Sampling and refinement protocols for template-based macrocycle docking: 2018 D3R Grand Challenge 4

Sergei Kotelnikov[1,2,3,*] and Andrey Alekseenko[1,2,4,*], Cong Liu[1,5], Mikhail Ignatov[1,2,4,6], Dzmitry Padhorny[1,2], Emiliano Brini[1], Mark Lukin[7], Evangelos Coutsias[1,2], Ken A. Dill[1,5,8], Dima Kozakov[1,2,4,†]

1 Laufer Center for Physical and Quantitative Biology, Stony Brook University, Stony Brook, NY, USA
2 Department of Applied Mathematics and Statistics, Stony Brook University, Stony Brook, NY, USA
3 Innopolis University, Innopolis, Russia
4 Institute of Computer Aided Design of the Russian Academy of Sciences, Moscow, Russia
5 Department of Chemistry, Stony Brook University, Stony Brook, NY, USA
6 Institute for Advanced Computational Sciences, Stony Brook University, Stony Brook, NY, USA
7 Department of Pharmacological Sciences, Stony Brook University, Stony Brook, NY, USA
8 Department of Physics and Astronomy, Stony Brook University, Stony Brook, NY, USA
* These authors contributed equally to this work
† Corresponding author: midas@laufercenter.org

Abstract

In this paper, we describe the methodology used by our group for the pose prediction task in the Drug Design Data Resource (D3R) Grand Challenge 4. We employed a template-based approach in which the initial placement of the ligand was based on the related structures in the Protein Data Bank. Since most of the targets in this round of D3R Grand Challenge were macrocyclic compounds, we used robotics-inspired BRIKARD algorithm to sample the conformations of the ligands. This was followed by filtering and several refinement protocols, including MELD-accelerated molecular dynamics and Monte Carlo on manifold minimization. The final ranking was performed semi-manually and was based on various scoring functions, clustering, and the similarity of the complex structure to known crystallographic complexes. Based on the assessment by D3R organizers, our approach was among the top-performing methods in the pose prediction challenge, both in cross-docking (1a) and self-docking (1b) stages.

Introduction

Over the last few years, our team has successfully participated [1, 2] in several rounds of the Drug Design Data Resource (D3R) community-wide blinded prediction challenge. Each round had a variety of tasks — prediction of poses, binding affinity, and free energy of small molecular compounds. In the last round, D3R 2019 Grand Challenge 4 (GC4), the participants were asked to predict poses and affinities for ligands of beta-secretase 1 (BACE1), a protease implicated in the production of beta-amyloid peptides in patients with Alzheimer's disease [3]. We took part in the first two stages of GC4: 1a and 1b. Stage 1a included prediction of the binding poses of 20

ligands and provided the participants with nothing but protein sequence and ligand SMILES strings. Stage 1b also included a pose prediction task for the same 20 ligands, but this time the participants were given the corresponding receptor cocrystal structures with water molecules and, in some cases, sulfate ions and glycerol molecules, but without the target ligand. While the D3R GC4 also included affinity and free energy prediction stages, as a team with expertise in molecular docking and structure prediction we focused on the pose prediction sub-challenges. Previous rounds of D3R motivated us to develop new tools and modify existing protocols to satisfy the needs of the proposed tasks; GC4 was no exception. In this round, we used a combination of template modeling, inverse kinematics sampling [4], restrained local minimization (RM), Monte-Carlo energy minimization on the manifold (MCMM) [2, 5, 6], conventional molecular dynamics (MD), and MELD-accelerated molecular dynamics (MELD × MD) simulation [7, 8].

BACE1 is a well-studied system, which, together with its sequence homologs, has several hundred crystallized structures deposited in PDB, many of which contain a bound small-molecule ligand. We collected and integrated the existing information on the bound compounds and used it throughout the whole pipeline: generation of initial poses, their refinement, and the final scoring.

Most of the proposed compounds were cyclic molecules (19 out of 20), which makes sampling difficult since the loop closure condition must be satisfied. To overcome this issue we used the inverse kinematics approach designed to exhaustively sample compound conformations, satisfying multiple closure conditions at the same time [4].

We sampled multiple structures for each target and then subjected them to preliminary filtering. The remaining structures were refined with several algorithms: full-atomic relaxation, Monte-Carlo on manifold minimization, conventional and accelerated MD. This pipeline resulted in very successful predictions with 0.76 Å average pose-1 RMSD poses, and sub-angstrom accuracy for 15 out of 20 compounds according to the official evaluation by the D3R organizers.

Methods

Workflow overview

For Stage 1a, the input data consisted of receptor sequence (as a FASTA string) and ligand structures (as SMILES strings). In total, 20 receptor-ligand pairs, or targets, were given, named BACE_1 to BACE_20. The BRIKARD algorithm [4] (for cyclic molecules BACE_1 to BACE_19) and the ETKDG [9] algorithm (for the only non-cyclic target, BACE_20) were used to generate multiple initial conformations for the ligands. The similarity search was performed in the PDB database to find templates — highly-homologous proteins with similar ligands. Out of all generated conformations, only those closest to the templates were retained. This way, for each target we created an ensemble of starting poses, containing between 4 and 402 structures. These structures were subjected to restrained full-atomic energy minimization (RM) to remove possible clashes and “relax” the ligand. Out of the resulting minimized poses, four to five were manually chosen for submission.

For *Stage 1b*, the D3R organizers provided X-Ray structures of the receptor cocrystallized with each of the ligands. We aligned the submitted results from Stage 1a to these receptor molecules and applied various refinement protocols: restrained full-atomic minimization (RM), Monte Carlo on manifold minimization (MCMM), molecular dynamics (MD), and Modeling Employing Limited Data accelerated molecular dynamics (MELD × MD). For each target, five structures were manually selected for submission.

The high-level overview of the workflow is presented in Fig. 1.

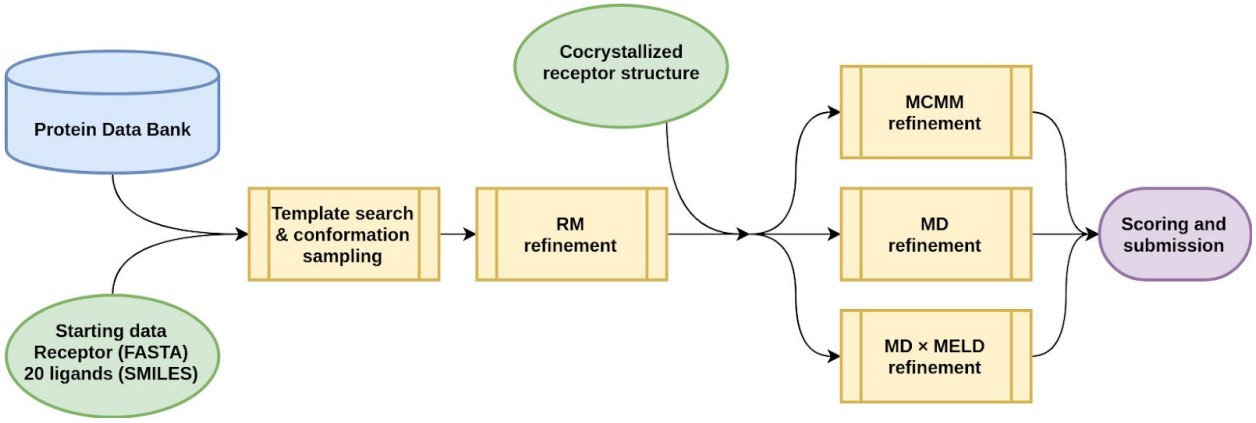

Fig. 1 The general outline of the protocol used by our team in D3R GC4. Yellow elements indicate the stages of the pipeline, the green ones indicate the data provided by the organizers, the blue ones indicate publicly-available databases, and the purple ones indicate the final submitted results. The MCMM, MD, and MD × MELD refinement stages were only used for refining with a cocrystallized receptor in Stage 1b and were skipped in Stage 1a

