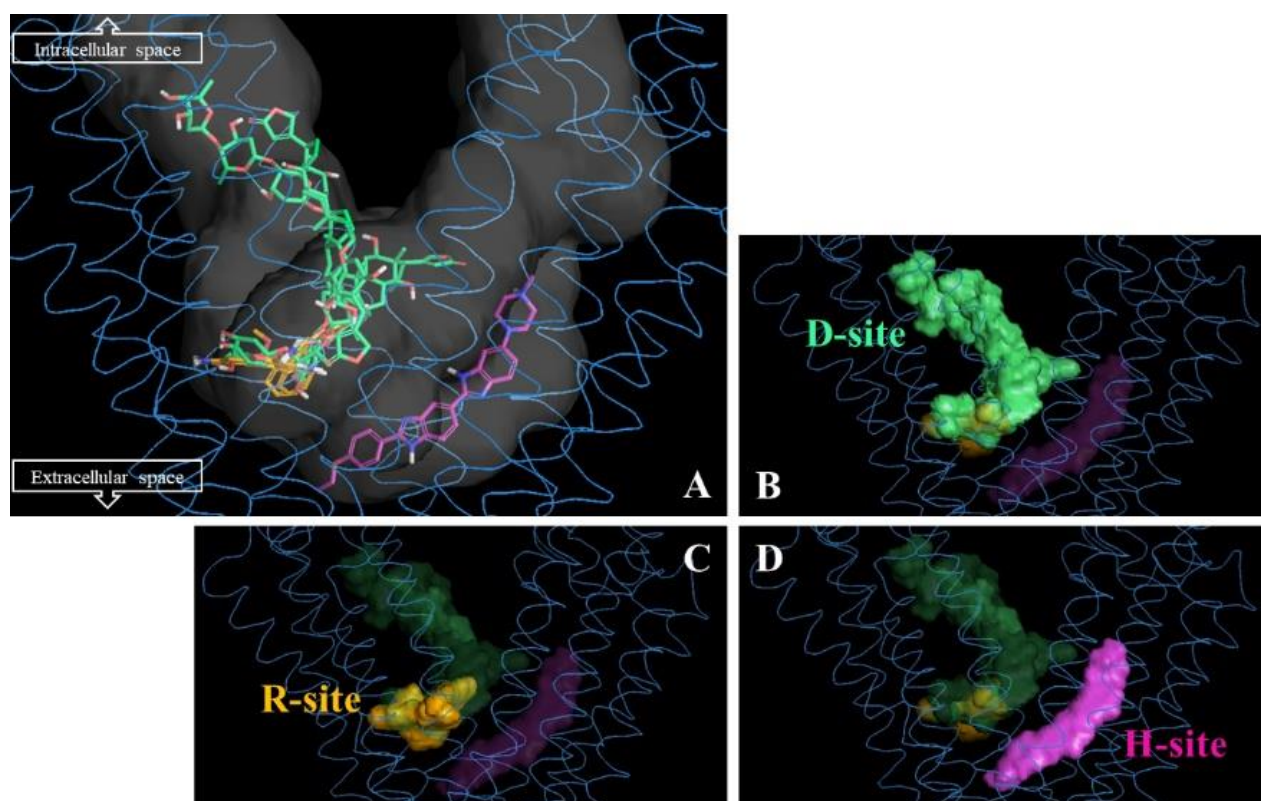

Fig. 3. (A) Top-ranked poses for the three probe substrates used: digoxin (green), Rh123 (orange) and H33342 (purple). **(B)** D-site depiction (green). **(C)** R-site depiction (orange). **(D)** H-site depiction (purple).
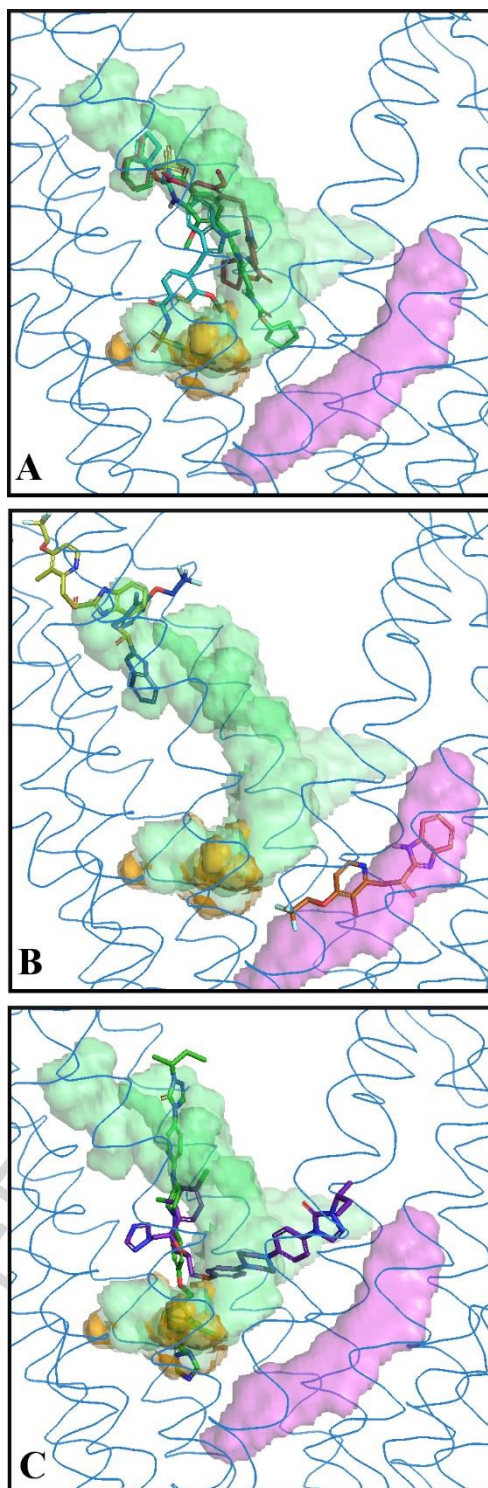
Fig. 4. Top-ranked binding poses for zafirlukast **(A)**, lansoprazole **(B)** and itraconazole **(C)**. The substrates binding sites are depicted transparently: green (D-site), orange (R-site) and purple (H-site).
