Fragment docking:
how to select the correct pose?

Célien Jacquemard, Priscila da S. F. C. Gomes, Guillaume Bret, Didier Rognan and Esther Kellenberger

Laboratory of Therapeutic Innovation, UMR-7200 CNRS University of Strasbourg Medalis Drug Discovery Center

Introduction to fragment docking

Drug discovery starts with the identification of molecules that binds the target, followed by the optimization of the binding and pharmacokinetics properties of the most promising molecule. The likelihood of success in the medicinal chemistry program is influenced by the quality of the lead compound. To that respect, fragments have a number of potential advantages over the larger drug-like compounds. In particular, fragments are easier to elaborate owing to their small size.

Fragment growing is largely facilitated by the knowledge of the binding mode to the target [1,2]. The best models come from X-ray crystallography. For targets which crystallize without difficulty, up to hundreds of three-dimensional (3D) structures are determined to enable the medicinal chemist to take decisions about optimization. When crystallography is not possible, NMR and computational approaches are good alternative methods.

Molecular docking is the most popular computational technique to predict the atomic coordinates of a protein-ligand complex [3,4]. Virtual screening of chemical libraries by high-throughput docking has shown to be a powerful tool for drug-like hit identification [5,6]. However, docking scoring functions are fairly able to prioritize the most relevant solutions in first place, making compound raking a serious limitation of the method [7]. Scoring problem is exacerbated for small ligands [8–11].

Several benchmarks and docking challenges have been proposed by the docking community to aid computational chemists to refine their practices and rescoring methods [7,12–15]. One successful approach to pose filtering is based on the principle that the knowledge of the binding mode of a ligand can help the prediction of the binding mode of another compound. A state-of-the-art rescoring method is GRIM [16], which converts protein-ligand interactions into graphs and score docking solutions by maximal similarity of predicted interaction graphs to that already visited in the Protein Data Bank (PDB) [17]. However there is a caveat when rescoring fragment poses: while drug-like molecule binding modes yield complex graphs that are easy to compare, fragment binding modes yield simple and potentially unselective graphs that are prone to generate irrelevant 3D-alignment of ligand/protein complexes upon graphs comparison.

In this tutorial, we will use a novel knowledge-based method which is applicable to fragment scoring. The method, named LID for Local Interaction Density, builds a consensus interaction map from all experimental observations. All reference complexes are 3D-aligned in a common frame, then the density of interactions is encoded into a grid. Docking pose are scored from the matching of fragment atoms to the grid.

Here we asked whether information on binding mode can help selecting the correct fragment pose of fragment docked into human carbonic anhydrase II (HCA II).

Human carbonic anhydrase II

The carbonic anhydrases comprise several well studied classes among which is the α class, including the human and animal enzymes. HCA II is the most extensively studied of these enzymes and it is essential for bone resorption and to maintain acid-base balance in kidney.

HCA II catalyzes the reversible hydration of carbon dioxide (CO_2), therefore helping its transport out of tissues. The interconversion of carbon dioxide and water to bicarbonate and protons (or vice versa) involves a zinc ion. The catalytic mechanism [18] and the intramolecular proton transfer is widely speculated to occur between the zinc-bound solvent and the side chain of His 64, that sits on the

rim of the active site, through a network of ordered active site waters to an 'in' conformer of His 64 (pointing towards the active site) which then rotates to an 'out' conformer (pointing towards the bulk solvent)[19,20] (Fig. 1).
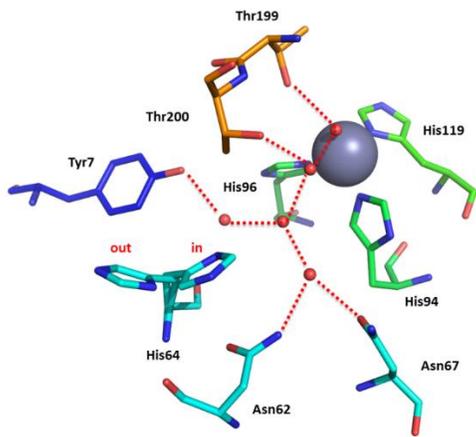

Figure 1: The active site of HCA II from the data of Fisher et al., 2005. The side chain of His64 is shown in both the inward and outward conformations. The red spheres represent oxygen including the oxygen atoms of ordered water molecules involved in the proton transfer. Image generated with PyMOL[25].

