Docking rigid macrocycles using Convex-PL, AutoDock Vina, and RDKit in the D3R Grand Challenge 4

Maria Kadukova, Vladimir Chupin, Sergei Grudinin

Noname manuscript No.
(will be inserted by the editor)

Docking rigid macrocycles using Convex-PL, AutoDock Vina, and RDKit in the D3R Grand Challenge 4

Maria Kadukova · Vladimir Chupin · Sergei Grudinin

Received: date / Accepted: date

Abstract The D3R Grand Challenge 4 provided a brilliant opportunity to test macrocyclic docking protocols on a diverse high-quality experimental data. We participated in both pose and affinity prediction exercises. Overall, we aimed to use an automated structure-based docking pipeline built around a set of tools developed in our team. This exercise again demonstrated a crucial importance of the correct local ligand geometry for the overall success of docking. Starting from the second part of the pose prediction stage, we developed a stable pipeline for sampling macrocycle conformers. This resulted in the subangstrom average precision of our pose predictions. In the affinity prediction exercise we obtained average results. However, we could improve these when using docking poses submitted by the best predictors. Our docking tools including the Convex-PL scoring function are available at `https://team.inria.fr/nano-d/software/`.

keywords : protein-ligand docking; ensemble docking; macrocycle modeling; Convex-PL; conformer generation; D3R; Drug Design Data Resource; scoring function;

Introduction

The Drug Design Data Resource (D3R, `www.drugdesigndata.org`) is a community initiative that hosts multiple blind challenges dedicated to modeling of proteins-ligand association events. Two subchallenges were suggested this time. Subchallenge 1 was focusing on pose and affinity predictions for the ligands binding the beta secretase 1 (BACE) receptor. In Subchallenge 2, participants were asked to predict the affinities of ligands that bind the cathepsin S (CatS) protein, which has already been a target of the previous Grand Challenge 3. Our team has only participated in the Subchallenge 1, which was divided into two stages, Stage 1 and Stage 2. The goal of Stage 1 was to predict the correct binding poses of the ligands. Later on, Stage 2 targeted affinity or free binding energy estimation for a larger set of ligands (compared to ligands in Stage 1). Following the ideas of the previous Grand Challenge 3, Stage 1 was split into Stage 1a and Stage 1b, where Stage 1b was a self-docking exercise allowing to utilize the revealed co-crystal receptor structures. It was also possible to participate in the affinity prediction in both substages of Stage 1. However, we only took part in pose prediction parts of Stage 1 substages, and in Stage 2.

This challenge provided interesting examples of macrocycle docking. Macrocycles are often described as large non-peptidic cyclic molecules. Modeling of cyclic molecules generally poses multiple computational tasks related to the preservation of molecular topology upon sampling of cycle conformations. When doing the sampling of cycles in torsion coordinates, one often has to solve the loop closure problem. There are efficient sampling methods specifically developed for cyclic peptides [1]. However, to the best of our knowledge, there are no free [2] methods for macrocycle sampling in torsion coordinates, which are essential for computationally efficient docking protocols.

M. Kadukova, S. Grudinin,
Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France
Tel.: +33 4 38 78 16 91

M. Kadukova, V. Chupin
Moscow Institute of Physics and Technology, 141700 Dolgoprudniy, Russia
E-mail: Sergei.Grudinin@inria.fr

In both stages of the exercise, we addressed the macrocycle docking problem using the classical fully structure-based sampling approach in torsional coordinates. This method keeps all the molecular cycles rigid. Therefore, we had to generate multiple starting conformations of each macrocycle. However, the cycle conformations we used in Stage 1a had unfavorable stereochemistry, which resulted in rather average RSMD values of our predictions. In the subsequent stages, we guided the cycle conformational sampling using additional constraints from the geometry of cyclic ligands crystallized with homologous receptors. This approach helped us to obtain low-RMSD predictions in Stage 1b. We have also participated in Stage 2, where we could only obtain average affinity prediction results.

Docking strategies in previous exercises

Several major docking challenges were organized during the past five years, namely CSAR 2013 [3], CSAR 2014 [4], D3R 2015-2016 [5], D3R Grand Challenge 2 [6], and D3R Grand Challenge 3 [7]. Some of them were remarkable for the exercise design or specific features of the receptor or ligands. For example, in Phase 1 of CSAR 2013 exercise participants were asked to find the best protein sequence that binds with the same compound, which involved extensive homology modeling. The target protein of the D3R Grand Challenge 2 was a flexible farnesoid X receptor (FXR). Its flexibility caused difficulties in pose predictions of several ligands, especially those of chemical series unrepresented in the crystallized homologous structures from the Protein Data Bank (PDB) [8]. Subchallenge 1 of D3R Grand Challenge 3 was focused on docking of chemically diverse ligand molecules to the CatS receptor. Although the receptor itself was fairly rigid, and a considerable number of homologous structures were available in the PDB, docking to its wide binding pocket exposed to the solvent turned out to be quite challenging for many classical structure-based approaches. The most successful strategies of ligand pose prediction for the CatS protein were structure-based methods with search space restricted with respect to known ligand structures crystallized with homologous proteins [9–13]. Two of these submissions included 3D similarity-based ligand placement into the binding pocket with a subsequent optimization of the ligand and the receptor sidechains conformations [9, 10]. Knowledge of ligand locations in homologous proteins can also be directly included into the scoring function used in docking [13]. Participants also reported on additional molecular dynamics-based refinement that improved the pose prediction quality [9, 14]. Explicit water molecules might be very important for proper estimation of interactions with the wide binding pockets [11]. Novel graph-based features for binding free energies prediction were proposed [12]. The two latest Grand Challenges are also remarkable for the first demonstrations of the 3D convolutional neural network-based methods [15]. Other approaches included molecular dynamics-based sampling and thermodynamic averaging [16] and implicit ligand theory [17] for binding free energy predictions.

