

Docking of small molecules to farnesoid X receptors using AutoDock Vina with the Convex-PL potential: lessons learned from D3R Grand Challenge 2

Maria Kadukova, Sergei Grudinin

▶ **To cite this version:**

Maria Kadukova, Sergei Grudinin. Docking of small molecules to farnesoid X receptors using AutoDock Vina with the Convex-PL potential: lessons learned from D3R Grand Challenge 2. Journal of Computer-Aided Molecular Design, 2018, 32 (1), pp.151-162. 10.1007/s10822-017-0062-1 . hal-01591157

HAL Id: hal-01591157
https://inria.hal.science/hal-01591157

Submitted on 20 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Copyright

This document is the edited Author's version of a Submitted Work that was subsequently accepted for publication in Journal of Computer-Aided Molecular Design after peer review. To access the published work see http://dx.doi.org/10.1007/s10822-017-0062-1.

Docking of small molecules to farnesoid X receptors using AutoDock Vina with the Convex-PL potential : lessons learned from D3R Grand Challenge 2

Maria Kadukova[1,2,3,4] and Sergei Grudinin*[1,2,3]

[1]Univ. Grenoble Alpes, LJK, F-38000 Grenoble, France
[2]CNRS, LJK, F-38000 Grenoble, France
[3]Inria, France
[4]Moscow Institute of Physics and Technology, Dolgoprudniy, Russia

The 2016 D3R Grand Challenge 2 provided an opportunity to test multiple protein-ligand docking protocols on a set of ligands bound to farnesoid X receptor that has many available experimental structures. We participated in the Stage 1 of the Challenge devoted to the docking pose predictions, with the mean RMSD value of our submission poses of 2.9 Å. Here we present a thorough analysis of our docking predictions made with AutoDock Vina and the Convex-PL rescoring potential by reproducing our submission protocol and running a series of additional molecular docking experiments.

We conclude that a correct receptor structure, or more precisely, the structure of the binding pocket, plays the crucial role in the success of our docking studies. We have also noticed the important role of a local ligand geometry, which seems to be not well discussed in literature. We succeed to improve our results up to the mean RMSD value of 2.15 – 2.33 Å dependent on the models of the ligands, if docking these to all available homologous receptors. Overall, for docking of ligands of diverse chemical series we suggest to perform docking of each of the ligands to a set of multiple receptors that are homologous to the target.

keywords : protein-ligand docking; ensemble docking; flexible docking; D3R; scoring function;

Introduction

The Drug Design Data Resource (D3R, www.drugdesigndata.org) is a successor of Community Structure-Activity Resource [1,2] whose aim is to advance the technology of computer-aided drug discovery through the interchange of protein-ligand datasets and workflows. It provides high-quality experimental data by hosting blinded unpublished datasets for testing and improving ligand-protein docking algorithms and their scoring protocols.

The D3R Grand Challenge 2 provided a blinded unpublished dataset containing the farnesoid X receptor (FXR) target, which was kindly contributed by Roche and curated by D3R. This dataset contained 36 crystal structures with resolution better than 2.6 Å supplied with binding affinity data (IC50s) for 102 compounds across five orders of magnitude, comprising four chemical series and 6 miscellaneous compounds. The D3R Grand Challenge 2 was oraganized in two stages. The oraganizers provided as input SMILES strings and SDfiles for 102 ligands, among which 36 were selected for docking prediction. The oraganizers also provided SMILES

*sergei.grudinin@inria.fr

strings and SDfiles of two subsets of 18 and 15 compounds for the calculation of relative binding affinities. In the first stage, the goals were to predict crystallographic binding poses of 36 selected ligands spanning all chemical series, predict binding affinities for all 102 ligands, and to predict the relative binding affinities for two designated free energy subsets of 18 and 15 compounds. After the end of the first stage, the organizers made available the blinded co-crystal structures of 36 selected FXR complexes. We should note that FXR_33 was excluded from the pose prediction evaluations because of crystal artifacts found in the co-crystal structure, and only 35 ligands were assessed. In the second stage, the goal was to predict binding affinities for all 102 ligands and the relative binding affinities for the two free energy subsets of 18 and 15 compounds using the released 36 co-crystal structures of protein-ligand complexes. After the end of the second stage, the experimental binding affinities (IC50s) were made available.

We participated only in the Stage 1 of D3R Grand Challenge 2 and our goal was to assess multiple docking pose prediction protocols. Some of these were introduced earlier [3, 4], others were designed specifically for the FXR targets. Like many other participants, we obtained sufficiently small RMSD values for most of the benzimidazole-containing ligands. However, we did not succeed in docking of most other groups of compounds. Below we provide a detailed explanation of our model used for the submission and several other docking protocols tested after the end of the challenge.

Docking strategies in previous exercises

Recently several major docking competitions were held, namely CSAR 2013 [5], CSAR 2014 [6] and D3R 2015-2016 [7], which became an opportunity for many teams to assess their protein-ligand pose and binding affinity prediction algorithms and protocols. In the course of these competitions various methods were used including classical docking methods [8–21], QSAR models [15, 22, 23], target-specific scoring functions [23–25], and sometimes combinations of these with more computationally expensive molecular dynamics-based methods [26–29]. The CSAR 2013 exercise also involved homology modeling to obtain proper receptors from the given sequences, while in CSAR 2014 protein structures were provided. In the D3R challenge, participants were encouraged to use existing crystallographic structures from the RCSB database, which could have been selected with ligand similarity-based methods.

Participants reported rather classical problems, such as the importance of taking into account the receptor flexibility, buried water molecules, presence of explicit solvent, lack of the homologous proteins present in databases, etc. In particular, several approaches were used to handle the receptor flexibility problem [30]. In case of a considerable number of complexes with homologous proteins available, their thorough inspection helped to choose either the proper docking target, or to identify a small number of residues suitable for flexible docking protocols [28, 31]. Side-chains repacking [19] and the induced-fit approach [27, 32] helped to optimize or widen the binding pockets. Flexible docking programs were also used [33]. The choice for the best strategy for picking the receptor structure, i.e. cross-docking or docking to a certain "closed"-form structure remains an open question, since the best strategy seems to be dependent on the availability of the co-crystal complexes with ligands of the chemical series of interest. Both D3R and CSAR organizers reported the difficulty of making conclusions about the best approaches, as the same methods used by different teams often lead to significantly different results. Notably, importance of visual inspection of predicted structures was mentioned in the D3R 2015 results description.

Methods

Docking dataset

The target receptor of the Grand Challenge 2 was a nuclear farnesoid X receptor (FXR). A set of compounds from several chemical series was provided, 36 of them were selected for the pose prediction exercise, however only 35 were used for the pose evaluation. These include benzimidazoles, isoxazoles, sulfonamides, spiranes and several other unclassified molecules. The RCSB database (also knows

as Protein Data Bank) [34] contains a large variety of FXR structures with a different degree of homology to the target apoprotein receptor. For example, for the homology threshold of 30%, the RCSB database provides about 300 of protein structures. In these structures most of the protein 3D organization is conserved, however, short helices near the binding site exhibit a high degree of conformational freedom, as it is shown in Fig. 1. We should specifically note that the ligands bound to FXRs from RCSB do not include spiranes and sulfonamides, so it was challenging to model the corresponding targets from the docking exercise.