Template search

The first step of the protocol in Stage 1a was finding known structures of closely related complexes. We did a BLAST search for sequence-similar (e-value = 10^{-20}, sequence identity >= 95%, resolution <= 3 Å) chain structures in the Protein Data Bank (PDB). Then, the following procedure was conducted independently for each of the 20 target ligands. For each sequence-similar structure, only the ligand with the best Tanimoto score located within 8 Å of the sequence-similar chain structure was retained, with the two thus forming a protein-ligand template. Tanimoto score was calculated using Daylight molecular fingerprint with RDKit [10]. All templates with Tanimoto score less than two-thirds of the maximal Tanimoto score for the current target were discarded. The Maximum Common Substructure (MCS) was calculated between the target ligand and each of the remaining template ligands. For each target, two MCS were calculated using RDKit with different tolerance criteria: “weak” MCS (atoms and valences should be the same, allows chiral centers to be different, single and aromatic bonds to match each other) and “strict” MCS (allowances mentioned before are prohibited). Several templates with the highest MCS coverage for given target were taken as final. A special case was BACE_20, where we additionally constructed a “chimera” built from two templates. Each of these two templates covers a different part of BACE_20 ligand, with some intersection between them forming a common “core”.

Starting poses preparation

Despite the advancements of refinement protocols, having a good starting pose is still a prerequisite for achieving a low-RMSD result. For the initial stage of the competition (Stage 1a), the starting poses were prepared using the following multi-stage approach. First, we generated 10^4 conformers for each target ligand.
For BACE_1-BACE_19 (macrocyclic molecules), conformer generation was carried out using the robotics-inspired BRIKARD [4] loop closure algorithm. Using inverse kinematics, BRIKARD rigorously samples “driver” torsions according to prescribed intervals and frequencies, while “driven” or “pivot” torsions [4] are computed by recursive ring closure consistently for all interconnected rings. This allows for uniform sampling of all dihedral angle values. BRIKARD allows manually specifying the order of solution of the rings in the system. However, for this study, we used ring perception algorithm [11] to carry out the recursive ring generation and find solutions automatically.
For BACE_20 (non-macrocyclic molecule), we used ETKDG method [9] from RDKit for conformer generation. All conformers were minimized in vacuum using Merck Molecular Force Field (MMFF) [12, 13]. Then, for each

template (a protein-ligand pair from the template PDB structure), for both “weak” and “strict” MCS, all possible MCS mappings were generated, taking into account 1) all possible impositions of the MCS on the template ligand and 2) internal symmetry of the MCS itself. For each mapping generated this way, we aligned all conformers of the target ligand to the template ligand. For each template and for each “flavor” of MCS we retained only the conformer with the best MCS-mapping RMSD. The resulting structures were used as starting poses for the refinement. The number of starting poses per target varied from 4 (BACE_10 and BACE_15) to 402 (BACE_7), depending on the number of templates found and the number of generated MCS mappings between target and template atoms.
In Stage 1b of the competition, we used our Stage 1a submissions as starting poses.

Refinement

In Stage 1a, the starting poses were subjected to a simplistic restrained energy minimization (RM), where harmonic restraints were pulling the ligand closer to the template. The RM results were scored and used for Stage 1a submission.
In Stage 1b, we started from the ligand structures (4 or 5 per target) we had submitted in Stage 1a. We aligned them to the X-Ray structures of the receptor provided by the organizers. In this stage, we employed four different refinement protocols — restrained minimization (RM, same as in Stage 1a), Monte Carlo on manifold minimization (MCMM), conventional Molecular Dynamics (MD), and MELD-accelerated molecular dynamics (MELD × MD). The resulting structures from all refinement approaches were pooled together and scored for the Stage 1b submission. The details of refinement protocols are as follows.

Restrained Minimization

In both Stages 1a and 1b, we employed a straightforward restrained minimization protocol to refine the starting poses. The protocol was based on full-atom energy minimization using a CHARMM-based energy function with GBSA (ACE) solvation term, described in Ref. [14] and implemented in libmol2 library (https://bitbucket.org/bu-structure/libmol2/). In Stage 1b, an explicit hydrogen bonding term [15, 16] was added to the energy function. During the minimization procedure, all receptor atoms except hydrogens were fixed, while ligand atoms matching the template were restrained with a harmonic potential to the positions of the corresponding template atoms. Applying the restraints allowed us to overcome the limitations of the general forcefield, and implicitly harness the details of known BACE1-ligand interactions. Minimization was carried out using the L-BFGS algorithm [17]. Two variations of minimization protocol, which we term RM_1 and RM_2, were used for each target.
RM_1 had three stages: (1) 500 steps without restraints to remove possible clashes; (2) 500 steps with harmonic restraints (10 kcal/mol/Å^2) ; (3) 500 steps without restraints to allow the structure to settle in the pocket.
RM_2 had five stages: (1) 500 steps of minimization with harmonic restraint (10 kcal/mol/Å^2), but without van der Waals, electrostatic, and dihedral potentials to pull the target close to the template structure; (2) same as (1), but dihedral potential is enabled; (3) same as (2), but all the receptor atoms in the interface are movable; (4) same as (3), but electrostatics and van der Waals potential are enabled; (5) same as (4), but without restraints. RM_2 tends to allow more drastic changes to the ligand structure, making it closer to the template, but sometimes it results in unnaturally twisted dihedrals.

Monte Carlo on Manifold Minimization

The Monte Carlo on manifold minimization algorithm was based on the protocol described previously [1, 2, 6, 14, 18]. MCMM relies on the assumption that covalent bonds and angles can be considered fixed, and thus the molecular flexibility is achieved through the rotation of dihedral angles alone. The molecule in this framework is described as a set of rigid molecular clusters connected by rotatable bonds. While full-atom minimization takes

place in a 3*N*-dimensional space, where *N* is the number of atoms, the manifold representation reduces this dimensionality to *D*+6, where *D* is the number of rotatable dihedrals, and 6 degrees of freedom are responsible for the rigid body movements of the molecule as a whole. Drastic reduction in the dimensionality of the problem allows significant speed-up [5].
In the current version of the protocol, we did not implement the rotations of the internal dihedrals of the cyclic part of the ligand, treating it as a single rigid cluster. However, since the cyclic part was sampled during the conformers generation, its flexibility was partially accounted for during the selection of starting poses and the following full-atom RM refinement. During the minimization procedure, we used the same energy function as for the RM refinement, including harmonic restraints [15, 16]. For each starting conformation, we performed 10,000 Monte Carlo steps (kT=2.0 kcal/mol), from which the pose with the lowest energy was selected, and additionally minimized without restraints. The MMCM approach was used only in Stage 1b.

Molecular Dynamics

This refinement procedure, used only in Stage 1b, was reserved for targets for which templates were similar to each other, suggesting a high degree of certainty in the starting model. The starting configuration for the refinement MD simulations was the top-1 pose from our Stage 1a submission, according to our ranking. We used Amber ff14SB force field [19] to model the protein, and the GAFF force field [20] to model the ligand. The protonation state of the ligand was determined based on the experimental conditions provided by the organizers. Ligand atom partial charges were assigned using the AM1-BCC method [21] implemented in the antechamber module of Amber [22]. Each structure (this includes protein, ligand, crystallographic water molecules, and, in some cases, sulfate ions and glycerol molecules) was solvated using Leap [17] with a TIP3P [23] octahedral water box and at least a 10 Å buffer region between any atom of the system and the edge of the box. Na+ or Cl− ions were added as needed in order to neutralize the total charge of the system [24].
The MD refinement procedure is similar to the one described in Ref. [2]. For each system, first, a multistage minimization and equilibration protocol is carried for 2.05 ns [25]. Then, an MD production run is carried for 200 ns with 4.0 fs timestep, at 300 K and 1 atm. *Hard* restraints (50 kcal/mol/Å^2) were applied to protein heavy atoms, sulfate ions, and glycerol molecules; crystallographic waters were restrained with stronger springs, k=100 kcal/mol/Å^2. *Soft* restraints (2.5 kcal/mol/Å^2) were used for the ligand. This kept the protein close to the crystallographic structure while allowing some degree of relaxation for the side-chains and the ligand.