Challenge data

BACE is a transmembrane aspartic-acid protease that is responsible for the cleavage of the amyloid precursor protein. This leads to amyloid-β peptide formation [18]. Beta amyloid is the main component of amyloid plaques found in brains of Alzheimer's disease patients, therefore activity regulation of beta-secretase is one of the promising Alzheimer's treatment strategies [19].

BACE substrate is normally a polypeptide in the extended β strand conformation. Potential BACE inhibitors are designed to mimic this property, which can be achieved with macrocyclization [20]. BACE binding pocket contains several sub-sites, which are partially or totally occupied by the inhibitor [21, 22]. One of the types of aspartic protease inhibitors are hydroxyethylamine-containig compounds, binding with hydrogen bonds to the aspartate residues.

This challenge focused on 158 hydroxyethylamine inhibitors provided by Novartis. 20 of them were used in the pose prediction of Stage 1. These were one acyclic and 19 macrocyclic compounds. Later on, 154 inhibitors were used in the affinity (IC50) prediction of Stage 2. Most of them were cyclic with cycle length varying between 14 and 17 atoms, with diverse substituents and cycle structures. In this paper we will refer to these compounds as to BACE_[ID], with ID ranging between 1 and 158.

Methods

This section briefly describes computational approaches that we have been using throughout the challenge. We were adapting the algorithms used for ligand conformer generation and some of the scoring function parameters between the stages based on the analysis of the previous results. Therefore, our structure preparation procedures and submission protocols will be described and analyzed in the Submission protocols and discussion section, along with the evaluation results discussion.

Pose sampling with AutoDock Vina and Convex-PL

Below we will describe the docking pipeline applied in all the stages. Binding pocket was centered on the co-crystal ligand geometrical center. Box sizes were set to (22, 22, 25) Å with respect to the orientation of the original structure. All ligand conformations were cross-docked to all the chosen receptors with an in-house modified version of AutoDock Vina [23] using the Convex-PL potential as an integrated scoring function [24] and the Knodle parametrization of small molecules [25]. More precisely, we generated 400 poses for each ligand conformation for the subsequent re-scoring. In the AutoDock Vina configuration files, the parameter *num_modes* was set to 400 and *exhaustiveness* to 10. Our in-house modifications also include the change of *num_saved_min* to a bigger value so that more conformations are outputted. PDBQT-formatted (the format is an extension of the PDB file format, which also allows representing a kinematic tree of a molecule) structures were generated in the AutoDockTools package [26], where we kept all rotatable bonds in the ligands to be flexible. Explicit hydrogens were removed from the molecules. In our parametrization, ligand protonation states are defined by the atom types, which are assigned according to the ligand 3D geometry. These were generated from the provided SMILES strings using RDKit functions, as it is explained in more detail below. Receptor atom types corresponded to those at neutral pH. Receptors were considered to be rigid.

Then, we re-scored the obtained poses with the Convex-PL potential [24] supplemented with additional descriptors that account for the solvation and ligand flexibility contributions to the binding free energy. Coefficients corresponding to these descriptors were trained with a linear ridge regression model to fit binding constants of a set of structures in the training set extracted from the PDBBind database [27]:

$$\min ||\mathbf{y} - \mathbf{X}\mathbf{w}||_2^2 + \alpha * ||\mathbf{w}||_2^2,$$

where $\mathbf{y}$ is a set of experimental binding constants, $\mathbf{X}$ is a set of vectors of descriptors, α is a regularization coefficient, and $\mathbf{w}$ is the unknown vector of weights. We used several versions of the enhanced Convex-PL scoring function, which differed from each other by the feature weights, distance cutoff, and omitting some of the descriptors. The features we chose to enhance Convex-PL were designed to take into account interactions with solvent and conformational ligand entropy. Protein-solvent and ligand-solvent interactions were computed using a grid representation of the solvent volume that was displaced upon binding. To do so, we constructed three solvent grids for the complex, standalone receptor, and standalone ligand using the linked-cell algorithm [28]. We marked all grid cells that are not occupied by the receptor or the ligand atoms as the solvent cells. Then, we superposed the receptor and the ligand grids on the complex grid and detected solvent cells overlapping with the receptor or the ligand cells. We used their centres as the positions of dummy atoms representing the displaced solvent molecules. Finally, we computed distance distribution functions between ligand atoms and solvent dummy atoms, and receptor atoms and solvent dummy atoms following the procedure described in [24], and used them as protein-solvent and ligand-solvent descriptors. We also used additional atomic solvent-accessible surface areas descriptors computed with the POWERSASA library [29, 30]. For the ligand *conformational entropy* we introduced a measure, called flexibility, which quantifies the conformational space a ligand molecule can adopt upon rotations about the rotatable bonds. More precisely, we assume the ligand conformational space to be discrete with its volume equal to the total number of ligand conformations. We then define the ligand flexibility as a logarithm of the conformational space volume, following the definition of entropy, as

$$\text{ligand flexibility} = \log \prod_{i}^{\#\ bonds} w_i,$$

where the product is taken over all the ligand bonds. Coefficients w_i specify the number of discrete rotations about the bonds, $w_i = 3$ for single bonds, $w_i = 2$ for double and conjugated bonds, and $w_i = 1$ for triple bonds. One of our submissions also included energy terms that approximated the conformational entropy of the receptor sidechains. We estimated the entropy using a volume accessible to each of the sidechains normalized by its solvent-accessible surface area. Then we computed a set of 20 descriptors, one per each of the amino acid types, using the following equation,

$$\text{receptor flexibility}_a = \log \prod_{i}^{\#\ residues_a} v_i \frac{s_{i,unbound}}{s_{i,single}},$$

where the product is taken over all amino acids of the same type located at the interface with the ligand. Here, a is a type of amino acid, v_i is a precomputed constant volume of a sphere that is obtained by the rotation of the sidechain of type a around its C_β carbon, $s_{i,unbound}$ is the solvent-accessible surface area of the residue i computed for the receptor molecule in the unbound state, and $s_{i,single}$ is the total surface area of the same residue, if it is extracted from the receptor.

