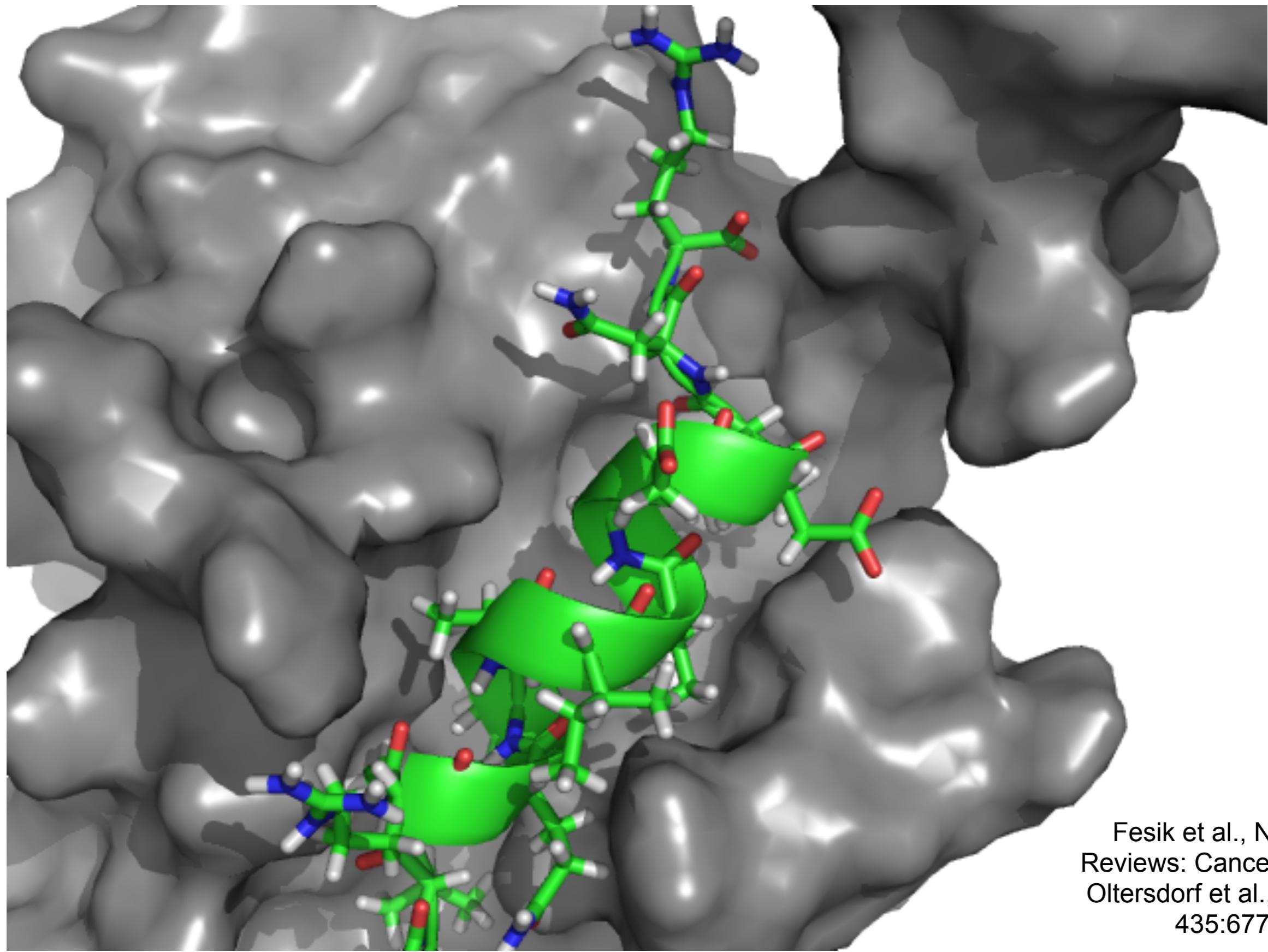
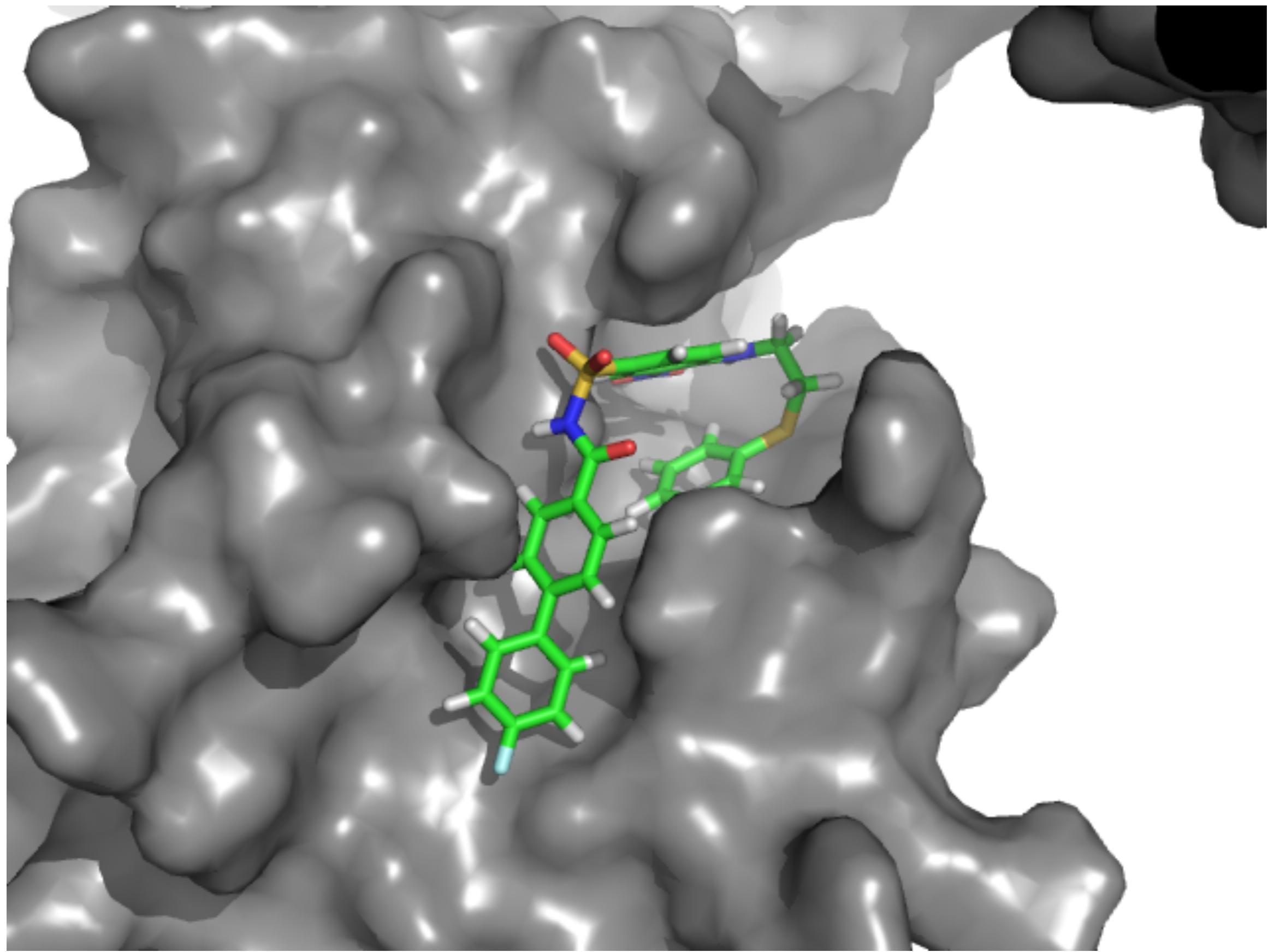


