



# Utility of protein structures in overcoming ADMET-related issues of drug-like compounds

Friederike Stoll, Andreas H. Göller and Alexander Hillisch

Bayer HealthCare AG, Global Drug Discovery, Medicinal Chemistry, Aprather Weg 18a, D-42096 Wuppertal, Germany

The number of solved X-ray structures of proteins relevant for ADMET processes of drug molecules has increased remarkably over recent years. In principle, this development offers the possibility to complement the quantitative structure–property relationship (QSPR)-dominated repertoire of *in silico* ADMET methods with protein-structure-based approaches. However, the complex nature and the weak nonspecific ligand-binding properties of ADMET proteins take structural biology methods and current docking programs to the limit. In this review we discuss the utility of protein-structure-based design and docking approaches aimed at overcoming issues related to plasma protein binding, active transport via P-glycoprotein, hERG channel mediated cardiotoxicity and cytochrome P450 inhibition, metabolism and induction.

## Introduction

Protein-structure-based drug design is, today, an integral part of the lead discovery and optimization process. It focuses on the identification of novel lead compounds and/or the improvement of binding affinity to a given pharmacological target protein. The method iteratively involves the determination of the binding mode of a lead structure to a target, preferentially by X-ray crystallography, the *in silico* design of novel derivatives that should improve binding and the subsequent synthesis and testing of these derivatives. The approach is especially powerful if the physicochemical properties of the lead are simultaneously taken into account to solve ADMET (absorption, distribution, metabolism, excretion and toxicity)-related issues. Several drugs have reached the market for which structure-based drug design played at least some role in their discovery and optimization [1,2]. Owing to significant advances in protein production and X-ray crystallography technologies, today, the 3D structures of many proteins involved in ADMET-relevant processes are known down to the atomic detail. It is expected that these structures are similarly helpful in drug design as are structures of pharmacological target proteins. Apart from affinity, ADMET issues are key obstacles during lead optimization and any rational approach that enables

the design of improved compounds is highly desirable. Traditionally, ADMET predictions have been relying on knowledge-based and quantitative structure–property relationship (QSPR) approaches in which the chemical structures of many drug-like molecules could be correlated with measured experimental data. The quality and, thus, the utility of the resulting prediction models depend, for example, on the quality of the experimental data, the number and selection of compounds, the choice of descriptors, and statistical methods. The advantage of such approaches is that complex biological processes involving various molecular targets can be modeled (e.g. cell permeation, Ames mutagenicity test). The approach is valuable especially in the early phase of lead discovery and the hit-to-lead process. However, in many cases these models are too inaccurate to predict subtle structural changes that frequently occur during lead optimization. For the latter application protein-structure-based approaches seem attractive. If a certain ADMET process can clearly be related to a single protein experimentally, and the 3D structure of this protein is accessible, it should be possible to design compounds that show, for example, attenuated binding to that protein. There are several proteins with distinguished relevance in processes such as plasma-protein-binding, active transport, cytochrome P450 (CYP) inhibition, metabolism and induction. In this review we will concentrate on such proteins, namely human serum albumin

Corresponding author: Hillisch, A. (alexander.hillisch@bayer.com)

(HSA), P-glycoprotein (P-gp), pregnane X receptor (PXR), constitutive androstane receptor (CAR), CYPs and the hERG (human Ether-à-go-go Related Gene) channel. This list of ADMET-relevant proteins is certainly not complete, but we will focus on those proteins that have attracted the most attention and research to date. The quality and nature of the available protein structure information is reviewed. If and how these protein structures can be utilized to address pressing ADMET problems in lead optimization will be considered.

### Active transport through membranes via P-gp

P-gp or multidrug resistance protein 1 (MDR1) is a member of the ATP-binding cassette (ABC) superfamily and, among other transporters, an active efflux pump that can prevent the influx of potentially harmful compounds into an organism by expelling them from cells in an energy-dependent manner. The protein recognizes predominantly lipophilic compounds with a molecular weight between 330 and 4000 g/mol [3,4]. It has been compared to a molecular 'hydrophobic vacuum cleaner', pulling substrates from the membrane and expelling them into the extracellular compartment. This property of P-gp interferes with absorption of drug-like molecules from the gastrointestinal tract and permeation into the central nervous system. Thus, it has received special attention in drug discovery research. Recent reviews summarize ligand- and structure-based prediction of P-gp transport in drug design [5–7].

P-gp is an integral membrane protein of 1280 amino acids arranged into two homologous halves, each comprising six transmembrane alpha helices (MD, membrane domain) and an ATP-binding domain (NBD, nucleotide-binding domain). The first more reliable homology models of human MDR1 (P-gp) [8,9] were built based on the structure of a putative multidrug transporter from *Staphylococcus aureus* (Sav1866) [10,11]. Earlier attempts to model the 3D structure of P-gp suffered from a low sequence identity to the template protein and incorrectly solved template X-ray structures such as the bacterial ABC transporter MsbA. In 2009, the first and, to date, only X-ray structures of murine P-gp (Fig. 1) in complex with two cyclopeptidic ligands were published [12]. The murine and human P-gp sequences are 87% identical. The structures show P-gp in an inward-facing stage of the transport cycle without nucleotides bound. A recent review summarizes the complexity of the transport cycle, especially the role of the interaction of NBDs with MDs [13]. Although the murine P-gp X-ray structure is now solved, it represents only a snapshot of one structural state in the transport cycle. This is certainly not enough to understand the entire transport phenomenon in detail. In particular, the dynamics of the interaction between the NBD and the MD are not completely understood. In several studies homology models were used to explain putative ligand-binding sites and interpret photo affinity labeling experiments on a more qualitative basis [8,14]. In just a few publications, the P-gp models were used to dock compounds into putative ligand-binding sites. Becker *et al.* [15] performed docking experiments into homology models based on the corrected MsbA template structure with four different ligands. The derived binding modes are in agreement with cross-linking and point mutagenesis experiments. A recent study used docking into the murine P-gp X-ray structure to explain binding of a compound series retrospectively [16]. Owing to the

complex nature of the P-gp transport process and the fragmentary structural understanding it has, so far, not been possible to use P-gp homology models or X-ray structures for prospective structure-based design approaches. The hope is that with more co-crystallized ligands one might apply structure-based design to overcome, for example, unwanted P-gp efflux.

### CYP induction: PXR and CAR

PXR and CAR are two nuclear hormone receptors involved in the detoxification of xenobiotic substances by upregulating CYP transcription, especially that of the CYP3A and CYP2B families. Because CYP3A4 is one of the most important drug-metabolizing enzymes, its upregulation can lead to undesirable drug–drug interactions (DDIs) [17,18].