The original Convex-PL is a knowledge-based scoring function, which we have already used in the previous D3R and CSAR challenges [31–33]. It is freely available on our website at `http://team.inria.fr/nano-d/convex-pl/`. The cutoff distance for the pairwise interactions in the original Convex-PL potential is 10 Å. In order to minimize potential overfitting, we reduced this value in most of the experiments with the enhanced versions of Convex-PL. The captions of evaluation tables list the description of the Convex-PL parameters we used during the computational experiments.

Finally, the best poses were clustered with the 0.5 Å threshold using the best-scored structures as seeds for the new clusters. The resulting scores in Stage 2 were averaged over the top 10 predictions for each compound.

Submission protocols and discussion

Stage 1a

For the first stage, we intended to use a simple and robust protocol with a minimal amount of user intervention, and also without using ligand-based approaches. Therefore we chose cross-docking of flexible ligands with multiple conformations of rigid cycles, to several receptor structures.

Structure preparation

Starting from the provided SMILES strings, we generated 1,000 3D conformations for each macrocyclic ligand using RDKit's [34] *EmbedMolecule* function [35] with default parameters. We then clustered these conformations with respect to the pairwise locations of the cycle atoms using hierarchical clustering from *scipy.cluster.hierarchy* with a threshold of 0.2 Å. One conformation from each cluster was then selected for docking. For the acyclic BACE_20 we generated one conformation using RDKit's *EmbedMolecule* function.

The Protein Data Bank contains more than 300 highly homologous structures of the BACE receptor, whose binding site seems to be rather conserved. Out of these 300 receptors, we selected 38 fully homologous structures for the acyclic BACE_20 docking. Nine of them were crystallized together with cyclic ligands and thus we chose them for the BACE_1-19 docking. Table 1 lists the PDB codes of selected structures. Apart from removing solvent molecules we did not do any other modifications of the selected structures.

Docking

Docking and scoring were performed according to the pipeline described above.

Evaluation results

It turned out that all cyclic ligand conformations generated by RDKit had an incorrectly sampled dihedral angle between the atoms of an amide group leading to a cis conformation instead of the native trans one. This angle is denoted as α in Figure 1, and is a part of all the cycle-containing ligands of Stage 1. This resulted in completely wrong geometry of the whole neighborhood of the amide group, which could not be fixed by docking due to the macrocycle rigidity. An example of an incorrectly predicted cycle conformation is shown in Figure 1, where the inclination of the cycle plane is different from the native geometry. In many cases this also lead to flipped and shifted ligand docking poses, which produced high RMSD values. We have noticed this amide bond sampling problem at the very end of the Stage 1a timeframe, and submitted two predictions where the flipped and shifted poses were rejected based on the cycle similarity with the co-crystallized ligands. One more submission also used visual inspection. Overall, improper cycle conformations lead to lower than average and average in case of the manual or automatic pose rejection results listed in Table 2. Using the automatic pipeline without rejection of unrealistic poses, we obtained satisfactory low RMSDs for only a few ligands, one of which was the acyclic BACE_20.

Stage 1b

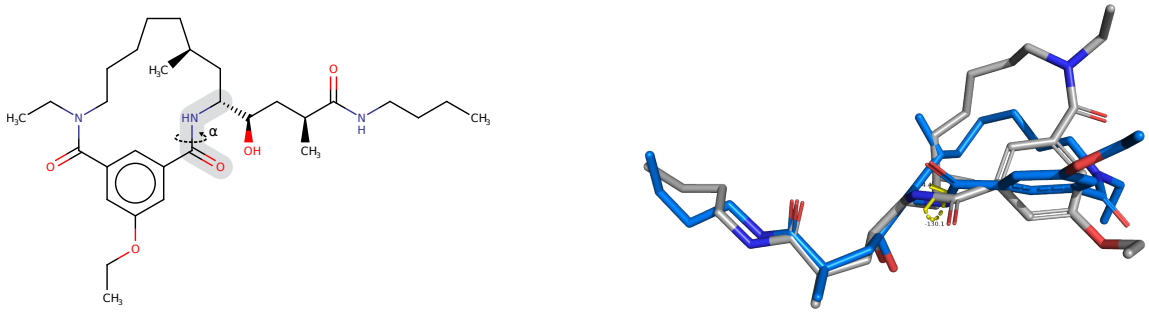
Structure preparation

For Stage 1b, crystallographic structures of all the receptors were revealed by the challenge organizers, and we used them to repeat the docking calculations. We removed the water molecules, and no other additional modifications were applied to the receptor structures.

Learning from the Stage 1a experience, we changed the way to sample ligand cycles. Initially we only tried to sample more conformations (up to 10,000). However, it turned out that in all of them RDKit produced the wrong α value of the dihedral angle despite different combinations of parameters in the *EmbedMolecule()* function. We then tried to minimize all conformers using a force field with a constraint on the wrongly predicted dihedral angle. The constraint applied with the UFF force field implemented in RDKit did not affect the final results. Also,

2f3e	2f3f	3dv1	3dv5	3k5c	4dpf	4dpi	4gmi	4k8s
2fdp	2g94	2hm1	2iqg	2p4j	2qk5	2qmd	2qmf	2qmg
2qp8	2zjn	3cib	3cic	3dm6	3duy	3i25	3ixj	3ixk
3k5d	3k5f	3k5g	3kyr	3l58	3l5e	3lnk	3veu	4gid
4k9h	5dqc							

Table 1: PDB codes of protein structures selected for Stage 1a docking. Structures highlighted in gray were used for docking of the acyclic BACE_20 ligand only.