# Inhibitors can make good drugs



Fesik et al., Nature  
Reviews: Cancer, 5:876;  
Oltersdorf et al., Nature  
435:677

# Small molecules can mimic binding partners

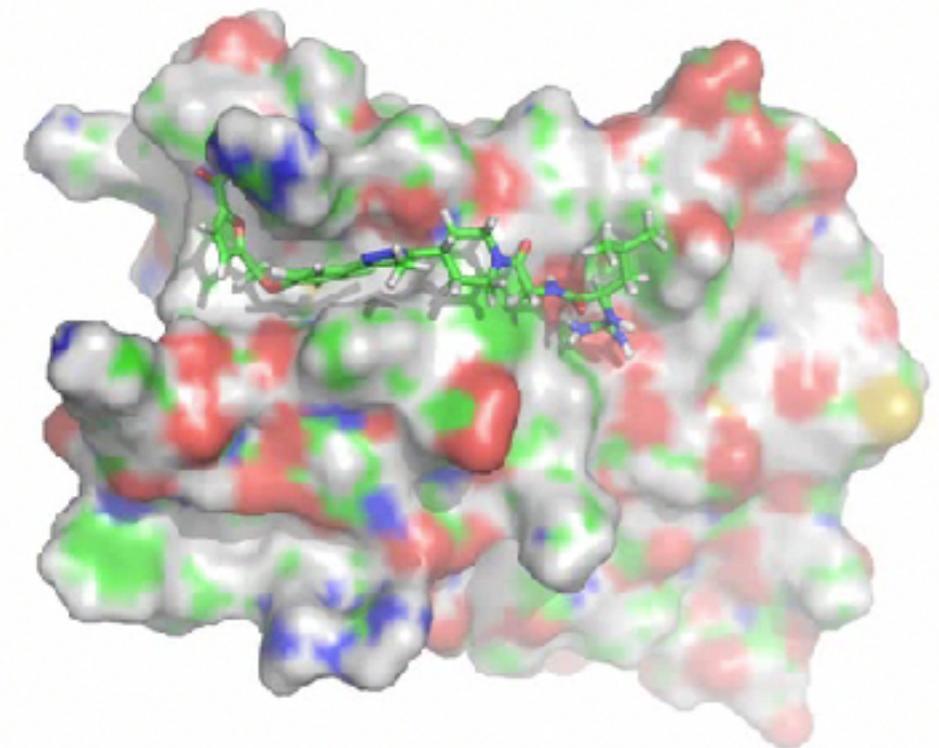


Pose prediction and scoring are probably  
not really separable

$$\Delta G = -k_B T \ln Q_{PL}/Q_P Q_L$$

Pose prediction and scoring are probably  
not really separable

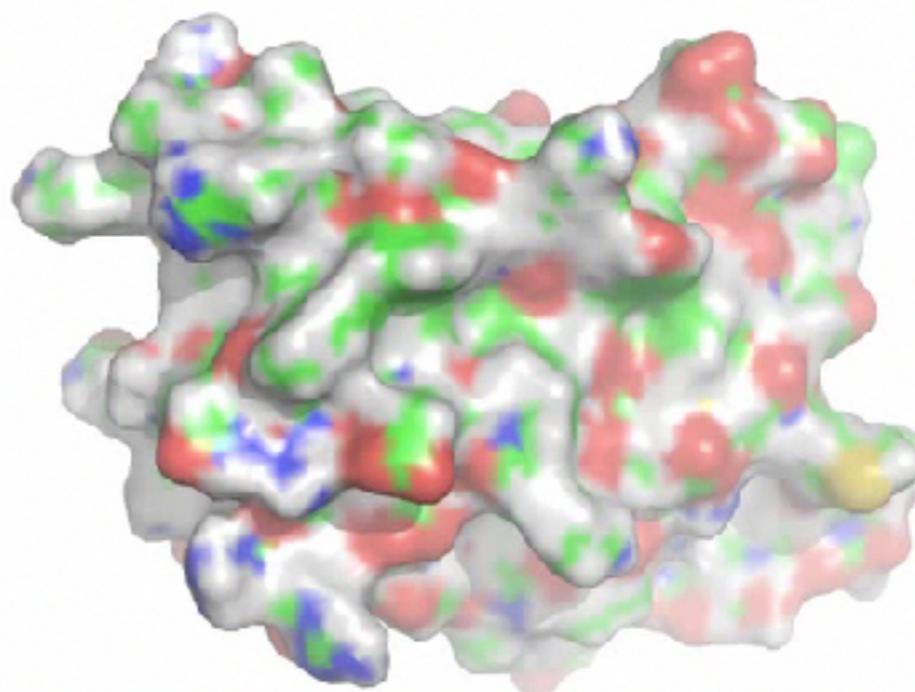
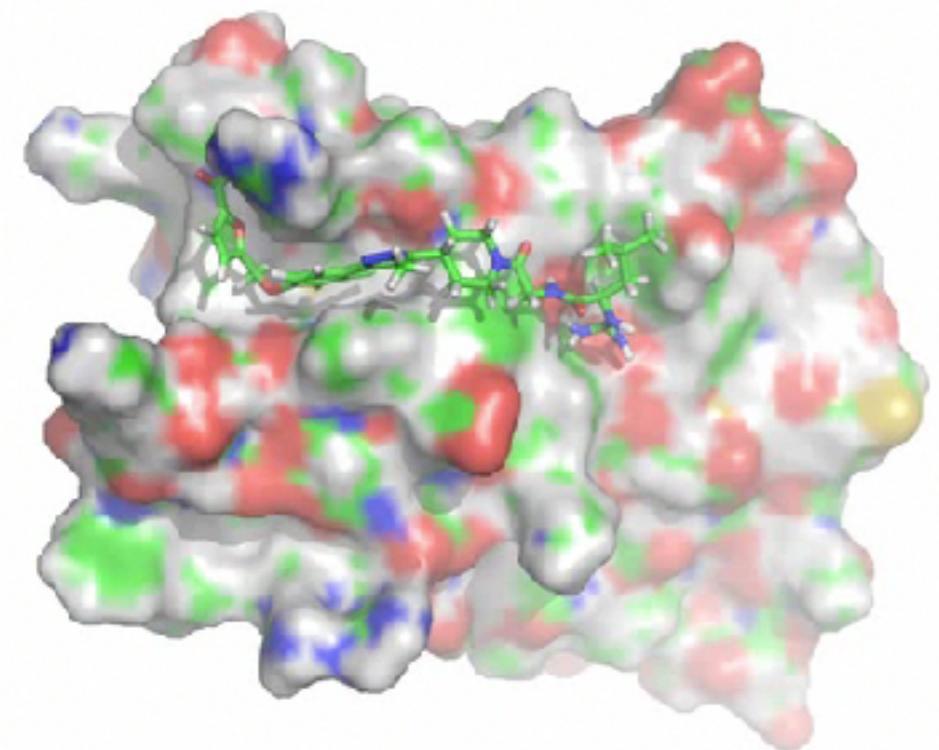
$$\Delta G = -k_B T \ln Q_{PL}/Q_P Q_L$$



