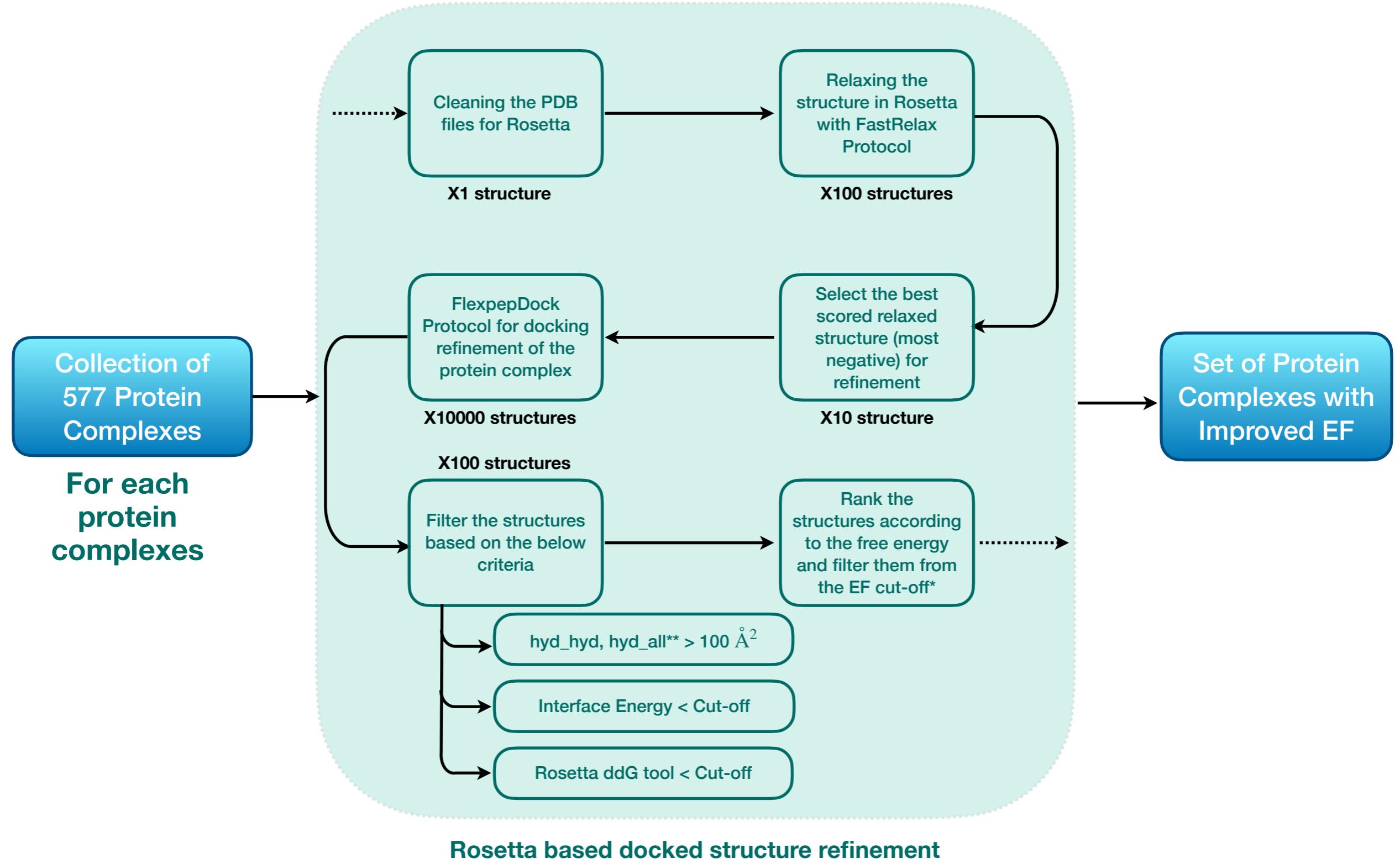


Proposed Pipelines for Improving PROTAC Ternary Complex Enrichment Factor

Nidhin Thomas

September 8th, 2022

Pipeline for Enrichment Factor Improvement



*Free energy cut-off based on 16.5% Enrichment Factor (EF) can be calculated from the input protein complex structures

**Overlap of hydrophobic patch area from receptor 1 and receptor 2 is calculated using GROMACS. Hydrophobic patch of receptor 1 with any patch of receptor 2 is also computed for filter.