(a) (b)

Fig. 1: BACE_1 ligand. (a) Incorrectly sampled torsion angle of the amide group present in most of the 158 compounds is highlighted in light gray. On average, the dihedral angle α's value differs by more than 100° from the ones found in crystallographic structures. (b) The native ligand conformation is shown in blue, our top-scored pose is shown in gray. It can be seen that the wrong α value leads to the incorrect conformation of the cycle.

id	scoring function	rejection of unrealistic conformations	visual inspection	mean RMSD, Å		
				average	closest	top-1
biw3a	enhanced Convex-PL	✓	✓	1.99	1.40	1.82
jit54	enhanced Convex-PL	✓	-	2.78	1.72	2.64
bsrv5	enhanced Convex-PL	✓	-	2.88	1.77	2.64
buck5	enhanced Convex-PL	-	-	3.90	2.52	3.99
maej5	enhanced Convex-PL	-	-	3.92	2.57	3.99
s4fu0	original Convex-PL	-	-	5.45	3.77	5.47

Table 2: Stage 1a evaluation results. Here we applied different versions of the enhanced Convex-PL function. The *jit54* and *buck5* submissions included the type-specific interactions with displaced solvent and Convex-PL score computed with a 5.2 Å distance cutoff. The *bsrv5* and *maej5* submissions included the solvent-accessible surface areas and the Convex-PL score computed with a 5.2 Å distance cutoff. In the *biw3a* submission, we chose the highest-ranked poses scored with the three versions of Convex-PL used in all the other Stage 1a submissions, and rejected some poses based on visual inspection.

constrained minimization using the MMFF94 [36] force field resulted in very distorted structures. Although at this stage it could have been possible to simply use another tool for conformer generation, not all of them are free, and we also felt being somewhat challenged to make RDKit generate better conformations. Finally, we decided to try the *coordMap* option of the *EmbedMolecule()* function, which rejects conformations where the distances between specified atoms' positions are different from those passed through the *coordMap* argument, up to a certain threshold. When using only the 4 dihedral angle atoms, conformational sampling results did not change and the angle was still wrongly sampled. We have tried to tweak internal threshold of this map-based reduction in the RDKit source code, but it did not improve the results. Therefore we increased the size of the map, pushing ourselves to a more ligand-based setup. Figure 2 schematically represents an algorithm for the map generation used for cyclic ligands.

We started with computing the maximum common substructures (MCS0) between the cycles (including non-rotatable cycle substituents) of each target ligand and the cycles of the 9 ligands co-crystallized with proteins listed in Table 1. We also computed the maximum common substructures between the entire ligands (MCS). For each target ligand, we chose a reference ligand based on the MCS0 size. Then, we selected 4 atoms corresponding to the wrongly predicted amide group, and two carbon atoms bound to them, including one from the hydroxyethylamine group. These are shown in yellow in Figure 3 and will be referenced as a "core set". The mapping of these 6 atom indices in the target ligand structure to the coordinates from the reference ligand structure were provided as a *coordMap* argument to the conformer generating function. We then computed α value of the generated conformers. If more than 10% α values were lying between $-25°$ and $25°$, we saved the conformers and proceeded to the next target ligand. If not, we iteratively increased the map based on a set of rules illustrated in Figure 3 until 10% of structures would have the correct amide bond conformation. If more than 80% of the MCS was included into the map without providing good conformers, we moved to the next reference structure. If three reference structures were not sufficient, we aligned them to each other based on the coordinates of the atoms of the "core set", and used the union of the MCSs of both reference molecules to create a new mapping. After at least 10% of good conformations was achieved, we stopped the algorithm and saved the molecules. If $\geq$ 70% of conformations were generated with α values inside the $[-25°, 25°]$ threshold interval, we squeezed this interval to $[-10°, 10°]$ and rejected outlying conformations.

Overall, even though we did not manage to find out what exactly led to the cycle sampling problems, this approach finally allowed us to create structures with correct α angle for all macrocyclic targets.

Evaluation results

This approach lead to low-RMSD results, summarized in Table 3. The mean RMSD of the closest pose of all our submissions was less than 1 Å. Figure 4 shows several examples of the poses we obtained in Stage 1b. The enhanced versions of Convex-PL on average predict binding poses more accurately compared to the original version. For example, the top-1 ranked pose of the BACE_12 ligand in the *dhueb* submission was considerably shifted and rotated with respect to the native pose, which resulted in the 10.53 Å RMSD. In the *nyrou* submission we obtained 0.80 Å RMSD. However, the biggest contribution to this performance improvement was driven not by the additional descriptors, but by the change of the interaction cutoff distance to 5.2 Å, which is smaller than the default value of 10 Å. This smaller cutoff value was used to train the enhanced versions of Convex-PL in the *nyrou* and *vfkn2* submissions. The low contribution of additional descriptors can be explained by the fact that all of them are related to the interactions that a molecule could have with displaced solvent. The BACE binding pocket is not very open to solvent, and the fraction of ligand surface that could be exposed to solvent does not change much even between the poses with 10 Å RMSD difference. Therefore, the sums of additional descriptors' contributions were very close to each other for the majority of ligand poses.

id	scoring function	mean RMSD, Å		
		average	closest	top-1
nyrou	enhanced Convex-PL	0.98	0.84	0.89
vfkn2	enhanced Convex-PL	0.99	0.84	0.89
mjevm	enhanced Convex-PL	1.14	0.79	1.00
dhueb	original Convex-PL	1.56	0.90	1.60

Table 3: Stage 1b evaluation results. Enhanced version of Convex-PL used in the *nyrou* submission was trained on the interactions with the volume displaced solvent and the original Convex-PL score computed with a 5.2 Å cutoff. The *vfkn2* submission included solvent-accessible surface area descriptors and the Convex-PL score computed with a 5.2 Å cutoff. Scoring function used in the *mjevm* submission included included interactions with the volume of displaced solvent and the original Convex-PL score computed with a 4.8 Å cutoff.

Stage 1a redocking

To check how did the macrocycle conformer quality influenced the results of Stage 1a, we repeated the ensemble docking of the BACE_1-19 ligand structures prepared for Stage 1b to the set of 9 receptors used in Stage 1a. As it could be expected, better ligand structures considerably improved the pose prediction. Without manual inspection or pose filtering we obtained the subangstrom mean RMSD value for the closest pose shown in the Table 4. Figure 5 illustrates the redocking pose of the BACE_7, which is superimposed with the one we submitted for Stage 1a. Here it can be clearly seen how did the bad initial conformation from our submission lead to a considerable shift of the ligand inside the pocket.