$Q_{PL}$

Pose prediction and scoring are probably  
not really separable

$$\Delta G = -k_B T \ln Q_{PL}/Q_P Q_L$$

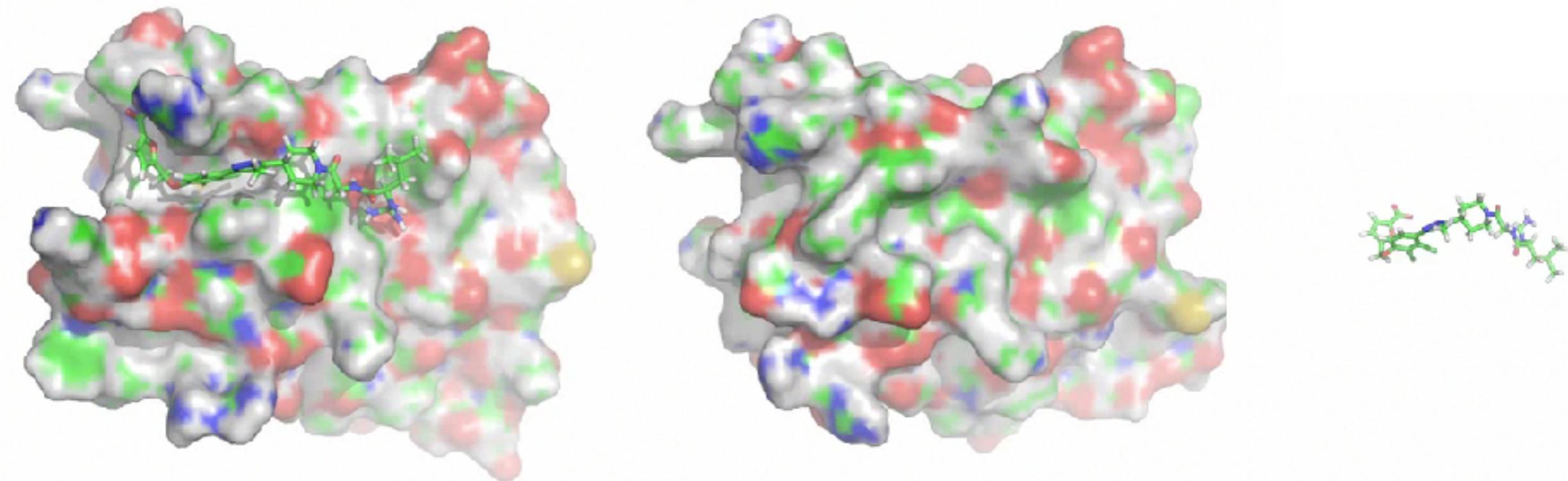


$Q_{PL}$

$Q_P$

# Pose prediction and scoring are probably not really separable

$$\Delta G = -k_B T \ln Q_{PL}/Q_P Q_L$$

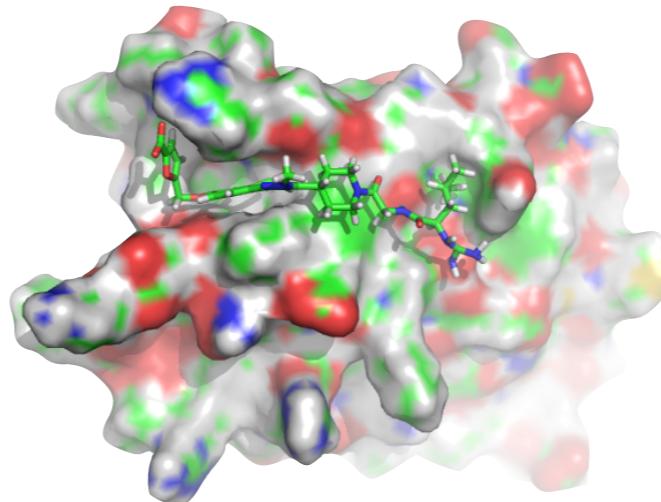


$Q_{PL}$

$Q_P$

$Q_L$

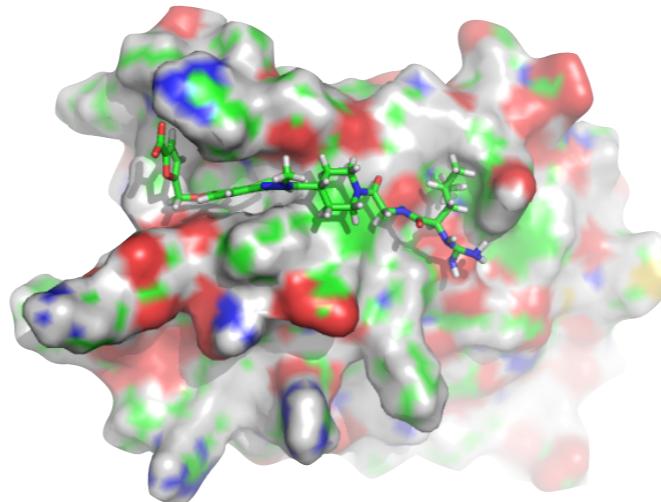
# Treating binding mode and score as separate problems requires single orientations



