J Comput Aided Mol Des (2012) 26:821–834
DOI 10.1007/s10822-012-9578-6

Multiple ligand docking by Glide: implications for virtual second-site screening

Márton Vass · Ákos Tarcsay · György M. Keserű

Received: 3 December 2011 / Accepted: 1 May 2012 / Published online: 26 May 2012

Abstract Performance of Glide was evaluated in a sequential multiple ligand docking paradigm predicting the binding modes of 129 protein–ligand complexes crystallized with clusters of 2–6 cooperative ligands. Three sampling protocols (single precision—SP, extra precision—XP, and SP without scaling ligand atom radii—SP hard) combined with three different scoring functions (GlideScore, Emodel and Glide Energy) were tested. The effects of ligand number, docking order and druglikeness of ligands and closeness of the binding site were investigated. On average 36 % of all structures were reproduced with RMSDs lower than 2 Å. Correctly docked structures reached 50 % when docking druglike ligands into closed binding sites by the SP hard protocol. Cooperative binding to metabolic and transport proteins can dramatically alter pharmacokinetic parameters of drugs. Analyzing the cytochrome P450 subset the SP hard protocol with Emodel ranking reproduced two-thirds of the structures well. Multiple ligand binding is also exploited by the fragment linking approach in lead discovery settings. The HSP90 subset from real life fragment optimization programs revealed that Glide is able to reproduce the positions of multiple bound fragments if conserved water molecules are considered. These case studies assess the utility of Glide in sequential multiple docking applications.

Electronic supplementary material The online version of this article (doi:10.1007/s10822-012-9578-6) contains supplementary material, which is available to authorized users.

M. Vass · Á. Tarcsay · G. M. Keserű (✉)
Discovery Chemistry, Gedeon Richter Plc., P.O.B. 27, 1475 Budapest, Hungary
e-mail: gy.keseru@richter.hu

Keywords Cooperative binding · Multiple ligand binding · Docking · Molecular modeling · Cytochrome P450 · Fragment screening

Introduction

Cooperativity is a powerful way of Nature to accelerate or regulate specific biological processes [1]. It is encountered on virtually all levels of biochemical complexity, from metal chelation through allosteric modulation and protein folding to communication between cells. Pharmaceutically relevant examples of cooperativity include multiple ligand binding to a single (typically orthosteric) binding site. This situation is typical for proteins that belong to the intricate defense mechanism of the body facilitating the metabolism or efflux of drugs and other xenobiotics. Cytochromes P450 (CYPs) are characteristic examples of these enzymes [2]. Particularly CYP3A4, the isoform responsible for the metabolism of the majority of marketed drugs [3], is involved in cooperative binding. Other metabolic enzymes such as UDP-glucuronosyltransferases (UGTs) [4, 5] and glutathione S-transferases (GSTs) [6] are able to bind ligands cooperatively. ATP-binding cassette (ABC) efflux transporters [7–10] such as the highly promiscuous P-gp can also bind multiple ligands simultaneously. Increasing number of evidences collected in these systems by site-directed mutagenesis experiments [11–13], deuterium isotope effect experiments [14], NMR T_1 paramagnetic relaxation studies [15–18] and the X-ray structure determination of protein–ligand complexes [19, 20] indicate that the binding of the substrate and effector compounds takes place at orthosteric sites. Recently published structures of CYP3A4 in complex with two ketoconazole molecules [21] and that of mouse P-glycoprotein with two cyclic peptide inhibitors bound [22] provided remarkable advances in the field.

The other important example of multiple ligand binding is fragment based drug discovery. The linking strategy involves screening and connecting fragments bound adjacent within the binding site keeping the fragments in their original binding modes [23]. An efficient method to identify fragments suitable for linking has been described, in which a fragment hit from the primary screen is added to the protein in high concentration so that it occupies the primary binding site and a second-site screen is performed to obtain proximally bound fragments. A subnanomolar inhibitor of Bcl-X_L [24] and a low micromolar inhibitor of HSP90 [25] are successful examples of this approach. A virtual analog of this procedure can be implemented by identifying first-site hits from a virtual fragment screening [26] and then docking another library into the predicted complexes of high-scoring ligands that explores further potential sites. Finding a suitable linker and finally docking the resulting compound as a whole to the receptor would then yield druglike hits.

Considering relevant applications of second site screening on both targets and anti-targets and also the increasing number of multiple ligand complexes solved in the last decade (Fig. 1) we concluded that computational prediction of cooperative binding could have significant impact on lead discovery and optimization.

To reach this goal we set out to investigate multiple ligand docking in a more general context. Traditional docking applications treat one site—one ligand interactions. Docking of multiple ligands to a single binding site has not yet been thoroughly investigated. Li et al. have recently implemented a Multiple Ligand Simultaneous Docking (MLSD) strategy into AutoDock4 using Lamarckian genetic algorithm and particle swarm optimization for the simultaneous conformational search of multiple ligands [27, 28]. However, this application has not yet been extensively tested. In this study we describe a simple sequential docking protocol using the commercial docking software Glide for reproducing experimental binding conformations of multiple ligands. A set of 129 X-ray crystal structures was collected from the RCSB Protein Data Bank (PDB) containing at least two non-cofactor type ligands in close proximity. Ligands were docked sequentially to their respective structure in a self-docking setup and the performance of the methodology was then analyzed. Results obtained for the pharmaceutically relevant subset of cytochrome P450 enzymes and structures coming from HSP90 fragment screens were further investigated.

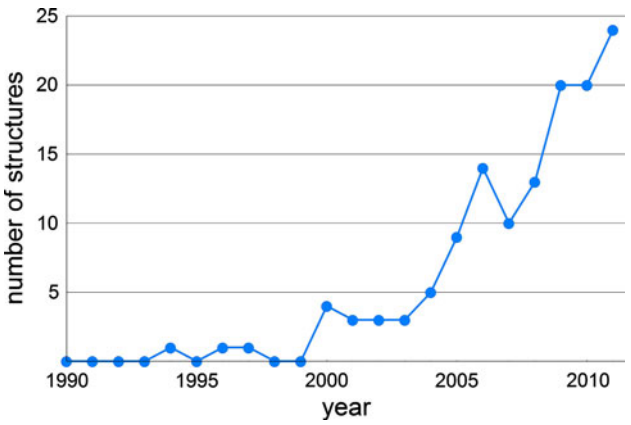

Fig. 1 Number of X-ray crystallographic structures with multiple ligands solved per year

Methods

Compilation of the data set

Crystal structures of protein–ligand complexes were selected from the PDB. Initial filters included a resolution of at least 2.5 Å, protein structures excluding DNA and RNA binding, and no appearance of words associated with photosynthesis or the words MEMBRANE and IMMUNE in the HEADER section of the pdb files. The number of non-covalently bound ligands in each structure was determined excluding water, common cations and anions, common solvents and crystallization agents including PEGs, buffer constituents, lipids, disulfide bond reducing agents etc., cofactors, common carbohydrates and carbohydrate-amines (e.g. NAG), modified residues, and the unknown species (UNL and UNX). Pairwise minimal interatomic distances (MIDs) between the ligands were calculated and clusters with 2–6 ligands having an MID within 6.0 Å were identified. Structures that did not contain cations, anions and cofactors except for heme in the 6.0 Å neighborhood of such a cluster were saved. Docking calculations were performed on clusters having at least two ligand centroids within the $14 \times 14 \times 14$ $Å^3$ cube centered on the common ligand centroid (the inner box in Glide docking). Protein binding sites were characterized using SiteMap [29] version 2.4 in single binding site region evaluation mode with a 6.0 Å buffer around the ligand cluster. Structures without a SiteMap recognized binding site were discarded. The remaining structures were finally visually inspected to eliminate cases with incorrectly defined connectivity or atom types not parametrized in the OPLS-2005 force field. This filtering of the PDB resulted in 129 structures as of 1 October 2011 (see the Supporting Information for PDB ID codes). These structures thus have good resolution and contain a cluster of at least two and at most six ligands in close proximity to each other suitable for docking. Two of these protein–ligand complexes (1e7c and 3g35) had two distinct, non-symmetry equivalent sites

where multiple ligands were present. The docking procedure in these cases was performed for both binding sites increasing the total number of sites to 131 representing 54 targets. These sites contain a total of 294 ligands to be docked, meaning an average of 2.24 ligands per binding site. Binding sites were then classified as sites containing multiple copies of the same ligand and those containing different ligands. Since in the latter case ligands can be docked in different orders the total number of docking calculations was 324.