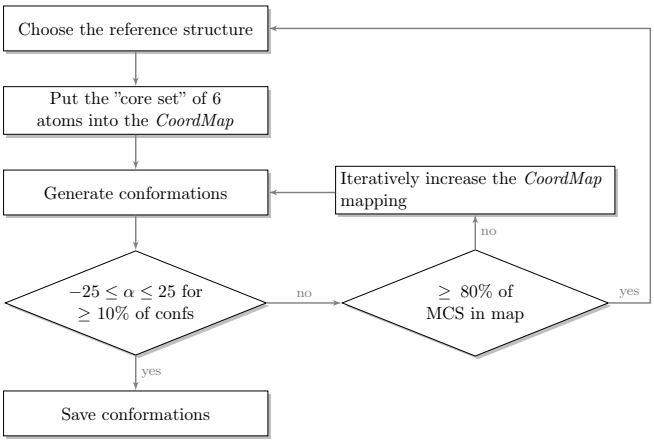
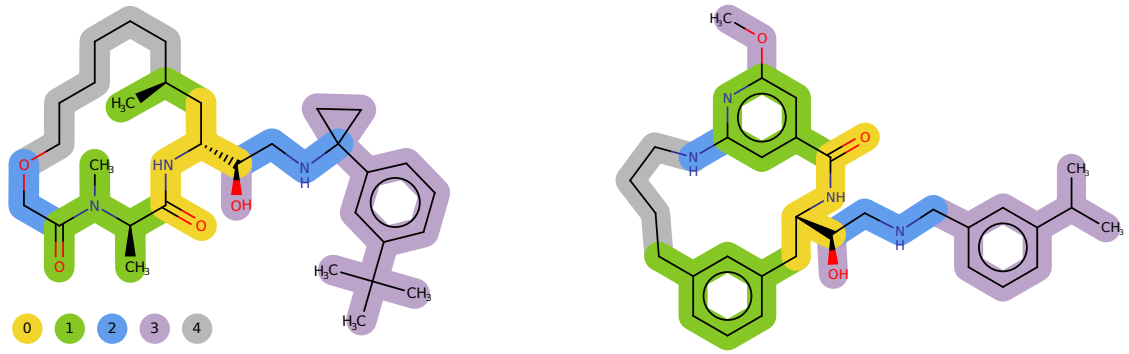


Fig. 2: Algorithm 1. Schematic description of an algorithm for conformer generation in RDKit driven by distance constraints. Please see main text for more details.



(a) BACE_4 (b) BACE_12

Fig. 3: Examples of ligand mapping priority. Each color represents a different priority, which are ranked from 0 to 4. On each iteration of the algorithm an atom (or a group of atoms in case of rings) was added to the map with the following priorities. (1) Atoms with minimal topological distance from the "core set", amide groups of the cycle, aromatic substituents topologically close to the "core set". (2) Carbons and nitrogen of the hydroxyethylamine group, non-carbon atoms of the cycle. (3) Atoms of the "tails", oxygen of the hydroxyethylamine group. (4) Rest of the macrocycle atoms topologically far from the "core set", hydroxyl and carboxyl substituents of the macrocycle. We tried to use as few of these atoms as possible since they adopt the most diverse conformations as compared between the cycles, and we would not like to occasionally freeze them.

id	scoring function	mean RMSD, Å		
		average	closest	top-1
-	enhanced Convex-PL	1.54	0.89	1.22

Table 4: Stage 1a redocking results. Here, we trained the scoring function using the interactions with the displaced solvent volume, atomic SASA values, and the original Convex-PL score computed with a 5.2 Å cut-off.

Stage 2

Stage 2 was dedicated to the scoring exercises. The goal was to correctly predict the relative binding affinities of the set of 154 molecules binding the BACE receptor. The 20 crystallographic structures of complexes from Stage 1 were already revealed for this stage.

Structure preparation

Since the amount of computations required for docking of all the 154 compounds was considerably higher compared to Stage 1, and more protein structures became available for docking, we first selected a set of target structures for each compound. The BACE_1 – BACE_20 ligands were docked into the co-crystal receptors. For the rest of the cyclic ligands we first extracted the fragments containing the macrocycle only,

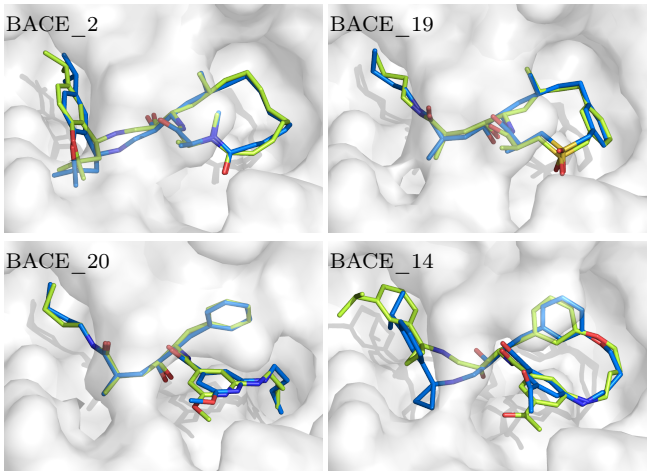


Fig. 4: Examples of the closest poses from our Stage 1b *nyrou* submission. Crystallographic structures are shown in blue, our predictions are shown in green. Bond orders are not shown.