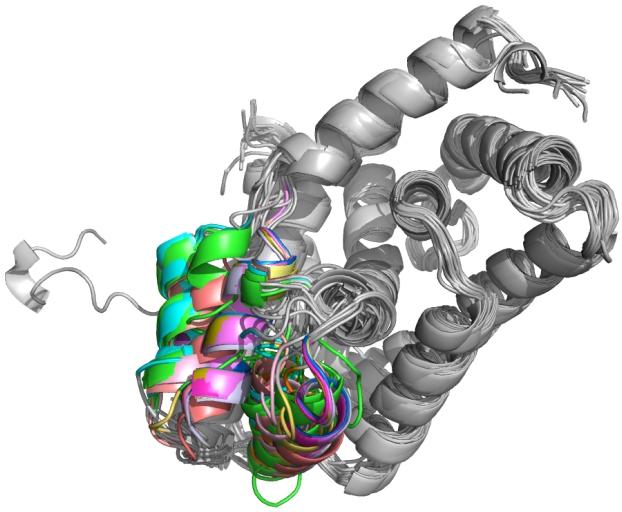


Figure 1: Proteins with a high sequence identity (higher than 90 %) to the target FXR apoprotein superposed to each other. Here only the A chains are shown. Two helices with a high conformational freedom are highlighted.

Submission protocol

For our submission, we generated 3D ligand structures starting from the corresponding SMILES strings using OpenBabel's *gen3d* command [35]. Then, we manually adjusted ligands with ambiguous symmetry using geometry of small molecules with similar functional groups from the RCSB database as an example. No additional structure optimization was performed at this step.

Regarding the receptors, we manually selected one model for each target from the RCSB database first using the homology threshold of 30 % and after selecting the best model using the similarity of the bound ligand as it was identified based on the SMILES substrings of the functional groups. This procedure identified 23 template receptors for 36 targets. Among these, 17 had homology to the target sequence higher than 90 %. No additional structure optimization of the receptors was used.

To generate putative binding poses, we used the AutoDock Vina software package with the default scoring function [36]. Using AutoDock Vina, we generated 1,000 docking solutions for the subsequent re-scoring. In the AutoDock Vina configuration files, the parameter *num_modes* was set to 1,000 and *exhaustiveness* to 100. We identified the receptor binding pocket based on the structures of homologous FXR proteins. We chose all the rotatable bonds in ligands to be flexible during the docking procedure, and we kept all the protein residues inside the binding pockets rigid. We assigned the Gasteiger atomic partial charges and converted all receptors and ligands to the PDBQT format using the AutoDockTools package [37]. We did not use explicit hydrogens either for the receptors or for the ligands.

Finally, we re-scored the obtained poses with the Convex-PL potential [38]. Convex-PL is a knowledge-based scoring potential, which we already used in the previous D3R and CSAR challenges [3, 4]. It is freely available on our web-site at `http://team.inria.fr/nano-d/convex-pl/`. No additional clustering of the final poses was made.

Self-docking protocol

Once the co-crystal FXR structures were released by the challenge organizers, we repeated the docking tests using crystallographic structures of 35 (since FXR_33 target was skipped from the assessment) receptors with the corresponding ligands. We used a docking protocol similar to the one for the submission. More precisely, we converted the input crystallographic structures to the PDBQT format and ran AutoDock Vina with the same parameters as described above. However, this time we used the Convex-PL scoring function inside AutoDock Vina without the subsequent rescoring. Atomic types were assigned using the Knodle method [39] and

no explicit hydrogens were used.

Additional protocols

To rigorously assess the performance of different docking strategies, in addition to the self-docking test, we designed seven other docking experiments. These are summarized in Table 3 and listed below. In the additional docking experiments, we optimized several ligands and also mutated the receptors whose sequence was different from the target as it is described below.

In the first experiment (exp. 1) we repeated the submission protocol but mutated all the receptors to the target sequence. This protocol may be seen as "one ligand to one receptor" docking.

In the second experiment (exp. 2) we exhaustively docked all the ligands used in the submission protocols to each of the receptors used in the submission plus to all other receptors with homology higher than 90 % to the target receptor sequence. In this experiment we mutated all the receptors to the target sequence. This protocol may be seen as "one ligand to multiple receptors" docking. The receptor PDB codes used in this experiment are listed in Table 1.

1osh	1osv	1ot7	3bej	3dct	3dcu
3fli	3fxv	3gd2	3hc5	3hc6	3l1b
3okh	3oki	3olf	3omk	3omm	3oof
3ook	3p88	3p89	3rut	3ruu	3rvf
4oiv	4qe6	4qe8	4wvd	5iaw	5ick
1s0x	3dy6	3peq	3w5p	4r06	5eit

Table 1: PDB codes of the 36 receptors chosen from the RCSB database for the "multiple receptors" docking protocols. Proteins with a low sequence identity are highlighted in grey.

In the third experiment (exp. 3) we repeated exp. 2 for each of the target with all the ligand models available as explained in more detail below. This protocol may be seen as "multiple ligands to multiple receptors" docking.

In the forth experiment (exp. 4) we repeated exp. 2, where for each of the target we used only a few receptors whose bound ligand had the most similarity to the target ligand. This protocol may be seen as "one ligand to some receptors" docking.

In the fifth experiment (exp. 5) we complemented exp. 2 with the co-crystal structures of the receptors. This protocol may be seen as "one ligand to multiple+self receptors" docking.

In the sixth experiment (exp. 6) for each of the targets we docked the ligand used in the submission to the co-crystal structure of the receptor. This protocol may be seen as "one ligand to self receptor" docking.

In the seventh experiment (exp. 7) for each of the targets we docked the co-crystal structure of the ligand to a number of receptors from exp. 2. This protocol may be seen as "self ligand to multiple receptors" docking.

In addition to the above-listed experiments, we also repeated exp.1, exp. 2, exp. 4, exp. 5, and exp. 6, where for each of the targets we complemented some difficult ligands from the submission with their optimized structures as explained below in more detail. These experiments are labeled "mod" in Table 3. The original experiments are respectively labeled "old".

In all the additional docking experiments we ran AutoDock Vina with the same parameters as described above for the self-docking protocol. More precisely, we used only the Convex-PL scoring function inside AutoDock Vina without the subsequent rescoring. Then, we selected the five top-scored structures after a simple clusterization procedure with a symmetry-adapted ligand RMSD threshold value of 1 Å.

Preparation of ligand structures

For all the additional experiments, we generated 3D ligand structures using OpenBabel *gen3d* command [35] complemented with fragments of similar small molecules in cases when OpenBabel was unable to provide reasonable structures. Regarding OpenBabel, its *gen3d* method generates ligands using a set of rules and templates, followed by a minimization in MMFF94 force-field [40]. Finally, it performs a conformational search followed once again by the ligand optimization.