PXR and CAR consist of 434 and 352 amino acids, respectively, that fold into a DNA- and a ligand-binding domain (DBD; LBD) and heterodimerize with RXR (9-*cis* retinoic acid receptor) to bind responsive DNA elements, controlling CYP3A and CYP2B transcription. In 2001 the first structure of the PXR LBD was elucidated [60]. The general fold (Fig. 2a) resembles that of other nuclear hormone receptors but the binding pocket is larger and more flexible. This is caused by a beta-sheet insert between helices 1 and 3 [17,18]. The pocket is lined by ~20 hydrophobic/aromatic residues supplemented by a small number of polar and charged amino acids. An analysis of the available PXR structures complexed to various ligands highlights Gln285, Ser247, and His407 as the most important H-bonding partners. Taking into account that most side-chains show a high degree of flexibility (Fig. 2b), and that there is additional variability regarding tautomeric and protonation states of the polar amino acids, it is comprehensible that PXR recognizes ligands of different sizes, shapes and properties.

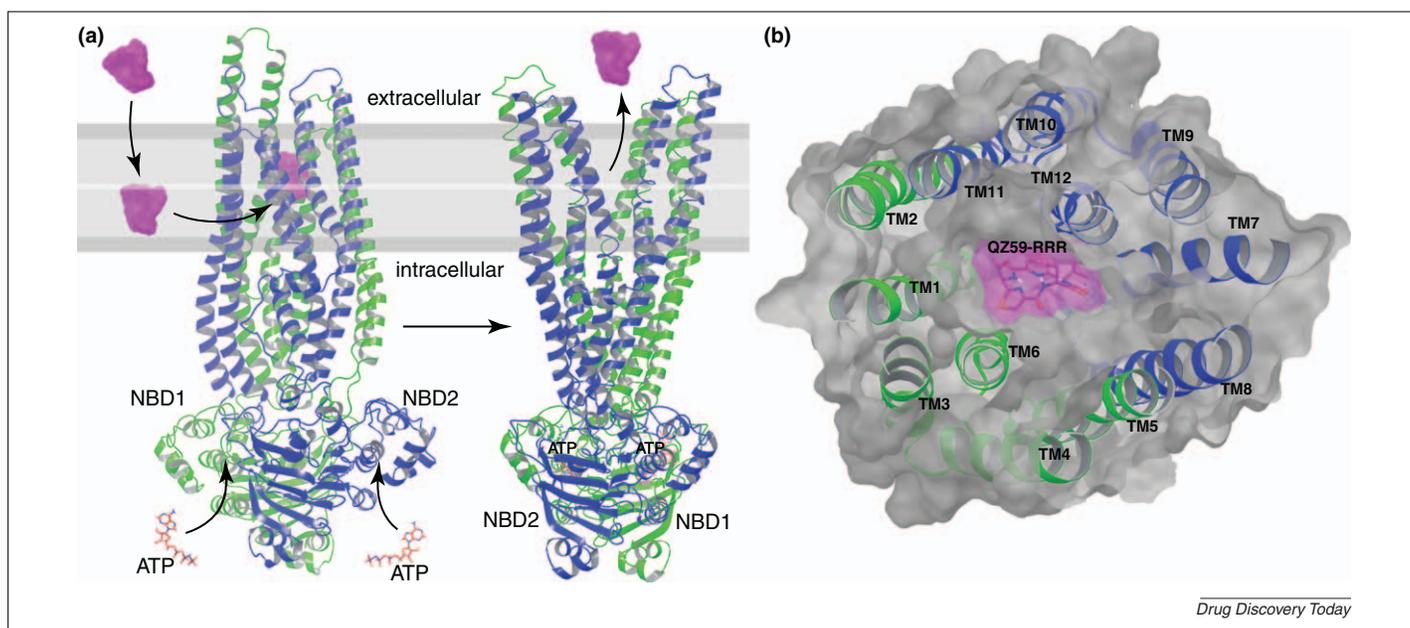
As of May 2011, nine PXR and four CAR X-ray structures were publicly available, most of them in complex with a ligand. A selection is listed in Table 1.

Several attempts have been made to classify CYP3A4 inducers and non-inducers by docking into the PXR LBD, often in combination with chemoinformatic models. Even for the narrow dataset used by Ekins *et al.* [19] the success rate was moderate, although subset-specific differences were observed. Similar results were obtained with other datasets [20,21]. Models that classify structurally diverse CYP3A4 inducers and non-inducers seem to be more robust when machine-learning methods or ligand-based similarity searches are applied without consideration of the protein structure [22,23].

However, it has been shown that PXR structural information can be used to design modifications that diminish PXR binding and activation to avoid putative DDIs. Because the binding mode of a compound has to be known in detail to suggest modifications that disturb crucial interactions, successful predictions will be much easier when a crystallized complex of a related compound is available. By contrast, it was even possible to attenuate PXR interaction based on two alternative plausible docking modes. The introduction of polar groups into apolar subpockets was shown successfully to reduce PXR activation and CYP3A4 induction [24,25]. This constitutes one of the rare reported cases in which protein structure information guided the design of compounds with improved ADMET properties.

**TABLE 1**  
**Selected X-ray structures of ADMET-relevant proteins**

ADMET protein	PDB code	Resolution (Å)	Ligand	Publication year	Reference
<b>P-gp</b>	3G5U	3.8	apo	2009	[12]
	3G60	4.4	QZ59-RRR	2009	[12]
	3G61	4.3	QZ59-SSS	2009	[12]
<b>PXR</b>	1NRL	2.0	Hyperforin	2003	[59]
	1SKX	2.8	Rifampicin	2005	[24]
	1ILH	2.8	SR12813	2001	[60]
<b>CAR</b>	1XV9	2.7	5 $\beta$ -Pregnane-3,20-dione	2004	[64]
<b>CYP1A2</b>	2HI4	2.0	$\alpha$ -Naphthoflavone	2007	[63]
<b>CYP2C9</b>	1OG2	2.6	apo	2003	[62]
	1OG5	2.6	Warfarin	2003	[62]
	1R9O	2.0	Flurbiprofen	2004	[61]
<b>CYP2D6</b>	2F9Q	3.0	apo	2006	[65]
<b>CYP3A4</b>	3WOE	2.8	apo	2004	[35]
	3WOF	2.6	Metyrapone	2004	[35]
	3WOG	2.7	Progesterone	2004	[35]
	2VOM	2.8	Ketoconazole	2006	[26]
	2J0D	2.8	Erythromycin	2006	[26]
	3NXU	2.0	Ritonavir	2010	[66]
<b>HSA</b>	1AO6	2.5	apo	1999	[67]
	2BXD	3.1	R-Warfarin	2005	[39]
	2BXE	2.7	Diflunisal	2005	[39]
	3B9L	2.6	Myristic acid, AZT	2008	[44]
	3B9M	2.7	Myristic acid, AZT, 2-hydroxybenzoic acid	2008	[44]



**FIGURE 1**  
**(a)** Schematic representation of the P-gp substrate transport cycle of a lipophilic compound (magenta). Two states of the P-gp are shown: left, P-gp with bound ligand and without ATP; right, P-gp without ligand and in complex with ATP. The two homologous halves of P-gp are further twisted around each other upon ATP binding and substrate release. **(b)** Ligand binding site on the inward facing conformation of murine P-gp. The cyclopeptidic inhibitor QZ59-RRR is co-crystallized and shown in magenta. Transmembrane helices (TM) are numbered.