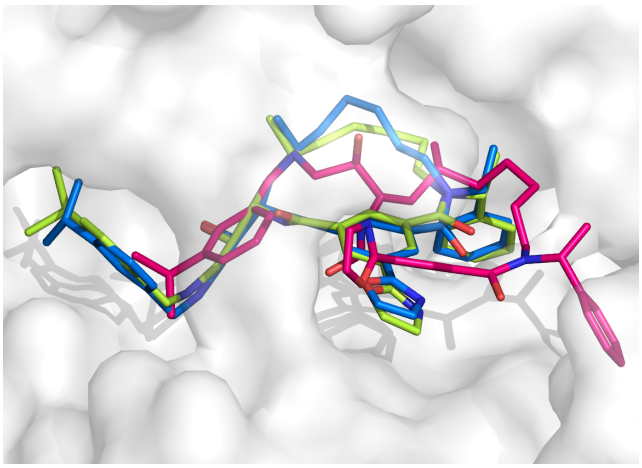


Fig. 5: BACE_7 ligand poses. Crystallographic structure is shown in blue, our initial Stage 1a prediction from the *buck5* submission is shown in red, the pose obtained with redocking is shown in green. Bond orders are not shown. Please note a considerable shift of the red ligand compared to the crystallographic (blue) one.

and the macrocycle with some substituents, such as aromatic rings. We then computed the maximum common substructures of these fragments with the ligands with known co-crystal structures, and selected the receptors with maximum MCS size resulting in 4 - 12 receptors per each compound. Receptors for the acyclic BACE_145 and BACE_146 ligands were chosen based on the overall MCS size.

To create the cyclic ligand structures, we followed the algorithm applied in Stage 1b with several modifications. The pool of reference ligands now included the 20 co-crystal structures from Stage 1. In some cases we visually inspected the results and supervised the process of macrocycle structure generation.

Evaluation results

We ran out of time and have not finished docking of all the conformations of macrocyclic molecules. We have submitted two sets of predictions containing about 60% and 80% of all docked conformations to see how the result will change depending on these numbers. This resulted in Kendall τ of 0.12 for the first subset's best prediction, and 0.14 for the second, which are listed in Table 5. We can see that regardless the cutoff value, the ligand flexibility descriptor, which estimates the conformational entropy change upon binding, improved the results in all the enhanced submissions. The scoring function used in the submission with the highest Kendall τ, *xx4i5*, was trained on both solvent-related and entropy-related descriptors. Unlike the Stage 1 pose prediction exercise, where solvent-related descriptors almost did not contribute to the comparison of the poses, here they do influence the results, since binding poses of different ligands are now compared to each other.

We have also evaluated the ability of our enhanced scoring function to predict binding affinities based on the docking poses generated by other predicting teams. To do so, we firstly rescored all the available submissions of structure-based predictor teams with the scoring function used in the *xx4i5* submission. Secondly, we also applied local optimization to the ligand positions in the binding sites using AutoDock Vina's algorithm and the basic version of the Convex-PL scoring function. We then recomputed the affinity scores. Figure 6 shows the rescoring results. We can see that our approach does not improve the predictions of the best submitters (those with Kendall $\tau > 0.15$). Local optimization improves the results from 0.09 to 0.11 τ averaged over all the predictions without and with local optimization, respectively. Our own submissions got also slightly improved after the re-scoring.

We have also found out that we obtain rather good affinity predictions with Kendall τ equal to 0.24 when using the docking poses submitted by the second-best structure-based affinity predictor *urt76*. However, this result gets worse if the local optimization is applied prior to computing the affinities.

id	scoring function	% initial conformations docked	Kendall's τ	Spearman's ρ
xx4i5	enhanced Convex-PL	80%	0.14	0.21
dzyxt	enhanced Convex-PL	80%	0.13	0.19
u7r6y	enhanced Convex-PL	80%	0.12	0.19
kzsv5	enhanced Convex-PL	60%	0.12	0.18
i88wa	original Convex-PL	80%	0.12	0.18
q6mvt	enhanced Convex-PL	60%	0.11	0.16

Table 5: Stage 2 affinity prediction results. Submissions *dzyxt* and *kzsv5* were scored only with two descriptors, the Convex-PL score computed with a 10 Å cutoff and the ligand flexibility. The *u7r6y* submission was scored using the ligand flexibility and the Convex-PL score computed with a 5.2 Å cutoff. The *xx4i5* and *q6mvt* submissions correspond to the scoring function trained on interactions with the volume of the displaced solvent, SASA values, ligand flexibility, flexibility of the interacting receptor residues, and the Convex-PL score computed with a 5.2 Å cutoff.

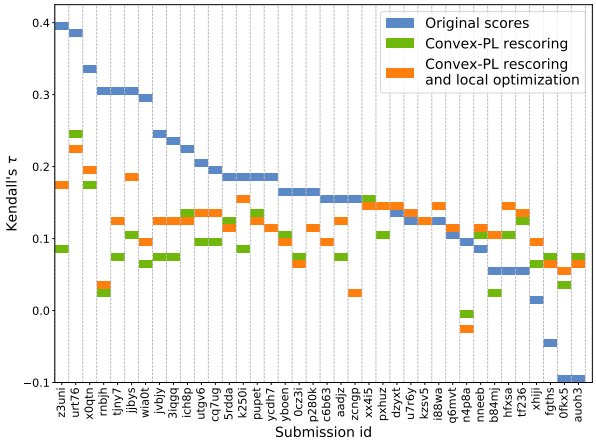


Fig. 6: Re-scoring of the available structure-based submissions computed with the scoring function that was used in the *xx4i5* submission. All scores were rounded up to the second digit, as in the evaluation results chart. Submissions *dxji8* and *pngkk* were excluded from the comparison due to the incorrect receptor structures. Submissions *6jyjp* and *ufr7g* were excluded from the comparison because the provided ligand chemical structures did not correspond to the original structures.

Technical details

We computed symmetry-adapted RMSD values with a modified GetBestRMS() function from the RDKit package [34]. The RMSD values we obtained corresponded to those reported in the official evaluation results. Receptor alignment was done with the PyMOL 1.8.6 [37] *align* function. Algorithm 1 was implemented in python3 using RDKit. Images were created with MarvinSketch, PyMOL 1.8.6, Matplotlib, and Inkscape.

Conclusion

This docking exercise provided us a unique opportunity to model macrocyclic ligands that bind to protein targets. The modeling part was challenging for us, as we aimed to use structure-based approaches and sampling in torsion coordinates. We have started with a fully structure-based and automated docking procedure. However, at the end of Stage 1a we analyzed the docking results and discovered a very poor generation of realistic ligand macrocycle conformations. Therefore, we supplemented the docking protocol with constraints based on the structure of similar ligands. Finally, we converged to a stable pipeline that resulted in sufficiently low (subangstrom) RMSD of binding poses. During the restricted challenge timeframe we have not tried other algorithms of fast ligand conformer generation besides the one implemented in RDKit. Yet, we believe that the problems we encountered with the amide bond conformation undersampling in cycles deserve further research and investigation.