Structure preparation

The most completely modeled biological assembly in the unit cell was retained from the crystal structures. If the biological assembly contained crystal mates, only chains in the vicinity of the docked ligands were added. In cases where there were more identical chains in the unit cell, the first chain containing the multiply ligated site was selected. Further phases of the work were automated using the Schrödinger Python API available in Schrödinger Suite 2010 (version 3.8). The structures were prepared for docking with the Protein Preparation Wizard [30] that includes assigning bond orders, adding hydrogens, treating metals, creating disulfide bonds, converting selenomethionines, deleting distant waters, assigning the H-bond network with water sampling and finally minimizing the structure up to 0.3 Å RMSD with the OPLS-2005 force field. All waters and ligands were then deleted from the structures before grid generation. In the second case study of HSP90 complexes structure preparation included retaining six conserved crystallographic water molecules in the structures, since these waters are known to play an important role in HSP90 binding [31]. Retained water molecules were selected by aligning the five HSP90 structures and identifying those present in all structures. Orientation of their hydrogens and protonation states of nearby residues were automatically assigned by the Protein Preparation Wizard. They form hydrogen bonds with the Asn51, Asp93 and Trp162 side chains, backbone carbonyls of Leu48 and Gly97, backbone NH of Gly137 and with each other. They all form two or three hydrogen bonds and have a relative B-factor of 0.4-0.9. Docked ligands were prepared by converting them first to 2D structures with the ChemAxon molconvert plugin [32] and then back to 3D with the Schrödinger LigPrep [33] version 2.4 retaining the configuration of the chiral centers. This procedure eliminated the conformational bias of using experimental binding modes. Epik [34–36] version 2.1 was used to generate tautomers and protomers at pH 7 ± 2. For proteins crystallized outside of this range we verified that LigPrep found no additional ligand protomers on the pH of crystallization. Physico-chemical properties, namely molecular weight, logP, number of H-bond donors and acceptors, number of rotatable bonds, molecular volume, hydrophobic and total accessible surface area (ASA) of the ligands were calculated with the ChemAxon cxcalc plugin [37] and their druglikeness was assessed using Lipinski's rule (MW $\leq$ 500 Da, logP $\leq$ 5.0, hydrogen bond acceptor count $\leq$ 10, hydrogen bond donor count $\leq$ 5). SiteMap was used to estimate binding site volumes and enclosure. The enclosure parameter ranges from 0.5 to 1, a higher value means a more buried site. The average value for tight-binding sites is 0.78 [38]. Figure 2 shows the distributions of crystallographic and calculated properties of binding sites and ligands considered in our dataset.

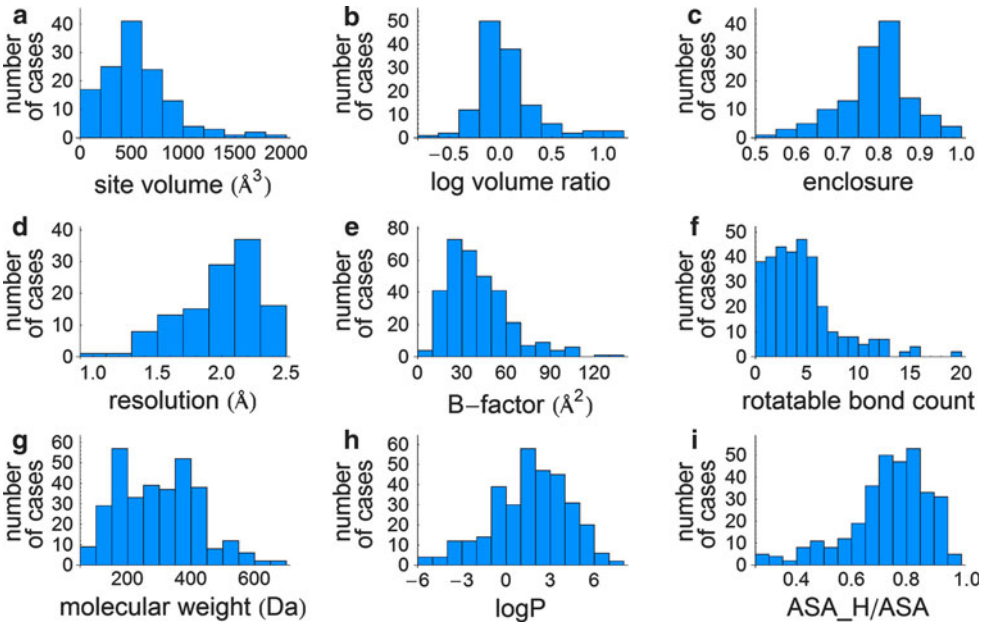

Fig. 2 Distribution histograms of binding site (**a**–**d**) and ligand (**e**–**i**) properties calculated with SiteMap (**a**–**c**), obtained from crystallographic data (**d**–**e**) and calculated with cxcalc (**f**–**i**). Log volume ratio is the logarithm of total ligand volume divided by the site volume to base 10. ASA is the accessible surface area and ASA_H is the hydrophobic accessible surface area

Docking protocol

Docking was performed by Glide 5.6 [39–42] using Single Precision (SP), Extra Precision (XP) and SP hard modes. The latter represents SP calculations without scaling down the van der Waals radii of nonpolar ligand atoms (using a scaling factor of 1.0 instead of the default 0.8). A docking run for a particular structure consisted of at most as many consecutive grid generation and docking steps as the number of ligands in the docked ligand cluster. The maximum available grid size ($36 \times 36 \times 36$ Å^3 outer and $14 \times 14 \times 14$ Å^3 inner box) was used. The grid was always centered on the common centroid of all the heavy atoms of the ligand cluster that allowed positioning the grid in the same way in each step of a docking run. The number of poses entered to post-docking minimization and saved was set to 30, other sampling parameters were set to their default. RMSDs between docked and experimental ligand conformations were calculated for heavy atom positions.

The sequential docking protocol used in this study contains multiple grid generations depending on the number of ligands evaluated for a particular binding site. The first grid was generated for the ligand-free receptor. In the first docking step each tautomer and protomer of the first ligand was docked. If different ligands were bound, both docking permutations were evaluated. The 30 saved poses were ranked by their respective Emodel scores. If any of the top three poses of any protomer or tautomer had an RMSD less than 2.0 Å to any of the experimental ligand conformations, that pose was selected and merged with the ligand-free starting structure (first poses had priority over second ones and second poses had priority over third ones; lower RMSDs had priority among different protomers and tautomers). If none of the top three poses had an RMSD less than 2.0 Å, but a higher ranked pose had, then the pose with the minimal RMSD was merged with the ligand-free starting structure, and the rank of the merged pose was recorded. If no poses with RMSD less than 2.0 Å were found, then the docking run was terminated since no satisfactory pose was generated.

In the second step a new grid was generated for the merged structure and all protomers and tautomers of the second ligand were docked. The pose to be merged with the one-ligand containing structure was selected the same way as in the first step and its rank was recorded. If the cluster contained more than two ligands, this procedure was repeated until all ligands were docked or the run was terminated because of no satisfactory RMSDs. After the docking run the ranks of good poses were also calculated using GlideScore and Glide Energy scoring functions by reordering the 30 saved poses for all protomers and tautomers (top three poses if any had an RMSD less than 2.0 Å, otherwise the pose with minimal RMSD). Our workflow is shown in Fig. 3.

Data analysis

RMSD values and Emodel, GlideScore and Glide Energy ranks of the merged poses were assigned. RMSDs of the first docking steps were analyzed in order to compare sequential multiple ligand docking to single ligand docking benchmark studies. Median and mean RMSDs and standard deviations were calculated for top scoring and lowest RMSD poses. Success rates are defined as the ratio of ligands where any pose had an RMSD less than 2.0 Å. Top pose rates were calculated as the ratio where the top pose had an RMSD less than 2.0 Å. RMSD and rank distributions of all docking steps for the different docking modes were subjected to Kolmogorov–Smirnov statistical analysis using STATISTICA 10.