$$\Delta G = \Delta H - T\Delta S$$

- Score  $\sim \Delta H$
- Sometimes add  $-\Delta G_{solv}$  to score
- Uses “optimal” single orientations

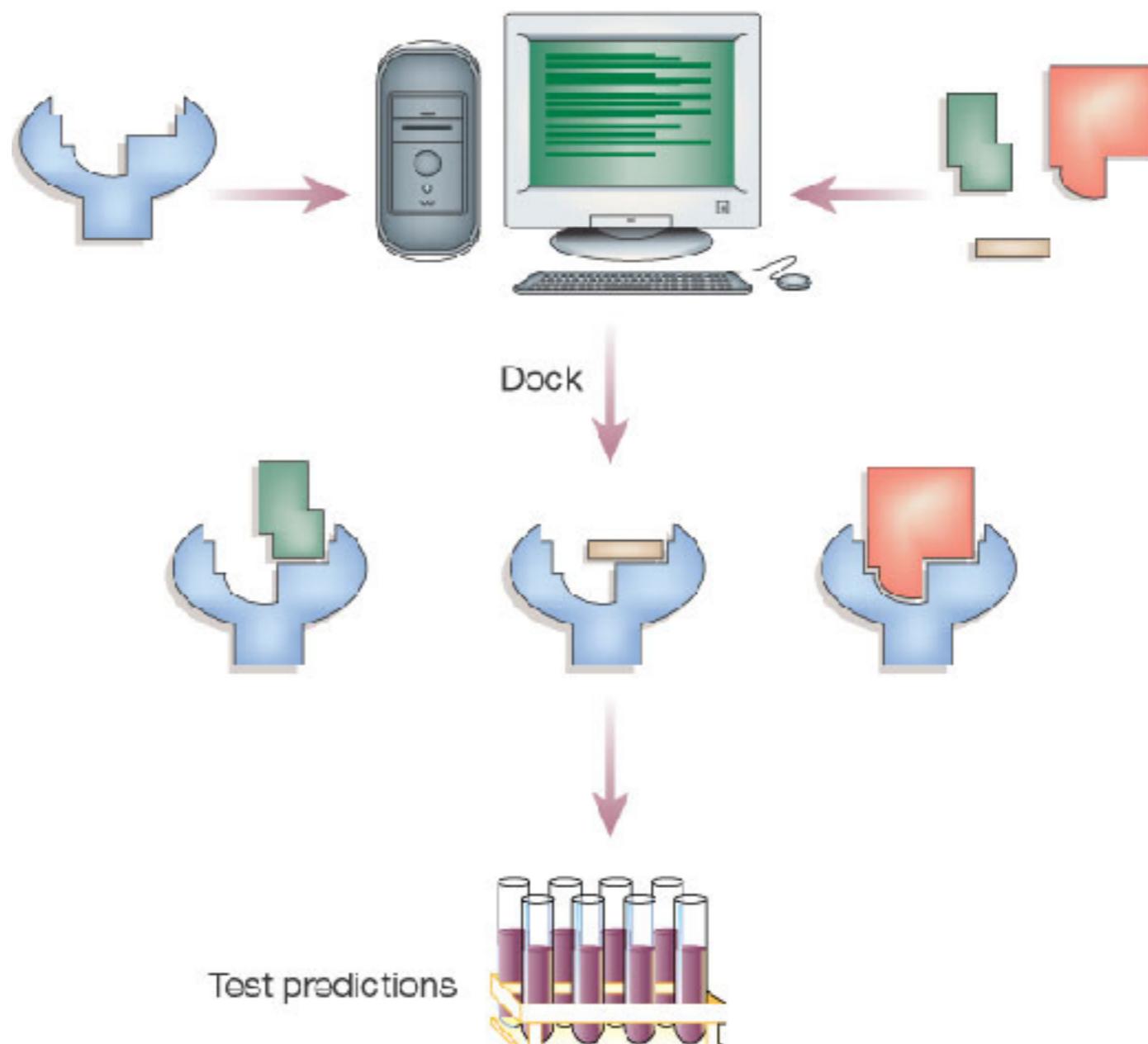
# Treating binding mode and score as separate problems requires single orientations