In this exercise we compared the performance of our original Convex-PL knowledge-based scoring function with its several enhanced versions that included additional terms and were trained with shorter cutoff values for the pairwise interactions. The additional descriptors accounted for interactions with solvent, and for ligand and receptor sidechain flexibility. Our results demonstrated a considerably better on average pose prediction power of the enhanced Convex-PL potential compared to its original version. For example, in Stage 1b we obtained the mean RMSD values averaged over top-5 best predictions of 0.98 Å for the enhanced Convex-PL versus 1.56 Å for the original version. However, this pose prediction improvement seems to be mostly caused by the change in the cutoff value.

In the affinity predictions we also relied on the values suggested by our scoring function. The resulting correlations turned out to be average compared to the other structure-based methods. We believe that we did not manage to obtain good binding poses for all the 154 ligands in Stage 2. For example, if we applied our scoring function to the pose predictions of some of the best submitters, we could considerably improve our own result. After rescoring of other predictors' submissions, we also noticed that local gradient-based pose optimization on average led to better binding affinity predictions.

Acknowledgement

The authors would like to thank Ivan Gushchin from MIPT Moscow for providing his expertise in crystallography. This work was partially supported by the Ministry of Education and Science of the Russian Federation (grant no. 6.3157.2017).

References

1. M. Jusot, D. Stratmann, M. Vaisset, J. Chomilier, and J. Cortes, "Exhaustive exploration of the conformational landscape of small cyclic peptides using a robotics approach," *Journal of chemical information and modeling*, vol. 58, no. 11, pp. 2355–2368, 2018.
2. D. Sindhikara, S. A. Spronk, T. Day, K. Borrelli, D. L. Cheney, and S. L. Posy, "Improving accuracy, diversity, and speed with prime macrocycle conformational sampling," *Journal of chemical information and modeling*, vol. 57, no. 8, pp. 1881–1894, 2017.
3. R. D. Smith, K. L. Damm-Ganamet, J. B. Dunbar Jr, A. Ahmed, K. Chinnaswamy, J. E. Delproposto, G. M. Kubish, C. E. Tinberg, S. D. Khare, J. 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.*, vol. 56, no. 6, pp. 1022–1031, 2015.
4. H. A. Carlson, R. D. Smith, K. L. Damm-Ganamet, J. A. Stuckey, A. Ahmed, M. A. Convery, D. O. Somers, M. Kranz, P. A. Elkins, G. Cui, C. E. Peishoff, M. H. Lambert, and J. B. Dunbar, Jr, "Csar 2014: A benchmark exercise using unpublished data from pharma," *J. Chem. Inf. Model.*, May 2016.
5. S. Gathiaka, S. Liu, M. Chiu, H. Yang, J. A. Stuckey, Y. N. Kang, J. Delproposto, G. Kubish, J. B. Dunbar, H. A. Carlson, *et al.*, "D3r grand challenge 2015: Evaluation of protein–ligand pose and affinity predictions," *J. Comput.-Aided Mol. Des.*, vol. 30, no. 9, pp. 651–668, 2016.
6. Z. Gaieb, S. Liu, S. Gathiaka, M. Chiu, H. Yang, C. Shao, V. A. Feher, W. P. Walters, B. Kuhn, M. G. Rudolph, *et al.*, "D3r grand challenge 2: blind prediction of protein–ligand poses, affinity rankings, and relative binding free energies," *Journal of computer-aided molecular design*, vol. 32, no. 1, pp. 1–20, 2018.
7. Z. Gaieb, C. D. Parks, M. Chiu, H. Yang, C. Shao, W. P. Walters, M. H. Lambert, N. Nevins, S. D. Bembenek, M. K. Ameriks, *et al.*, "D3r grand challenge 3: blind prediction of protein–ligand poses and affinity rankings," *Journal of computer-aided molecular design*, vol. 33, no. 1, pp. 1–18, 2019.
8. P. W. Rose, A. Prlić, A. Altunkaya, C. Bi, A. R. Bradley, C. H. Christie, L. Di Costanzo, J. M. Duarte, S. Dutta, Z. Feng, *et al.*, "The rcsb protein data bank: integrative view of protein, gene and 3d structural information," *Nucleic Acids Research*, vol. 45, no. D1, pp. D271–D281, 2017.
9. M. Ignatov, C. Liu, A. Alekseenko, Z. Sun, D. Padhorny, S. Kotelnikov, A. Kazennov, I. Grebenkin, Y. Kholodov, I. Kolosvari, *et al.*, "Monte carlo on the manifold and md refinement for binding pose prediction of protein–ligand complexes: 2017 d3r grand challenge," *Journal of computer-aided molecular design*, vol. 33, no. 1, pp. 119–127, 2019.
10. A. Kumar and K. Y. Zhang, "Shape similarity guided pose prediction: lessons from d3r grand challenge 3," *Journal of computer-aided molecular design*, vol. 33, no. 1, pp. 47–59, 2019.
11. P. I. Koukos, L. C. Xue, and A. M. Bonvin, "Protein–ligand pose and affinity prediction: Lessons from d3r grand challenge 3," *Journal of computer-aided molecular design*, vol. 33, no. 1, pp. 83–91, 2019.
12. D. D. Nguyen, Z. Cang, K. Wu, M. Wang, Y. Cao, and G.-W. Wei, "Mathematical deep learning for pose and binding affinity prediction and ranking in d3r grand challenges," *Journal of computer-aided molecular design*, vol. 33, no. 1, pp. 71–82, 2019.
13. P. C.-H. Lam, R. Abagyan, and M. Totrov, "Hybrid receptor structure/ligand-based docking and activity prediction in icm: development and evaluation in d3r grand challenge 3," *Journal of computer-aided molecular design*, vol. 33, no. 1, pp. 35–46, 2019.
14. L. Chaput, E. Selwa, E. Elisee, and B. I. Iorga, "Blinded evaluation of cathepsin s inhibitors from the d3rgc3 dataset using molecular docking and free energy calculations," *Journal of computer-aided molecular design*, vol. 33, no. 1, pp. 93–103, 2019.
15. J. Sunseri, J. E. King, P. G. Francoeur, and D. R. Koes, "Convolutional neural network scoring and minimization in the d3r 2017 community challenge," *Journal of computer-aided molecular design*, vol. 33, no. 1, pp. 19–34, 2019.
16. X. He, V. H. Man, B. Ji, X.-Q. Xie, and J. Wang, "Calculate protein–ligand binding affinities with the extended linear interaction energy method: application on the cathepsin s set in the d3r grand challenge 3," *Journal of computer-aided molecular design*, vol. 33, no. 1, pp. 105–117, 2019.
17. B. Xie and D. D. Minh, "Alchemical grid dock (algdock) calculations in the d3r grand challenge 3," *Journal of computer-aided molecular design*, vol. 33, no. 1, pp. 61–69, 2019.
18. R. Vassar, D. M. Kovacs, R. Yan, and P. C. Wong, "The β-secretase enzyme bace in health and alzheimer's disease: regulation, cell biology, function, and therapeutic potential," *Journal of Neuroscience*, vol. 29, no. 41, pp. 12787–12794, 2009.
19. F. Prati, G. Bottegoni, M. L. Bolognesi, and A. Cavalli, "Bace-1 inhibitors: From recent single-target molecules to multitarget compounds for alzheimer's disease: Miniperspective," *Journal of medicinal chemistry*, vol. 61, no. 3, pp. 619–637, 2017.
20. S. Hanessian, G. Yang, J.-M. Rondeau, U. Neumann, C. Betschart, and M. Tintelnot-Blomley, "Structure-based design and synthesis of macroheterocyclic peptidomimetic inhibitors of the aspartic protease β-site amyloid precursor protein cleaving enzyme (bace)," *Journal of medicinal chemistry*, vol. 49, no. 15, pp. 4544–4567, 2006.
21. J. B. Jordan, D. A. Whittington, M. D. Bartberger, E. A. Sickmier, K. Chen, Y. Cheng, and T. Judd, "Fragment-