For the additional computational experiments described above, we used the same ligand structures

PDB code (chain)	#	values for mutated residues in the binding pocket: mean molecular weight		mean pH at isoelectric point		mean index of hydrophobicity	
		before	after	before	after	before	after
1osv	5	129.34	106.51	6.34	5.81	18	25
1ot7 (AC)	4	118.12	110.87	5.23	5.76	26	21
1ot7 (BDE)	5	129.34	106.51	6.34	5.81	18	25
3bej	1	147.13	89.10	3.22	6.00	-31	41
3rvf	3	142.14	89.10	6.55	6.00	-17	41
3dct	2	126.11	89.10	4.45	6.00	-18	41
3dcu	2	126.11	89.10	4.45	6.00	-18	41
3gd2	2	126.11	89.10	4.45	6.00	-18	41
3hc5	2	126.11	89.10	4.45	6.00	-18	41
3hc6	2	126.11	89.10	4.45	6.00	-18	41
3p88	2	126.11	89.10	4.45	6.00	-18	41
3p89	2	126.11	89.10	4.45	6.00	-18	41
3rut	2	126.11	89.10	4.45	6.00	-18	41
3ruu	2	126.11	89.10	4.45	6.00	-18	41
4oiv	2	160.66	89.10	6.99	6.00	-23	41
4qe6	3	142.14	89.10	6.55	6.00	-17	41
4qe8 (AC)	2	160.66	89.10	6.99	6.00	-23	41
4qe8 (BD)	1	147.13	89.10	3.22	6.00	-31	41
5iaw	1	147.13	89.10	3.22	6.00	-31	41
1s0x	23	138.12	131.53	6.04	5.99	59	38
3dy6 (A)	23	133.76	139.46	5.86	5.89	49	50
3dy6 (B)	22	134.61	142.39	5.84	5.89	53	53
3peq	21	134.77	141.30	5.83	5.91	51	51
3w5p	42	138.04	138.45	5.91	6.10	51	43
4r06 (A)	22	142.89	140.26	5.71	5.89	52	60
4r06 (B)	17	134.68	127.38	5.79	5.64	19	41
5eit (A)	26	136.04	133.27	5.88	5.96	47	35
5eit (B)	25	136.20	132.71	5.90	6.06	50	38

Table 2: List of receptors with the mutated amino acids in the binding pocket used in the experiments. The binding pocket was defined as all the residues having at least one heavy atom within the distance of 4 Å from any atom of ligands that are available for the 36 receptor proteins and the co-crystal FXR protein targets. Proteins with a low sequence identity are highlighted in grey. Thirteen proteins had 100% sequence identity in the binding pocket and were excluded from the table.

as in our submission. However, for "difficult" ligands, i.e. those whose docking to the co-crystal receptors resulted in RMSD values greater than 1.5 Å in all the top-5 poses without clusterization, we performed an additional structure optimization. More precisely, we optimized these ligands using force-field-based structure optimization with the MMFF94 force-field [40] implemented in the Avogadro package [41]. If the gradient-based optimization with MMFF94 did not converge, we used the UFF force-field [42]. Although we used the MMFF94 force-field optimization both in OpenBabel (implicitly) and in Avogadro, the resulting conformations were not the same, perhaps due to the differences between the 3D structure generation algorithms, and also due to the fact that we used different input files, 2D SDF files for OpenBabel and SMILES strings for Avogadro. Finally, all the structures of ligands were converted to the PDBQT format using AutoDockTools with the default identification of rotatable bonds.

Preparation of receptor structures

To represent structural heterogeneity of the receptors, we selected all the structures from the RCSB database with the identity to the target sequence higher than 90 %. The identity was computed with the *blastp* program from the BLAST+ package [43]. This resulted in a set of 30 receptors. For the submission, as mentioned above, we used 6 additional low-homologous receptors chosen based on the ligand similarity to the target ligands. These

were also added to the receptors set, resulting in a pool of 36 structures used in the additional experiments. For each of the receptors, we mutated all the residues different from the apoprotein sequence using the SCWRL4 tool [44] as it is described below. We did not model residues that were missing in the crystallographic structures. To structurally align the receptors to the provided model, we use the *align* command from PyMOL [45] with the default options. Finally, all the structures of receptors were converted to the PDBQT format using AutoDockTools. We treated all the receptors' bonds as rigid, i.e. no flexible residues were selected.

Mutations

Table 2 lists the summary of the total number and properties of the mutated residues in the binding pocket. More precisely, we list the mean molecular weights, pI and hydrophobicity indices of the mutated residues before and after the mutations. The reference values were taken from the Sigma-Aldrich webpage [46]. For all the highly homologous proteins listed in Table 2 only 1-5 residues were mutated in the binding site consisting of 39-60 residues. Three of these mutations occurred in all of these proteins and they considerably changed at least one property of the residues. These are the polar Ser342 and the hydrophilic Glu350 and Arg455 residues that were replaced by the hydrophobic alanine residues[1]. Ser342 is located close to most of the available ligands, but its mutation preserves the residue's size and charge. Arg455 is located relatively close to several aromatic ligand groups, and thus its turn into a more hydrophobic alanine might be energetically favorable. Glu350 is situated rather close to several other ligands, but its hydrophilic glutamate sidechain is rotated outward from the binding pocket. Therefore, we did not expect these mutations to strongly influence the docking results. However, we found it better to apply these to the protein structures. As for the low-identity proteins, although the mean values of the pI, size and hydrophobicity index remain similar, as it can be seen from Table 2, about a half of the mutations for each receptor lead to significant changes of some of the residue's properties.

Missing residues

As it was mentioned above, we did not model missing residues in the binding sites of the protein structures that we used for the submission. For example, in several proteins an entire helix (Gln271–Glu285, Arg268–Phe282, Gln257–Phe282 in 1osh, 3l1b, and 4oiv, respectively) or a small loop (Lys343-Gly347, Lys343-Pro345, Leu344-Ala346 in 3fxv, 4qe6, and 3bej, respectively) were missing. For these structures, inserting the helix model might have been a good solution, as its conformation is conserved in all the 30 highly homologous receptor models. This helix also appeared to have the same conformation in a number of reference co-crystal FXR complexes and it has no steric clashes with any of the correct ligand poses. The missing loop has a different conformation compared to the apoprotein in most of the co-crystal receptor structures, and its modeling would have been more challenging.

RMSD computation

We computed symmetry-adapted RMSD values with a modified GetBestRMS() function from the RDKit package [47]. Originally, this function aligns molecules taking into account the symmetry matchings by iterating over them to find the best alignment transformation with the lowest RMSD value. In our modification, we preserved the iteration over all symmetry matchings and removed the alignment part to make the function only compute the smallest RMSD value among various symmetry matchings without superposing the molecules. We should note that we obtained RMSD values of our submission slightly larger compared to those reported in the D3R challenge results, i.e. we computed the mean value for our submission of 2.96 Å versus 2.90 Å reported by the challenge organizers. This may happen due to the differences in RMSD computation and alignment procedures.

[1] Residue numbers are given with respect to the 3rvf receptor and may differ slightly for other proteins. The corresponding apoprotein residues are Ala346, Ala354, Ala459.

Results and discussion

Description of the docking results

Table 3 lists the top-5 pose prediction results for all the docking experiments with Convex-PL as a scoring function. As it is clearly seen from the table, we obtained sufficiently good results for the self-docking experiment with the mean RMSD value to the solutions of 0.58 Å. This indicates that our method is always able to predict near-native ligand poses given as input co-crystal receptor conformations along with the correct local geometries of the ligands. The local geometry includes bond lengths and angles, because these are kept unchanged during the docking procedure.

To figure out why the mean RMSD value of our submission is about 5 times larger that the result of the self-docking experiment, we ran another experiment with docking ligands to the co-crystal receptors, which is summarized in the exp. 6 column of Table 3. It can be seen that only in a few cases we are not able to predict a binding pose within 3 Å RMSD. However, the co-crystal receptor structure does not guarantee the highest docking score if there are other receptor structures present in the docking pool. This can be seen in the exp. 5 column, where the co-crystal receptor structures were complemented with the homologous FXR structures from the RCSB database. For example, for the ligands of FXR_4, FXR_10, FXR_11, and FXR_16 targets, our scoring function preferred non co-crystal receptors with considerably different ligand conformations.

Re-docking of the ligands to the same receptors that we used for the submission did not affect the mean RMSD, as it is seen from the column exp. 1 of Table 3. For some targets RMSD improved, for others it became worse. Considerable RMSD differences that can be seen for only several ligands may occur due to mutation of the residues, some of which were situated near the binding site (FXR_20, FXR_13, FXR_15). These differences remained even when we re-generated the ligand structures with a force-field-based optimization.