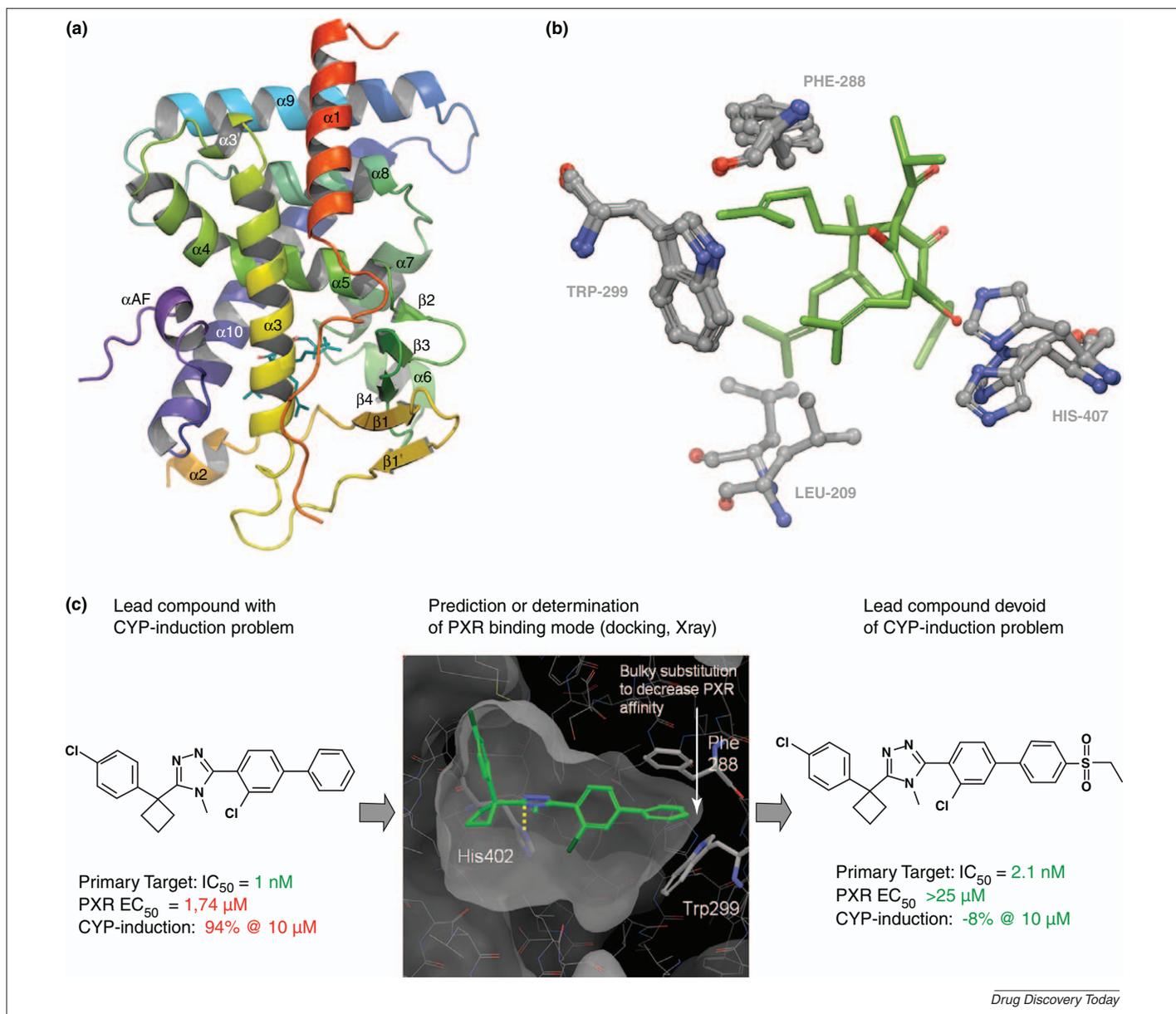


FIGURE 2

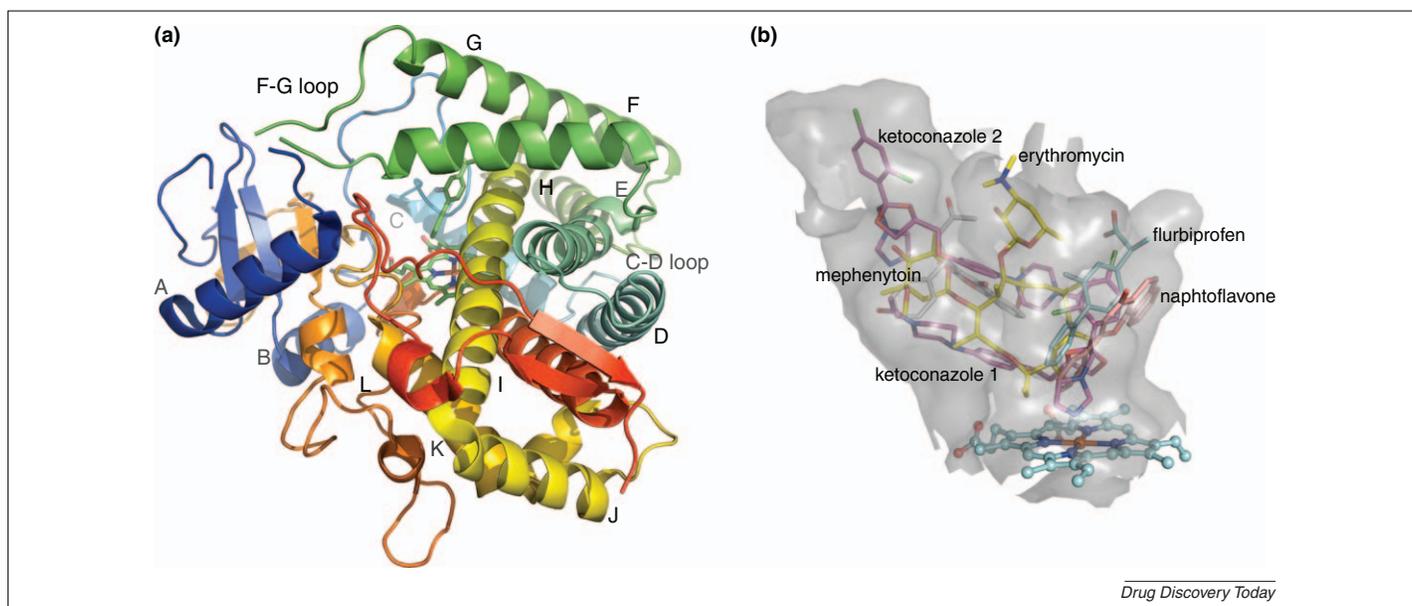
**(a)** The ligand-binding domain of PXR with the activator hyperforin [59]. Secondary structure elements are labeled. **(b)** Hyperforin–PXR complex (PDB code: 1M13 [59]) superimposed with the X-ray complexes of SR12813 (1ILH [60]) and rifampicin (1SKX [24]), ligands not shown. Selected amino acids represent rigid and flexible areas of the binding site. **(c)** Example for a successful structure-based improvement of unwanted PXR binding and cytochrome P450 3A4 induction [25]. The lead structure having a Cyp-induction problem was docked into the binding pocket of PXR. On the basis of the derived binding mode an ethylsulfone group was introduced to form repulsive interactions with rigid sidechains Phe 288 and Trp299. The resulting compound did not lose the primary target affinity and was devoid of PXR binding.