CA inhibitors (CAIs) are currently used in the treatment of glaucoma, high blood pressure, epilepsy, altitude sickness, gastric and duodenal ulcers, neurological disorders and osteoporosis[21,22]. CAIs can be divided into two classes: those who bind to the enzyme active site anchoring to the catalytic zinc ion, and those who do not interact directly with the zinc. Four groups of zinc binding inhibitors have been studied by X-ray crystallography: ureates/hydroxamates, the mercaptophenols, the metal-complexing anions and the sulfonamides with their bioisosteres, such as sulfamates and sulfamides[23].

To date, the sulfonamide group is the most largely used zinc binding function for the design of CAIs and constitute the majority of the clinically used drugs. One of the earliest and most frequently prescribed CAI is acetazolamide[24] (AZM; 5-acetamido-1,3,4-thiadiazole-2-sulfonamide) (Fig. 2), marketed as Diamox and indicated to treat glaucoma.

Acetazolamide has a Ki of 10 nM for CA II. Sulfonamides mechanism of binding lies on the coordination of the negatively charged deprotonated sulfonamide nitrogen to the catalytic Zn^{2+}, with consequent substitution of the zinc-bound water molecule, and by two H-bonds of the sulfonamide moiety with residue Thr199[22,23] (scene 4, at *Intro.pse* and Fig. 2).
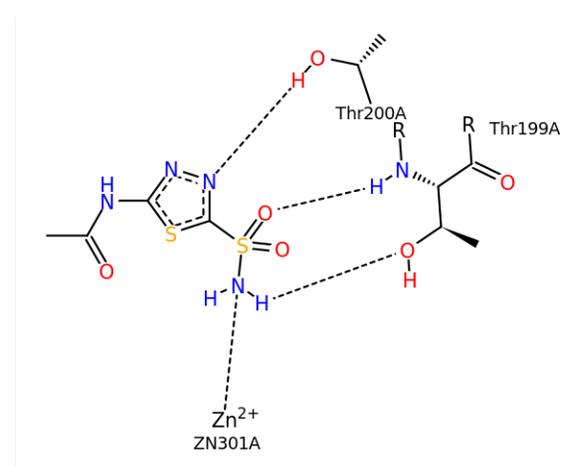

Figure 2: 2D chemical representation of the drug acetazolamide (AZM) bound to the active site residues (Thr199 and Thr200) of HCA II. Dashed lines represent H-bonds/metal coordination. PDB ID: 3HS4. Image generated with Poseview[26].

INSTRUCTION: OBSERVATION OF HCA II 3D-STRUCTURE

At the PyMOL session file *Intro.pse*, at folder *PyMOL_sessions*, you can notice at scenes 1 and 2 the HCA II structure and active site, with the conserved water molecules and the zinc ion that participates on the proton transfer. In addition, notice His64 with its side-chain in both states (in/out). Scene 3 shows the superposition of some of the many available protein structures for HCA II, notice the conservation of the active site residues side-chains.

Purpose of the tutorial

Here we compare pose ranking by the empirical Chemplp score and by the LID method. We consider two possible contexts:

1. There are several reference binding modes to HCA II. All the reference ligands are drug-like molecules sharing a common substructure interacting with the protein binding site, here a sulfonamide group interacting with the zinc cation.
2. There are no reference ligands and therefore we use as references all the additive molecules crystallized with the apo-protein (e.g. glycerol which is a cryoprotectant).