Columns exp. 2 and exp. 3 of Table 3 list the results of docking of each ligand (either single or multiple) to a set of homologous FXR structures found in the RCSB database. Overall, these experiments considerably improved the docking accuracy as compared to exp. 1. In case of "easy" ligands such as benzimidazoles, all the three strategies (exp. 1-3) performed well. However, for more complicated targets when the procedure of choosing a proper receptor is unclear, docking to multiple receptors allows to achieve better poses compared to the results of exp. 1, where we performed docking to a single receptor.

Unfortunately, in exp. 3, which as we believed was the most rigorous blind experiment, we obtained a lower mean RMSD value of 2.33 Å than in both parts of exp. 2 (2.15 Å and 2.21 Å). FXR_11 target contributed the most to this RMSD difference, as the top-2 predictions for the "modified" and top-4 predictions for the "old" ligands consist only of poses with RMSD values greater than 8 Å, which are not rejected by the subsequent clusterization procedure in exp. 3. The best pose was chosen properly for only 6 out of 13 "modified" ligands (FXR_2, FXR_3, FXR_8, FXR_17, FXR_22, FXR_34). Although the number of the ligands that we re-optimized is rather small to make rigorous conclusions, we would suggest to be careful with the "multiple ligand" docking protocols, as these may require better clusterization algorithms and a visual inspection of the obtained poses for each of the ligands.

Docking to the smaller sets of receptors, that we chose for each ligand manually by the ligand similarity, aided to improve pose prediction for FXR_2 and the submission version of the FXR_11 ligand as it is shown in the exp. 4 column of Table 3. On the other hand, this strategy excluded some potentially good candidates from the docking set of receptors. For example, in the case of FXR_15 target in exp. 2, for four receptors we were able to obtain rather low-scored poses within RMSD $<$ 3.5 Å and the top-scored conformation with RMSD of 4.87 Å, which preserved the sulphonamide group in the near-native position. Nevertheless, only one of these four receptors was in the set that we chose for the FXR_15 target in this experiment.

Below we would like to discuss several cases of the molecules that turned out to be the most challenging for our docking algorithms.

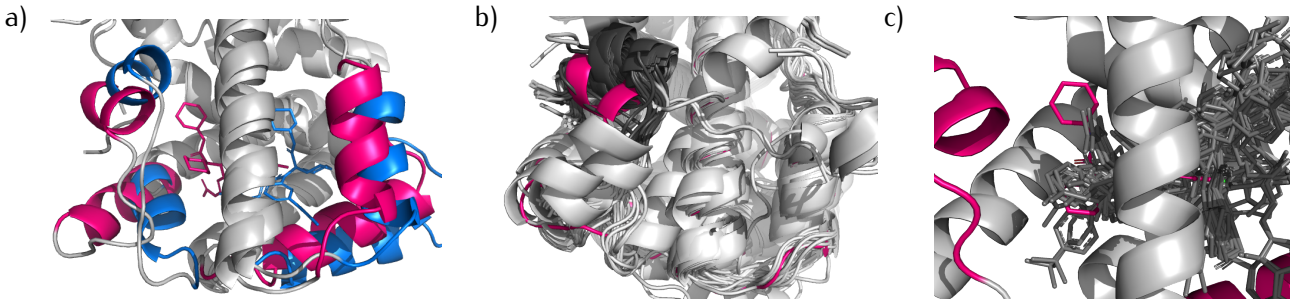


Figure 2: (a) Reference structure of the FXR apoprotein in complex with the FXR_18 ligand (red) and the result of docking to the 3oki protein (blue). Conformationally conserved chains are shown in grey. The two helices on the right have a high conformational freedom in all the homologous structures of FXR. (b) Co-crystal structure of apoprotein superposed with 36 proteins used in "multiple receptors" docking protocols. Here, the trp473–thr466 helix (the residue numbers may differ for different proteins) is shown in red for the reference structure and in dark grey for the other proteins. (c) Co-crystal ligand conformation and 30 ligands found in the highly homologous proteins (dark grey).

Importance of the receptor conformation

The most difficult target for our protocols appeared to be FXR_18. Its ligand contains the same functional groups (chlorine, amide, aromatic and cyclohexane rings, planar trigonal nitrogen in the center of the molecule) as some of the benzimidazole ligands from the complexes available at RCSB, namely, 3oki, 3omk, 3oof and several other structures of human FXRs obtained by Roche in 2011 [48]. Our scoring protocol correctly identifies all these proteins from a larger set of receptor candidates, however, the co-crystal ligand binding pose differs from our predictions by RMSD of more than 9 Å.

This can be the result of the high conformational mobility of four of the binding pocket helices, as highlighted in Fig. 2 (a) for the highest scored receptor model 3oki, which differ by 6.34 Å RMSD from the co-crystal structure.

More precisely, besides the two helices that are known to be flexible for the farnesoid X receptors, as we have mentioned in the dataset description, two other helices of the binding site of our top-scored receptor 3oki have different conformations compared to the co-crystal receptor structure. These are connected with a loop and located on the opposite site of the binding pocket, as it is shown at the left side of Fig. 2 (a). This region seems to be relatively conserved in all of the proteins that we used for the "multiple receptors" docking protocols. The correct receptor conformation superposed with these proteins can be seen in Fig. 2 (b), where it is "buried" under the the other proteins' conformations except for a short trp473–thr466 helix, which is located near one of the rings of the FXR_18 ligand. In the co-crystal structure, this cyclohexane ring has a different 3D configuration from all the other ligands from the targets we used for the docking experiments, as it is shown in Fig. 2 (c). Thus, the correct docking poses of this part of the ligand result in steric clashes with the flexible helices of the non co-crystal receptor structures.

A similar situation is observed for another difficult FXR_4 target. Here, our docking algorithm prefers the above-mentioned receptors found in complex with benzimidazole ligands, in particular, the 3oki one, which has a benzole ring with a chlorine substitute and two cyclohexane rings. Fig. 3 shows the best predicted poses for FXR_4 and FXR_18 ligands superposed with the ligands of the 3oki structure. These, unfortunately, are not the right solutions, and the conformation of one of the 3oki helices differs from the co-crystal structure of FXR_4. Thus, the structure of the 3oki receptor would once again cause clashes with the co-crystal pose of the ligand.

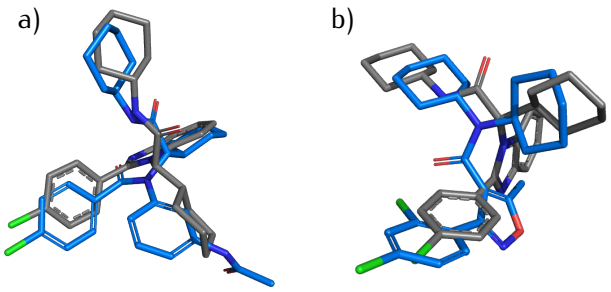


Figure 3: (a) The best predicted ligand pose for FXR_18 (RMSD=9.34, blue) superposed with the 3oki ligand (chain A, grey). (b) The best predicted ligand pose for FXR_4 (RMSD=6.74, blue) superposed with the 3oki ligand (chain C, grey).