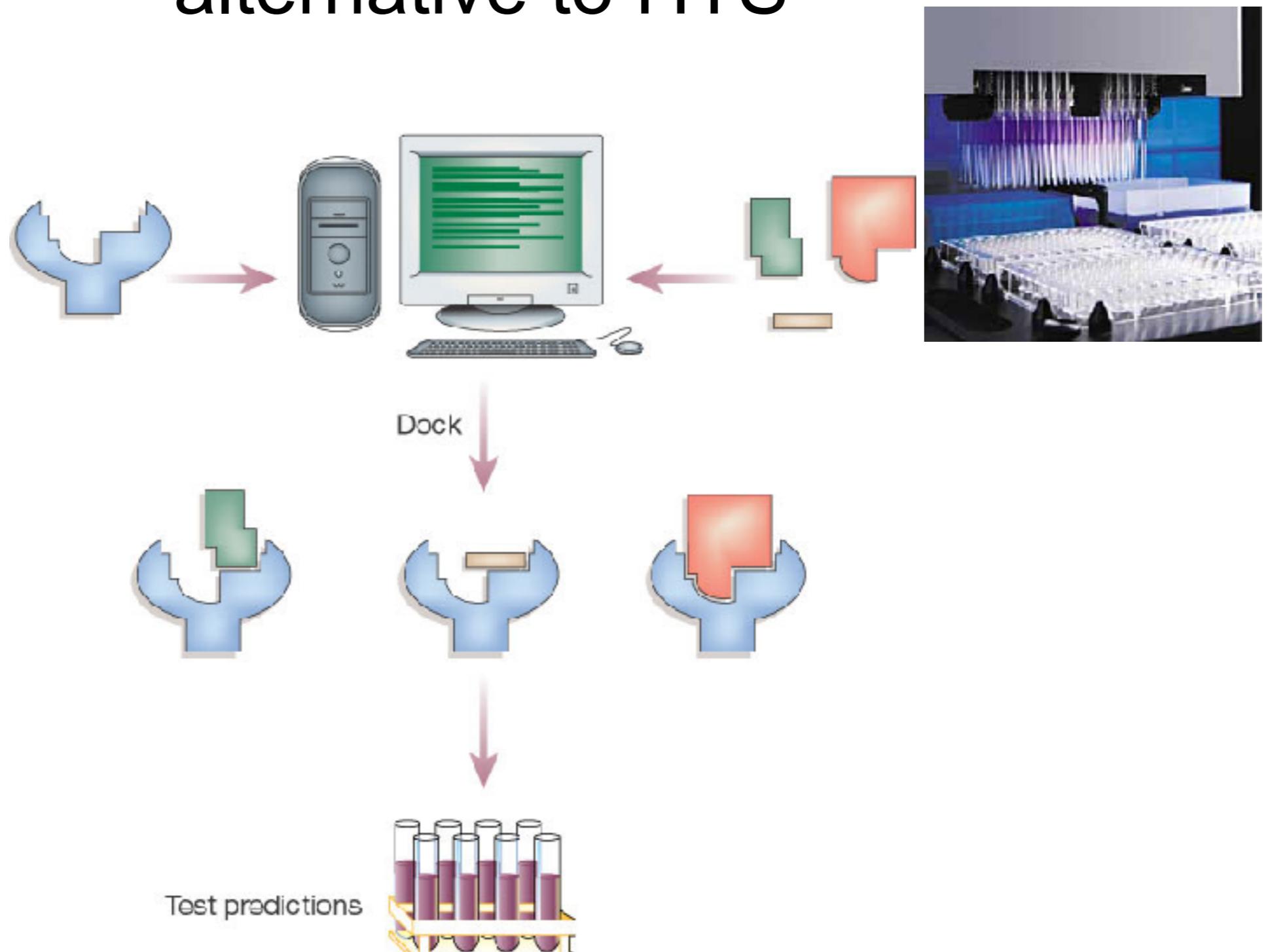
$$\Delta G = \Delta H - \cancel{T\Delta S}$$

- Score  $\sim \Delta H$
- Sometimes add  $-\Delta G_{solv}$  to score
- Uses “optimal” single orientations

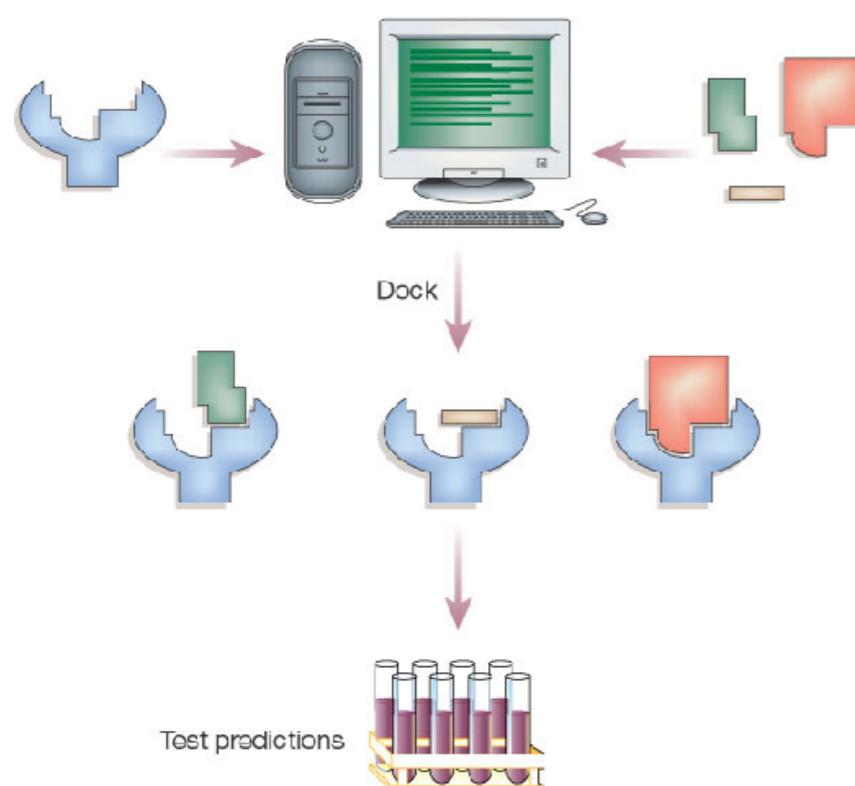
# Virtual screening provides the computational alternative to HTS