The rescoring exercise is exemplified with two chemically unrelated fragments docked into HCA II (Fig. 3). The first fragment, 1H-benzimidazole-2-sulfonamide (EVE), contains the sulfonamide group required to anchor the zinc cation of HCA II. Note that the sulfonamide group of EVE has been negatively charged to facilitate docking. The second ligand, 2-(4-phenylmethoxyphenyl)ethanoic acid (IO2) has a different chemical structure, which is not similar to any of the reference drug-like ligand structures.
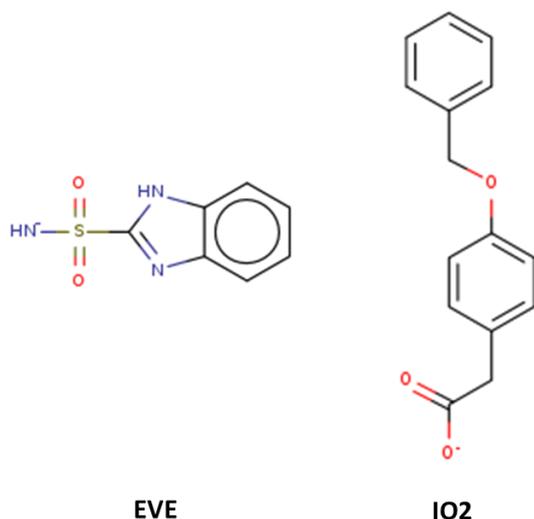

Figure 3: 2D chemical structures of fragments 1H-benzimidazole-2-sulfonamide (EVE) and 2-(4-phenylmethoxyphenyl)ethanoic acid (IO2). The ionization states of docking input are shown.

INSTRUCTION: OBSERVATION THE REFERENCE LIGANDS AND THEIR BINDING MODE TO HCA II

Open the file *References.html* at folder *html* into your favorite web browser and observe the chemical structure of the 48 drug-like ligands and the 10 additives crystallized with HCA II. Note the number and type of interactions made by the two ligand classes.

Go back to the PyMOL session file *Intro.pse*. Scene 4 and 5 show the co-crystallized drug-like ligands (magenta) and additives (green), respectively, that will be used as reference ligands on the exercises. Use the keyboard arrows to move through the different ligands. The interaction points are also represented for each drug-like and additive ligand.

Although the 48 drug-like ligands all include a sulfonamide group, they represent a chemically diverse ensemble. Their molecular weight ranges from 304 to 473 Da. By comparison, fragment molecular weights are significantly smaller (197 and 242 Da for EVE and IO2, respectively). Note that the AZM drug shown in Figure 1 is also a fragment (molecular weight = 222 Da), which has a very high ligand efficiency. There are 7 different additives crystallized with HCA II apo-enzyme: carbonic dioxide, glycerol and small acids (cyanic, acetic, carbonic, sulfuric and perchloric acids) which are in their deprotonated form at physiological pH.

The drug-like ligands all anchor to the zinc cation of HCA II. Many of them also sits into the hydrophobic pocket formed by Phe 131, Val 135 and Pro 202. Nevertheless, considered as a whole, the drug-like ligands set explore the full protein active site, each ligand being engaged in both hydrophobic contact and at least one hydrogen bond. By contrast additives are principally clustered near the zinc cation. They also reveal five H-bonded amino acids (Asn 62, His 64, His 94, Asn 67, and Thr 199) and a small hydrophobic patch near Val 121.

Material and methods

For the sake of simplicity, you are provided with the docked poses pre-generated with PLANTS. In addition, the interaction points for references and docked poses were also previously calculated using IChem. Folder organization with a description of the inputs and output files is available at Table 1.

Materials

Inputs

- Reference interaction files: 48 for drug-like molecules and 10 for additive molecules;
- Docking interaction files: generated for each docked conformation of EVE and IO2 into the multiple PDB structures.

Methods

1. Preparation of PDB structures. Hydrogens were added to all protein-ligand complexes using Protoss[28]. Protoss identifies missing hydrogen atoms in a protein-ligand complex by a detection of free valences of all heavy atoms; Protein structures preparation also involved the removal of all bound molecules, including water; Protein structures are aligned with the CE algorithm[29] against a reference (determined by a hierarchical clustering, average linkage, Cα-RMSD as distance);

2. Docking. Docking was performed using PLANTS[30], v1.2 with the following parameters: scoring function Chemplp; accuracy set to "speed1" (most accurate); binding site center -0.631 x 4.763 x 13.579 Å; binding site radius 10.0 Å, cluster_rmsd 2.0 Å, 10 poses.