We have previously mentioned the FXR_15 ligand, for which we were able to find only three receptors, docking to which allowed to obtain poses with RMSD values of about 3–3.5 Å (in exp. 2). For each of these receptors, the sulfonamide-containing part of the ligand is predicted correctly, while the flexible "tails" adopt other conformations to avoid steric clashes with the protein side chains. Overall, from the "self ligand to multiple receptors" experiments we can see that we would not be able to predict neither of FXR_4, FXR_15, and FXR_18 docking poses without the co-crystal protein structures even if we use the co-crystal geometries of the ligands.

Finally, we would like to mention the importance of the correct receptor sequence in the binding pocket. As it can be seen from a comparison of Table 3 with Table S1 from Supporting Information, the mean RMSD values for the experiments with the mutated proteins are lower except for exp. 1. In this experiment, docking several ligands to low-homologous receptors became less successful after mutating a big portion of their residues.

Importance of the ligand geometry

If we consider the results obtained for docking the ligand models to the co-crystal receptors (exp. 6), the spiro and sulfonamide-containing FXR_11 ligand docking poses have the lowest quality with RMSD of more than 10 Å. Figure 4 (a) shows the 2D structure of the FXR_11 ligand. As the self-docking results are close to the co-crystal poses, it seems that we have certain problems with the 3D structures of the FXR_11 ligand that we generated from its SMILES string. We should note that visually these structures look very similar to each other and to the co-crystal ligand.

Although the clusterization procedure helped to achieve a near-native pose for the ligand that we used in submission, the best scored poses of this ligand are 'flip-flopped' in comparison to the co-crystal pose resulting in high RMSD values of about 9 Å. Therefore, we re-generated its 3D structure once again in Avogadro, which lead to even worse docking pose of 10 Å RMSD and caused us to thoroughly examine the structure of this ligand. A closer look on the reference co-crystal ligand conformation revealed rather small differences in valence and dihedral angles that, however, seem to be crucial for the correct docking. In particular, we have found two regions where small changes in geometry lead to improvement of, at least, docking of the ligand generated in Avogadro to the reference receptor. More precisely, the C25 atom of the FXR_11 ligand has sp3 hybridization, although its geometry may be somehow distorted by the adjoined aromatic ring, while the value of the C26-C25-N13 angle in the reference structure shown in Fig. 4 (b) is more than 120°. However, we were able to only generate structures with this angle's values varying from 111° to 115° depending on the force-field we used (MMFF94 or UFF), optimization precision, and the algorithm (OpenBabel or Avogadro). The second part of the molecule, whose geometry seems to be important, was the N1 nitrogen and atoms connected to it that are shown in Figs. 4 (c)-(d). Here, the nitrogen atom has rather planar geometry, probably due to the vicinity of a sulfonamide group, while the structures that we generated in Avogadro were more tetrahedral. What is interesting, although OpenBabel uses the MMFF94 force-field for the ligand optimization, it produced even more flat geometries. More precisely, the normalized triple product of the bonds formed by N1 and its neighbors equals to -0.07 for the ligand generated in OpenBabel, while for the cases shown in Fig. 4 these are equal to 0.24 and 0.52 for (c) and (d), correspondingly. Making changes in these two regions of the molecule lead to an improvement of

the docking poses, as it can be seen in Figs. 4 (e)-(f) and their relative scores compared to the results of docking to 3oki and 3oof receptors. The latter poses were highly scored but had conformations very different from the reference.

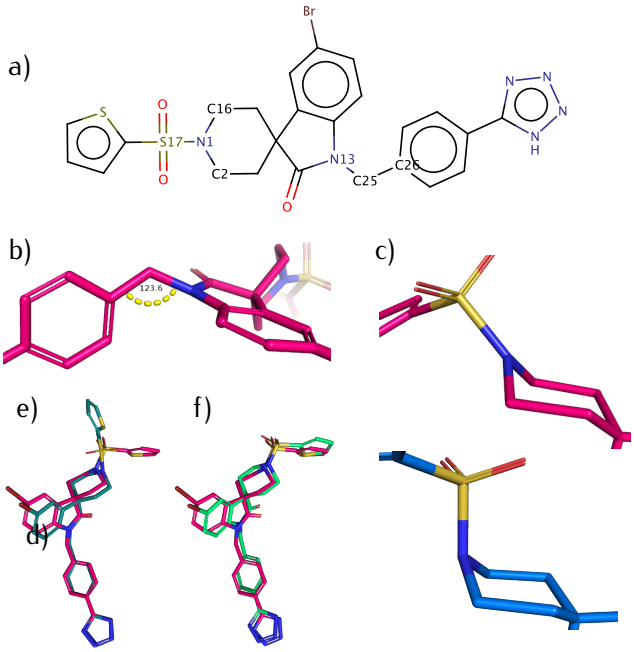


Figure 4: FXR_11 docking examples. Experimentally obtained reference structure is shown in red, the ligand that we generated in Avogadro is shown in blue. (a) 2D structure. (b) C26-C25-N13 valence angle of the co-crystal structure. (c)-(d) Planar geometry of the nitrogen atom. (e) The best-scored pose of docking of the structure with corrected C26-C25-N13 angle. (f) The best-scored pose of docking of the structure with corrected C26-C25-N13 angle and corrected N1 atom geometry.

Although this section emphasizes the importance of the initial ligand structure, we should not forget about the significance of the correct receptors' conformations. For example, the exp. 7 column of Table 3 lists docking results of the co-crystal ligand structures to multiple receptor models. Here, docking of the FXR_11 structure to 36 homologous receptors found in RCSB database did not yield any low-RMSD results.
Even though the 3gd2 receptor has a very similar conformation to the co-crystal structure of the FXR_11 receptor, one of its residues is too close to the co-crystal ligand structure, and thus we are only able to obtain a pose with RMSD value of 2.84 Å, where a half of the ligand conformation is predicted correctly, while the rest is distorted to avoid a clash. Docking with flexible side chains might be a solution in this situation, however the large size of the binding pocket entangles the choice of proper flexible residues in the blind experiment.

While we were able to detect two specific regions of the FXR_11 ligand that improve the docking results, we could not do the same for, for example, the FXR_2 ligand, although it is smaller than FXR_11. More precisely, the FXR_2 ligand has only three rotatable bonds and aromatic rings that have rather classical, or "template", geometry. Moreover, all the structures that we generated with OpenBabel and Avogadro had some differences in valence angle values from the experimental structure, and the fact that one of our structures improved the docking results seems in some sense to be rather random.

Figure 5 shows a histogram representing differences of valence angle and bond length values between the 35 FXR structures that we generated using OpenBabel and the co-crystal structures. The majority of angles and bond lengths optimized by OpenBabel appear to be close to the co-crystal values. Indeed, as it can be computed from the training data of our Knodle method [39], the mean standard deviation of the value of the valence angle for a given hybridization (sp2 or sp3, for example) is about 5°. The mean standard deviation of a bond length between atoms of the same type is about 5 pm. Figure 5 shows that only a few differences in angles and bond lengths between the modeled and co-crystal structures exceed these values. In particular, as it was already shown in Fig. 4 (b), it happens for some angles formed by atoms connected to aromatic rings, while the differences for the aromatic rings themselves never exceed 5 Å or 5 pm. Considerable differences of terminal oxygens' local geometry may occur either due to ambiguous protonation states, or due to the inaccuracies in structural alignment. Other atoms with highly diverging geometries include several sulfurs, oxygens from ester and ether functional groups, and several atoms belonging to non-aromatic cycles.

