Binding Data Selection

"Training test set" (~20) (to perform explicit simulation)

Our compiled "Test set" (~150 complexes with known binding affinity and structure)
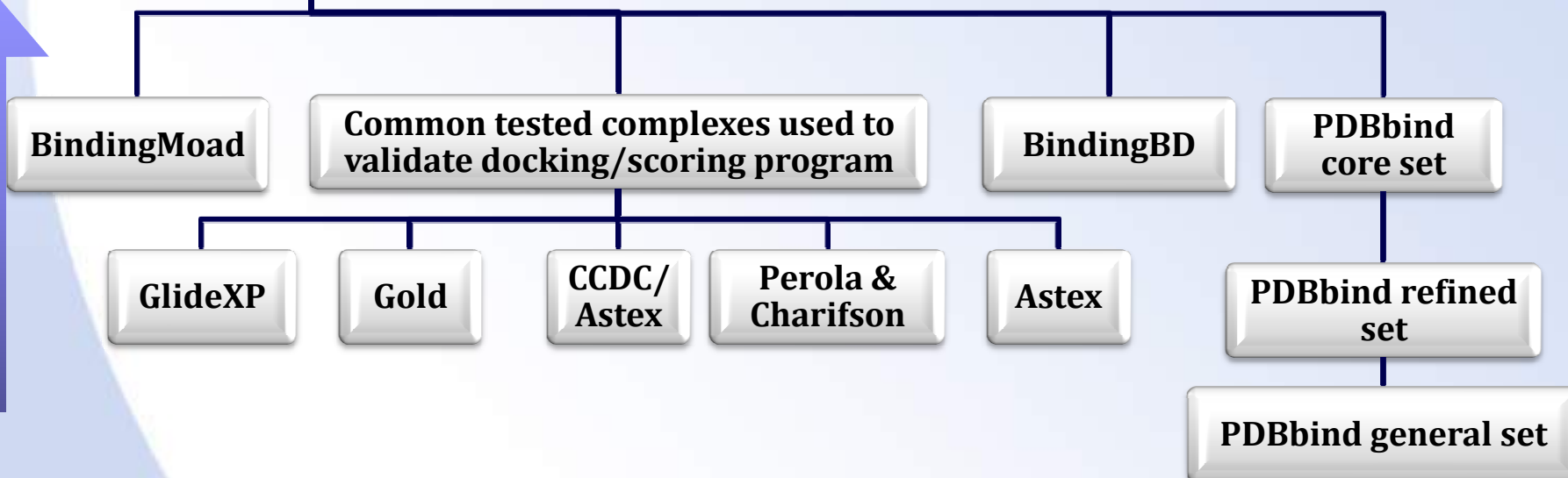
Examples of "filtering" criterions
1. Complexes are grouped into clusters by protein sequence similarity.
2. Complexes with minority binding constant type (Kd or Ki) in one cluster will be removed.
3. Complexes containing big ligands will be removed.
4. The complexes with pKd or pKi < 3 will be removed.
5. One entry for each structure.

BindingMoad

Common tested complexes used to validate docking/scoring program

BindingBD

PDBbind core set

GlideXP

Gold

CCDC/Astex

Perola & Charifson

Astex

PDBbind refined set

PDBbind general set

1

Ligand Parameterization Protocol

- At the level of HF/6-31G*
- Gaussian 03

1. Optimize ligand structure quantum mechanically

2. Perform single point energy calculation to compute molecular dipole moments and density matrix

- At the level of MP2/6-31++g(2D,2P)
- Gaussian 03

- GDMA v2.2

3. Derive electrostatic parameters from the density matrix

4. Transfer or derive vdw, bond, angle, torsion, atomic polarizability or other parameters