3. Detection of ligand/protein interactions. Interaction graphs were generated using the module GRIM, calculated with IChem v5.9.2. Each interaction is encoded by a triplet of pseudo-atoms: one matching the protein atom, one matching the ligand atom and one at the center of these two. Depending on ligand and protein atom typing, distance and angle, five interaction types could be detected (Hydrophobic, π-stacking, hydrogen bond, salt bridge and metal chelation). IChem is freely distributed to academia upon license request to Dr. Didier Rognan (mail to rognan@unistra.fr).

4. Rescoring of docking poses. LID rescoring involves two steps:

(1) Generation of the hashed map. In practice, space is discretized into 0.1 Å-binned grid. Each cube is annotated with the count of all the reference pseudo-atoms it contains (one annotation per interaction type and mode). The cube density is then computed considering annotation of all neighboring cubes within a radius of 0.5 Å.

(2) Rescoring of docking poses by placing predicted pseudo-atoms in the grid and summing density of cubes which are hit.

Free energy of binding were predicted using HYDE[31]. HYDE estimates binding free energy based on two terms for dehydration and hydrogen bonding only. The essential feature of this scoring function is the integrated use of log P-derived atomic increments for the prediction of free dehydration energy and hydrogen bonding energy[31]. HYDE is developed by BioSolveIt and a trial license for HYDE can be requested upon registration at their website (https://www.biosolveit.de/).

5. Evaluation of rescoring. The performance of the rescoring method is evaluated by the calculation of the Root Mean Square Deviation (RMSD) of the selected poses against the 3D coordinates of the crystallized target ligand. A prediction is considered good when RMSD between the docked pose and the crystallographic pose is inferior to 2.0 Å. RMSD values were calculated using all non-hydrogen atoms.

Program files and installation

For this training, we will use PyMOL, an molecular visualization system on an open-source foundation, maintained and distributed by Schrödinger. To install PyMOL:

Windows
Download PyMOL for Windows: https://PyMOL.org/installers/PyMOL-2.1.1_0-Windows-x86_64.exe. Do not check the cases for "Advanced options"
Before starting the tutorial, open the *Windows Command prompt* and type:
doskey python=C:\Users\<username>\PyMOL\python.exe $*

For those who do not want to install PyMOL and prefer using another molecular visualization program, be aware that a python-interpreter is needed to run LID. It is possible to download the installation file: https://www.python.org/ftp/python/2.7.15/python-2.7.15.amd64.msi

Linux
cd *<target directory>* # a PyMOL directory will be created here
wget https://PyMOL.org/installers/PyMOL-2.1.1_0-Linux-x86_64.tar.bz2
tar -xjf PyMOL-2.1.1_0-Linux-x86_64.tar.bz2

MacOS
Download PyMOL for MacOS
https://PyMOL.org/installers/PyMOL-2.1.1_0-MacOS.dmg

Folders organization

lid_tutorial	Content	Description
docs	Tutorial.pdf	Text file for the tutorial
html	References.html	2D visualization of the drug-like and additive references
outputs	intgrid/ lid/	Exercises outputs will be written here
progs	rescoring.py, help.txt	The script necessary to generate maps and run LID and a text file with the command lines
PyMOL_sessions	intro.pse, exercise_1.pse, exercise_2.pse	PyMOL session files.
interactions	references/[additives/druglike]/PDBID_ints.mol2 docked/ [EVE/IO2]/ PDBID_conf_[1to10]_ints.mol2	Input interaction files used for LID scoring
results	intgrid/druglike_all.iva, druglike_all.gri, druglike_all.mol2 intgrig/additive_all.iva, additive_all.gri, additive_all.mol2 lid/ [druglike/additive]_all_EVE.csv, [druglike/additive]_all_IO2.csv hyde/hyde_EVE.csv, hyde_IO2.csv	Pre-calculated exercises and HYDE outputs
xray_structures	additives/PDBID_[protein/ligand].mol2 druglikes/ PDBID_[protein/ligand].mol2	Input structures used for docking and interaction points generation
docking	inputs/PDBID_protein.mol2, EVE.mol2, IO2.mol2 outputs/[EVE/IO2]/PDBID/PDBID_conf_[1to10]_protein.mol2, outputs/[EVE/IO2]/PDBID/PDBID_conf_[1to10]_ligand.mol2	Inputs used for docking; output structures generated with PLANTS for each fragment and interaction points files generated for each docked pose

Exercise 1: EVE pose prediction

The purpose of this tutorial is the rescoring of docking generated poses using the LID method.