MELD-accelerated Molecular Dynamics

This refinement procedure was used only in those Stage 1b cases where the chosen templates were significantly different from each other. MELD-accelerated MD (MELD × MD) uses external information to reduce the phase space of physics-based simulations [7, 8]. This is achieved by energetically penalizing the regions of the phase space that do not agree with the information. Since no energetic bias is applied to areas of the phase space that agree with external information, the relative population of these basins is consistent with the relative population of unbiased simulations and can, therefore, be used as a proxy for free energy. In order to jump between the different basins created by the introduction of the information, energy-bias replica exchange simulations are necessary. For each MELD × MD refinement target, structure information of all five starting poses was incorporated into the simulation in two ways:

1. Ligand heavy atoms positions shared by all the poses were restrained using MELD cartesian restraints (delta=1 Å, k=5 kcal/mol/Å^2). Other ligand atoms were left unrestrained. Visual inspection was used to identify which ligand heavy atoms are shared between all the poses (and therefore are restrained during the MELD × MD simulation).
2. Protein and crystallographic water heavy atoms were restrained using hard MELD cartesian restraints (delta=0.5 Å, k=5 kcal/mol/Å^2).

To reduce the convergence time of simulations, the five starting poses were seeded along the replica ladder. To further reduce the simulation time, only part of the receptor was simulated in MELD. This part was selected by searching for receptor residues, cofactors, or crystal waters having a heavy atom within 10 Å from the ligand. The system was subjected to the same minimization and equilibration protocol as in the MD protocol before refining with MELD. The MELD simulation was run in TIP3P explicit solvent environment with REST2 solute-tempering technique [26]. The effective temperature was scaled from 300 K to 400 K with MELD Geometric Temperature Scaler. Hydrogen mass repartition was applied. The simulation time step was 4.5 fs. An 8 Å cut-off was used for all interactions.

Clustering Protocol
The following procedure was used to analyze the MD and MELD trajectories. The trajectories were clustered without using any information about starting poses, and the cluster centroids were chosen as final predictions. In the case of MD refinements, the trajectories (one per target) were clustered using the whole 200 ns of trajectory. In the case of MELD simulations, only trajectories of the three lowest replicas were clustered. We used the DBSCAN clustering algorithm [27] implemented in scikit-learn [28]. The distance cut-off was 5.0 Å and the population cut-off for identifying core points was 20. The distance metric used was the ligand RMSD (LRMSD) computed on all ligand heavy atoms after aligning the receptor Cα's.

Scoring and ranking

For each target, the results of all refinement methods (RM for Stage 1a; RM, MCMM, MD, and MELD for Stage 1b) were combined in a single pool and scored together. The results were clustered using the Butina algorithm [29] from RDKit. For each obtained cluster centroid, AutoDock Vina-based score [30], CHARMM-based energy score, and cluster size were calculated. The final model selection was done manually based on these scores and the fit of the model to known crystallographic ligands binding to the same pocket.

Results and discussion

In the analysis of the results below, we used native crystallographic poses provided by the organizers as a reference for RMSD calculation. A custom comparison tool was used, which caused minor discrepancies between the RMSD values reported here and in the official D3R GC4 rankings.

Macrocycle sampling

One of the main challenges of the current D3R round was a conformational sampling of 19 macrocyclic target ligands (BACE_1-BACE_19). Their main ring contained 14 to 16 atoms, one (BACE_8-BACE_10, BACE_12-BACE_19) or two (BACE_1-BACE_7, BACE_11) peptide bonds, up to two fused planar aromatic benzene rings and has up to three flexible sidechains. Besides that, in two target ligands (BACE_2 and BACE_13) one of the sidechains additionally included independent flexible hex-ring.

Although it is important to have an overall good conformer as a refinement starting point, we wanted to focus most of our efforts on a conformational sampling of the main rings of the target molecules. The main reason to do so is the relative complexity (because of the loop-closure condition) of the conformational space and energy landscape of the macrocyclic parts wrt sidechain parts. Unlike the macrocyclic parts, the sidechains can be efficiently additionally sampled with the refinement protocols that we have, thus increasing the value of conformational sampling success for macrocyclic parts.

We compared two conformer generators, ETKDG from RDKit [10] and BRIKARD [4] to determine their effectiveness for sampling macrocycles. Specifically, we generated 10^5 conformers using both tools, minimized them in vacuum using Merck Molecular Force Field (MMFF), and calculated RMSDs to the native structures only for macrocyclic parts (all atoms rigidly attached to the main ring) of the molecules (macrocyclic RMSD).