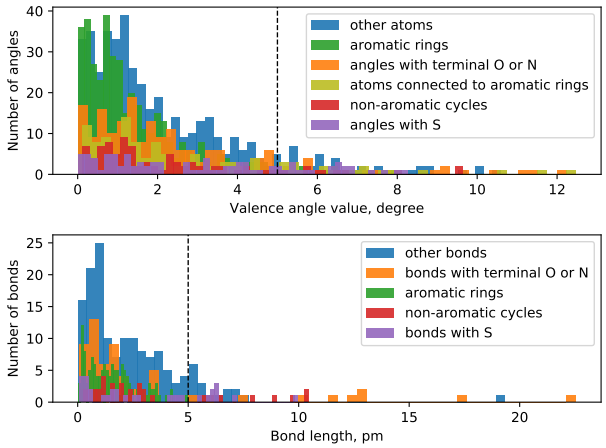


Figure 5: Valence angle and bond length differences between the ligands generated with OpenBabel and co-crystal structures. Dashed lines at 5° and 5 pm provide a visual indication of mean standard deviations of valence angles and bond length values, correspondingly, expected at the same hybridizations and bond orders.

Conclusion

Overall, the CG2 exercise provided us an excellent opportunity to assess the performance of multiple docking protocols on a set of ligands bound to the FXR receptor, whose multiple structures in different bound conformations can be found in the RCSB database. For the initial submission we used the closest receptor template from RCSB as measured by the similarity of the bound ligand and obtained the mean RMSD value of 2.96 Å.

Later on, we ran a series of additional computational experiments to evaluate the overall performance of different protocols. These include self-docking and docking of ligands that we used in the submission, several ligands generated with another algorithms, and co-crystal ligand structures to the receptors used for the submission, proteins with high sequence identity to the target apoprotein, and experimentally obtained co-crystal apoprotein structures.

We have discovered two major hurdles in the successful predictions. Definitely, the major one is finding the correct structure of the receptor. Its prediction in the bound form remains a challenge for molecular docking software developers and users. For example, in case of a known receptor structure we are able to predict the correct docking poses within a mean RMSD of about 1.26 Å, which is more than twice better than our submission results. The wide binding pocket of the FXR apoprotein allows, on the one hand, docking of ligands of different sizes without any special procedures designed to enlarge or open the pocket. On the other hand, a scoring function may prefer selecting a smaller ligand in a wrong place of the binding pocket, especially if it is the binding site for a similar ligand, as we have seen in the example of the FXR_18 target. Applying a small number of binding site mutations for the highly homologous proteins turned out to be effective. Conversely, mutating of nearly a half of the binding site for the low homologous proteins worsens the docking results. The second hurdle is the correct local geometry of the modeled ligands. Indeed, AutoDock Vina, as well as many other methods, can only sample torsion angles of the ligand, preserving the local geometry such as valence angles and bond lengths. We have shown in the example of the FXR_11 target that even small changes in the local geometry of the ligand may considerably affect the docking results. However, predicting the regions with important local geometry seems to be a difficult task in a blind experiment. Our results demonstrate that docking protocols with the correct structures of ligands allow to obtain docking poses with better mean RMSD values than other approaches. However, the choice of a correct receptor still makes the major contribution to the quality of the final docking results. We should note that the importance of 3D ligand structures is not a widely discussed topic in docking literature, and in the current study we have tried only two 3D ligand structure building algorithms. A more thorough comparison of other methods on a wider dataset will definitely provide useful results.

Overall, for docking of ligands of diverse chemical series, we would suggest to rely on the docking program and its scoring capabilities, and perform docking of each ligand to a set of multiple receptors that are homologous to the target. As it can be seen from the low RMSD results of docking to the co-crystal receptor structures, much success depends on the structures of the chosen receptor models. Therefore, since modeling conformational changes

FXR id	self docking	submission	exp. 1		exp. 2		exp. 3	exp. 4		exp. 5		exp. 6		exp. 7
			mod	old	mod	old		mod	old	mod	old	mod	old	
1	0.347	5.467	5.898		3.050		3.050	3.050		0.549		0.549		4.076
2	0.304	5.166	3.088	3.083	7.931	7.453	7.453	3.088	3.083	0.447	7.453	0.447	5.158	6.448
3	0.466	5.196	6.928	6.905	3.694	3.678	3.679	3.985	3.990	1.982	3.678	1.982	5.032	1.281
4	1.346	3.499	6.119		6.741		6.741	6.741		6.741		1.372		6.638
5	0.157	0.773	0.678		0.678		0.678	0.678		0.673		0.673		0.620
6	0.313	0.474	0.412		0.425		0.425	0.425		0.391		0.423		0.456
7	0.865	1.212	1.176		0.883		0.883	0.883		0.748		0.748		0.991
8	0.789	0.668	0.976	0.706	0.951	0.723	0.723	0.951	0.723	0.951	0.723	0.883	1.631	0.603
9	0.462	0.455	0.482		0.495		0.495	0.495		0.530		0.645		0.396
10	0.561	4.558	4.616		3.199		3.199	3.648		3.199		1.158		2.227
11	1.027	6.355	7.164	7.010	3.268	5.223	8.929	10.007	2.840	3.268	5.223	10.014	1.047	6.905
12	1.177	4.811	6.311		3.187		3.187	3.385		1.308		1.308		2.568
13	0.337	5.686	1.179		0.679		0.679	0.679		0.392		0.392		0.502
14	0.503	0.729	0.588		0.553		0.553	0.553		0.617		0.635		0.434
15	0.265	4.146	9.677		4.875		4.875	9.088		1.277		1.277		4.300
16	0.333	7.622	9.341	9.909	2.313	3.751	3.751	2.313	3.751	2.085	3.751	2.085	1.620	1.064
17	0.709	9.562	9.528	6.820	2.224	2.158	2.158	2.224	2.158	2.224	1.931	1.667	1.931	1.381
18	0.405	6.411	7.504		9.207		9.207	9.207		0.653		0.653		8.805
19	0.359	0.626	0.704		0.695		0.695	0.695		0.653		0.653		0.623
20	0.627	4.006	0.848		0.824		0.824	0.824		0.969		0.749		0.700
21	0.478	0.527	0.717		0.680		0.680	0.680		0.660		0.786		0.518
22	0.259	1.876	2.332	1.922	2.248	1.956	1.956	2.248	1.956	0.783	1.746	0.783	2.016	2.075
23	0.325	7.285	7.350	7.281	2.745	3.726	3.600	2.745	3.557	2.745	1.488	2.806	1.488	2.542
24	0.428	0.805	0.683		0.629		0.629	0.629		0.627		0.584		0.424
25	0.390	0.670	0.533		0.556		0.556	0.556		0.556		0.486		0.416
26	0.447	0.451	1.039		0.509		0.509	0.509		0.634		0.634		0.849
27	0.434	0.929	1.371		1.395		1.395	1.395		1.389		1.389		0.412
28	0.608	0.774	1.004		0.503		0.503	0.503		0.439		0.439		0.386
29	0.528	0.812	1.027		0.704		0.704	0.704		0.660		0.660		0.853
30	0.575	1.358	2.164	1.948	1.696	1.056	1.630	1.696	1.056	1.530	1.056	1.530	1.815	1.608
31	0.626	1.702	0.992	1.658	1.003	1.068	1.068	1.003	1.068	0.998	1.068	0.998	0.937	0.893
32	0.456	2.019	1.968	2.015	2.046	2.032	2.082	2.046	2.032	1.995	2.018	1.983	1.994	2.131
34	1.552	4.906	4.568	4.551	2.886	2.390	2.390	2.866	2.751	1.729	1.422	1.305	1.422	2.270
35	0.774	0.883	0.895		0.826		0.826	0.826		0.544		0.559		0.841
36	0.774	1.159	1.352		0.952		0.952	0.952		0.952		0.974		1.029
mean	0.572	2.959	3.177	3.103	2.150	2.213	2.333	2.351	2.173	1.311	1.621	1.264	1.253	1.950