In this first exercise, we are going to predict the binding mode of the sulfonamide containing fragment EVE (Fig. 3), using two different scenarios:

(1) using drug-like sulfonamide containing molecules as references;
(2) using miscellaneous additive molecules as references.

INSTRUCTION: OBSERVATION OF CRYSTALLOGRAPHIC STRUCTURE AND DOCKING POSES

Open the PyMOL session file *exercise_1.pse* located at *PyMOL_sessions*. Scenes 1 and 2 shows you the crystallographic structure of the fragment EVE (PDB ID: 3S72), represented as sticks colored in cyan, complexed with HCA II. Scene 3 shows you the interaction points generated with IChem.

At scene 4 you can see the top ranked pose predicted by the docking scoring function Chemplp (colored in green) and the pose with the best RMSD (colored in marine blue). At scene 5 you can see all the docked poses colored by the Chemplp score, ranging from red to blue (worse to better).

Observing the crystal structure of the complex between EVE and HCAII, you can notice the interactions made with the active site residues, and the metal coordination by the sulfonamide and histidine residues. You can see that it accounts not only the polar interactions (hydrogen bonds and metal coordination) but also hydrophobic contacts with residue Leu 197.

The docking proposed multiple binding modes of EVE in HCA II. Chemplp scoring function did not identify the best solution as the top ranked pose.

INSTRUCTION: RESCORING POSES

Go to the root folder (*lid_tutorial/*)

a) Generate the interaction map from 48 drug-like sulfonamide.

Execute the scripts:

Note: for Windows users, use "\" instead of "/".
```
python progs/rescoring.py intgrid
 –r druglike
```
AND
```
python progs/rescoring.py intgrid
 –r additive
```

b) Score the docking poses

Execute the scripts:
```
python progs/rescoring.py lid
 –p EVE
 -r druglike
```
AND
```
python progs/rescoring.py lid
 –p EVE
 -r additive
```

c) Visualize the results at the *outputs/lid* folder

Take a look at the output files *druglike_all_EVE.csv* and *additivs_all_EVE.csv*.

At the file *druglike_all_EVE.csv* you can see the ligands named according to their docking rank (conf_01,02, etc) and re-classified by their LID score, followed by the corresponding RMSD value in reference to the crystallographic structure of EVE. You can notice that LID is capable of retrieving a near-native conformation of the target ligand: the best pose, 3N0N_EVE_conf_03, with score 0.45 has a RMSD of 0.54 Å in comparison to the crystal structure. Also, see that in the most cases, the top ranked pose predicted by docking (conf_01) are not the ones with the lowest RMSD to the crystallographic structure.

As for the drug-like scenario, LID is capable of finding the correct pose also by using as references the additive molecules. At the file *additives_all_EVE.csv*, you see that the best pose, 2H15_EVE_conf_06, with score 0.57 has an RMSD of 0.293 Å in comparison to the crystal structure. This exercise highlights the use of additive molecules when there are no other small molecules complexed with the protein of interest.

INSTRUCTION: OBSERVATION OF POSES RANKING BY LID

Go back to the PyMOL session, at scenes 6 and 7 you can see all the generated docking poses for EVE, colored by the LID score, using the drug-like and additive references, respectively. They are colored ranging from red to blue (worse to better). Higher the LID score, the closest are the poses to the crystallographic structure of EVE, showing the powerfulness of the method.

At scene 8 you can find the top ranked pose predicted by LID, using drug-like references (colored in magenta), the top ranked pose predicted by LID, using additives as references (colored in orange) to compare with the lowest RMSD and crystallographic poses.

