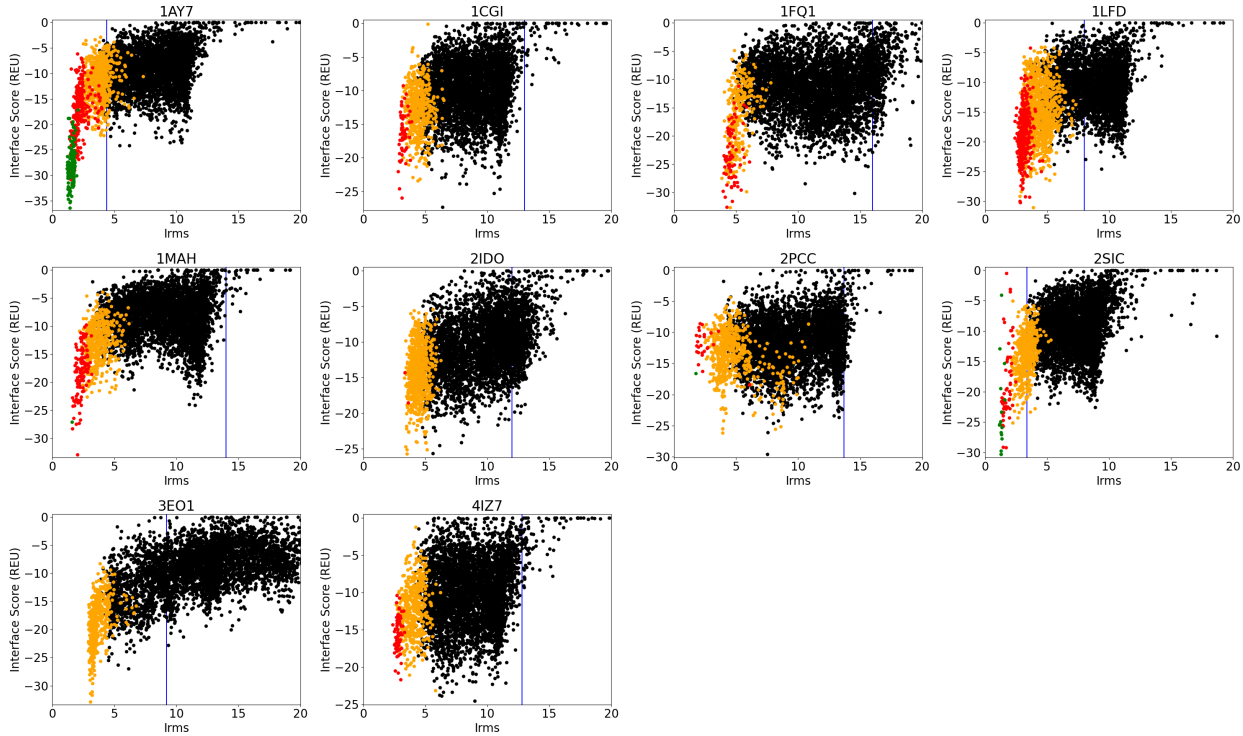
Scientific test: docking_ensemble

FAILURES

None

RESULTS



AUTHOR AND DATE

Adapted for the current benchmarking framework by Ameya Harmalkar (harmalkar.ameya24@gmail.com; Gray Lab), March 2021

PURPOSE OF THE TEST

This benchmark is meant to test how well we discriminate native protein-protein binding orientations from decoys based on the interface score term by performing flexible protein-protein docking experiments with conformer selection across a diverse set of protein-protein complexes.

BENCHMARK DATASET

The dataset consists of 3 protein-protein complexes extracted from the Docking Benchmark 5.0 (Vreven, T. et al. J. Mol. Biol., 2015). The set contains 1 rigid (conformational change < 1.5 Ang), 1 medium-flexible (conformational change between 1.5 and 2.2 Ang), and 1 difficult (conformational change > 2.2 Ang).