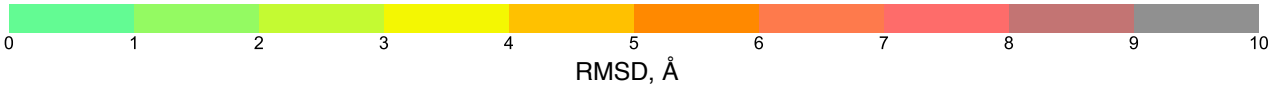


Table 3: Top-5 docking pose prediction results. The 'mod' column represents ligands that were created in Avogadro and minimized with either UFF or MMFF94 force-field. The 'old' one represents ligands that we used for the submission. All the RMSD values are computed with RDKit. The following experiments are listed in the table: self docking – docking of the co-crystal ligands to the co-crystal receptors; submission – results of the submission, where one ligand was docked to a single receptor for each of the targets; exp. 1 – "one ligand to one receptor"; exp. 2 – "one ligand to multiple receptors"; exp. 3 – "multiple ligands to multiple receptors"; exp. 4 – "one ligand to some receptors"; exp. 5 – "one ligand to multiple+self receptors"; exp. 6 – "one ligand to self receptor"; exp. 7 – "self ligand to multiple receptors". Please see the main text for more details.

of a receptor's backbone and side-chains is still a challenging problem, docking a single or multiple ligands to multiple receptors seems to be the right strategy. With such a protocol we came up with RMSD values of 2.15 – 2.33 Å depending on the chosen structures of the ligands without any visual inspection of the input structures and the obtained results.

Acknowledgement

The authors thank Vladimir Chupin from MIPT Moscow for helpful discussions during the development of the Convex-PL potential. The authors also thank Andreas Eisenbarth from the University of Kaiserslautern for the development of docking protocols. This work was partially supported by the Ministry of Education and Science of the Russian Federation (No. 6.3157.2017/PP).

References

[1] Richard D Smith, James B Dunbar Jr, Peter Man-Un Ung, Emilio X Esposito, Chao-Yie Yang, Shaomeng Wang, and Heather A Carlson. Csar benchmark exercise of 2010: combined evaluation across all submitted scoring functions. *J. Chem. Inf. Model.*, 51(9):2115–2131, 2011.

[2] Kelly L Damm-Ganamet, Richard D Smith, James B Dunbar Jr, Jeanne A Stuckey, and Heather A Carlson. CSAR benchmark exercise 2011–2012: Evaluation of results from docking and relative ranking of blinded congeneric series. *J. Chem. Inf. Model.*, 53(8):1853–1870, 2013.

[3] Sergei Grudinin, Petr Popov, Emilie Neveu, and Georgy Cheremovskiy. Predicting binding poses and affinities in the csar 2013–2014 docking exercises using the knowledge-based convex-pl potential. *J. Chem. Inf. Model.*, 56(6):1053–1062, 2015.

[4] Sergei Grudinin, Maria Kadukova, Andreas Eisenbarth, Simon Marillet, and Frédéric Cazals. Predicting binding poses and affinities for protein-ligand complexes in the 2015 d3r grand challenge using a physical model with a statistical parameter estimation. *J. Comput.-Aided Mol. Des.*, 30(9):791–804, 2016.

[5] Richard D Smith, Kelly L Damm-Ganamet, James B Dunbar Jr, Aqeel Ahmed, Krishnapriya Chinnaswamy, James E Delproposto, Ginger M Kubish, Christine E Tinberg, Sagar D Khare, Jiayi Dou, et al. Csar benchmark exercise 2013: evaluation of results from a combined computational protein design, docking, and scoring/ranking challenge. *J. Chem. Inf. Model.*, 56(6):1022–1031, 2015.

[6] Heather A Carlson, Richard D Smith, Kelly L Damm-Ganamet, Jeanne A Stuckey, Aqeel Ahmed, Maire A Convery, Donald O Somers, Michael Kranz, Patricia A Elkins, Guanglei Cui, Catherine E Peishoff, Millard H Lambert, and James B Dunbar, Jr. Csar 2014: A benchmark exercise using unpublished data from pharma. *J. Chem. Inf. Model.*, May 2016.

[7] Symon Gathiaka, Shuai Liu, Michael Chiu, Huanwang Yang, Jeanne A Stuckey, You Na Kang, Jim Delproposto, Ginger Kubish, James B Dunbar, Heather A Carlson, et al. D3r grand challenge 2015: Evaluation of protein–ligand pose and affinity predictions. *J. Comput.-Aided Mol. Des.*, 30(9):651–668, 2016.

[8] Chengfei Yan, Sam Z Grinter, Benjamin Ryan Merideth, Zhiwei Ma, and Xiaoqin Zou. Iterative knowledge-based scoring functions derived from rigid and flexible decoy structures: evaluation with the 2013 and 2014 csar benchmarks. *J. Chem. Inf. Model.*, 56(6):1013–1021, 2015.

[9] Virginie Y Martiny, François Martz, Edithe Selwa, and Bogdan I Iorga. Blind pose prediction, scoring, and affinity ranking of the csar 2014 dataset. *J. Chem. Inf. Model.*, 56(6):996–1003, 2015.

[10] Philip Prathipati and Kenji Mizuguchi. Integration of ligand and structure based approaches for csar-2014. *J. Chem. Inf. Model.*, 56(6):974–987, 2015.

[11] Sheng-You Huang, Min Li, Jianxin Wang, and Yi Pan. Hybriddock: a hybrid protein–ligand docking protocol integrating protein- and ligand-based approaches. *J. Chem. Inf. Model.*, 56(6):1078–1087, 2015.

[12] Veronica Salmaso, Mattia Sturlese, Alberto Cuzzolin, and Stefano Moro. Dockbench as docking selector tool: the lesson learned from d3r grand challenge 2015. *J. Comput.-Aided Mol. Des.*, 30(9):773–789, 2016.

[13] Woong-Hee Shin, Gyu Rie Lee, and Chaok Seok. Evaluation of galaxydock based on the community structure–activity resource 2013 and 2014 benchmark studies. *J. Chem. Inf. Model.*, 56(6):988–995, 2015.

[14] Jocelyn Sunseri, Matthew Ragoza, Jasmine Collins, and David Ryan Koes. A d3r prospective evaluation of machine learning for protein-ligand scoring. *J. Comput.-Aided Mol. Des.*, 30(9):761–771, 2016.

[15] Philip Prathipati, Chioko Nagao, Shandar Ahmad, and Kenji Mizuguchi. Improved pose and affinity predictions using different protocols tailored on the basis of data availability. *J. Comput.-Aided Mol. Des.*, 30(9):817–828, 2016.

[16] Hervé Hogues, Traian Sulea, and Enrico O Purisima. Evaluation of the wilma-sie virtual screening method in community structure–activity resource 2013 and 2014 blind challenges. *J. Chem. Inf. Model.*, 56(6):955–964, 2015.

[17] Stefano Piotto, Luigi Di Biasi, Riccardo Fino, Raffaele Parisi, Lucia Sessa, et al. Yada: a novel tool for molecular docking calculations. *J. Comput.-Aided Mol. Des.*, 30(9):753–759, 2016.

[18] Xiaolei Zhu, Woong-Hee Shin, Hyungrae Kim, and Daisuke Kihara. Combined approach of patch-surfer and pl-patchsurfer for protein–ligand binding prediction in csar 2013 and 2014. *J. Chem. Inf. Model.*, 56(6):1088–1099, 2015.