## CYP inhibition and metabolism

CYPs constitute a large and ubiquitous family of heme-containing redox enzymes. They are the most important enzyme class involved in phase 1 metabolism of xenobiotic substances. CYPs transform compounds to more polar substrates by oxidation and more-seldom reduction reactions, leading to improved elimination from the organism. Drugs interact directly with CYPs either as inhibitors or as substrates. Inhibition by the drug itself, or by its metabolites (time-dependent inhibition, TDI), blocks its own metabolism or the metabolism of co-administered drugs leading to unwanted DDIs. CYPs are predominantly expressed in the liver with, currently, 57 isoforms identified in humans. The most

important isoforms are 1A2, 2C8, 2C9, 2C19, 2D6 and 3A4 which together metabolize 75% of drugs on the market, with 2D6 and 2C9 being highly polymorphic and variable in their metabolic rates.

All crystal structures available (for a selection see Table 1) show the same general fold (Fig. 3a) with a conserved core region. Other regions such as the active site, the loops between helices C and D, and the regions between F and G (with their connection loop) are more flexible. Non-Michaelis–Menten kinetics, cooperativity effects and a CYP3A4 crystal structure with two bound ketocozazole molecules [26] prove the possibility of having more than one ligand in the highly flexible large pocket (Fig. 3b). Similarly, a

**FIGURE 3**

**(a)** X-ray structure of cytochrome P450 2C9 (PDB code 1R9O [61]); the ligand flurbiprofen and the heme moiety are shown as sticks with green carbons; the protein is represented as ribbons color-coded by sequence – helices and important loops are labeled. **(b)** Superposition of five P450 X-ray structures with shaded active site cavity for cytochrome P450 3A4 (PDB code 2V0M [26]) demonstrates the promiscuity of ligand binding in this enzyme class. Superimposed are the ligands ketoconazole (PDB code 2V0M, [26], two ligands co-crystallized), erythromycin (CYP3A4, PDB code 2JOD [26]), flurbiprofen (CYP2C9, PDB code 1R9O [61]), (S)-mephenytoin (CYP2C9, PDB code 1OG5 [62]) and  $\alpha$ -naphthoflavone (CYP1A2, PDB code 2HI4 [63]). Ligands are shown as stick representations, heme as ball-and-stick.

2C9 crystal structure with warfarin bound  $\sim 10$  Å from the iron center allows to model a second warfarin in productive orientation into the pocket, suggesting either an allosteric mechanism for activation, a fixation in the entry channel or substrate cooperativity. Erythromycin, by contrast, co-crystallized with CYP3A4 [26] in a rotated non-productive binding mode for the observed metabolism, is indicative of flexibility via rotation in the binding pocket or via substrate re-entry. All described observations considerably limit the chance of success of any docking approaches. Many ligand- and target-based models for substrate metabolism and CYP inhibition are summarized in a series of reviews [27–30].

Afzelius *et al.* [31] compared three ligand-based approaches, two docking methods and one mixed approach for the prediction of sites of metabolism (SOMs) in CYP2C9 and CYP3A4 for 82 public and 40 Astra Zeneca compounds. The docking methods were generally about 15% worse than the other approaches reported. They concluded that for both isoforms metabolism is more driven by reactivity than by pocket constraints.

Tarcsay *et al.* [32] docked metabolites instead of substrates. This innovative concept allowed the application of an iron-metabolite distance pose filter of maximally 6 Å to decide between productive and non-productive docking poses. Owing to the limited performance of the knowledge-based approach used for metabolite creation yielding only 74% of the experimental metabolites, the overall success rate was limited to 69%. Binding site flexibility and four different substrate orientations were derived by Seifert *et al.* [33] via a series of molecular dynamic simulations starting with warfarin-bound and apo CYP2C9 structures.

Teixeira *et al.* [34] provide a direct measure of the binding site promiscuity of CYP3A4 by combined molecular dynamics and

docking. They identified productive binding modes by docking 16 known substrates into 125 protein conformations derived as snapshots from five extensive molecular dynamic simulations with CYP3A4 [35] (PDB-ID: 1W0F), whereas all substrates failed to dock into the binding site in the original X-ray structure. They showed that the cavity volumes change by  $>1000$  Å<sup>3</sup> between snapshots. The downside of this approach is that productive orientations can only be identified if the metabolites are known beforehand.

Although not currently feasible in drug-discovery time frames, molecular dynamic studies can provide insights into substrate regioselectivity and migration to the active site and, based on prior knowledge about one substrate, could give hints as to how to reduce metabolism in alternative derivatives.

CYP inhibition is estimated to be somewhat more difficult to predict than metabolism owing to the lack of reactive sites on the inhibitors that can be used as anchor positions and the lack of direct experimental information [36]. Additionally, CYP inhibitors can chelate to the heme site, bind to the entrance or exit channel, or a pocket far from the heme, or even undergo irreversible reactions. General belief is that most inhibitors will bind in similar ways as substrates. de Groot *et al.* [36] created a combined pharmacophore, homology and reactivity model for 2D6 metabolism and inhibition. They provide two examples for a general strategy to design out heme-binder and non-heme-binder 2D6 inhibition, but without any structural information. An almost quantitative approach for the prediction of pIC<sub>50</sub> values was obtained from combined flexible docking and multi-dimensional QSAR (quantitative structure–activity relationship) with flexibility introduced by an adaptive minimum and a multi-conformer maximum binding site concept with a predictive  $r^2$  of 0.66 [37].