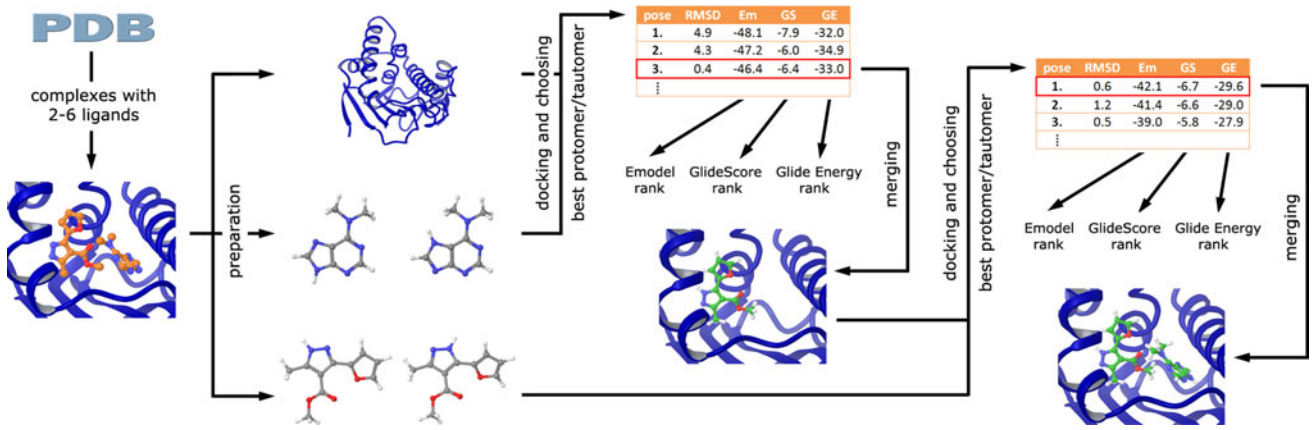

Fig. 3 The sequential docking workflow used in this study

Results and discussion

Docking performance for the first ligand

Sampling efficiency of docking programs is most often assessed by RMSD values calculated between docking poses and experimental binding modes. The RMSD distribution of top scoring poses and ranks of the best poses are informative of scoring performance. Evaluating the sampling efficiency of Glide in multiple ligand docking we first calculated the median and mean RMSDs for the top scoring and the best pose of the first ligands and compared the results to those obtained in single ligand docking for 68 high resolution X-ray complexes [43] (see Table 1). We found that median and mean RMSDs of top poses were typically higher in our data set than those for single ligand docking as expected for docking ligands significantly smaller than their respective binding sites. Although XP mode is usually considered being more precise, the difference in top pose median RMSDs is smaller for the SP protocol than that for the XP (0.09 and 0.75 Å, respectively). Best pose median RMSDs were, however, somewhat smaller than those of Cross et al. (0.18 and 0.03 Å for SP and XP, respectively). This might not be specific to multiple ligand docking but might even be attributed to the difference in Glide versions. Standard deviations for the SP protocol are pretty much similar to those reported for single ligand docking, while for XP we obtained significantly higher values (with the difference of 0.8 Å and 1.0 Å for top poses and best poses, respectively). SP and SP hard protocols performed similarly when docking the first ligand. Although the top pose rates (the percentage of top poses within 2.0 Å RMSD relative to the experimental binding mode) reported by Cross et al. were identical for SP and XP protocols (69.1 %) [43], for our dataset we found that SP based protocols outperformed the XP (top pose rates were 58.0, 51.1 and 58.0 %, for SP, XP and SP hard protocols, respectively). Other parameters such as the resolution of the X-ray structure, relative B-factors (the ligand's mean B-factor divided by the whole structure's mean B-factor) and physico-chemical descriptors of the ligands showed virtually no impact on docking accuracy (data not shown).

Docking performance for the next ligands

Our objective in this step was the evaluation of the sequential docking methodology. To achieve this goal the number of successful docking steps (defined by any docking pose within 2.0 Å RMSD) among the 30 saved poses and the impact of the number of consecutively docked ligands was investigated. Analysis of the expectation values revealed that the successive docking of more than two ligands is highly unlikely to give reliable results (Table 2). The calculated values were 1.37, 1.03 and 1.40 for SP, XP and SP hard protocols, respectively, reflecting

Table 1 Statistical results for docking the first ligands to multiple ligand binding sites and for single ligand docking

RMSD	This study			Cross et al. [43]		
	Median (Å)	Mean (Å)	SD (Å)	Median (Å)	Mean (Å)	SD (Å)
SP top pose	1.31	2.42	2.49	1.22	2.08	2.49
SP best pose	0.54	1.23	1.50	0.72	1.30	1.40
XP top pose	1.90	2.96	2.93	1.15	1.97	2.13
XP best pose	0.76	2.16	2.56	0.79	1.34	1.53
SP hard top pose	1.34	2.40	2.48	–	–	–
SP hard best pose	0.49	1.29	1.66	–	–	–

Table 2 Number of successfully docked ligands (RMSD < 2.0 Å) per binding site using SP, XP and SP hard protocols as the function of the number of ligands

Number of ligands at the site	Number of sites	Number of docking runs with n successful consecutive docking steps											
		SP				XP				SP hard			
		n = 0	n = 1	n = 2	n = 3	n = 0	n = 1	n = 2	n = 3	n = 0	n = 1	n = 2	n = 3
2	110	20	28	62	-	37	31	42	–	23	25	62	–
3	13	4	2	3	4	6	4	3	0	3	1	4	5
4	6	3	1	1	1	1	2	2	1	2	1	2	1
5	1	0	0	1	0	0	1	0	0	0	0	1	0
6	1	1	0	0	0	1	0	0	0	1	0	0	0
Total	131	28	31	67	5	45	38	47	1	29	27	69	6

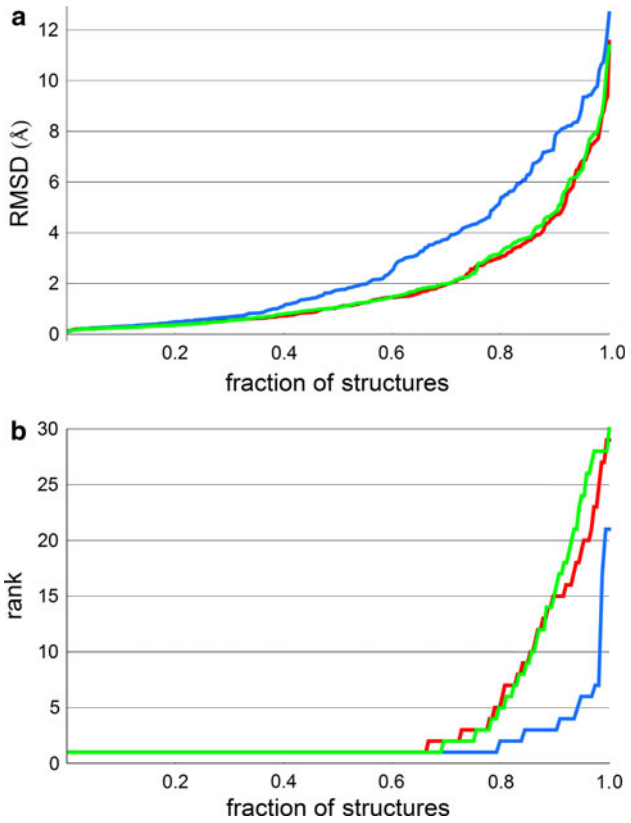

Fig. 4 RMSD (**a**) and rank distributions (**b**) using the Emodel scoring function of all docking steps obtained by SP (*red*), XP (*blue*) and SP hard (*green*) protocols

that the chance to recover at least two experimental binding conformations is 55, 37 and 57 %, respectively. In conjunction to that observed for docking the first ligand, we found that SP and SP hard performed similar and much better relative to the XP protocol in multiple docking situations.

RMSD distributions for all docking steps and rank distributions for all successful (RMSD < 2.0 Å) steps were analyzed by Kolmogorov–Smirnov statistics (Fig. 4). RMSD distributions were significantly higher for the XP protocol as compared to that of SP and SP hard ($p < 0.005$). Rank distributions obtained by the default Emodel scoring function showed the reverse trend; XP provided significantly lower ranks relative to the other two protocols ($p < 0.05$), while SP and SP hard were similar. Rank distributions obtained by the three different scoring functions (Emodel, GlideScore and Glide Energy) showed no statistically significant differences in any of the docking protocols (SP, XP, SP hard). In successful docking steps RMSDs were usually similar or lower for XP than that of the SP. However, it was interesting to see that the higher the rank of the selected pose (typically if larger than the 6th), the more challenging successful docking was for XP. Thus, it seems that XP typically sorts out high-scoring poses only, and is unable to score better than SP in docking multiple ligands. RMSDs of the second docking steps were somewhat higher than those of the first docking steps, however, their distributions were not significantly different from that of the first steps. Median RMSDs of second docking steps were 0.26 Å, 0.59 Å and 0.28 Å higher than first ones for SP, XP and SP hard, respectively. We found that if the first ligand was docked successfully (RMSD ≤ 2.0 Å), then the second docking step was also successful in 70 % of the cases. If the first RMSD was over 2.0 Å, 86 % of the second ligands also failed to dock successfully (in the SP protocol when the docking run was not terminated after an RMSD > 2.0 Å).

Performance on druglike ligands and closed sites

Finally, the impact of ligand and binding site properties were investigated. Considering that scoring functions were typically optimized against a set of known binders of pharmaceutical targets and decoys with drug- and lead-like features [40, 41], we expected that the multiple docking of druglike ligands would be more efficient. This hypothesis was tested by ligands which did not violate any of Lipinski's rules providing a 110 membered druglike subset. Using the SP protocol with Emodel scoring the success rate calculated in at least two consecutive and successful (RMSD < 2.0 Å) docking steps increased from 55 to 59 % (see Fig. 5, all numerical data are available in the Supporting Information). The number of cases where the selected pose was actually the top ranking pose (top pose rate) also increased from 31 to 33 %. Thus we concluded that druglike ligands performed only slightly better in multiple docking settings. XP and SP hard protocols showed similar characteristics as for the whole dataset. The success rates increased from 37–40 % to 57–61 %, respectively, while only marginal improvements were seen in top pose rates (from 25–26 % to from 34–35 %, respectively). Success rates and top pose rates obtained by all the protocols with GlideScore and Glide Energy score were slightly lower.