# Virtual screening provides the computational alternative to HTS

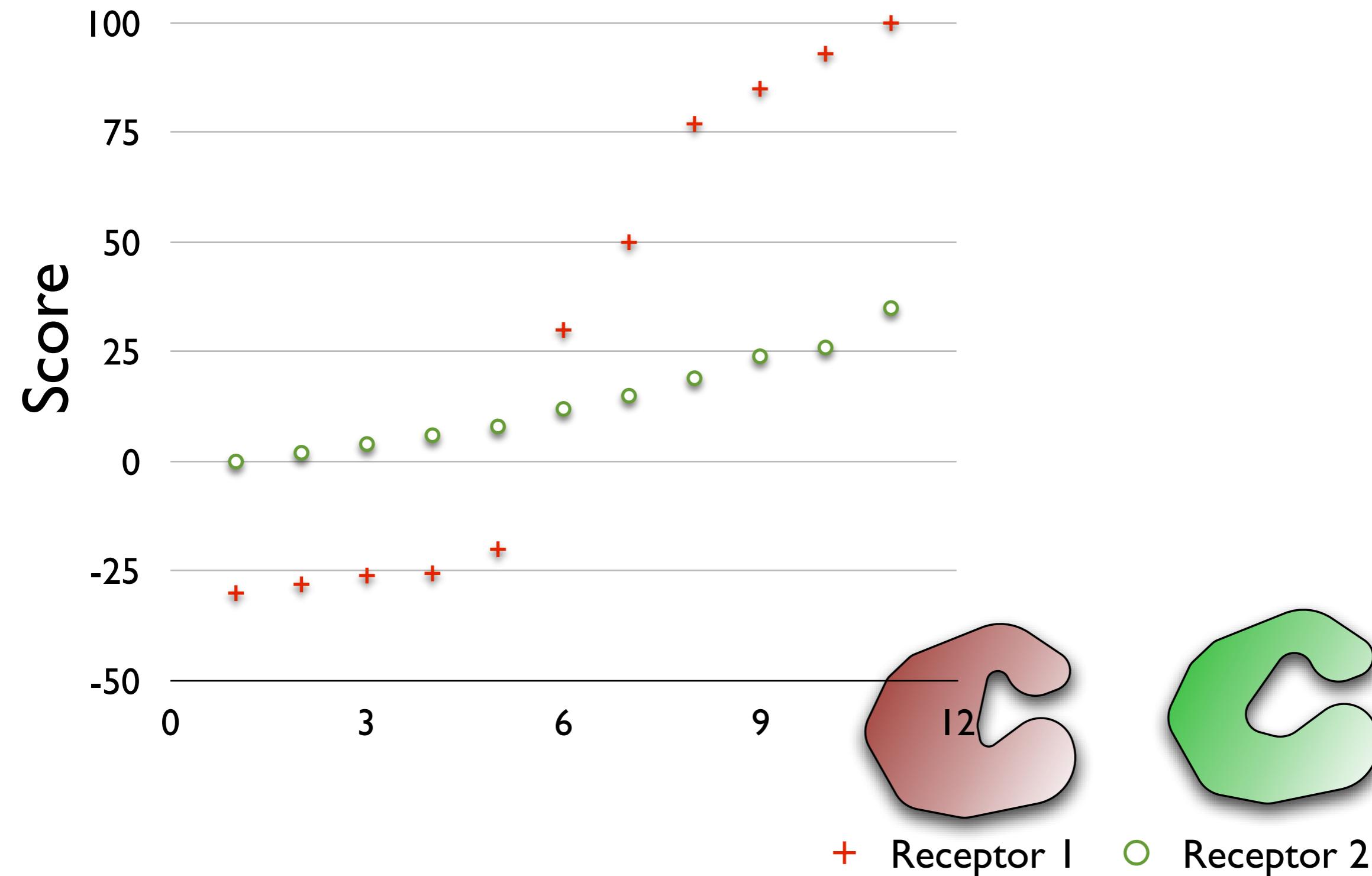


# Virtual screening has different strengths and weaknesses

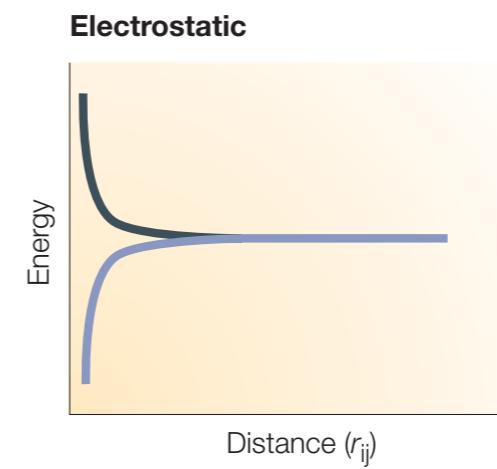
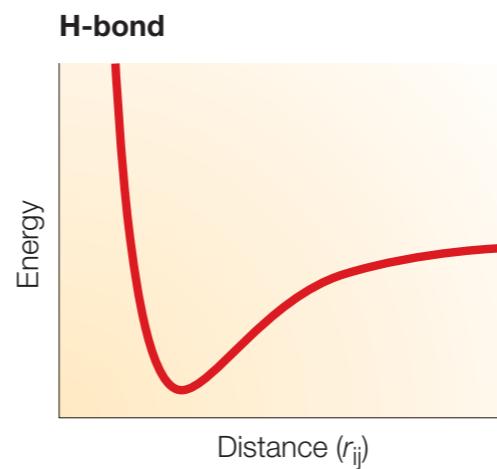
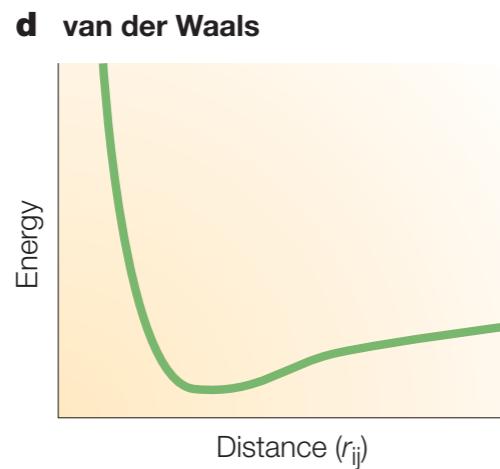


- False negatives and false positives
- Requires experimental confirmation
- Strengths:
  - New chemistry, new compounds!
  - Different hits from HTS
  - Fast
  - Higher hit rates than HTS?

# Scores are somewhat arbitrary

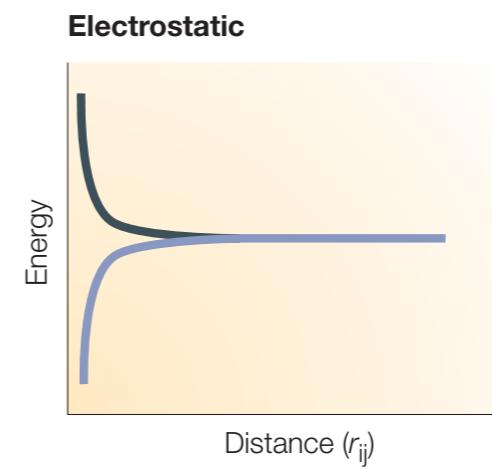
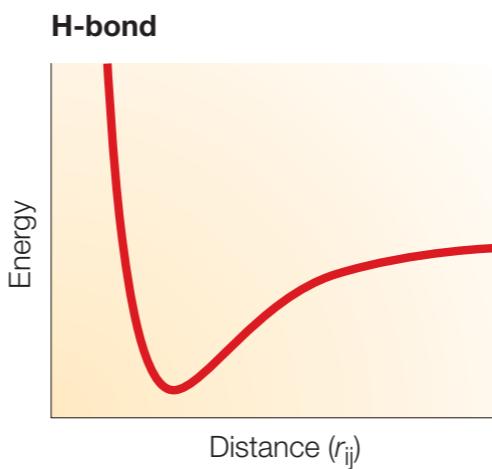
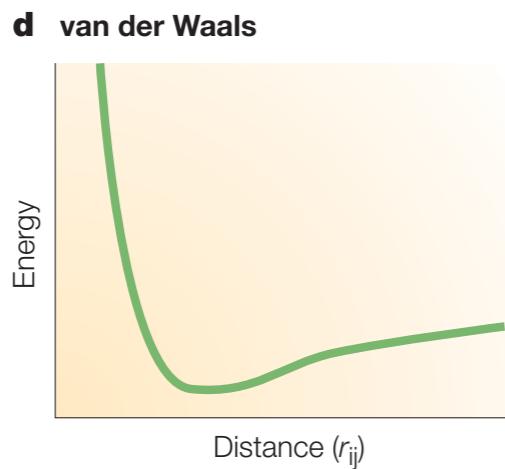