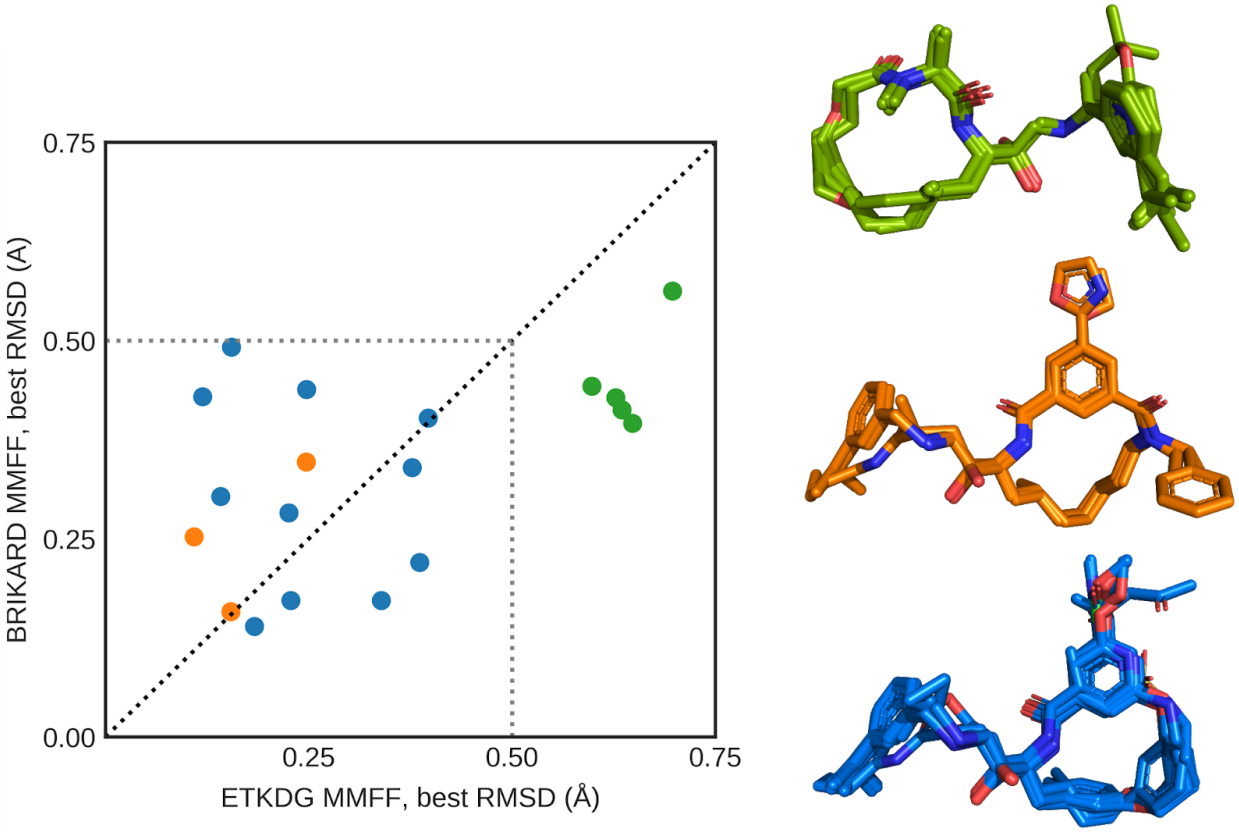

Fig. 1 Left: Comparison of the best macrocyclic RMSD generated by ETKDG MMFF or BRIKARD MMFF for each macrocyclic target ligand (BACE_1-19). Right: Native structures of the macrocyclic target ligands. Green structures (BACE_2-BACE_5, BACE_11) have two peptide bonds (one of them has methyl nitrogen-sidechain), orange structures (BACE_1, BACE_6-7) have two peptide bonds (one of them has massive and flexible nitrogen-sidechain), blue structures (BACE_8-10, BACE_12-19) has one peptide bond. The color of the points on the left corresponds to the carbon-color of the structures on the right.

The result of the comparison is presented in Fig. 1, where we have divided all macrocyclic target ligands into 3 colored subgroups based on the number and type of peptide bonds within the macrocycle. These peptide bonds significantly complicate the sampling of macrocycles by introducing energetic barriers in addition to the already nontrivial energy landscape of macrocycles. The blue group macrocycles have only one plain peptide bond in trans native conformation. The orange group has an additional peptide bond in cis native conformation with massive and flexible nitrogen-sidechain. This sidechain prevents the second peptide bond from being in trans conformation because of possible steric clashes with the macrocycle itself, thus simplifying the sampling problem. In terms of best macrocyclic RMSD performance of both ETKDG MMFF and BRIKARD MMFF for both subgroups of macrocycles mentioned before is comparable (see Fig. 1).

A significantly different situation is for the green subgroup macrocycles, which also contain second peptide bonds but with methyl nitrogen-sidechain and in trans native conformation (Fig. 1). For these peptide bonds, we found that the structures produced by ETKDG highly over-represented cis conformation. The distribution of the dihedral angle

(highlighted in Fig. 2a) is shown in Fig. 2b, with most of the samples being in a near-cis state. We attribute it to the relaxation associated with loop closure via fragment assembly, employed by ETKDG, so the sampling exhaustiveness is sacrificed in favor of a loop-closure condition satisfaction. After MMFF minimization, almost all generated conformers were in a cis state with trans state being undersampled. In Fig. 2c we can see that the inverse-kinematic approach of BRIKARD allows broad uniform sampling of all dihedral rotations. After performing MMFF minimization, both cis and trans states are populated.

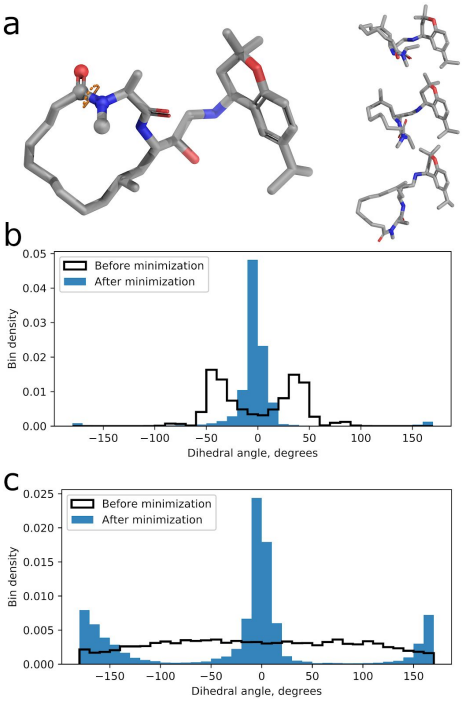

Fig. 2 (a) Left: Structure of one of the submitted BACE_2 ligand poses. One of the peptide bond dihedrals is highlighted. Right: Three of the alternative structures for BACE_2 sampled by BRIKARD. (b) The distribution of the peptide bond dihedral in structures generated by ETKDG, before (white bars) and after (blue bars) MMFF energy minimization. (c) The distribution of the peptide bond dihedral in structures generated by BRIKARD in broad sampling mode, before (white bars) and after (blue bars) MMFF energy minimization.

We believe that for even more complex macrocyclic systems with multiple dihedral barriers and multiple rings fusion the ability of BRIKARD to uniformly sample all pivot macrocyclic dihedrals intrinsically taking into account multiple loops closure condition is of high significance and helps to improve results as for example was demonstrated in [4].

We're planning to make BRIKARD publicly available as an automated server. Users will be able to upload the starting cyclic structure or amino acid sequence (in case of peptidic macrocycle) and specify the desired number of conformers (up to 10^5). Users will also be able to minimize the energy of each generated conformation, optionally including distance restraints derived from NMR spectrum or using direct optimization of 2D NMR spectrum (NOESY). After the job is completed, the user will be able to download all the sampled minimized/non-minimized conformers or only the cluster representatives.

Restrained Minimization

Despite its simplicity, the restrained minimization protocol did significantly improve starting poses in Stage 1a. The results can be seen in Fig. 3, where we show the RMSD values for the best-refined pose (with the lowest RMSD to reported native structure) and the corresponding starting pose. While in all cases except BACE_12 starting pose has RMSD over 1 Å, in most cases energy-based RM succeeds in lowering it into the sub-angstrom range.

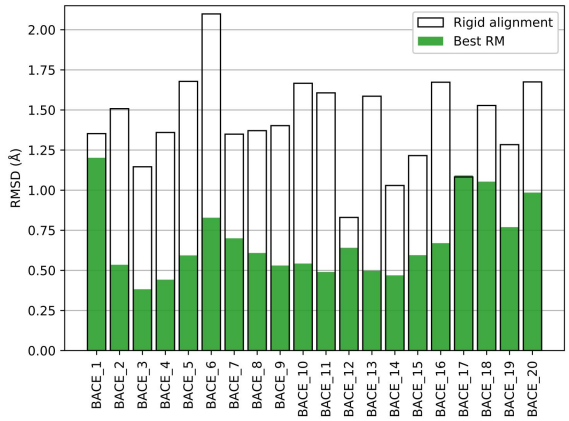

Fig. 3 Comparison of RM refinement results and starting conformation. The RMSD of the best RM refined pose (green bar) and the corresponding starting pose (white boxes) against the native pose

As expected for a template-based method, the accuracy of refined structure tends to improve with higher template similarity. This effect can be seen in Fig. 4, where we compare RMSD for the best (lowest RMSD) Stage 1a predictions versus corresponding MCS coverages. Filled markers correspond to the refined structures, while empty markers with the same MCS coverages correspond to the starting structures. We also visually separated "weak" MCS (shown as circles) and "strict" MCS (shown as squares). We see that finding a good template, with high MCS coverage, is an important step for obtaining high-quality models with the protocol used.
Insets (a) and (b) in Fig. 4 showcase two examples of refinement achieved by the RM approach, for targets BACE_6 and BACE_11, respectively. In the case of BACE_6 (inset a), significant changes in both "tail" and macrocycle are observed, leading to RMSD reduction from 2.10 Å to 0.83 Å. In the case of BACE_11 (inset b), the tail is relatively constant, but a ring flip happens in the macrocyclic part, reducing RMSD from 1.61 Å to 0.49 Å.