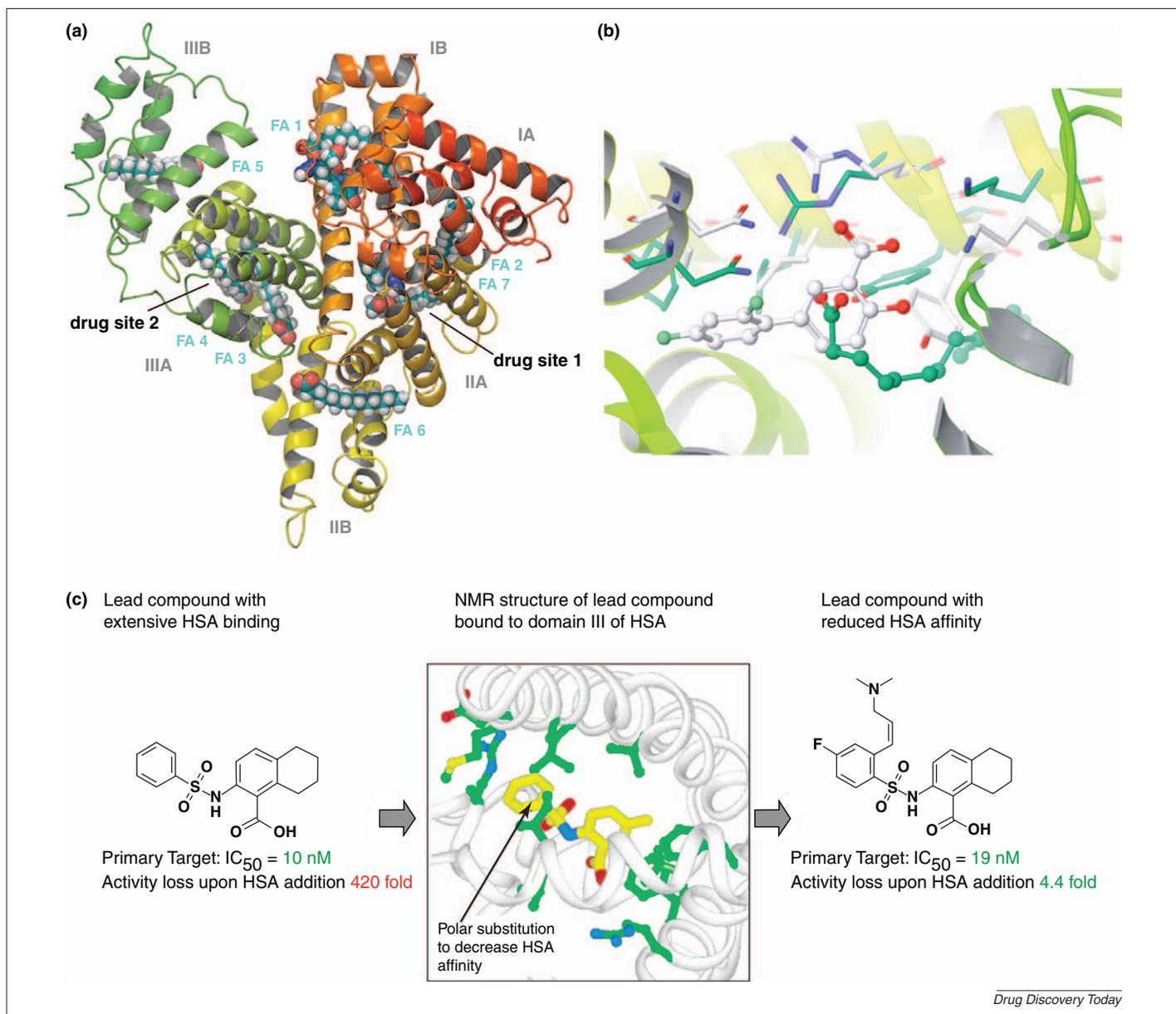


FIGURE 4

(a) HSA has multiple binding sites for drugs and fatty acids (PDB code 3B9L complexed with AZT and myristate [44]); all bound ligands are shown as atom-colored volumes. Subdomains, fatty acid binding sites and the main drug binding sites are labeled. (b) Diflunisal (white carbons, PDB code 2BXE [39]) and myristic acid (green carbons, PDB code 3B9M [44]) binding to drug site 2 on HSA. The shape of the protein-binding pocket and the side chain orientations adjust to the ligand; in addition, different subpockets can be used for binding. (c) Example for a successful structure-based improvement of extensive HSA binding [41]. The binding mode of a lead structure on domain III of HSA was determined by NMR spectroscopy (reproduced, with permission from ACS). On the basis of this binding mode basic/polar substituents were designed that attenuate HSA binding. The resulting compound did not lose the primary target (MetAP2) affinity and had better cellular activity.

Isoform selectivity and metabolic clearance finally add additional complexity to the prediction and only QSAR approaches seem to be appropriate. Overall, CYP inhibition modeling guided by X-ray structures for a specific class during lead optimization might be feasible. SOMs, by contrast, are reasonably predicted by reactivity prediction methods [31].

### HSA binding

HSA (human serum albumin) is the most abundant protein in plasma. It is responsible for the binding and transport of many endogenous and exogenous substances such as fatty acids, por-

phyrins and drugs, and it is an important regulator of osmolarity in plasma and interstitial fluid [38].

HSA is a heart-shaped protein composed of 585 amino acids that fold into three domains (I–III), which consist of two subdomains (A and B) each (Fig. 4a). There are seven fatty acid binding sites but most drugs bind to one of two primary binding sites located on subdomains IIA and IIIA (Fig. 4a,b), although there are numerous secondary binding sites. To complicate matters, there are allosteric effects between the different binding sites, for example fatty acids can compete and cooperate with drugs binding to HSA [39].

HSA binding is important for drug distribution within the body. Although bound and unbound drug concentrations are predominantly in equilibrium, problems might occur if the HSA affinity of a drug is extremely high. The unbound drug concentration might be too low to produce the desired pharmacodynamic effect. Because a wealth of structural information is available, structure-based approaches are an appealing way to design the appropriate binding affinity for a drug (Table 1). Successful applications using NMR structural and binding studies were reported by Mao *et al.* for diflunisal analogs [40], by Sheppard *et al.* for MetAP2 inhibitors [41] (Fig. 4c) and by Wendt *et al.* for antagonists of B-cell lymphoma 2 family proteins [42]. In all these studies, it was possible to design analogs with high affinity for the primary target and with significantly reduced HSA binding. The available NMR and X-ray structures of HSA guided the design process (Fig. 4c).

There are few examples where predicted binding modes could be checked later in comparison with the crystallized complexes. Zidovudine (AZT) was placed into drug site 1 using docking and molecular dynamics simulations [43]. When two complex X-ray structures became available, different binding modes for AZT were observed, both diverging from the predicted mode [44]. A closer comparison reveals that AZT in the presence of myristate binds to a new subsite of drug site 1. Flexibility of protein domains and the pronounced side chain movements make a correct prediction very difficult. In addition, most crystallized drugs do not show the typical good fit familiar from other protein–ligand complexes but rather weak interactions with unusual binding poses. For warfarin, a similar comparison can be made and the docking/molecular dynamics results are also not convincing [39,45].

Modifying HSA interactions by structure-based design currently only seems possible if a confirmed binding mode for the lead

compound is known. One has to keep in mind, however, that for some drugs multiple binding modes were observed [44]. For the tuning of properties in lead optimization, detailed knowledge about the HSA binding mode is expected to be highly valuable.

### hERG channel inhibition

The inhibition of the hERG K<sup>+</sup> channel by drugs is different to the previously described off-target effects in a sense that the channel is just accidentally blocked by exogenous substances. hERG is evolutionarily not trained to be part of a defense system against xenobiotic substances.