# There are several major classes of scoring functions which see common use



- Force field-based
- Knowledge-based
- Empirical
- Consensus scoring

# There are several major classes of scoring functions which see common use



## Force-field-based

- D-Score<sup>26</sup>
- G-Score<sup>26</sup>
- GOLD<sup>47</sup>
- AutoDock<sup>45</sup>
- DOCK<sup>24</sup>

## Empirical

- LUDI<sup>92,93</sup>
- F-Score<sup>50</sup>
- ChemScore<sup>49</sup>
- SCORE<sup>131,132</sup>
- Fresno<sup>51</sup>
- X-SCORE<sup>60</sup>

## Knowledge-based

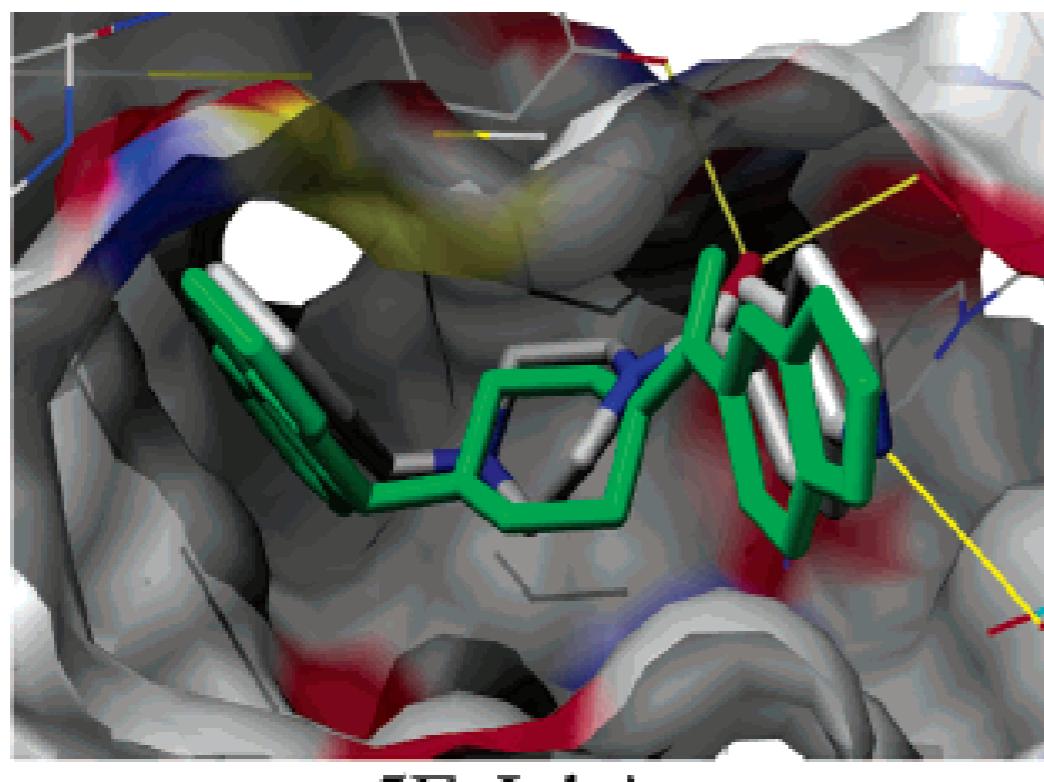
- PMF<sup>54–56</sup>
- DrugScore<sup>57</sup>
- SMoG<sup>58</sup>