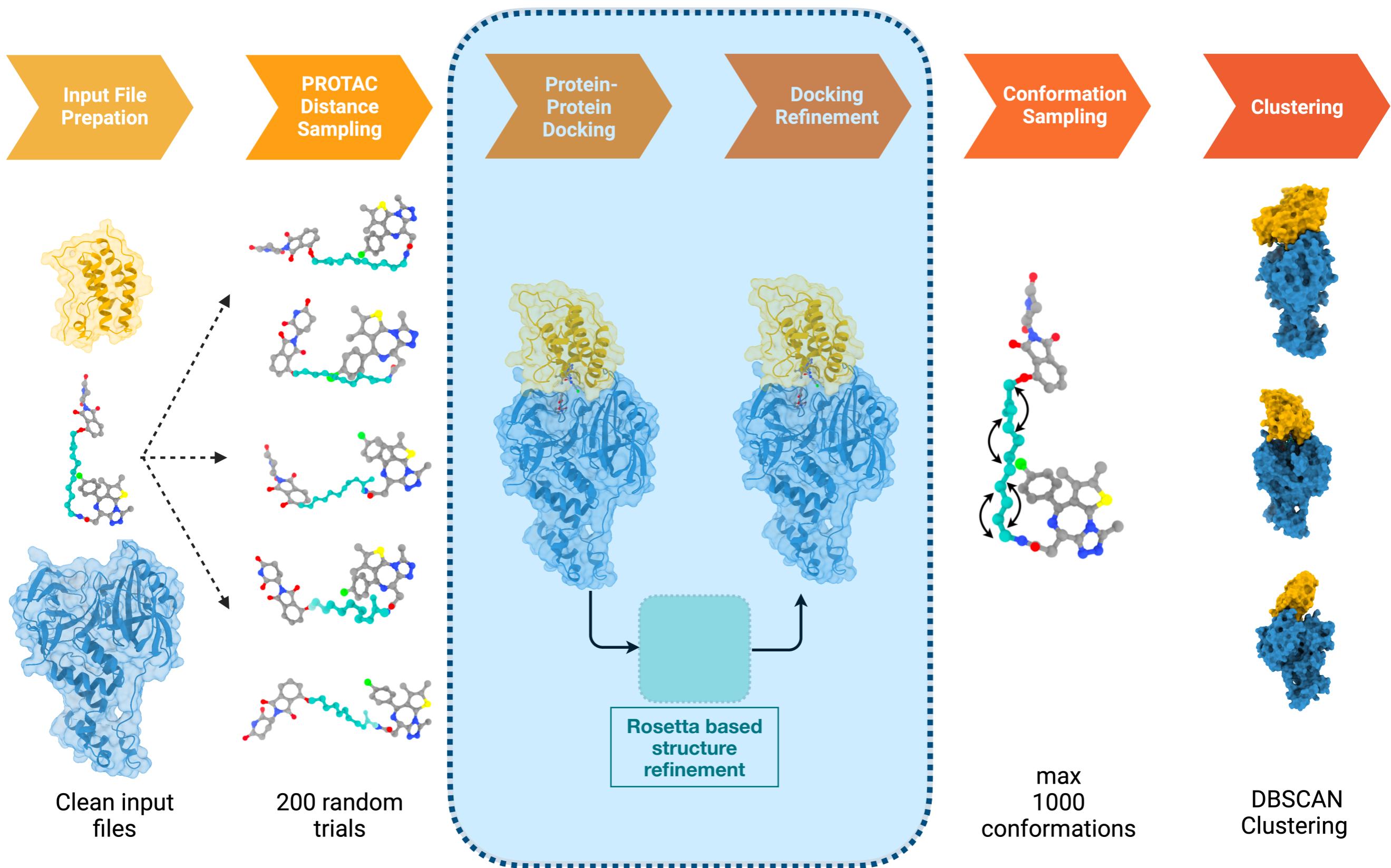
Brief description of the pipeline:

- The enrichment factor (EF) can be computed based on the protein complex alone.
- Several criteria can be used to filter the protein complexes such that the enrichment factor can be improved.
 1. Hydrophobic patch area overlap on the target protein (BCL-XL) and E3 ligase protein (VHL) $> 100 \text{ \AA}^2$.
 2. Hydrophobic patch area with any patch area overlap on the target protein (BCL-XL) and E3 ligase protein (VHL) $> 100 \text{ \AA}^2$ (GROMACS can be used).
 3. Rosetta based binding free energy for the complex from ddG protocol.
 4. Rosetta interface scores from Interface_Analyzer protocol.
 5. PROTAC total energy as a measure of it's strain energy compared to solution structure of PROTAC.
 6. Docking score.
- We can also generate more structures based on the available docked structures such that enrichment factor can be improved.
- The challenge with these near crystallographic structures from computational models is that docking of PROTAC ligand after their filtering may cause problems.
- Filtering the protein complexes with PROTAC present at the docking site would be more reasonable in this case.