The long QT syndrome (LQTS; prolongation of the QT interval in the electrocardiogram) caused as a side-effect of non-antiarrhythmic drugs has been implicated as a predisposing factor for torsades de pointes and can cause sudden death. From >20 different ion channels identified in the heart, only the hERG channel (*I<sub>Kr</sub>*) is involved in nearly every clinical case. Therefore, hERG channel inhibition experiments as a surrogate of potential cardiac effects are required now by regulatory authorities and reliable *in silico* predictions of hERG inhibition would be highly appreciated. The following discussion will omit basic simulations on channel gating and voltage sensing itself [46], as well as otherwise reviewed pharmacophore [47,48] and QSAR models [49], and will focus on target-based [50–52] approaches and on medicinal chemistry optimization strategies [53].

Almost all published models rely on the same low-diversity set of <200 mostly high-affinity blockers with the classic motif of a basic nitrogen flanked by aromatic or hydrophobic groups attached by flexible linkers, with binding data from different assay conditions on different cell types. Recent publications underline the importance and show the lack of prediction models for diverse,

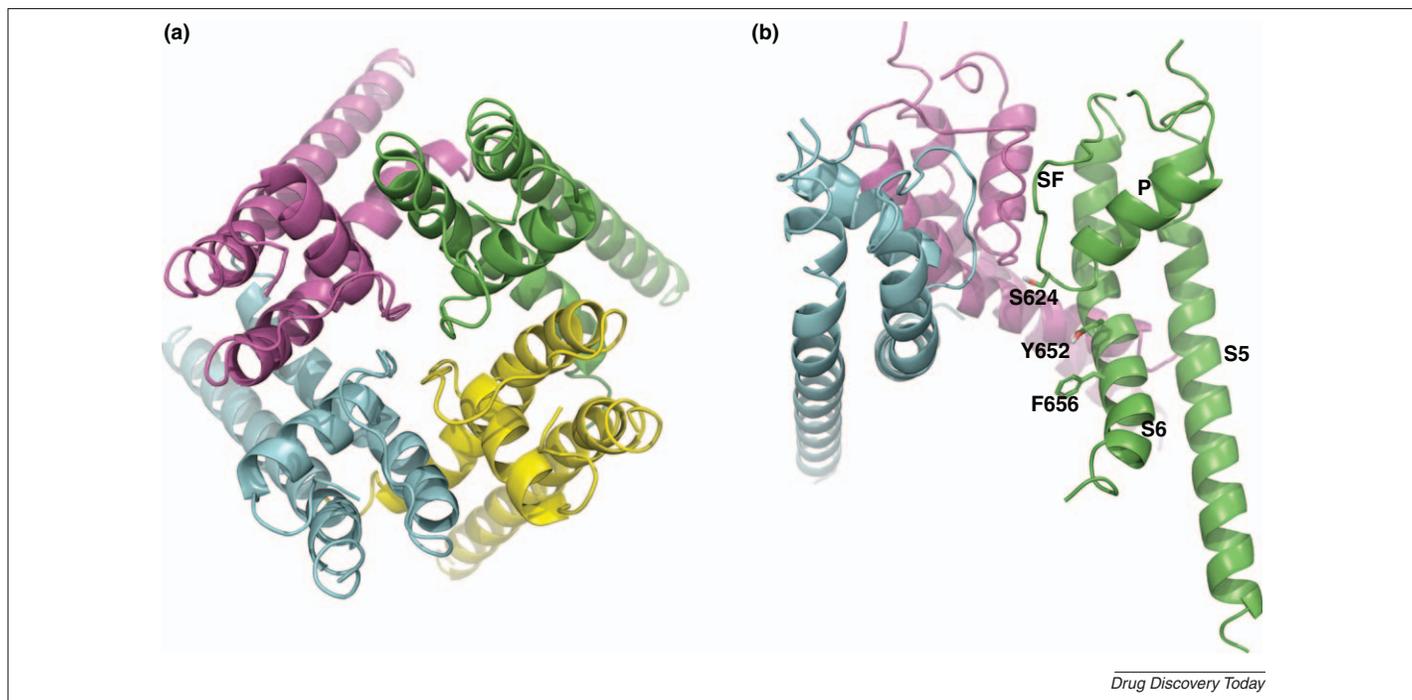


FIGURE 5

(a) Top view from the outer side of the membrane onto hERG channel homology model of helices S5 and S6 with parts of the loop regions missing. (b) Side view with one of the four monomers removed. Indicated by labels are the main ligand-interacting residues Phe656, Tyr652, and Ser624 identified by mutation experiments, P-loop and selectivity filter SF. The homology model [58] was downloaded from [www.schrodinger.com](http://www.schrodinger.com).

neutral, non-classical weak inhibitors [47] with  $IC_{50}$ s between 1 and 20  $\mu$ M [54]. These compounds probably have divergent binding modes or might even bind outside the pore. The potassium channel Kv11.1, the gene product of hERG, is a heterotetrameric protein with six transmembrane helices (S1–S6) in each subunit consisting of 1160 amino acids each. The first four helices comprise the voltage-sensor domain and the remaining two are the pore-domain. Between helices S5 and S6 the turret helix, the pore helix and the P-loop (with the selectivity filter) are located. In the absence of a hERG crystal structure one has to rely on homology models (Fig. 5) based on a fairly low sequence identity to the open or closed bacterial  $K^+$  channels KcsA (closed form), MthK (open) or KvAP (open), KirBac1.1 and mammalian channel Kv1.2 [48]. Whereas alignment of P-loop and S6 is facilitated by some signature sequences, alignments of S5 vary widely [46,50] and the turret helix is fully omitted in most models. Following the work of Mitcheson *et al.* [55], a variety of models on rigid or flexible docking into homology models of the open, closed or partially open forms, induced fit docking, or molecular dynamics were described. Different alignment and refinement methods lead to the key residues such as Tyr652, Phe656 or Ser624 pointing to the ligand or outward to the membrane, depending on the approach. Several model validity measures are applied by the different authors such as fit to SAR [51], matching to experimental mutation data [46,48], conventional geometrical/packing/normality validation methods [50] or molecular dynamics trajectories [56].

Most of the models partially match mutation data but still differ significantly making even a qualitative and descriptive interpretation questionable. Recanatini [46] directly compared docking solutions of sertindole into three differently generated hERG channel homology models in a parallel [56], parallel upside down [57] or a curled orientation to the pore axis [58]. Zachariae *et al.* [51] derive the same conclusions on the SAR of CCR5 (C–C chemokine receptor 5) antagonists or 5HT<sub>A1</sub> agonists based on perpendicular poses as the original authors, who argue with parallel ligand placements, clearly emphasizing that the receptor-based homology models are not predictive but at best qualitative and descriptive.

Empirically, medicinal chemists have developed strategies to overcome hERG liabilities via discrete structural modifications, reducing lipophilicity and the  $pK_a$  of the basic center, or creating

zwitterionic species. Rules for risk assessment have been described [53].

In summary, the best strategy to overcome hERG problems to date is to systematically apply chemical changes that do not hamper the primary target SAR based on the rules by Jamieson *et al.* [53]. Substance-class-specific QSAR models in this respect can drive prospective synthesis planning, and docking or pharmacophore models consistent with hERG inhibition SAR may be enlightening to focus on certain areas of the molecules.