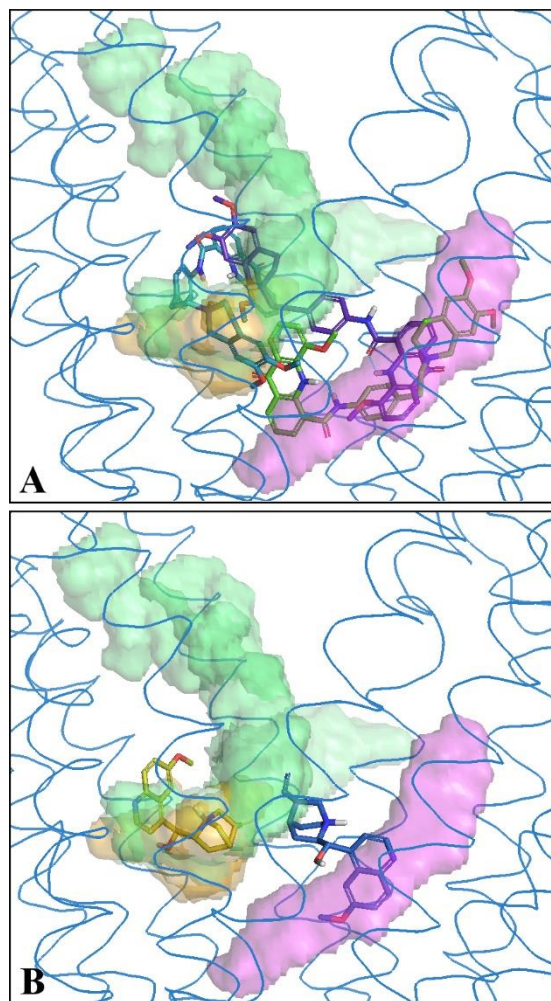
Fig. 5. Top-ranked poses for elacridar **(A)** and quinidine **(B)**. The substrates binding sites are depicted transparently: green (D-site), orange (R-site) and purple (H-site).
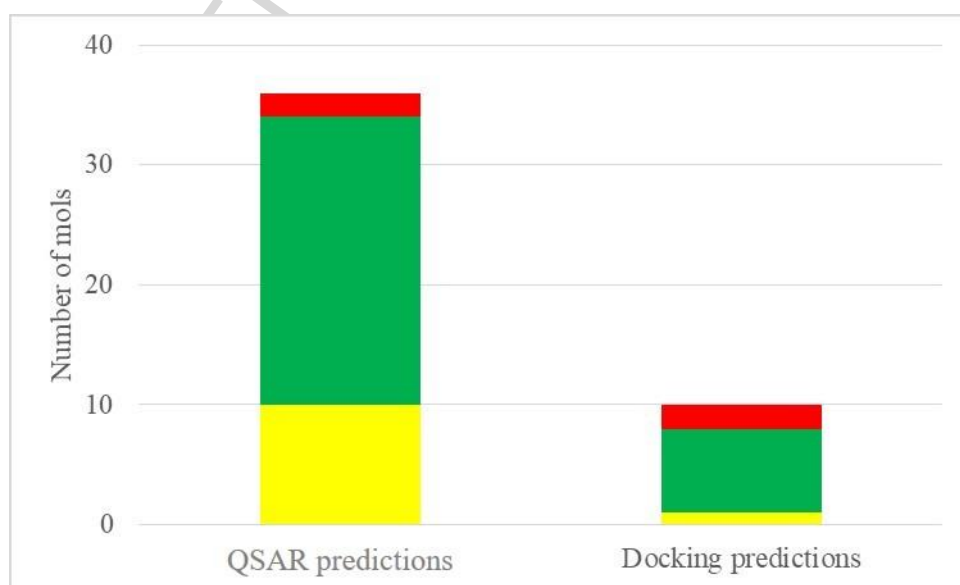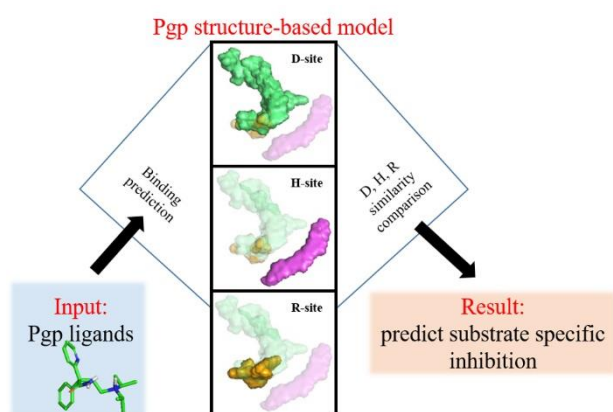
Fig. 6. Schema of predictions for the Keorgh's dataset. Correct predictions are reported in green, wrong predictions in red and uncertain in yellow. The right bar (docking) predictions were generated only for those molecules not predicted by QSAR (yellow left bar).

Graphical abstract