Softwares used for Pipeline 1

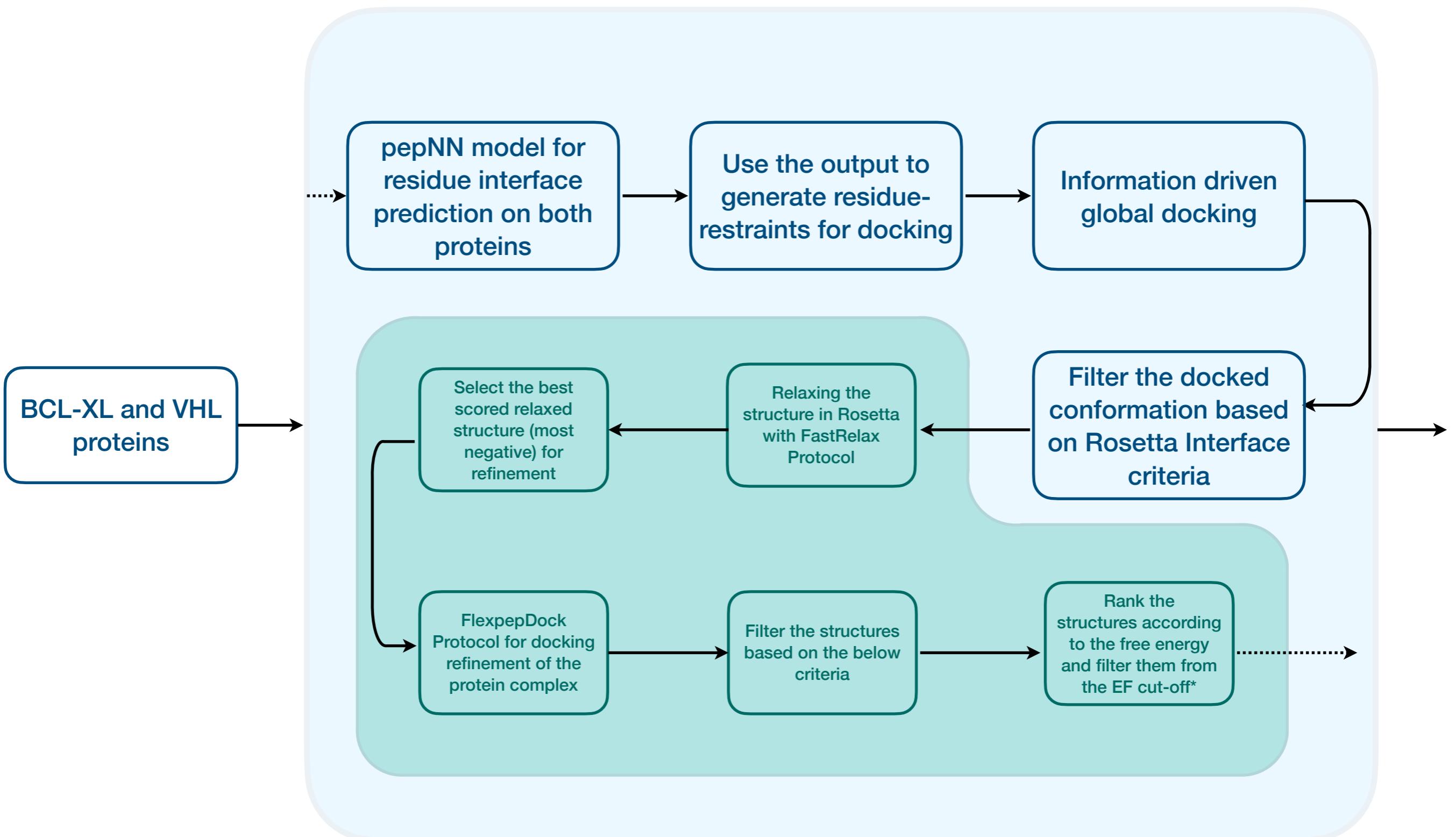
- Rosetta 3.3 (Docking refinement and Interface relaxation)
- GROMACS (SASA area calculation)
- RDKit (PROTAC conformation generation)
- PyMol (RMSD calculations)
- ChimeraX (Visualization)
- DBSCAN for clustering (if needed)
- Python 3 (Post Processing)
- Bash scripting (Automation)

Rosetta based PROTAC design and optimization (PRosettaC modified):



Zaidman, Daniel et.al. "PRosettaC: Rosetta based modeling of PROTAC mediated ternary complexes." Journal of chemical information and modeling (2020).

Protein-protein Docking and Refinement in Pipeline 2 in detail:



Softwares used for Pipeline 2

- Rosetta 3.3 (Docking refinement and Interface relaxation)
- pepNN deep learning model (residue-restraint prediction)
- HADDOCK/LightDock (Information driven or blind global docking)
- GROMACS (SASA area calculation)
- RDKit (PROTAC conformation generation)
- OMEGA or BCL (Ligand conformation sampling)
- PyMol (RMSD calculations)
- ChimeraX (Visualization)
- DBSCAN for clustering (if needed)
- Python 3 (Post Processing)
- Bash scripting (Automation)