We also expected that docking to shallow and open binding sites would be more challenging than to the closed sites. Since in the former case ligands can exploit less binding interactions, scoring would be more difficult. Binding sites were therefore classified using their enclosure values calculated by SiteMap as 48 open and 83 closed sites. Out of the 83 closed sites 53 (64 %) provided a docking pose within 2.0 Å RMSD relative to the experimental binding conformation in at least two successive steps. Docking to 29 out of the 48 open sites (60 %), however, was unsuccessful yielding poses with RMSD larger than 2.0 Å. Finally combining the enclosure and druglike filters we investigated docking druglike ligands to

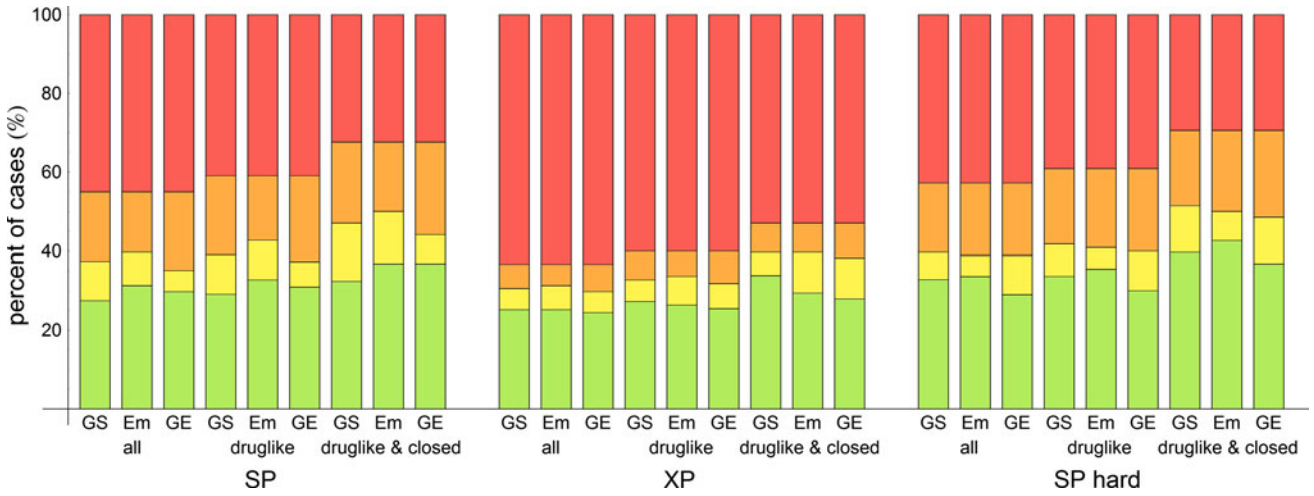

Fig. 5 Cumulative success rates obtained in at least two successful consecutive docking steps for SP, XP and SP hard protocols combined with GlideScore, Emodel and Glide Energy scoring functions and filters (all sites, sites with druglike ligands and closed sites with druglike ligands). From *bottom* to *top*: fraction of cases where both of the top ranking poses had RMSD < 2.0 Å (*green*), where any of the top three poses had RMSD < 2.0 Å (*green* + *yellow*), where any of the poses had RMSD < 2.0 Å (*green* + *yellow* + *orange*) and all structures in the subset (*green* + *yellow* + *orange* + *red*)

closed sites. Evaluating at least two consecutive docking steps in this subset of 68 complexes we obtained the success rates of 68, 47 and 71 % for the SP, XP and SP hard protocols with Emodel scoring, respectively. Top pose rates were 37, 29 and 43 %, respectively.

Independent analysis of each docking step required the selection of a near native binding mode for the first ligand that mimics the best case scenario in a real life screening situation. Since typically this information is not available in prospective applications the binding mode of the first ligand should be selected based on docking scores to merge with the receptor for the next round of docking. The simplest method is to merge the top pose identified by a single scoring function. Top pose rates calculated in two consecutive steps for our dataset show the usefulness of this method. Thus, using the SP hard protocol with Emodel scoring on druglike ligands and closed binding sites binding modes obtained for 43 % of the compounds are expected to be accurate. A computationally more demanding method would be to merge multiple (N) poses of the first docked ligand with the receptor and perform the second docking step on all of the acquired complexes saving multiple (N) poses again. However, in this method the selection of the best complex from the resulting $N \times N$ pool is not trivial. Instead, we evaluated the performance for $N = 3$ by considering well-docked poses among the top three binding modes. This estimates the upper limit of the performance since the second docking steps usually failed when the first merged pose had an RMSD over 2.0 Å. Selecting the well-docked poses from a 30×30 pool would provide significant improvement on the expense of 31 grid generation and docking steps.

Case study 1: Application to cooperative CYP binding

Cytochromes P450 (CYPs) are the most studied promiscuous metabolic enzymes with the ability to bind multiple ligands in their active site as demonstrated by X-ray crystallography [19, 21, 44–49]. In vitro CYP assays are generally used to predict in vivo pharmacokinetic properties and drug interaction potential of compounds. Determination of binding constants, however, is sometimes not straightforward as these assays often show non-Michaelis–Menten kinetic profiles [50] indicating the cooperative binding of substrates. There is a large body of evidence published on cooperative binding to CYP3A4 [51] and CYP2C9 [52] but similar findings have been reported for CYP2A6 [14], CYP1A2 [53] isoforms and the bacterial CYPeryF [54] as well. Heterotropic cooperativity in these isozymes may lie in the background of drug–drug interactions, though only a few studies were able to connect in vitro and in vivo observations [55–58]. Predicting drug interactions or metabolic activation by computational methods is a challenging task since metabolic enzymes usually have broad substrate specificities and their heteroactivation profile is substrate dependent. Up to now only two CYP3A4 and CYP2C9 heteroactivator pharmacophore models [59, 60] have been published and a single structure based docking method was developed to predict activators of CYP2C9-mediated flurbiprofen hydroxylation [61].

Since our data set contains seven CYP structures it was straightforward to evaluate the performance of our sequential docking protocol on this set of pharmaceutically relevant enzymes. Two further CYP structures (human CYP2C8 with two bound retinoic acids with a resolution of

2.60 Å and human CYP3A4 with two bound ketoconazole molecules with a resolution of 3.80 Å) were added to this assembly as they fulfilled all but the resolution criterion, that were set up in the compilation of our data set. The resolution of the latter is higher than typically required for docking but we used the structure since double occupancy is evident from the electron density map, however, the binding mode of the distal ligand may be ambiguous. Our final CYP dataset composed six bacterial, a rabbit and two human CYP isoforms.

Selected crystallographic properties and docking results are summarized in Table 3. It can be seen that except for the triple occupancy 3g5n structure, all other CYP sites are characterized as closed sites due to high enclosure values. Calculated site volumes confirm that bacterial CYP isoforms comprise much more compact sites, while mammalian CYPs exhibit more spacious ones that can accommodate compounds of various sizes. The ligands present in the bacterial and the rabbit 2B4 isoforms fulfill Lipinski's rule of druglikeness, while in human CYPs a logP value of 5.01 was calculated for retinoic acid and ketoconazole has a molecular weight of 531 Da, both rather close to corresponding Lipinski limits. Most B-factors of the ligands are relatively low indicating that the experimental binding modes are realistic. Based on our previous experience these observations suggested good docking results. In fact, SP and SP hard protocols with Emodel scoring provided successful poses in all cases but one that was the third ligand in the only triple ligand complex. Again, the performance of the XP protocol was inferior; 6 of the 19 docking steps were unsuccessful. SP provided 6 of the 9 structures with good poses found among the top three poses in both steps, while XP performed similarly for only 3 complexes.

The aminophenantrene ligand coordinating the heme iron in 1egy was docked perfectly as the top pose by both SP and SP hard protocols. The top pose of XP had the aromatic rings flipped but occupied almost the same space. The second ligand made apolar contacts with the aromatic and aliphatic side chains of the active site and its amino group was not involved in any hydrogen bonds in the crystal structure. Interestingly, Glide identified a hydrogen bond in most of the top ranked poses to the hydroxyl group of Tyr75 or to the backbone carbonyl of Phe86. As a result a flipped pose or even an irrelevant binding mode was predicted.

The proximal androstenedione in 1eup forms a hydrogen bond with both of its carbonyl groups. One of them is with Asn89 that was lost in the docking runs due to flipped asparagine side chain created by the Protein Preparation Wizard when optimizing the H-bond network. Despite this Glide identified the experimental binding mode as top.