## Conclusions

Although high-resolution structures for HSA, PXR, several CYP enzymes and P-gp have been available for some time now, reported successes in rationally translating this information into compounds with improved ADMET properties are disappointingly rare. It might be speculated that fundamental differences between ADMET-relevant proteins and pharmacological targets exist. Often, ADMET-relevant proteins are evolutionarily optimized to bind and recognize a broad variety of compounds with multiple (e.g. HSA) and/or flexible (HSA, CYPs, PXR, and P-gp) binding sites. Affinities are typically in the  $\mu$ M range. This is in contrast to pharmacological targets that often recognize only very few compounds at well-defined binding sites with high affinity. The weak nonspecific binding properties of ADMET-relevant proteins potentiate inherent problems in docking approaches. Reliable solutions are often not guaranteed, especially if many protein side chains within a binding pocket are defined as flexible or backbone flexibility is essential. This limits the usefulness of those protein structures in a broader sense. However, some successful structure-based applications are reported for HSA and PXR, both being proteins with several high-resolution X-ray or NMR structures and different ligands determined experimentally. In the absence of reliable flexible docking approaches it currently seems mandatory for protein-structure-based predictions of ADMET properties to rely on experimentally determined (X-ray, NMR) binding modes. In addition, structure-based approaches might be combined with QSAR or ligand-based methods. Both approaches are orthogonal to each other in a sense that they rely on completely different data. A well balanced combination of both approaches might offer robust ADMET prediction models for some selected endpoints.

## References

- Hardy, L.W. and Malikayil, A. (2003) The impact of structure-guided drug design on clinical agents. *Curr. Drug Discov.* 12, 15–20
- Congreve, M. *et al.* (2005) Structural biology and drug discovery. *Drug Discov. Today* 10, 895–907
- Ramachandra, M. *et al.* (1998) Human P-glycoprotein exhibits reduced affinity for substrates during a catalytic transition state. *Biochemistry* 37, 5010–5019
- Lam, F.C. *et al.* (2001) beta-Amyloid efflux mediated by p-glycoprotein. *J. Neurochem.* 76, 1121–1128
- Ecker, G.F. *et al.* (2008) Computational models for prediction of interactions with ABC-transporters. *Drug Discov. Today* 13, 311–317
- Li, Y. *et al.* (2010) The structure and functions of P-glycoprotein. *Curr. Med. Chem.* 17, 786–800
- Demel, M.A. *et al.* (2009) Predicting ligand interactions with ABC transporters in ADME. *Chem. Biodivers.* 6, 1960–1969
- Globisch, C. *et al.* (2008) Identification of putative binding sites of P-glycoprotein based on its homology model. *ChemMedChem.* 3, 280–295
- Ravna, A.W. *et al.* (2007) Molecular model of the outward facing state of the human P-glycoprotein (ABCB1), and comparison to a model of the human MRP5 (ABCC5). *Theor. Biol. Med. Model.* 4, 33
- Dawson, R.J. and Locher, K.P. (2007) Structure of the multidrug ABC transporter Sav1866 from *Staphylococcus aureus* in complex with AMP-PNP. *FEBS Lett.* 581, 935–938
- Dawson, R.J. and Locher, K.P. (2006) Structure of a bacterial multidrug ABC transporter. *Nature* 443, 180–185
- Aller, S.G. *et al.* (2009) Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. *Science* 323, 1718–1722
- Seeger, M.A. and van Veen, H.W. (2009) Molecular basis of multidrug transport by ABC transporters. *Biochim. Biophys. Acta* 1794, 725–737
- Pleban, K. *et al.* (2005) P-glycoprotein substrate binding domains are located at the transmembrane domain/transmembrane domain interfaces: a combined photoaffinity labeling-protein homology modeling approach. *Mol. Pharmacol.* 67, 365–374