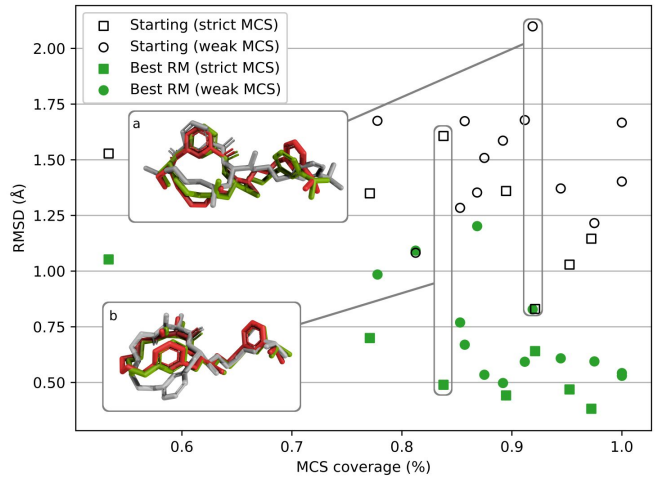

Fig. 4 The dependency between obtained RMSD and corresponding MCS coverage. For each of 20 target ligands, two points are shown: a filled one for the best (lowest RMSD) RM-refined pose, and an empty one for the corresponding starting pose. The shape of a point indicates used MCS "flavor": square for "strict", and circle for "weak". Inset (a) shows an example of RM refinement for BACE_6, while inset (b) shows RM refinement for BACE_11. In both cases, the native structure is shown in red, the starting structure in gray, and the RM-refined structure in green

Since the starting poses were built by aligning conformer to the template without taking into account the pocket environment, clashes between the small molecule and protein could occur. One of the structures for BACE_18 might serve as an example of such a situation, as shown in Fig. 5. The clash between a protein loop and starting macrocycle pose was resolved by RM, reducing RMSD from 4.10 Å to 1.80 Å.

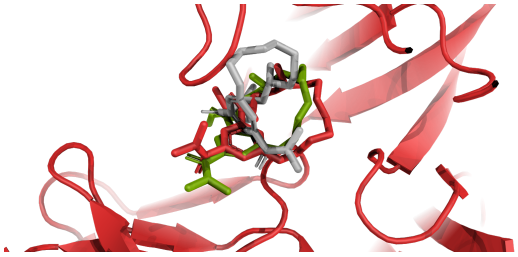

Fig. 5. Example of macrocycle refinement with RM protocol for target BACE_18. Starting pose, obtained by rigid alignment of sampled structure to the template, is shown in gray. The native pose and RM refined pose are shown in red and green, respectively. Receptor backbone in native structure is shown in red and is very similar to the backbone in the refined and starting structures

Monte Carlo on Manifold Minimization

While RM refinement could perform a significant conformational change, it was limited to exploring only local minima. Although perturbations on manifold didn't affect macrocycle, they did extensively sample side-chains, and subsequent full-atom relaxation led to adjustments of the macrocycle structure as well. For example, Fig. 6 shows the comparison of MCMM and RM refinement for BACE_1 ligand in Stage 1b of the competition. While RM reduced RMSD from 1.20 Å (starting pose, shown in gray licorice) to 0.94 Å (green licorice), it failed to establish hydrogen bonds with nearby water molecules, despite hydrogen-bonding terms being included in the forcefield. MCMM, on the other hand, sampled ligand side-chains enough to find the conformation (teal licorice) with these hydrogen bonds, reducing RMSD to 0.64 Å.

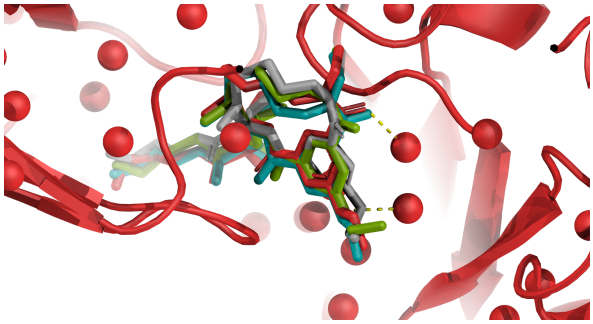

Fig. 6 Comparison of MCMM refinement result (teal) and RM refinement result (green) for BACE_1 in the presence of crystallographic water. Starting pose is shown in gray; native pose, including receptor and water oxygens, is shown in red. Two hydrogen bonds, formed by native and MCMM-refined poses, and not formed by starting and RM-refined poses, are shown as yellow dashed lines

However, extensive conformational sampling combined with general forcefield often led to suboptimal results, and MCMM outperformed RM, in terms of best RMSD, only on 6 targets.

Molecular Dynamics

During the competition, conventional MD refinement was not used for systems BACE_8,9,17,18,19,20. However, to make the discussion of the results more complete, the MD refinement protocol was run afterward for the MELD-refined systems as well. We can see that using MD we can slightly refine (most) poses.
Fig. 7 shows the quality of the initial structures (*i.e.*, the top-1 structure from Stage 1a submission), and how restrained MD can improve the structure of 14 systems. Of these, BACE_2,6,7,10,17,20 have improvement of 0.2 Å or more. Only in two cases (BACE_14 and BACE_17) of the seven remaining systems the quality of the refined structure is slightly worse than the starting structures. This shows that, in most cases, restrained MD simulations can slightly improve the quality of good starting structures. Since the positions of the heavy atoms of the ligand are restrained, the final structures of our simulations can not deviate significantly from the initial ones. Improvement using this approach are therefore limited to fractions of angstrom. This is still a helpful step for systems where there is a consensus between starting poses since it "relaxes" the physics of the system.

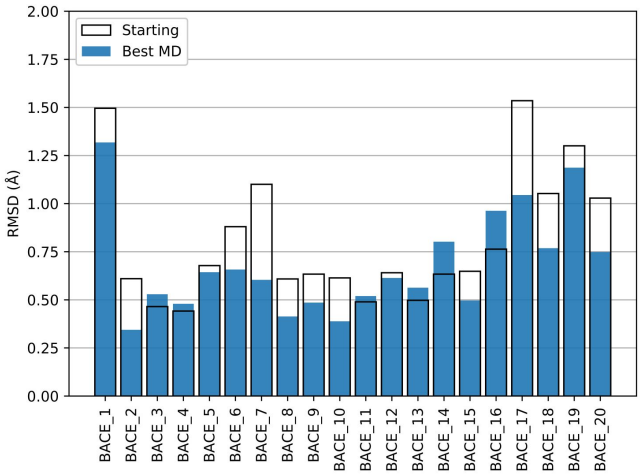

Fig. 7 Comparison of MD refinement results and starting conformation. The RMSD of the best MD refined pose (blue bars) and the starting pose (white boxes) against native poses

MELD-accelerated Molecular Dynamics

During Stage 1b of the competition, MELD was run only for targets where starting poses had relatively different structures, namely BACE_8,9,12,17,18,19,20. In these cases, MELD × MD is able to recognize the best structure and refine it. Fig. 8 shows that, for the seven systems for which alternative poses are available, MELD × MD simulations allow to consistently identify and refine the best pose. For two of the systems, BACE_8 and BACE_9, the top-1 starting pose is already pretty close to native. MELD correctly identifies it and refines it. *I.e.*, both MD refinement and MELD refinement yield poses that are similar to each other and the top-1 starting structure. For BACE_17, the top-1 pose (gray licorice in the structural representation) is not the correct one. MELD × MD correctly identifies and refines the correct pose (orange licorice). In this case, MD is also able to relax the incorrect top-1 pose to one closer (blue licorice) to the native one (red licorice), since we leave that portion of the ligand unrestrained. For the BACE_19 system, the ring in the top-1 pose is flipped (gray licorice) compared to the native pose (red licorice). MD is not able to relax its orientation (blue licorice). MELD is able to identify the pose with the

correct ring orientation and refine it (orange licorice). For BACE_18, MELD gives slightly worse pose compared to MD. Visualizing the whole trajectory shows that MELD does sample better poses, but the clustering protocol fails to extract them (Fig. S1). Clustering with a more strict cut-off might yield better poses.