Structure preparation:

Proteins were extracted directly from the PDB according to the PDB ID of the native structure. They were then cleaned to ensure that both the bound and the unbound structures have the same residues. The unbound structures were superimposed on the bound and then the smaller partner (ligand) was moved away by 15 Ang and rotated by 60 degrees to scramble the interface. The ensembles were generated with relax, NMA and backrub protocols for each ligand and receptor chain/s respectively. Prepacking was performed with docking prepack protocol.

PROTOCOL

Protocol:

The ensemble docking protocol is based on the conformational-selection mechanism of protein docking. It utilizes an ensemble of the protein partners pre-generated with Rosetta Relax, Backrub and NMA to sample diverse backbone conformations while docking. In the low resolution stage, each docking run performs rigid-body translation and rotation around the protein partner while swapping backbones from the pre-generated ensemble. This allows the sampling of diverse backbone conformations while docking. In the high-resolution stage, an all-atomistic refinement is performed over the generated encounter complex and side-chains at the interface are packed for optimal binding. The output decoy is then evaluated with Rosetta score function and the best binding complex structure is determined on the basis of Interface Score.

Publication: The methodological details of the protocol, RosettaDock 4.0 and the performance on the benchmark have been thoroughly discussed in Marze, N. A., Roy Burman, S. S. et al. Bioinfo., 2018.

CPU hours:

The benchmark takes 3000-3500 CPU hrs for 5000 decoys. The debug mode takes roughly 1h 15 mins.

PERFORMANCE METRICS

Usually, to assess the performance of a docking simulation, the number of structures with a CAPRI-acceptable or better ranking are analyzed. CAPRI model rankings are based on a combination of factors like fraction of native contacts, ligand RMSD, and interface RMSD and are described in detail in Lensink, Wodak et al. 2019 PSFBI - Table 3. The protocol computes the CAPRI rankings for each model (as well as some of the metrics it is based on), which are written into the score file ("CAPRI_rank"). Rankings are the following:

0 - incorrect model (black)

1 - acceptable model (yellow)

2 - medium quality model (red)

3 - high quality model (green)

The output models are resampled via bootstrap to remove possible sampling biases. Docking complexes can also be classified via the N5 metric, classifying N5 >= 3 as successful. N5 >= 3 means that at least 3 out of the top 5 scoring models should be acceptable or better according to CAPRI metrics. In this test, we report this metric in the result.txt file but don't use it for a pass/fail criterion because most targets would fail according to it. This scientific test passes if all targets pass the following metrics:

(1) the highest CAPRI ranking sampled for any model should be equal or higher than the cutoff ranking (as computed in the first run) AND

(2) the interface RMSD of the top-scoring model should be equal or lower than the cutoff I_rmsd (as computed in the first run + 2A)

KEY RESULTS

Unbound structures are docked and compared to the bound native structure. Docking targets are classified into rigid, moderate and difficult based on the extent of backbone motion that the proteins undergo while transitioning from an Unbound to Bound state. This is estimated with Interface-RMSD (Unbound-to-Bound) such that, Rigid Targets : I-RMSD < 1.2 Angstrom; Medium Targets : 1.2 < I-RMSD < 2.2 A; Difficult Targets : I-RMSD > 2.2 A.

Out of the targets benchmarked, we have 4 rigid targets (1AY7, 1MAH, 2SIC, 2PCC), 4 medium targets (1CGI, 1LFD, 3EO1, 4IZ7) and 2 difficult targets (1FQ1, 2IDO)

DEFINITIONS AND COMMENTS

Note that the ensemble docking protocol swaps structures from a pre-generated ensemble of 100 structures (for each receptor and ligand). For adequate sampling and robust results, it might be necessary to generate upto 5000 decoy structures.

LIMITATIONS

This benchmark set includes a few targets from a much larger standard docking benchmark used in the docking community (Vreven, T. et al. J. Mol. Biol., 2015). Ensemble docking requires much more CPU hours than rigid body docking, hence, we have a limit on the size of the benchmark.