- 15 Becker, J.P. *et al.* (2009) Molecular models of human P-glycoprotein in two different catalytic states. *BMC Struct. Biol.* 9, 3
- 16 Pajeva, I.K. *et al.* (2009) Combined pharmacophore modeling, docking, and 3D QSAR studies of ABCB1 and ABCG1 transporter inhibitors. *ChemMedChem*, 4, 1883–1896
- 17 Willson, T.M. and Kliewer, S.A. (2002) PXR. CAR and drug metabolism. *Nat. Rev. Drug Discov.* 1, 259–266
- 18 Orans, J. *et al.* (2005) The nuclear xenobiotic receptor pregnane X receptor: recent insights and new challenges. *Mol. Endocrinol.* 19, 2891–2900
- 19 Ekins, S. *et al.* (2009) Challenges predicting ligand–receptor interactions of promiscuous proteins: the nuclear receptor PXR. *PLoS Comput. Biol.* 5, e1000594
- 20 Yasuda, K. *et al.* (2008) A comprehensive in vitro and in silico analysis of antibiotics that activate pregnane X receptor and induce CYP3A4 in liver and intestine. *Drug Metab. Dispos.* 36, 1689–1697
- 21 Lemaire, G. *et al.* (2007) Discovery of a highly active ligand of human pregnane X receptor: a case study from pharmacophore modeling and virtual screening to ‘in vivo’ biological activity. *Mol. Pharmacol.* 72, 572–581
- 22 Ung, C.Y. *et al.* (2007) In silico prediction of pregnane X receptor activators by machine learning approaches. *Mol. Pharmacol.* 71, 158–168
- 23 Jyrkkarinne, J. *et al.* (2008) Insights into ligand-elicited activation of human constitutive androstane receptor based on novel agonists and three-dimensional quantitative structure–activity relationship. *J. Med. Chem.* 51, 7181–7192
- 24 Chrencik, J.E. *et al.* (2005) Structural disorder in the complex of human pregnane X receptor and the macrolide antibiotic rifampicin. *Mol. Endocrinol.* 19, 1125–1134
- 25 Gao, Y.D. *et al.* (2007) Attenuating pregnane X receptor (PXR) activation: a molecular modelling approach. *Xenobiotica* 37, 124–138
- 26 Ekroos, M. and Sjogren, T. (2006) Structural basis for ligand promiscuity in cytochrome P450 3A4. *Proc. Natl. Acad. Sci. USA* 103, 13682–13687
- 27 de Graaf, C. *et al.* (2005) Cytochrome p450 in silico: an integrative modeling approach. *J. Med. Chem.* 48, 2725–2755
- 28 Crivori, P. and Poggesi, I. (2006) Computational approaches for predicting CYP-related metabolism properties in the screening of new drugs. *Eur. J. Med. Chem.* 41, 795–808
- 29 Stjenschantz, E. *et al.* (2008) Computational prediction of drug binding and rationalisation of selectivity towards cytochromes P450. *Exp. Opin. Drug Metab. Toxicol.* 4, 513–527
- 30 Czodrowski, P. *et al.* (2009) Computational approaches to predict drug metabolism. *Exp. Opin. Drug Metab. Toxicol.* 5, 15–27
- 31 Afzelius, L. *et al.* (2007) State-of-the-art tools for computational site of metabolism predictions: comparative analysis, mechanistical insights, and future applications. *Drug Metab. Rev.* 39, 61–86
- 32 Tarcsay, A. *et al.* (2010) Site of metabolism prediction on cytochrome P450 2C9: a knowledge-based docking approach. *J. Comput. Aided Mol. Des.* 24, 399–408
- 33 Seifert, A. *et al.* (2006) Multiple molecular dynamics simulations of human p450 monooxygenase CYP2C9: the molecular basis of substrate binding and regioselectivity toward warfarin. *Proteins* 64, 147–155
- 34 Teixeira, V.H. *et al.* (2010) Analysis of binding modes of ligands to multiple conformations of CYP3A4. *Biochim. Biophys. Acta – Proteins Proteomics* 1804, 2036–2045
- 35 Williams, P.A. *et al.* (2004) Crystal structures of human cytochrome P450 3A4 bound to metyrapone and progesterone. *Science* 305, 683–686
- 36 De Groot, M.J. *et al.* (2009) Understanding CYP2D6 interactions. *Drug Discov. Today* 14, 964–972
- 37 Lill, M.A. *et al.* (2006) Prediction of small-molecule binding to cytochrome P450 3A4: flexible docking combined with multidimensional QSAR. *ChemMedChem*, 1, 73–81
- 38 Peters, T., ed. (1996) *All About Albumin: Biochemistry, Genetics, and Medical Applications*, Academic Press
- 39 Ghuman, J. *et al.* (2005) Structural basis of the drug-binding specificity of human serum albumin. *J. Mol. Biol.* 353, 38–52
- 40 Mao, H. *et al.* (2001) Rational design of diflunisal analogues with reduced affinity for human serum albumin. *J. Am. Chem. Soc.* 123, 10429–10435
- 41 Sheppard, G.S. *et al.* (2006) Discovery and optimization of anthranilic acid sulfonamides as inhibitors of methionine aminopeptidase-2: a structural basis for the reduction of albumin binding. *J. Med. Chem.* 49, 3832–3849
- 42 Wendt, M.D. *et al.* (2006) Discovery and structure-activity relationship of antagonists of B-cell lymphoma 2 family proteins with chemopotential activity in vitro and in vivo. *J. Med. Chem.* 49, 1165–1181
- 43 Quevedo, M.A. *et al.* (2008) Binding to human serum albumin of zidovudine (AZT) and novel AZT derivatives. Experimental and theoretical analyses. *Bioorg. Med. Chem.* 16, 2779–2790
- 44 Zhu, L. *et al.* (2008) A new drug binding subsite on human serum albumin and drug–drug interaction studied by X-ray crystallography. *J. Struct. Biol.* 162, 40–49
- 45 Deeb, O. *et al.* (2010) Exploration of human serum albumin binding sites by docking and molecular dynamics flexible ligand–protein interactions. *Biopolymers* 93, 161–170
- 46 Recanatini, M. *et al.* (2008) Modeling HERG and its interactions with drugs: recent advances in light of current potassium channel simulations. *ChemMedChem*, 3, 523–535
- 47 Aronov, A.M. (2008) Tuning out of hERG. *Curr. Opin Drug Discov. Devel.* 11, 128–140
- 48 Aronov, A.M. (2005) Predictive in silico modeling for hERG channel blockers. *Drug Discov. Today* 10, 149–155
- 49 Aronov, A.M. (2006) Common pharmacophores for uncharged human ether-a-go-go-related gene (hERG) blockers. *J. Med. Chem.* 49, 6917–6921
- 50 Stary, A. *et al.* (2010) Toward a consensus model of the HERG potassium channel. *ChemMedChem*, 5, 455–467
- 51 Zachariae, U. *et al.* (2009) Side chain flexibilities in the human ether-a-go-go related gene potassium channel (hERG) together with matched-pair binding studies suggest a new binding mode for channel blockers. *J. Med. Chem.* 52, 4266–4276
- 52 Du, L. *et al.* (2009) The interactions between hERG potassium channel and blockers. *Curr. Top. Med. Chem.* 9, 330–338
- 53 Jamieson, C. *et al.* (2006) Medicinal chemistry of hERG optimizations: highlights and hang-ups. *J. Med. Chem.* 49, 5029–5046
- 54 Redfern, W.S. *et al.* (2003) Relationships between preclinical cardiac electrophysiology, clinical QT interval prolongation and torsade de pointes for a broad range of drugs: evidence for a provisional safety margin in drug development. *Cardiovasc. Res.* 58, 32–45
- 55 Mitcheson, J.S. *et al.* (2000) A structural basis for drug-induced long QT syndrome. *Proc. Natl. Acad. Sci. USA* 97, 12329–12333
- 56 Stansfeld, P.J. *et al.* (2007) Drug block of the hERG potassium channel: insight from modeling. *Proteins* 68, 568–580
- 57 Osterberg, F. and Aqvist, J. (2005) Exploring blocker binding to a homology model of the open hERG K<sup>+</sup> channel using docking and molecular dynamics methods. *FEBS Lett.* 579, 2939–2944
- 58 Farid, R. *et al.* (2006) New insights about HERG blockade obtained from protein modeling, potential energy mapping, and docking studies. *Bioorg. Med. Chem.* 14, 3160–3173
- 59 Watkins, R.E. *et al.* (2003) Coactivator binding promotes the specific interaction between ligand and the pregnane X receptor. *J. Mol. Biol.* 331, 815–828
- 60 Watkins, R.E. *et al.* (2001) The human nuclear xenobiotic receptor PXR: structural determinants of directed promiscuity. *Science* 292, 2329–2333
- 61 Wester, M.R. *et al.* (2004) The structure of human cytochrome P450 2C9 complexed with flurbiprofen at 2.0-Å resolution. *J. Biol. Chem.* 279, 35630–35637
- 62 Williams, P.A. *et al.* (2003) Crystal structure of human cytochrome P450 2C9 with bound warfarin. *Nature* 424, 464–468
- 63 Sansen, S. *et al.* (2007) Adaptations for the oxidation of polycyclic aromatic hydrocarbons exhibited by the structure of human P450 1A2. *J. Biol. Chem.* 282, 14348–14355
- 64 Xu, R.X. *et al.* (2004) A structural basis for constitutive activity in the human CAR/RXRα heterodimer. *Mol. Cell* 16, 919–928
- 65 Rowland, P. *et al.* (2006) Crystal structure of human cytochrome P450 2D6. *J. Biol. Chem.* 281, 7614–7622
- 66 Sevrioukova, I.F. and Poulos, T.L. (2010) Structure and mechanism of the complex between cytochrome P4503A4 and ritonavir. *Proc. Natl. Acad. Sci. USA* 107, 18422–18427
- 67 Sugio, S. *et al.* (1999) Crystal structure of human serum albumin at 2.5 Å resolution. *Protein Eng.* 12, 439–446