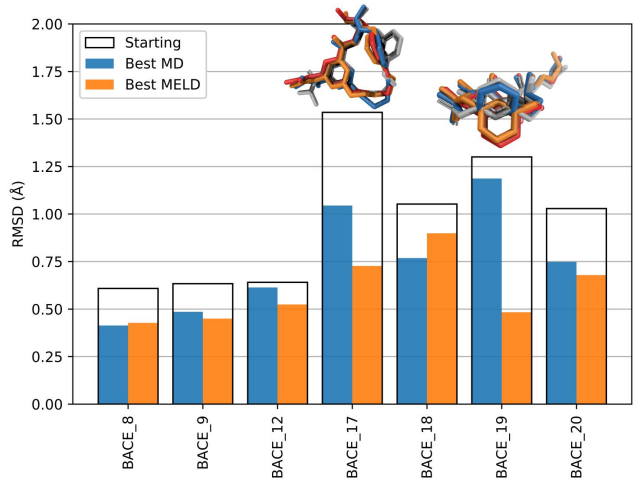

Fig. 8 Comparison of MELD refinement results, MD refinement results, and top-1 starting conformations. The best MELD refined pose, best MD refined pose, and top-1 starting pose are shown in orange bars, blue bars, and white boxes, respectively. The corresponding structures for BACE_17 and BACE_19 are shown above the bars in the same colors; the native structure is shown in red

Summary

Template search and restrained minimization were enough to achieve low-RMSD scores in Stage 1a of D3R GC4. As evaluated by the organizers, our mean closest RMSD (best of five for each target) value was 0.65 Å, with 18 targets having a sub-angstrom accuracy, and our mean pose-1 RMSD was 0.77±0.34 Å, with 15 targets in sub-angstrom range (Fig. 9a). This placed our group first in official Stage 1a rankings by both mean closest and mean pose-1 RMSD (Fig. 10a).

In Stage 1b, with receptor crystallographic structures and a higher number of refinement protocols, the accuracy was further improved. MCMM, MD, and MELD refinement protocols were used to refine the ligand starting conformation, other than strongly perturb it, which often resulted in drastic RMSD improvements. Unfortunately, due to technical issues, we were unable to analyze and submit all refinement results in the time allotted. Therefore, not all of the discussed results were included in the official submission, and the overall improvement was marginal compared to Stage 1a. Still, for some targets, we did improve RMSD drastically. *E.g.*, closest RMSD for BACE_20 was 0.97 Å in Stage 1a, and 0.65 Å in Stage 1b submission. According to the official evaluation, we achieved mean closest RMSD 0.63 Å, and mean pose-1 RMSD 0.76±0.32 Å (Fig. 9b). Among all participants, we were ranked second in terms of mean closest RMSD, and third in terms of mean pose-1 RMSD (Fig. 10b).

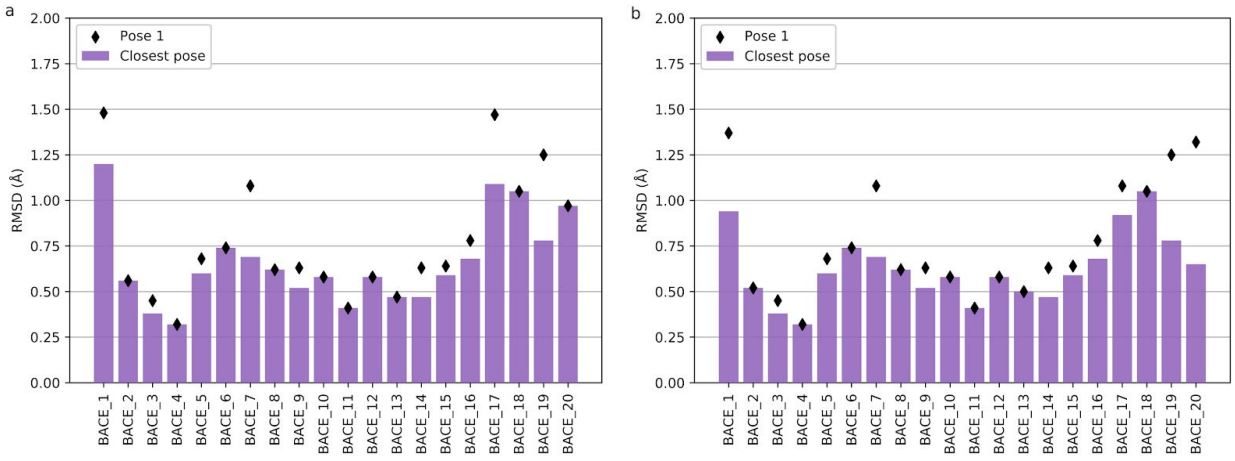

Fig. 9 The closest RMSD (bars) and pose-1 RMSD (diamonds) of submitted structures in Stages 1a (subfigure a, left) and 1b (subfigure b, right), as calculated by the D3R GC4 organizers

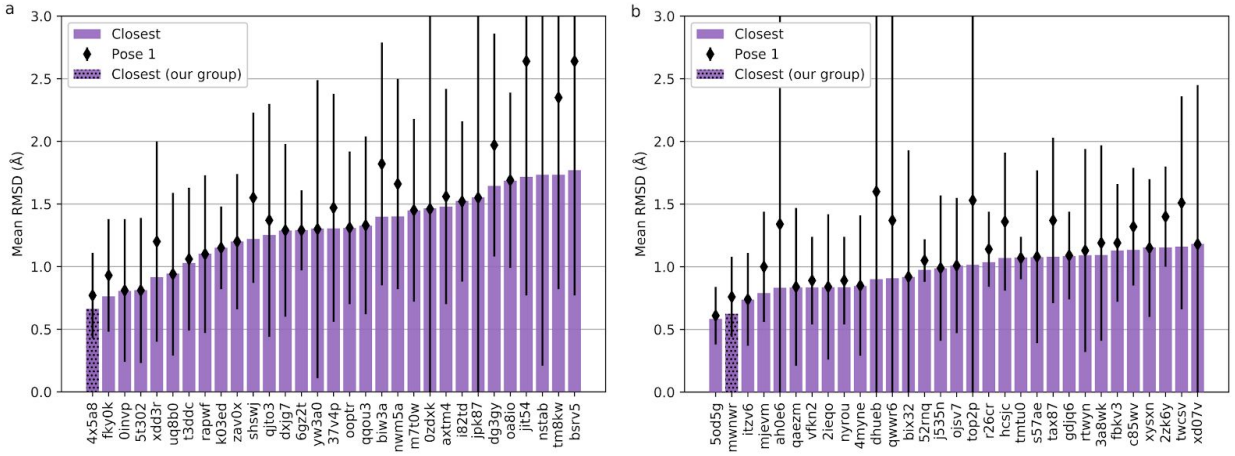

Fig. 10 The top-30 predictors by mean closest RMSD (bars) in official D3R GC4 rankings. Corresponding mean pose-1 RMSD's with standard deviations are shown as diamonds with whiskers above the bars. Results for Stage 1a are shown in subfigure a (left); results for Stage 1b are shown in subfigure b (right). The scores or our group are shown as hatched bars

Conclusion

Template-based approach has been used in previous rounds of D3R Grand Challenge for protein-ligand docking [31–33], as well as in community-wide protein-protein docking competitions [34–37], and for protein-RNA docking [38, 39].

