Protein Preparation and Docking Methods:

Protein Preparation:

PDB files of the targets were fetched from RCSB using PDB-Tools. For entries with multiple biological assemblies in their PDB, the assembly with the highest degree of coverage and lowest disorder was retained for use.

The co-crystal structures were prepared using Protein Prep Wizard[1] in Schrödinger Suites (version 2021-1) using the Command Line Interface. Bond orders were assigned, explicit hydrogens were added, and non-bridging waters (fewer than two H-bonds to non-waters) were removed. Het states were generated for pH 7.4 using Epik[2]. Hydrogen bonds were assigned and the protein was ionized using PROPKA at pH 7.4. A restrained minimization was performed using the OPLS4 force field[3] to a heavy atom RMSD convergence of 0.3Å. Alternative residue locations with the highest average occupancy were selected for use. If the alternative residue locations had equal average occupancy, the first presented location was arbitrarily selected.

CLI commands for protein pre-processing, preparation, structure splitting, and export can be found in:

protein-ligand-benchmark/preparation/CLI-commands.txt

Ligand Preparation:

Ligand preparation was performed using LigPrep[4] with the OPLS4 force field. Ionization and tautomeric states were generated using Epik at pH 7.4 $\pm$ 0.0.

Ligand preparation scripts can be found in:

protein-ligand-benchmark/preparation/ligand-prep

Grid Generation:

Using the Schrödinger Suites' receptor grid generator, the receptor-binding site was defined as the cubic centroid around the co-crystallized ligand. Nonpolar parts of the receptor were softened using Van der Waals radius scaling (scaling factor of 1.0 with a partial charge cutoff of 0.25).

Docking grid generation scripts can be found in:

protein-ligand-benchmark/preparation/glide-grids

Ligand Docking:

Ligands were docked into the prepared crystal structures using Glide SP[6] with flexible ligand sampling. Nitrogen inversions and ring conformations were sampled and Epik state penalties were added to the docking score. Core constraints were used to restrain the docked ligand to a position that overlaps with the maximum common substructure of the co-crystallized ligand within an RMSD of 0.25Å Post docking minimization was performed for all poses. Ligands were docked both with and without the constraints described above.

Docking scripts can be found in:

protein-ligand-benchmark/data/preparation/docking

References: (for methods)

1. **Schrödinger Release 2022-1**: Protein Preparation Wizard; Epik, Schrödinger, LLC, New York, NY, 2021; Impact, Schrödinger, LLC, New York, NY; Prime, Schrödinger, LLC, New York, NY, 2021.

2. **Schrödinger Release 2022-1**: Epik, Schrödinger, LLC, New York, NY, 2021.

3. **Schrödinger Release 2022-1**: LigPrep, Schrödinger, LLC, New York, NY, 2021.

4. *J. Chem. Theory Comput.* 2021, 17, 7, 4291–4300

5. **Schrödinger Release 2022-1**: Glide, Schrödinger, LLC, New York, NY, 2021.