Table 3 Crystallographic and docking data of cytochrome P450 structures with multiple bound ligands

PDB ID	CYP isoform	res. (Å)	vol. (Å^3)	encl.	Ligands	lig. ID	B-fact. (Å^2)	RMSD (Å) SP	Em	RMSD (Å) XP	Em	RMSD (Å) SP hard	Em
1egy	107	2.35	489.8	0.931	9-Aminophenantrene	1	25.73	0.19	1	0.18	3	0.20	1
						2	35.40	0.36	3	0.32	1	0.33	2
1eup	107	2.10	615.0	0.910	Androstenedione	1	29.95	0.57	1	0.68	4	0.45	1
						2	47.27	0.73	1	0.62	1	0.93	1
2whf	130	1.58	644.8	0.759	1-(3-Methylphenyl)-1H-benzimidazol-5-amine	1	43.67	0.77	2	>2.00	-	1.97	2
						2	38.02	0.68	5	>2.00	-	1.25	1
2d0e	158	2.15	562.5	0.908	2-Hydroxynaphthoquinone	1	47.45	0.50	3	0.76	3	0.25	12
						2	71.84	0.42	1	>2.00	-	0.35	1
1t93	158	1.62	358.1	0.948	Flaviolin	1	16.50	0.48	3	0.48	4	0.45	2
						2	16.22	0.52	9	>2.00	-	0.27	8
2z3u	245	2.40	353.3	0.905	Chromopyrrolic acid	1	14.79	0.20	1	0.22	1	0.18	1
						2	33.02	0.40	1	0.40	1	0.34	1
3g5n	2B4	2.50	633.1	0.672	1-(Biphenyl-4-ylmethyl)-1H-imidazole	1	64.48	0.25	1	0.53	1	0.94	1
						2	40.36	0.48	1	>2.00	-	0.41	1
						3	73.21	>2.00	-	>2.00	-	>2.00	-
2nnh	2C8	2.60	772.1	0.869	Retinoic acid	1	57.23	0.49	1	0.39	1	0.40	1
						2	54.19	0.52	14	1.94	1	0.55	1
2v0m	3A4	3.80	1247.1	0.808	Ketoconazole	1	34.35	0.89	1	0.86	1	1.17	1
						2	71.84	1.94	2	1.20	6	1.54	30

res. = resolution, vol. = site volume calculated by SiteMap, encl. = enclosure calculated by SiteMap, lig. ID = ligand number, B-fact. = mean B-factor of the ligand, Em = rank by Emodel

Many misdocked poses featured a perpendicular orientation of the androstenedione molecule to the heme probably because of the overemphasized electrostatic interactions between the heme iron and one of the carbonyl groups. The distal ligand was well-docked with all docking protocols.

In 2whf the scoring function ranked a flipped iron coordinating pose to the top thus rendering the well-docked distal pose second. It was interesting to see that the distal ligand could penetrate deeper into the active site with its aromatic end, retaining only one of its hydrogen bonds that resulted in a high RMSD (see Fig. 6). Surprisingly no heme iron coordinating pose could be found with the XP protocol.

2d0e and 1t93 are structures of the same isoform cocrystallized with very similar ligands, thus the two binding sites are nearly identical. Flaviolin and hydroxynaphthoquinone molecules form multiple hydrogen bonds. Active site Arg288 anchors the ligands in both experimental and docked poses in addition to aromatic stacking with the heme and each other. Interestingly this was not preserved in the top ranking poses of the first docking steps instead a binding mode with three hydrogen bonds was found for both compounds. For the second docked flaviolin molecule the stacking interaction was captured well but poses involved in more hydrogen bonds were still enforced for hydroxynaphthoquinone even at the expense of distorting the planarity of its rings.

In 2z3u the two chromopyrrolic acid residues (the natural substrate of CYPStaP) are held very firmly by multiple hydrogen bonds, π-π and cation-π interactions that were predicted successfully as the top ranked pose in all docking protocols.

3g5n was one of the cases where the inner grid box contained only two out of the three ligand centroids. Two molecules were correctly docked by SP based protocols (see Fig. 6). The semi-distal ligand forms contacts mostly

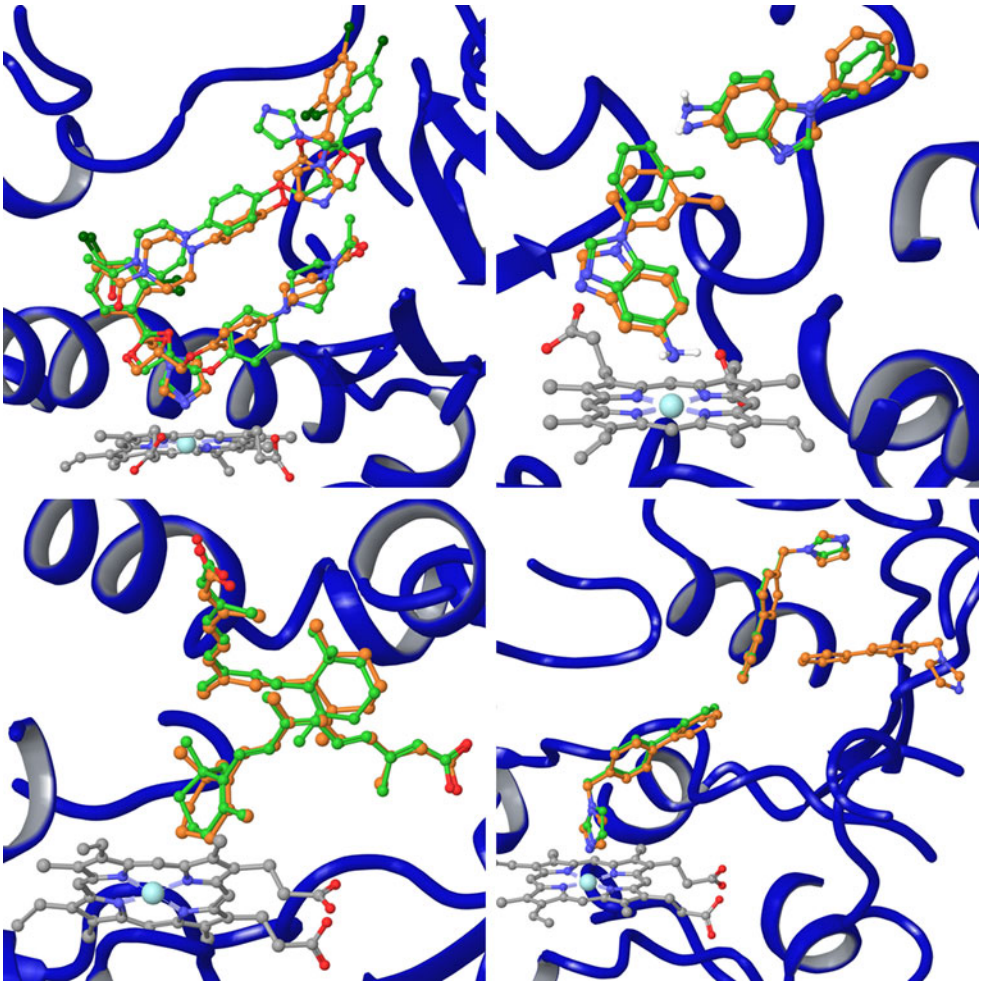

Fig. 6 Representative binding modes found in cytochrome P450 complexes with multiple ligands by the SP protocol and Emodel scoring. *Top left* 2v0m, *top right* 2whf, *bottom left* 2nnh, *bottom right* 3g5n. Heme carbon, docked ligand carbon, docked ligand polar hydrogen, co-crystallized ligand carbon, oxygen, nitrogen, chlorine and iron atoms are *colored grey*, *green*, *white*, *orange*, *red*, *blue*, *dark green* and *cyan*, respectively

with aromatic and aliphatic side chains that were correctly identified in the first docking step. Its imidazole part is encased in a polar environment but is not involved in specific interactions. The iron coordinating binding mode was found second by the SP protocols, but XP missed the iron coordinating pose again. The binding conformation of the third ligand, which is partially exposed to the solvent with its imidazole ring, could not be reproduced by either of the protocols. Instead poses exhibiting aromatic stacking with the iron coordinating ligand and involved in a hydrogen bond with the amide hydrogen of Gly99 were obtained.

The retinoic acid molecules in 2nnh are both involved in two strong hydrogen bonds with the backbone amides of Gly98 and Ser100 for the proximal and with Asn204 and Arg241 for the distal ligand. These interactions were recovered in all docked poses that resulted in good ranks in the first docking step. On the other hand, however, the ring of the distal ligand was flipped in many poses in the second docking step that was responsible for RMSDs greater than 2.0 Å, though the binding motif is essentially the same (see Fig. 6).

The binding mode of the proximal ketoconazole molecule in 2v0m was remarkably well reproduced with all docking protocols. The acetylpiperazine moiety of the distal ligand was positioned well in the second step; however, the dichlorophenyl and imidazole rings were typically misdocked that increased the RMSDs significantly (see Fig. 6). The chlorine atom of the distal ligand forms interactions with the backbone amide of Leu216 and the aromatic ring of Phe213 while its imidazole ring faces a polar environment in the crystal structure. These interactions were replaced by a hydrogen bond between the backbone of Asp217 and the rotated imidazole ring in many of the docked poses resulting high ranks in the second step. It should be noted, however, that the electron density at the head region of the distal ligand was poorly defined [21] and the experimental binding mode might also be questionable.