Exercise 2: IO2 pose prediction

In this exercise, we are going to predict the binding mode of the fragment IO2. Like in Exercise 1, we are going to use drug-like and additive molecules as references in the rescoring procedure. However, in this case the drug-like references don't share structural similarity with IO2.

INSTRUCTION: OBSERVATION OF CRYSTALLOGRAPHIC STRUCTURE AND DOCKING POSES

Open the PyMOL session file *exercise_2.pse* located at *PyMOL_sessions*. Scenes 1 and 2 shows you the crystallographic structure of the fragment IO2 (PDB ID: 5FLQ), represented as sticks colored in cyan, complexed with HCA II. Scene 3 shows you the interaction points generated with IChem.

At scene 4 you can see the top ranked pose predicted by the docking scoring function Chemplp (colored in green) and the pose with the best RMSD (colored in marine blue). At scene 5 you can see all the docked poses colored by the Chemplp score, ranging from red to blue (worse to better).

You can notice the interactions made by IO2 with the Zn ion through its carboxylic portion. Apart from the metal coordination, the interactions are predominately hydrophobic, with residues Val 134 and Leu 197.

The docking proposed multiple binding modes of IO2 in HCA II. Again, the docking scoring function failed in correctly ranking the best solution as the top ranked pose.

INSTRUCTION: RESCORING POSES

Note that the maps required for LID scoring have already been computed during exercise 1.

Go to the root folder (*lid_tutorial/*) and execute the scripts:

```
python progs/rescoring.py lid
 –p IO2
 -r druglike
```

AND

```
python progs/rescoring.py lid
 –p IO2
 -r additive
```

Looking at the file *druglike_all_IO2.csv*, you can notice that the best pose 4IWZ_IO2_conf_06, with score 0.47 has low RMSD value (1.24 Å) in comparison with the crystallographic IO2 conformation. Nevertheless, in comparison with the rescoring for EVE, there are lower RMSD poses that were not rescored as the top ranked poses.

The same can be observed for rescoring using additives. At the file additives_all_IO2.csv, you see that the top ranked pose, 3N4B_IO2_conf_08 with score 0.52 has a RMSD of 3.386 Å, while the following poses in the rank have lower RMSD values (under 2.0 Å).

INSTRUCTION: OBSERVATION OF POSES RANKING BY LID

Go back to the PyMOL session, at scene 6 and 7 you can see all the generated docking poses for EVE, colored by the LID score, using the drug-like and additive references, respectively. They are colored ranging from red to blue (worse to better).

Poses bearing a high LID score (blue) show all the ionic interaction between IO2 carboxylic acid group and the Zn ion, yet position of phenyl rings is variable. The drug-like map positively scores the well-positioned aromatics rings, but also tilted ones because in this pocket sub-site the map is mainly populated by hydrophobic points, which do not encode directional interaction by contrast to aromatic points. In the additive map the information in this subpocket is limited, therefore not sufficient to correctly place an aromatic group in this part of the pocket.

In order to identify the correct poses, we are going to combine LID with another rescoring method, based on the calculation of the free energy of binding. We took the three top ranked poses predicted by LID and submitted to the energy estimation. The output file (*hyde_IO2.csv*) is at the *outputs* folder. By using a combined rescoring approach, the pose previously ranked at the third position by LID, now is the best predicted pose, according to HYDE.

INSTRUCTION: OBSERVATION OF POSES RANKING BY LID AND HYDE

At scene 8 of PyMOL session you can find the top ranked pose predicted by LID, using drug-like references (colored in magenta) and the best three poses predict by LID, using additive references (colored from orange to yellow, respectively) and the best pose predicted by HYDE (colored in wheat).

The top ranked LID pose using additives (orange) has a larger deviation from the crystal structure when compared to the poses 2 and 3 by LID ranking. The best pose predicted using HYDE corresponds to pose 3 at LID ranking, representing a good match with the crystallographic conformation of IO2 (cyan).

Conclusion

LID is efficient in pose prediction and can be used for scaffold hoping in a medicinal chemistry campaign. If experimental information on binding mode is scarce, LID can still be advantageously used to prioritize poses for scoring using a slower but more accurate approach such as free energy calculation.