Our team has developed a novel approach for template-based small molecular docking and demonstrated its efficiency in D3R GC4, where we predicted the poses of 20 compounds with 0.76 Å mean pose-1 RMSD and 0.66 Å mean closest (best of five) RMSD. In terms of pose-1, 15 out of 20 compounds were predicted with sub-angstrom accuracy. This has placed us among the best performers in the pose prediction challenge.

In the approach, we incorporated ample existing information on the bound ligands for the system of interest and the tools previously developed by us. Armed with this data, we were able to produce low-RMSD poses even with a

simplistic local refinement method. More advanced structure refinement protocols were able to reduce RMSD even further in Stage 1b, although no single method is clearly superior to others.
Besides, while in many cases sampling and refinement steps produced very high-quality structures, they were not chosen for submission, indicating that the scoring and ranking protocol has room for improvement. In particular, a better scoring function would be crucial for ranking the obtained poses. While, for example, the Vina scores did not show a high correlation with RMSD even within multiple poses of the same ligand (see Supplementary Fig. S2), we believe that an automated scoring method can be devised instead of relying on human experts as we did in this competition. Of particular interest are knowledge-based scoring functions due to their good performance in D3R ranking stages in this and previous years [40–43]. Another possible direction for scoring function optimization is the inclusion of a density-based score, which measures the similarity of the structure to known crystallographic complexes of related compounds.

Acknowledgments

This work was supported by National Institutes of Health grants R21 GM127952 and 1R01GM125813-01; National Science Foundation grants AF 1816314, AF 1645512, DBI 1759277; Russian Science Foundation grant 19-74-00090; and the National Science Foundation PRAC Award ACI-1713695. Part of this research has been enabled by the Blue Waters sustained-petascale computing project, which is supported by the National Science Foundation (Awards OCI-0725070 and ACI-1238993) and the state of Illinois. Blue Waters is a joint effort of the University of Illinois at Urbana−Champaign and its National Center for Supercomputing Applications.

References

1. Padhorny D, Hall DR, Mirzaei H, et al (2018) Protein–ligand docking using FFT based sampling: D3R case study. J Comput Aided Mol Des 32:225–230. https://doi.org/10.1007/s10822-017-0069-7

2. Ignatov M, Liu C, Alekseenko A, et al (2019) Monte Carlo on the manifold and MD refinement for binding pose prediction of protein–ligand complexes: 2017 D3R Grand Challenge. J Comput Aided Mol Des 33:119–127. https://doi.org/10.1007/s10822-018-0176-0

3. Vassar R, Bennett BD, Babu-Khan S, et al (1999) Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. Science 286:735–741. https://doi.org/10.1126/science.286.5440.735

4. Coutsias EA, Lexa KW, Wester MJ, et al (2016) Exhaustive Conformational Sampling of Complex Fused Ring Macrocycles Using Inverse Kinematics. J Chem Theory Comput 12:4674–4687. https://doi.org/10.1021/acs.jctc.6b00250

5. Mirzaei H, Beglov D, Paschalidis IC, et al (2012) Rigid Body Energy Minimization on Manifolds for Molecular Docking. J Chem Theory Comput 8:4374–4380. https://doi.org/10.1021/ct300272j

6. Mirzaei H, Zarbafian S, Villar E, et al (2015) Energy Minimization on Manifolds for Docking Flexible Molecules. J Chem Theory Comput 11:1063–1076. https://doi.org/10.1021/ct500155t

7. Perez A, MacCallum JL, Dill KA (2015) Accelerating molecular simulations of proteins using Bayesian inference on weak information. Proc Natl Acad Sci U S A 112:11846–11851.

https://doi.org/10.1073/pnas.1515561112

8. MacCallum JL, Perez A, Dill KA (2015) Determining protein structures by combining semireliable data with atomistic physical models by Bayesian inference. Proc Natl Acad Sci U S A 112:6985–6990. https://doi.org/10.1073/pnas.1506788112

9. Riniker S, Landrum GA (2015) Better Informed Distance Geometry: Using What We Know To Improve Conformation Generation. J Chem Inf Model 55:2562–2574. https://doi.org/10.1021/acs.jcim.5b00654

10. Landrum G (2019) RDKit: Open-source cheminformatics https://rdkit.org/

11. Coutsias EA, Wester MJ (2019) Multiple Loop Closure and Ring Perception (unpublished)

12. Halgren TA (1996) Merck molecular force field. I. Basis, form, scope, parameterization, and performance of MMFF94. J Comput Chem 17:490–519

13. Tosco P, Stiefl N, Landrum G (2014) Bringing the MMFF force field to the RDKit: implementation and validation. J Cheminform 6:37. https://doi.org/10.1186/s13321-014-0037-3

14. Moghadasi M, Mirzaei H, Mamonov A, et al (2015) The impact of side-chain packing on protein docking refinement. J Chem Inf Model 55:872–881. https://doi.org/10.1021/ci500380a

15. Kortemme T, Morozov AV, Baker D (2003) An Orientation-dependent Hydrogen Bonding Potential Improves Prediction of Specificity and Structure for Proteins and Protein–Protein Complexes. J Mol Biol 326:1239–1259. https://doi.org/10.1016/S0022-2836(03)00021-4

16. Wedemeyer WJ, Baker D (2003) Efficient minimization of angle-dependent potentials for polypeptides in internal coordinates. Proteins 53:262–272. https://doi.org/10.1002/prot.10525

17. Liu DC, Nocedal J (1989) On the limited memory BFGS method for large scale optimization. Math Program 45:503–528. https://doi.org/10.1007/BF01589116

18. Mamonov AB, Moghadasi M, Mirzaei H, et al (2016) Focused grid-based resampling for protein docking and mapping. J Comput Chem 37:961–970. https://doi.org/10.1002/jcc.24273

19. Maier JA, Martinez C, Kasavajhala K, et al (2015) ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. J Chem Theory Comput 11:3696–3713. https://doi.org/10.1021/acs.jctc.5b00255

20. Wang J, Wolf RM, Caldwell JW, et al (2004) Development and testing of a general amber force field. J Comput Chem 25:1157–1174. https://doi.org/10.1002/jcc.20035

21. Jakalian A, Bush BL, Jack DB, Bayly CI (2000) Fast, efficient generation of high-quality atomic charges. AM1-BCC model: I. Method. J Comput Chem 21:132–146

22. Case DA, Betz RM, Cerutti DS, et al (2016) AMBER 2016 Reference Manual

23. Jorgensen WL, Chandrasekhar J, Madura JD, et al (1983) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79:926–935. https://doi.org/10.1063/1.445869

24. Joung IS, Cheatham TE 3rd (2008) Determination of alkali and halide monovalent ion parameters for use in explicitly solvated biomolecular simulations. J Phys Chem B 112:9020–9041. https://doi.org/10.1021/jp8001614

25. Hornak V, Okur A, Rizzo RC, Simmerling C (2006) HIV-1 protease flaps spontaneously open and reclose in molecular dynamics simulations. Proc Natl Acad Sci U S A 103:915–920. https://doi.org/10.1073/pnas.0508452103

26. Wang L, Friesner RA, Berne BJ (2011) Replica exchange with solute scaling: a more efficient version of replica exchange with solute tempering (REST2). J Phys Chem B 115:9431–9438. https://doi.org/10.1021/jp204407d

27. Ester M, Kriegel H-P, Sander J, Xu X (1996) A density-based algorithm for discovering clusters in large spatial databases with noise. In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining. AAAI Press, pp 226–231

28. Pedregosa F, Varoquaux G, Gramfort A, et al (2011) Scikit-learn: Machine Learning in Python. J Mach Learn Res 12:2825–2830