Case study 2: Application to virtual fragment screening against HSP90

One of the most important drug discovery applications of multiple ligand docking would be fragment based hit identification. Our approach provides the virtual equivalent of experimental second site screening in order to find fragments suitable for linking. Demonstarting this capability we picked up five complexes of the heat shock protein 90-alpha (HSP90) from our dataset that were obtained in fragment based discovery programs [25, 31, 62, 63]. Each of these complexes contains two different fragments mostly with low B-factors and fulfilling Lipinski's criteria by definition. The binding site of the protein is quite small and closed as indicated by the corresponding enclosure values. These features implicated successful docking but high RMSDs and ranks were obtained. With the exception of the 2xdu complex, protocols were able to reproduce experimental binding modes, but poses within 2.0 Å RMSD were typically not ranked into the top three poses (the average rank was 8.3). However, structural waters are known to play an important role in ligand binding to HSP90 [31]. Visual inspection indeed revealed that top ranked high RMSD poses overlap with conserved crystal water molecules that were omitted by the Protein Preparation Wizard in our original protocol.

The total of six conserved water molecules were identified that form hydrogen bonds with the Asn51, Asp93 and Trp162 side chains, backbone carbonyls of Leu48 and Gly97, backbone NH of Gly137 and with each other as described in the methods section. Repeating the docking runs with the six conserved waters included yielded top ranked poses with RMSDs less than 2.0 Å for all of the docking steps in the SP protocol, and most of the docking steps in the XP and SP hard runs. Selected crystallographic properties and docking results of this protocol are summarized in Table 4. The binding site volumes calculated with SiteMap decreased by 60-110 $Å^3$ while enclosure values did not change significantly. The better ranks of the well-docked poses, however, are probably not only the results of the decreased site volumes and different binding site shapes. Instead without waters included, docked fragments interact with the same side chains of the receptor that form hydrogen bonds in the water mediated complex. Docking the phenylaminofuranone first to 2qfo without waters the top pose forms direct hydrogen bonds with Asp93 and Asn51 taking the place of the pyrimidinamine moiety. In 3hz1 and 2yej top poses of all ligands form direct hydrogen bonds with the carbonyl of Leu103 without waters while in the crystal structures the pyrazole ring forms a water mediated bridge with the same amino acid. In top poses identified in 2yei direct hydrogen bonds to Leu103 and Trp162 were found witout considering structural waters. Thus, the region around the Leu103 carbonyl group is a docking hot spot of HSP90. Examples of docking results are shown in Fig. 7. In this system the SP protocol with Emodel ranking proved to be the best in binding mode reproduction but even the XP and SP hard protocols failed at the ranking of only one well-docked pose.

Conclusions

The objective of this work was to evaluate the performance of a simple sequential docking methodology in reproducing

Table 4 Crystallographic and docking data of HSP90 complexes with multiple fragments bound

PDB ID	res. (Å)	vol. (Å^3)	vol. with waters (Å^3)	encl.	lig ID	B-fact. (Å^2)	RMSD (Å) SP	Em	RMSD (Å) XP	Em	RMSD (Å) SP hard	Em
2qfo	1.68	388.3	324.5	0.761	1	15.95	0.38	1	0.38	1	0.43	1
					2	17.41	0.16	1	0.16	1	0.26	1
3hz1	2.30	476.1	364.6	0.795	1	25.84	0.46	1	1.10	1	0.29	1
					2	18.53	0.47	1	1.25	1	0.38	1
2xdu	1.74	411.9	350.9	0.811	1	24.60	0.21	1	0.21	1	0.20	1
					2	33.47	0.29	1	0.30	1	0.24	1
2yei	2.20	418.8	362.2	0.800	1	47.47	1.04	1	0.28	1	1.14	3
					2	42.31	0.90	1	1.13	1	0.95	1
2yej	2.20	447.3	346.8	0.779	1	53.74	1.02	1	1.02	1	1.04	1
					2	82.07	0.42	1	0.40	2	0.25	1

Conserved waters were included in all docking runs. res. = resolution, vol. = site volume calculated by SiteMap, encl. = enclosure calculated by SiteMap, lig. ID = ligand number, B-fact. = mean B-factor of the ligand, Em = rank by Emodel

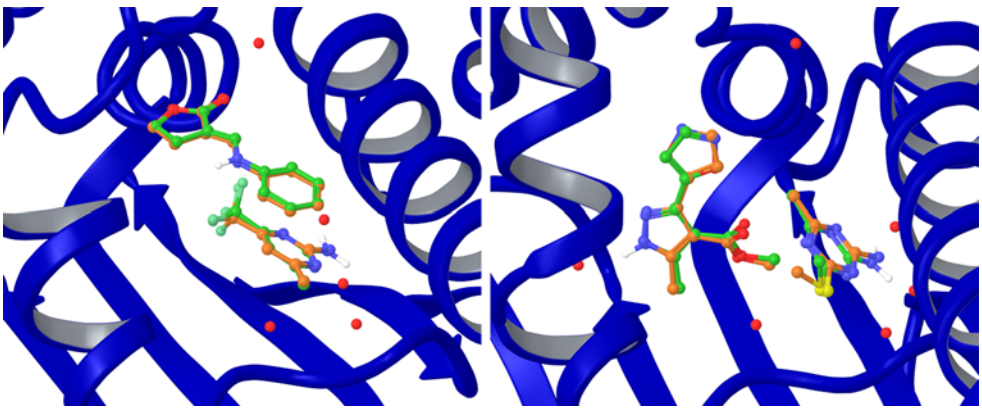

Fig. 7 Representative binding modes of ligands in HSP90 complexes obtained with the SP protocol using Emodel scoring with conserved waters included. *Left* 2qfo, *right* 2yej. Docked ligand carbon, docked ligand polar hydrogen, co-crystallized ligand carbon, oxygen, nitrogen, sulfur and fluorine atoms are *colored green*, *white*, *orange*, *red*, *blue*, *yellow* and *turquoise* respectively. Only the oxygen atoms of waters are shown

experimental binding modes of ligands in higher stoichiometry protein–ligand complexes. This phenomenon has relevance to the linking strategy of fragment evolution in fragment based drug discovery and to the study of drug–drug interactions, which are frequently mediated by metabolic enzymes or transporters that can bind multiple ligands in their active site. The performance of Glide in docking multiple ligands to their native binding sites was evaluated on a set of 129 protein–ligand complexes from the PDB. Three different docking protocols were tested: default single (SP) and extra precision (XP), and single precision without scaling down the van der Waals radii of nonpolar ligand atoms (SP hard). Each of the protocols was used in conjunction with three scoring functions: GlideScore, Emodel and Glide Energy.

It was seen that docking to multiply ligated binding sites is more difficult than single ligand docking indicated by higher top pose RMSDs and lower success rates for all protocols. Docking of the second ligand fails with a chance of 86 % if the RMSD of the first docked ligand is over 2.0 Å, thus a well-docked first ligand is a prerequisite to dock the second ligand successfully. Docking more than two ligands with such a sequential protocol seems to be even more challenging. SP and SP hard protocols did not provide significant difference in RMSD and rank distributions. In contrast, XP yielded significantly lower success rates and lower ranks. This indicates that the sampling algorithm in XP usually doesn't find different poses than that of the SP protocol. More likely it is able to separate reasonable poses from false solutions but it often rejects good poses as well. Though excessive differences between the performances of the different scoring functions were not discovered, Emodel gave somewhat higher ratios of top ranked well-docked poses in the SP and SP hard protocols than the two other functions. For the XP protocol, however, GlideScore provided the highest success rates.

The fractions of structures with at least two well-docked ligands ranked among the top three poses, which is a preferred scenario when docking compounds with unknown binding modes, were 40 % for the SP protocol, 31 % for the XP protocol and 39 % for the SP hard protocol with Emodel scoring. Specific subsets of the structures were examined to find criteria for a higher success rate of the methodology. Success rates increased both for the subset of closed sites with enclosure values greater than 0.78, and for sites containing only Lipinski compliant ligands. Sites fulfilling both criteria provided ratios of 50, 40 and 50 % using SP, XP and SP hard protocols with Emodel scoring, respectively. The highest ratios of successful docking runs are encountered with the SP hard protocol. On the other hand, the fact that XP gave the lowest ratios and it rarely produced well-docked poses when SP did not find one either, is likely the consequence of its sampling algorithm, which uses specific parts of the ligand from the SP poses as starting cores [42], thus not being able to sample substantially different binding modes from those found by SP.