Bibliography

(1)   Bienstock, R. J. Computational Methods for Fragment-Based Ligand Design: Growing and Linking; Humana Press, New York, NY, 2015; pp 119–135.

(2)   Lamoree, B.; Hubbard, R. E. Current Perspectives in Fragment-Based Lead Discovery (FBLD). *Essays Biochem.* **2017**, *61* (5), 453–464.

(3)   Brooijmans, N.; Kuntz, I. D. Molecular Recognition and Docking Algorithms. *Annu. Rev. Biophys. Biomol. Struct.* **2003**, *32*, 335–373.

(4)   Rognan, D. The Impact of in Silico Screening in the Discovery of Novel and Safer Drug Candidates. *Pharmacol. Ther.* **2017**, *175*, 47–66.

(5)   Ripphausen, P.; Nisius, B.; Peltason, L.; Bajorath, J. Quo Vadis, Virtual Screening? A Comprehensive Survey of Prospective Applications. *J. Med. Chem.* **2010**, *53* (24), 8461–8467.

(6)   Zhu, T.; Cao, S.; Su, P.-C.; Patel, R.; Shah, D.; Chokshi, H. B.; Szukala, R.; Johnson, M. E.; Hevener, K. E. Hit Identification and Optimization in Virtual Screening: Practical Recommendations Based on a Critical Literature Analysis. *J. Med. Chem.* **2013**, *56* (17), 6560–6572.

(7)   Smith, R. D.; Dunbar, J. B.; Ung, P. M.-U.; Esposito, E. X.; Yang, C.-Y.; Wang, S.; Carlson, H. A.; Carlson, H. A. CSAR Benchmark Exercise of 2010: Combined Evaluation across All Submitted Scoring Functions. *J. Chem. Inf. Model.* **2011**, *51* (9), 2115–2131.

(8)   Marcou, G.; Rognan, D. Optimizing Fragment and Scaffold Docking by Use of Molecular Interaction Fingerprints. *J. Chem. Inf. Model.* **2007**, *47* (1), 195–207.

(9)   Hubbard, R. E.; Chen, I.; Davis, B. Informatics and Modeling Challenges in Fragment-Based Drug Discovery. *Curr. Opin. Drug Discov. Devel.* **2007**, *10* (3), 289–297.

(10)   Klebe, G. Virtual Ligand Screening: Strategies, Perspectives and Limitations. *Drug Discov. Today* **2006**, *11* (13–14), 580–594.

(11)   Chen, Y.; Shoichet, B. K. Molecular Docking and Ligand Specificity in Fragment-Based Inhibitor Discovery. *Nat. Chem. Biol.* **2009**, *5* (5), 358–364.

(12)   Carlson, H. A.; Smith, R. D.; Damm-Ganamet, K. L.; Stuckey, J. A.; Ahmed, A.; Convery, M. A.; Somers, D. O.; Kranz, M.; Elkins, P. A.; Cui, G.; et al. CSAR 2014: A Benchmark Exercise Using Unpublished Data from Pharma. *J. Chem. Inf. Model.* **2016**, *56* (6), 1063–1077.

(13)   Li, Y.; Han, L.; Liu, Z.; Wang, R. Comparative Assessment of Scoring Functions on an Updated Benchmark: 2. Evaluation Methods and General Results. *J. Chem. Inf. Model.* **2014**, *54* (6), 1717–1736.

(14)   Slynko, I.; Da Silva, F.; Bret, G.; Rognan, D. Docking Pose Selection by Interaction Pattern Graph Similarity: Application to the D3R Grand Challenge 2015. *J. Comput. Aided. Mol. Des.* **2016**, *30* (9), 669–683.

(15)   da Silva Figueiredo Celestino Gomes, P.; Da Silva, F.; Bret, G.; Rognan, D. Ranking Docking Poses by Graph Matching of Protein-Ligand Interactions: Lessons Learned from the D3R Grand Challenge 2. *J. Comput. Aided. Mol. Des.* **2018**, *32* (1), 75–87.