linking approach using 19f nmr spectroscopy to obtain highly potent and selective inhibitors of β-secretase," *Journal of medicinal chemistry*, vol. 59, no. 8, pp. 3732–3749, 2016.
22. S. Butini, S. Brogi, E. Novellino, G. Campiani, A. K Ghosh, M. Brindisi, and S. Gemma, "The structural evolution of β-secretase inhibitors: a focus on the development of small-molecule inhibitors," *Current topics in medicinal chemistry*, vol. 13, no. 15, pp. 1787–1807, 2013.
23. O. Trott and A. J. Olson, "AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading," *J. Comput. Chem.*, vol. 31, no. 2, pp. 455–461, 2010.
24. M. Kadukova and S. Grudinin, "Convex-pl: a novel knowledge-based potential for protein-ligand interactions deduced from structural databases using convex optimization," *Journal of computer-aided molecular design*, vol. 31, no. 10, pp. 943–958, 2017.
25. M. Kadukova and S. Grudinin, "Knodle: A support vector machines-based automatic perception of organic molecules from 3d coordinates," *J. Chem. Inf. Model.*, vol. 56, pp. 1410–9, Aug 2016.
26. G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, and A. J. Olson, "Autodock4 and autodocktools4: Automated docking with selective receptor flexibility," *J. Comput. Chem.*, vol. 30, no. 16, pp. 2785–2791, 2009.
27. Z. Liu, M. Su, L. Han, J. Liu, Q. Yang, Y. Li, and R. Wang, "Forging the basis for developing protein–ligand interaction scoring functions," *Accounts of chemical research*, vol. 50, no. 2, pp. 302–309, 2017.
28. S. Artemova, S. Grudinin, and S. Redon, "A comparison of neighbor search algorithms for large rigid molecules," *Journal of Computational Chemistry*, vol. 32, no. 13, pp. 2865–2877, 2011.
29. K. V. Klenin, F. Tristram, T. Strunk, and W. Wenzel, "Derivatives of molecular surface area and volume: Simple and exact analytical formulas," *Journal of computational chemistry*, vol. 32, no. 12, pp. 2647–2653, 2011.
30. K. Klenin, F. Tristram, T. Strunk, and W. Wenzel, "Achieving numerical stability in analytical computation of the molecular surface and volume," *From Computational Biophysics to Systems Biology (CBSB11)–Celebrating Harold Scheraga's 90th Birthday*, vol. 8, p. 75, 2012.
31. S. Grudinin, P. Popov, E. Neveu, and G. Cheremovskiy, "Predicting binding poses and affinities in the csar 2013–2014 docking exercises using the knowledge-based convex-pl potential," *J. Chem. Inf. Model.*, vol. 56, no. 6, pp. 1053–1062, 2015.
32. S. Grudinin, M. Kadukova, A. Eisenbarth, S. Marillet, and F. 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," *Journal of computer-aided molecular design*, vol. 30, no. 9, pp. 791–804, 2016.
33. M. Kadukova and S. 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*, vol. 32, no. 1, pp. 151–162, 2018.
34. G. Landrum, "Rdkit: Open-source cheminformatics." http://www.rdkit.org.
35. S. Riniker and G. A. Landrum, "Better informed distance geometry: using what we know to improve conformation generation," *Journal of chemical information and modeling*, vol. 55, no. 12, pp. 2562–2574, 2015.
36. P. Tosco, N. Stiefl, and G. Landrum, "Bringing the mmff force field to the rdkit: implementation and validation," *Journal of cheminformatics*, vol. 6, no. 1, p. 37, 2014.
37. Schrödinger, LLC, "The PyMOL molecular graphics system, version 1.3," 2011.