Further sources of high RMSDs in docking multiple ligands may be the reward and penalty terms of the scoring functions for filling hydrophobic pockets by hydrophobic groups, or inadequate solvation of groups forming hydrogen bonds. During the first docking step the ligand may partially occupy the space or even important interaction points needed for the binding of the other ligand(s), because those contacts are scored more favorable than the ones formed in the multiply ligated structure. The most common errors in the docking runs identified by inspection of the top poses and experimental binding conformations conform to this hypothesis. In many cases the first docked pose either occupied such specific interaction points of the other ligand, or if it was mostly lipophilic—it appeared to maximize its apolar contact surface with the binding site. A special case of the latter was when two planar ligands were aligned parallel in the X-ray structure (e.g. in 1eb9, 2cbt, 2vq5, 3e85, 3etg and 3g35), and the docked poses were also parallel to each other but different to the experimental binding modes, the two ligands together filling essentially the same space. Either way, the second ligand was partially excluded from the place it should have been docked into, which resulted in misdocked poses and high RMSDs in the docking steps after the first one. A possible remedy for this problem would be to allow the previously docked ligands to move when docking a new one, which would mean a special induced fit approach considering translations and rotations of ligands. Furthermore a ligand might need to hop from one interaction point to another, involving greater displacement, or passing through a higher energy barrier than is usually allowed. Attempts were made to refine misdocked poses of two docked ligands simultaneously using low-mode conformational sampling including rotations and translations of the ligands in the binding site. However, this method did not provide significantly lower RMSDs on a small subset of our data set.

In the cytochromes P450 case study we investigated 9 complexes in total. SP and the SP hard protocols were able to find well-docked poses for all the ligand pairs with generally low ranks. Two-thirds of the docking steps provided well-docked top ranked poses in the SP hard protocol. These results are especially encouraging as binding modes in the human 2B4, 2C8 and 3A4 isoforms were reproduced well by the default SP protocol. Thus, it seems that this important family of enzymes is a promising target for multiple ligand docking. Misdocked poses could be mainly related to scoring problems. They most often arose from enforcing hydrogen bonds not present in the experimental structures or inaccurate scoring of interactions with the heme iron.

Results omitting conserved crystal waters were less than promising for HSP90 complexes, but their inclusion yielded excellent docking results. Two top ranked well-docked poses were obtained for all complexes with all protocols. The binding motifs of the fragments in these structures usually comprise only one specific interaction. Misdocked poses obtained with the water-free structures arose when the sterically feasible alternative position of the fragment was stabilized by a hydrogen bond formed with a residue originally involved in a hydrogen bond with a water molecule.

In summary, the performance of Glide was investigated on 129 multiple ligand bound protein complexes in a sequential docking setup with three different protocols and three different scoring functions. On average 36 % of the whole set of structures with at least two well-docked ligands ranked among the top three poses were obtained. The introduction of the druglikeness filter for ligands and closeness filter for binding sites resulted in higher performance, and the ratio of well-reproduced structures increased to 50% for the SP hard protocol. XP was found to provide lower success rates but with a higher probability of the top scoring poses being well-docked. Three ligands could be docked in only a few cases. Two pharmaceutically relevant subsets, that of cytochromes P450 and HSP90 complexes from fragment screens, were examined in more detail. In these cases higher success rates were observed than in the druglike and closed site subset. These results show the limitations of large-scale screening applications in the sequential docking of multiple ligands but indicate their utility in screening against drug interactions and in virtual fragment screening. Further efforts with induced fit protocols, however, seem to be validated to provide more precise binding mode information for the increasing number of cooperative binding events.

Acknowledgments The authors thank Dóra K. Menyhárd for critically reading the manuscript.

References

1. Whitty A (2008) Cooperativity and biological complexity. Nat Chem Biol 4:435–439
2. Denisov IG, Frank DJ, Sligar SG (2009) Cooperative properties of cytochromes P450. Pharmacol Ther 124:151–167
3. Guengerich FP (2008) Cytochrome P450 and chemical toxicology. Chem Res Toxicol 21:70–83
4. Wong H, Tong V, Riggs KW, Rurak DW, Abbott FS, Kumar S (2007) Kinetics of valproic acid glucuronidation: evidence for in vivo autoactivation. Drug Metab Dispos 35:1380–1386
5. Uchaipichat V, Galetin A, Houston JB, Mackenzie PI, Williams JA, Miners JO (2008) Kinetic modeling of the interactions between 4-methylumbelliferone, 1-naphthol, and zidovudine glucuronidation by UDP-glucuronosyltransferase 2B7 (UGT2B7) provides evidence for multiple substrate binding and effector sites. Mol Pharmacol 74:1152–1162
6. Caccuri AM, Antonini G, Ascenzi P, Nicotra M, Nuccetelli M, Mazzetti AP, Federici G, Lo Bello M, Ricci G (1999) Temperature adaptation of glutathione S-transferase P1-1. A case for homotropic regulation of substrate binding. J Biol Chem 274:19276–19280
7. Martin C, Berridge G, Higgins CF, Mistry P, Charlton P, Callaghan R (2000) Communication between multiple drug binding sites on P-glycoprotein. Mol Pharmacol 58:624–632
8. Kondratov RV, Komarov PG, Becker Y, Ewenson A, Gudkov AV (2001) Small molecules that dramatically alter multidrug resistance phenotype by modulating the substrate specificity of P-glycoprotein. Proc Natl Acad Sci USA 98:14078–14083
9. Higgins CF (2007) Multiple molecular mechanisms for multidrug resistance transporters. Nature 446:749–757
10. Sterz K, Möllmann L, Jacobs A, Baumert D, Wiese M (2009) Activators of P-glycoprotein: structure-activity relationships and investigation of their mode of action. Chem Med Chem 4:1897–1911
11. Harlow GR, Halpert JR (1998) Analysis of human cytochrome P450 3A4 cooperativity: construction and characterization of a site-directed mutant that displays hyperbolic steroid hydroxylation kinetics. Proc Natl Acad Sci USA 95:6636–6641
12. Domanski TL, He YA, Khan KK, Roussel F, Wang Q, Halpert JR (2001) Phenylalanine and tryptophan scanning mutagenesis of CYP3A4 substrate recognition site residues and effect on substrate oxidation and cooperativity. Biochemistry 40:10150–10160
13. Khan KK, He YQ, Domanski TL, Halpert JR (2002) Midazolam oxidation by cytochrome P450 3A4 and active-site mutants: an evaluation of multiple binding sites and of the metabolic pathway that leads to enzyme inactivation. Mol Pharmacol 61:495–506
14. Harrelson JP, Atkins WM, Nelson SD (2008) Multiple-ligand binding in CYP2A6: probing mechanisms of cytochrome P450 cooperativity by assessing substrate dynamics. Biochemistry 47:2978–2988
15. Hummel MA, Gannett PM, Aguilar JS, Tracy TS (2004) Effector-mediated alteration of substrate orientation in cytochrome P450 2C9. Biochemistry 43:7207–7214
16. Roberts AG, Díaz MD, Lampe JN, Shireman LM, Grinstead JS, Dabrowski MJ, Pearson JT, Bowman MK, Atkins WM, Campbell AP (2006) NMR studies of ligand binding to P450eryF provides insight into the mechanism of cooperativity. Biochemistry 45:1673–1684
17. Cameron MD, Wen B, Allen KE, Roberts AG, Schuman JT, Campbell AP, Kunze KL, Nelson SD (2005) Cooperative binding of midazolam with testosterone and alpha-naphthoflavone within the CYP3A4 active site: a NMR T1 paramagnetic relaxation study. Biochemistry 44:14143–14151
18. Cameron MD, Wen B, Roberts AG, Atkins WM, Campbell AP, Nelson SD (2007) Cooperative binding of acetaminophen and caffeine within the P450 3A4 active site. Chem Res Toxicol 20:1434–1441
19. Schoch GA, Yano JK, Sansen S, Dansette PM, Stout CD, Johnson EF (2008) Determinants of cytochrome P450 2C8 substrate binding: structures of complexes with montelukast, troglitazone, felodipine, and 9-cis-retinoic acid. J Biol Chem 283:17227–17237
20. Schumacher MA, Miller MC, Brennan RG (2004) Structural mechanism of the simultaneous binding of two drugs to a multidrug-binding protein. EMBO J 23:2923–2930
21. Ekroos M, Sjögren T (2006) Structural basis for ligand promiscuity in cytochrome P450 3A4. Proc Natl Acad Sci USA 103:13682–13687
22. Aller S, Yu J, Ward A, Weng Y, Chittaboina S, Zhuo R, Harrell PM, Trinh YT, Zhang Q, Urbatsch IL, Chang G (2009) Structure of P-glycoprotein reveals a molecular basis for poly-specific drug binding. Science 323:1718–1722
23. Congreve M, Chessari G, Tisi D, Woodhead AJ (2008) Recent developments in fragment-based drug discovery. J Med Chem 51:3661–3680
24. Petros AM, Dinges J, Augeri DJ, Baumeister SA, Betebenner DA, Bures MG, Elmore SW, Hajduk PJ, Joseph MK, Landis SK, Nettesheim DG, Rosenberg SH, Shen W, Thomas S, Wang X, Zanze I, Zhang H, Fesik SW (2006) Discovery of a potent inhibitor of the antiapoptotic protein Bcl-xL from NMR and parallel synthesis. J Med Chem 49:656–663
25. Huth JR, Park C, Petros AM, Kunzer AR, Wendt MD, Wang X, Lynch CL, Mack JC, Swift KM, Judge RA, Chen J, Richardson PL, Jin S, Tahir SK, Matayoshi ED, Dorwin SA, Ladror US, Severin JM, Walter KA, Bartley DM, Fesik SW, Elmore SW, Hajduk PJ (2007) Discovery and design of novel HSP90 inhibitors using multiple fragment-based design strategies. Chem Biol Drug Des 70:1–12
26. Sandor M, Kiss R, Keseru GM (2010) Virtual fragment docking by Glide: a validation study on 190 protein-fragment complexes. J Chem Inf Model 50:1165–1172
27. Li H, Li C (2010) Multiple ligand simultaneous docking: orchestrated dancing of ligands in binding sites of protein. J Comput Chem 31:2014–2022
28. Li H, Liu A, Zhao Z, Xu Y, Lin J, Jou D, Li C (2011) Fragment-based drug design and drug repositioning using multiple ligand simultaneous docking (MLSD): identifying celecoxib and template compounds as novel inhibitors of signal transducer and activator of transcription 3 (STAT3). J Med Chem 54:5592–5596
29. SiteMap, version 2.4, Schrödinger, LLC, New York, NY, 2010
30. Schrödinger Suite 2010 Protein Preparation Wizard; Epik version 2.1, Schrödinger, LLC, New York, NY, 2010; Impact version 5.6, Schrödinger, LLC, New York, NY, 2010; Prime version 2.2, Schrödinger, LLC, New York, NY, 2010
31. Roughley SD, Hubbard RE (2011) How well can fragments explore accessed chemical space? A case study from heat shock protein 90. J Med Chem 54:3989–4005
32. Molecule File Converter, version 5.4.1.1, © 1999-2011 ChemAxon Ltd
33. LigPrep, version 2.4, Schrödinger, LLC, New York, NY, 2010
34. Epik, version 2.1, Schrödinger, LLC, New York, NY, 2010
35. Shelley JC, Cholleti A, Frye LL, Greenwood JR, Timlin MR, Uchiyama M (2007) Epik: a software program for pKa prediction and protonation state generation for druglike molecules. J Comput Aided Mol Des 21:681–691
36. Greenwood JR, Calkins D, Sullivan AP, Shelley JC (2010) Towards the comprehensive, rapid, and accurate prediction of the favorable tautomeric states of drug-like molecules in aqueous solution. J Comput Aided Mol Des 24:591–604