(16)   Desaphy, J.; Raimbaud, E.; Ducrot, P.; Rognan, D. Encoding Protein–Ligand Interaction Patterns in Fingerprints and Graphs. *J. Chem. Inf. Model.* **2013**, *53* (3), 623–637.

(17)   Berman, H. The Protein Data Bank: A Retrospective and Prospective. *Biophys. J.* **2000**, *78* (1), 267a–267a.

(18)   Lindskog, S. Structure and Mechanism of Carbonic Anhydrase. *Pharmacol. Ther.* **1997**, *74* (1), 1–20.

(19)   Tu, C. K.; Silverman, D. N.; Forsman, C.; Jonsson, B. H.; Lindskog, S. Role of Histidine 64 in the Catalytic Mechanism of Human Carbonic Anhydrase II Studied with a Site-Specific Mutant. *Biochemistry* **1989**, *28* (19), 7913–7918.

(20)   Fisher, S. Z.; Maupin, C. M.; Budayova-Spano, M.; Govindasamy, L.; Tu, C.; Agbandje-McKenna, M.; Silverman, D. N.; Voth, G. A.; McKenna, R. Atomic Crystal and Molecular Dynamics Simulation Structures of Human Carbonic Anhydrase II: Insights into the Proton Transfer Mechanism [†,‡]. *Biochemistry* **2007**, *46* (11), 2930–2937.

(21)   Supuran, C. T. Carbonic Anhydrases: Novel Therapeutic Applications for Inhibitors and Activators. *Nat. Rev. Drug Discov.* **2008**, *7* (2), 168–181.

(22)   Aggarwal, M.; Kondeti, B.; McKenna, R. Insights towards Sulfonamide Drug Specificity in α-Carbonic Anhydrases. *Bioorg. Med. Chem.* **2013**, *21* (6), 1526–1533.

(23)   Alterio, V.; Di Fiore, A.; D'Ambrosio, K.; Supuran, C. T.; De Simone, G. Multiple Binding Modes of Inhibitors to Carbonic Anhydrases: How to Design Specific Drugs Targeting 15 Different Isoforms? *Chem. Rev.* **2012**, *112* (8), 4421–4468.

(24)   BREININ, G. M.; GORTZ, H. Carbonic Anhydrase Inhibitor Acetazoleamide (Diamox); a New Approach to the Therapy of Glaucoma. *AMA. Arch. Ophthalmol.* **1954**, *52* (3), 333–348.

(25)   DeLano, W. L.; Lam, J. W. PyMOL: A Communications Tool for Computational Models. *Abstr. Pap. Am. Chem. Soc.* **2005**, *230*, U1371–U1372.

(26)   Stierand, K.; Rarey, M. From Modeling to Medicinal Chemistry: Automatic Generation of Two-Dimensional Complex Diagrams. *ChemMedChem* **2007**, *2* (6), 853–860.

(27)   Desaphy, J.; Da Silva, F.; Rognan, D. IChem. 2014.

(28)   Bietz, S.; Urbaczek, S.; Schulz, B.; Rarey, M.; Berman, H.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.; Weissig, H.; et al. Protoss: A Holistic Approach to Predict Tautomers and Protonation States in Protein-Ligand Complexes. *J. Cheminform.* **2014**, *6* (1), 12.

(29)   Shindyalov, I. N.; Bourne, P. E. Protein Structure Alignment by Incremental Combinatorial Extension (CE) of the Optimal Path. *Protein Eng. Des. Sel.* **1998**, *11* (9), 739–747.

(30)   Korb, O.; Stützle, T.; Exner, T. E. PLANTS: Application of Ant Colony Optimization to Structure-Based Drug Design; Springer, Berlin, Heidelberg, 2006; pp 247–258.

(31)   Schneider, N.; Lange, G.; Hindle, S.; Klein, R.; Rarey, M. A Consistent Description of HYdrogen Bond and DEhydration Energies in Protein-Ligand Complexes: Methods behind the HYDE Scoring Function. *J. Comput. Aided. Mol. Des.* **2013**, *27* (1), 15–29.