29. Butina D (1999) Unsupervised Data Base Clustering Based on Daylight's Fingerprint and Tanimoto Similarity: A Fast and Automated Way To Cluster Small and Large Data Sets. J Chem Inf Comput Sci 39:747–750. https://doi.org/10.1021/ci9803381

30. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 31:455–461. https://doi.org/10.1002/jcc.21334

31. Duan R, Xu X, Zou X (2018) Lessons learned from participating in D3R 2016 Grand Challenge 2: compounds targeting the farnesoid X receptor. J Comput Aided Mol Des 32:103–111. https://doi.org/10.1007/s10822-017-0082-x

32. da Silva Figueiredo Celestino Gomes P, Da Silva F, Bret G, Rognan D (2018) Ranking docking poses by graph matching of protein-ligand interactions: lessons learned from the D3R Grand Challenge 2. J Comput Aided Mol Des 32:75–87. https://doi.org/10.1007/s10822-017-0046-1

33. Fradera X, Verras A, Hu Y, et al (2018) Performance of multiple docking and refinement methods in the pose prediction D3R prospective Grand Challenge 2016. J Comput Aided Mol Des 32:113–127. https://doi.org/10.1007/s10822-017-0053-2

34. Lensink MF, Velankar S, Kryshtafovych A, et al (2016) Prediction of homoprotein and heteroprotein complexes by protein docking and template-based modeling: A CASP-CAPRI experiment. Proteins 84 Suppl 1:323–348. https://doi.org/10.1002/prot.25007

35. Kundrotas PJ, Anishchenko I, Badal VD, et al (2018) Modeling CAPRI targets 110-120 by template-based and free docking using contact potential and combined scoring function. Proteins 86 Suppl 1:302–310. https://doi.org/10.1002/prot.25380

36. Peterson LX, Shin W-H, Kim H, Kihara D (2018) Improved performance in CAPRI round 37 using LZerD docking and template-based modeling with combined scoring functions. Proteins 86 Suppl

1:311–320. https://doi.org/10.1002/prot.25376

37. Porter KA, Padhorny D, Desta I, et al (2019) Template-Based Modeling by ClusPro in CASP13 and the Potential for Using Co-evolutionary Information in Docking. Proteins: Struct Funct Bioinf

38. Zheng J, Kundrotas PJ, Vakser IA, Liu S (2016) Template-Based Modeling of Protein-RNA Interactions. PLoS Comput Biol 12:e1005120. https://doi.org/10.1371/journal.pcbi.1005120

39. Zheng J, Hong X, Xie J, et al (2019) P3DOCK: a protein-RNA docking webserver based on template-based and template-free docking. Bioinformatics. https://doi.org/10.1093/bioinformatics/btz478

40. Kadukova M, Grudinin S (2017) Convex-PL: a novel knowledge-based potential for protein-ligand interactions deduced from structural databases using convex optimization. J Comput Aided Mol Des 31:943–958. https://doi.org/10.1007/s10822-017-0068-8

41. Jiménez J, Škalič M, Martínez-Rosell G, De Fabritiis G (2018) KDEEP: Protein–Ligand Absolute Binding Affinity Prediction via 3D-Convolutional Neural Networks. J Chem Inf Model 58:287–296. https://doi.org/10.1021/acs.jcim.7b00650

42. Nguyen DD, Cang Z, Wu K, et al (2019) Mathematical deep learning for pose and binding affinity prediction and ranking in D3R Grand Challenges. J Comput Aided Mol Des 33:71–82. https://doi.org/10.1007/s10822-018-0146-6

43. Sunseri J, King JE, Francoeur PG, Koes DR (2019) Convolutional neural network scoring and minimization in the D3R 2017 community challenge. J Comput Aided Mol Des 33:19–34. https://doi.org/10.1007/s10822-018-0133-y

Supplementary Material

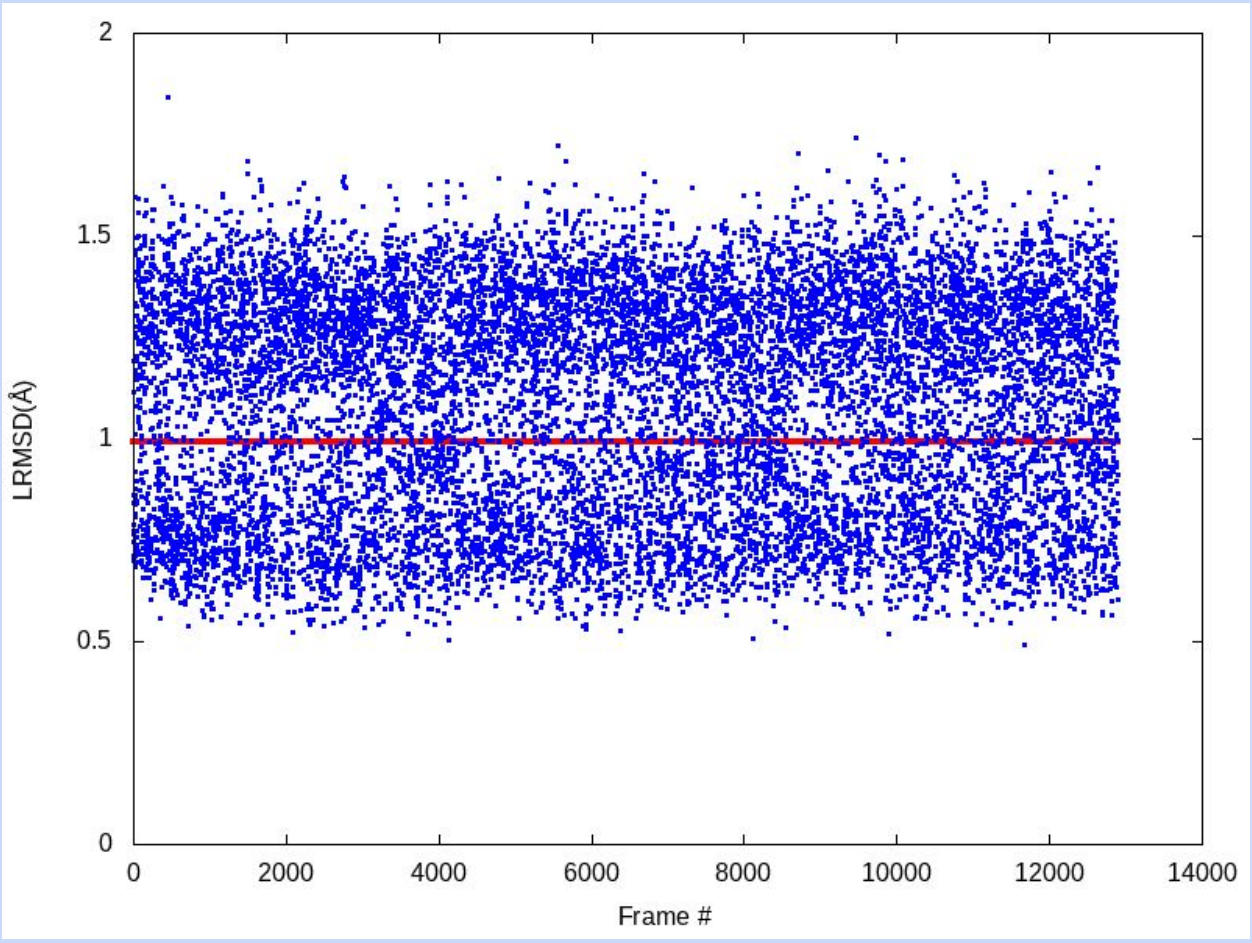

Fig. S1 LRMSD of MELD trajectory for BACE_18. The LRMSDs of all poses sampled during MELD simulation and the final best MELD refined pose are depicted as blue dots and a red line, respectively. MELD samples poses closer to native, but the current clustering protocol fails to extract them