[19] Ashutosh Kumar and Kam YJ Zhang. Prospective evaluation of shape similarity based pose prediction method in d3r grand challenge 2015. *J. Comput.-Aided Mol. Des.*, 30(9):685–693, 2016.

[20] Ashutosh Kumar and Kam YJ Zhang. Application of shape similarity in pose selection and virtual screening in csardock2014 exercise. *J. Chem. Inf. Model.*, 56(6):965–973, 2015.

[21] Inna Slynko, Franck Da Silva, Guillaume Bret, and Didier Rognan. Docking pose selection by interaction pattern graph similarity: application to the d3r grand challenge 2015. *J. Comput.-Aided Mol. Des.*, 30(9):669–683, 2016.

[22] Denis Fourches, Eugene Muratov, Feng Ding, Nikolay V Dokholyan, and Alexander Tropsha. Predicting binding affinity of csar ligands using both structure-based and ligand-based approaches. *J. Chem. Inf. Model.*, 53(8):1915–1922, 2013.

[23] Regina Politi, Marino Convertino, Konstantin Popov, Nikolay V Dokholyan, and Alexander Tropsha. Docking and scoring with target-specific pose classifier succeeds in native-like pose identification but not binding affinity prediction in the csar 2014 benchmark exercise. *J. Chem. Inf. Model.*, 56(6):1032–1041, 2016.

[24] Denis Fourches, Regina Politi, and Alexander Tropsha. Target-specific native/decoy pose classifier improves the accuracy of ligand ranking in the csar 2013 benchmark. *J. Chem. Inf. Model.*, 55(1):63–71, 2014.

[25] Diogo Santos-Martins. Interaction with specific hsp90 residues as a scoring function: validation in the d3r grand challenge 2015. *J. Comput.-Aided Mol. Des.*, 30(9):731–742, 2016.

[26] Nanjie Deng, William F Flynn, Junchao Xia, RSK Vijayan, Baofeng Zhang, Peng He, Ahmet Mentes, Emilio Gallicchio, and Ronald M Levy. Large scale free energy calculations for blind predictions of protein–ligand binding: the d3r

grand challenge 2015. *J. Comput.-Aided Mol. Des.*, 30(9):743–751, 2016.

[27] Majda Misini Ignjatović, Octav Caldararu, Geng Dong, Camila Muñoz-Gutierrez, Francisco Adasme-Carreño, and Ulf Ryde. Binding-affinity predictions of hsp90 in the d3r grand challenge 2015 with docking, mm/gbsa, qm/mm, and free-energy simulations. *J. Comput.-Aided Mol. Des.*, 30(9):707–730, 2016.

[28] Sergio Ruiz-Carmona and Xavier Barril. Docking-undocking combination applied to the d3r grand challenge 2015. *J. Comput.-Aided Mol. Des.*, 30(9):805–815, 2016.

[29] Antonia SJS Mey, Jordi Juárez-Jiménez, Alexis Hennessy, and Julien Michel. Blinded predictions of binding modes and energies of hsp90-α ligands for the 2015 d3r grand challenge. *Bioorganic & Medicinal Chemistry*, 24(20):4890–4899, 2016.

[30] Matthew P Baumgartner and Carlos J Camacho. Choosing the optimal rigid receptor for docking and scoring in the csar 2013/2014 experiment. *J. Chem. Inf. Model.*, 56(6):1004–1012, 2015.

[31] Edithe Selwa, Virginie Y Martiny, and Bogdan I Iorga. Molecular docking performance evaluated on the d3r grand challenge 2015 drug-like ligand datasets. *J. Comput.-Aided Mol. Des.*, 30(9):829–839, 2016.

[32] Zhaofeng Ye, Matthew P Baumgartner, Bentley M Wingert, and Carlos J Camacho. Optimal strategies for virtual screening of induced-fit and flexible target in the 2015 d3r grand challenge. *J. Comput.-Aided Mol. Des.*, 30(9):695–706, 2016.

[33] Praveen Nedumpully-Govindan, Domen B Jemec, and Feng Ding. Csar benchmark of flexible medusadock in affinity prediction and nativelike binding pose selection. *J. Chem. Inf. Model.*, 56(6):1042–1052, 2015.

[34] Peter W Rose, Andreas Prlić, Ali Altunkaya, Chunxiao Bi, Anthony R Bradley, Cole H Christie, Luigi Di Costanzo, Jose M Duarte, Shuchismita Dutta, Zukang Feng, et al. The rcsb protein data bank: integrative view of protein, gene and 3d structural information. *Nucleic Acids Research*, 45(D1):D271–D281, 2017.

[35] Noel M O'Boyle, Michael Banck, Craig A James, Chris Morley, Tim Vandermeersch, and Geoffrey R Hutchison. Open Babel: an open chemical toolbox. *J. Cheminform.*, 3:33, 2011.

[36] Oleg Trott and Arthur J Olson. AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.*, 31(2):455–461, 2010.

[37] Garrett M Morris, Ruth Huey, William Lindstrom, Michel F Sanner, Richard K Belew, David S Goodsell, and Arthur J Olson. Autodock4 and autodocktools4: Automated docking with selective receptor flexibility. *J. Comput. Chem.*, 30(16):2785–2791, 2009.

[38] Maria Kadukova and Sergei Grudinin. Convex-PL: a novel knowledge-based potential for protein-ligand interactions deduced from structural databases using convex optimization. *J. Comput.-Aided Mol. Des.*, 9 2017.

[39] Maria Kadukova and Sergei Grudinin. Knodle: A support vector machines-based automatic perception of organic molecules from 3d coordinates. *J. Chem. Inf. Model.*, 56(8):1410–9, Aug 2016.

[40] Thomas A Halgren. Merck molecular force field. i. basis, form, scope, parameterization, and performance of mmff94. *J. Comput. Chem.*, 17(5-6):490–519, 1996.

[41] Marcus D Hanwell, Donald E Curtis, David C Lonie, Tim Vandermeersch, Eva Zurek, and Geoffrey R Hutchison. Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.*, 4(1):17, 2012.

[42] Anthony K Rappé, Carla J Casewit, KS Colwell, WA Goddard Iii, and WM Skiff. Uff,

a full periodic table force field for molecular mechanics and molecular dynamics simulations. *Journal of the American chemical society*, 114(25):10024–10035, 1992.

[43] Grzegorz M Boratyn, Alejandro A Schäffer, Richa Agarwala, Stephen F Altschul, David J Lipman, and Thomas L Madden. Domain enhanced lookup time accelerated blast. *Biology direct*, 7(1):12, 2012.

[44] Georgii G Krivov, Maxim V Shapovalov, and Roland L Dunbrack. Improved prediction of protein side-chain conformations with scwrl4. *Proteins: Struct., Funct., Bioinf.*, 77(4):778–795, 2009.

[45] Schrödinger, LLC. The PyMOL molecular graphics system, version 1.3. 2011.

[46] Sigma-Aldrich Amino Acids Reference Chart.

[47] Greg Landrum. Rdkit: Open-source cheminformatics.

[48] Hans GF Richter, GM Benson, KH Bleicher, D Blum, E Chaput, N Clemann, S Feng, C Gardes, U Grether, P Hartman, et al. Optimization of a novel class of benzimidazole-based farnesoid x receptor (fxr) agonists to improve physicochemical and adme properties. *Bioorganic & medicinal chemistry letters*, 21(4):1134–1140, 2011.