37. Calculator, version 5.4.1.1, © 1998–2011 ChemAxon Ltd
38. SiteMap 2.4 User Manual, http://www.schrodinger.com/supportdocs/18/20/
39. Glide, version 5.6, Schrödinger, LLC, New York, NY, 2010
40. Friesner RA, Banks JL, Murphy RB, Halgren TA, Klicic JJ, Mainz DT, Repasky MP, Knoll EH, Shelley M, Perry JK, Shaw DE, Francis P, Shenkin PS (2004) Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. J Med Chem 47:1739–1749
41. Halgren TA, Murphy RB, Friesner RA, Beard HS, Frye LL, Pollard WT, Banks JL (2004) Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. J Med Chem 47:1750–1759
42. Friesner RA, Murphy RB, Repasky MP, Frye LL, Greenwood JR, Halgren TA, Sanschagrin PC, Mainz DT (2006) Extra precision Glide: docking and scoring incorporating a model of hydrophobic enclosure for protein-ligand complexes. J Med Chem 49:6177–6196
43. Cross JB, Thompson DC, Rai BK, Baber JC, Fan KY, Hu Y, Humblet C (2009) Comparison of several molecular docking programs: pose prediction and virtual screening accuracy. J Chem Inf Model 49:1455–1474
44. Cupp-Vickery J, Anderson R, Hatziris Z (2000) Crystal structures of ligand complexes of P450eryF exhibiting homotropic cooperativity. Proc Natl Acad Sci USA 97:3050–3055
45. Podust LM, Ouellet H, von Kries JP, de Montellano PR (2009) Interaction of Mycobacterium tuberculosis CYP130 with heterocyclic arylamines. J Biol Chem 284:25211–25219
46. Zhao B, Guengerich FP, Voehler M, Waterman MR (2005) Role of active site water molecules and substrate hydroxyl groups in oxygen activation by cytochrome P450 158A2: a new mechanism of proton transfer. J Biol Chem 280:42188–42197
47. Zhao B, Guengerich FP, Bellamine A, Lamb DC, Izumikawa M, Lei L, Podust LM, Sundaramoorthy M, Kalaitzis JA, Reddy LM, Kelly SL, Moore BS, Stec D, Voehler M, Falck JR, Shimada T, Waterman MR (2005) Binding of two flaviolin substrate molecules, oxidative coupling, and crystal structure of Streptomyces coelicolor A3(2) cytochrome P450 158A2. J Biol Chem 280:11599–11607
48. Makino M, Sugimoto H, Shiro Y, Asamizu S, Onaka H, Nagano S (2007) Crystal structures and catalytic mechanism of cytochrome P450 StaP that produces the indolocarbazole skeleton. Proc Natl Acad Sci USA 104:11591–11596
49. Gay SC, Sun L, Maekawa K, Halpert JR, Stout CD (2009) Crystal structures of cytochrome P450 2B4 in complex with the inhibitor 1-biphenyl-4-methyl-1H-imidazole: ligand induced structural response through α-helical repositioning. Biochemistry 48:4762–4771
50. Atkins WM (2005) Non-Michaelis-Menten kinetics in cytochrome P450-catalyzed reactions. Annu Rev Pharmacol Toxicol 45:291–310
51. Houston JB, Galetin A (2005) Modelling atypical CYP3A4 kinetics: principles and pragmatism. Arch Biochem Biophys 433:351–360
52. Zhou SF, Zhou ZW, Yang LP, Cai JP (2009) Substrates, inducers, inhibitors and structure-activity relationships of human cytochrome P450 2C9 and implications in drug development. Curr Med Chem 16:3480–3675
53. Sohl CD, Isin EM, Eoff RL, Marsch GA, Stec DF, Guengerich FP (2008) Cooperativity in oxidation reactions catalyzed by cytochrome P450 1A2: highly cooperative pyrene hydroxylation and multiphasic kinetics of ligand binding. J Biol Chem 283:7293–7308
54. Khan KK, Liu H, Halpert JR (2003) Homotropic versus heterotopic cooperativity of cytochrome P450eryF: a substrate oxidation and spectral titration study. Drug Metab Dispos 31:356–359
55. Lasker JM, Huang MT, Conney AH (1984) In vitro and in vivo activation of oxidative drug metabolism by flavonoids. J Pharmacol Exp Ther 229:162–170
56. Tang W, Stearns RA, Kwei GY, Iliff SA, Miller RR, Egan MA, Yu NX, Dean DC, Kumar S, Shou M, Lin JH, Baillie TA (1999) Interaction of diclofenac and quinidine in monkeys: stimulation of diclofenac metabolism. J Pharmacol Exp Ther 291:1068–1074
57. Hutzler JM, Frye RF, Korzekwa KR, Branch RA, Huang SM, Tracy TS (2001) Minimal in vivo activation of CYP2C9-mediated flurbiprofen metabolism by dapsone. Eur J Pharm Sci 14:47–52
58. Egnell AC, Houston B, Boyer S (2003) In vivo CYP3A4 heteroactivation is a possible mechanism for the drug interaction between felbamate and carbamazepine. J Pharmacol Exp Ther 305:1251–1262
59. Egnell AC, Eriksson C, Albertson N, Houston B, Boyer S (2003) Generation and evaluation of a CYP2C9 heteroactivation pharmacophore. J Pharmacol Exp Ther 307:878–887
60. Egnell AC, Houston JB, Boyer CS (2005) Predictive models of CYP3A4 Heteroactivation: in vitro-in vivo scaling and pharmacophore modeling. J Pharmacol Exp Ther 312:926–937
61. Locuson CW, Gannett PM, Ayscue R, Tracy TS (2007) Use of simple docking methods to screen a virtual library for heteroactivators of cytochrome P450 2C9. J Med Chem 50:1158–1165
62. Barker JJ, Barker O, Courtney SM, Gardiner M, Hesterkamp T, Ichihara O, Mather O, Montalbetti CA, Muller A, Varasi M, Whittaker M, Yarnold CJ (2010) Discovery of a novel Hsp90 inhibitor by fragment linking. Chem Med Chem 5:1697–1700
63. Murray CW, Carr MG, Callaghan O, Chessari G, Congreve M, Cowan S, Coyle JE, Downham R, Figueroa E, Frederickson M, Graham B, McMenamin R, O'Brien MA, Patel S, Phillips TR, Williams G, Woodhead AJ, Woolford AJ (2010) Fragment based drug discovery applied to Hsp90. Discovery of two lead series with high ligand efficiency. J Med Chem 53:5942–5955