- Force field-based
- Knowledge-based
- Empirical
- Consensus scoring

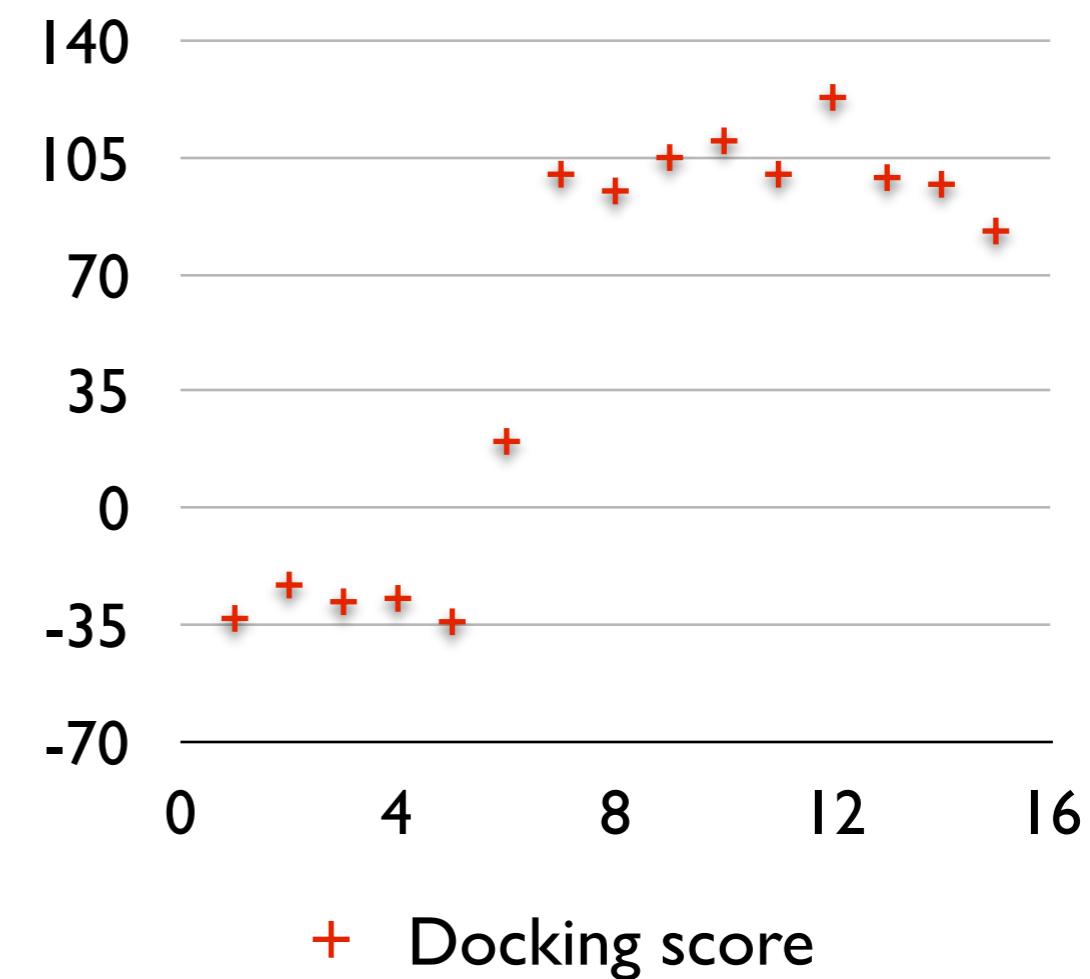
# Docking works well for filtering

Docking works well for filtering out compounds that simply cannot fit in a binding site

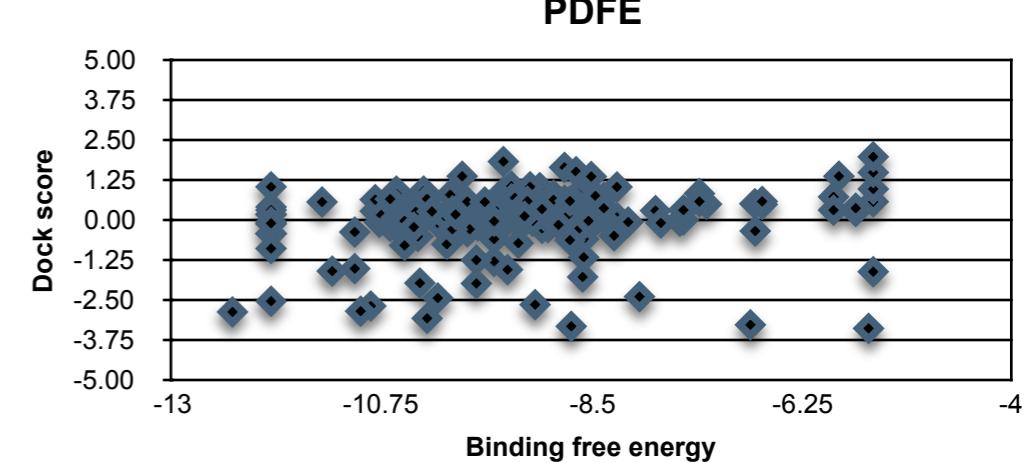
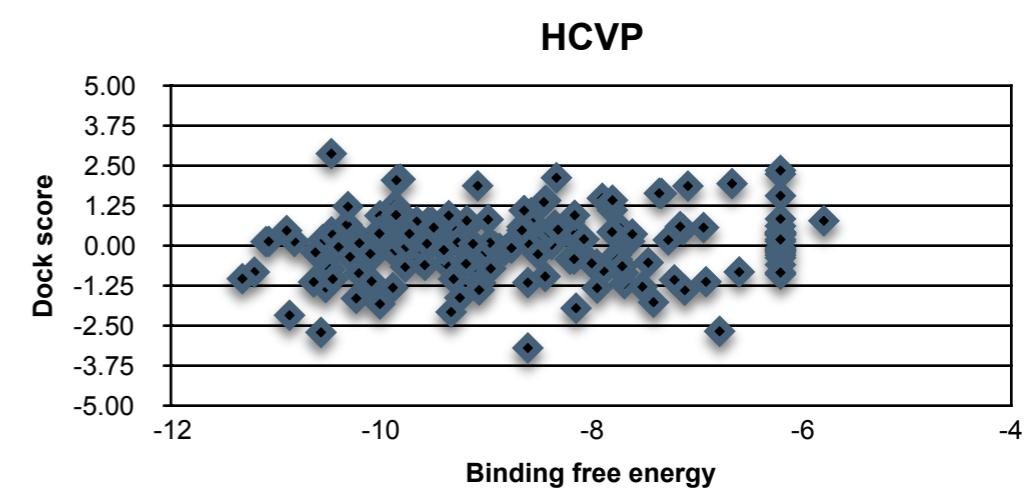
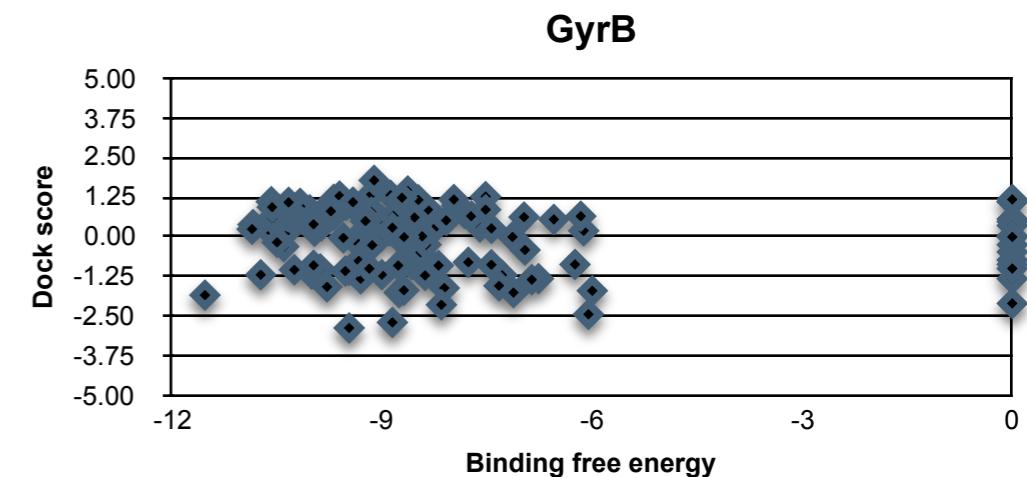
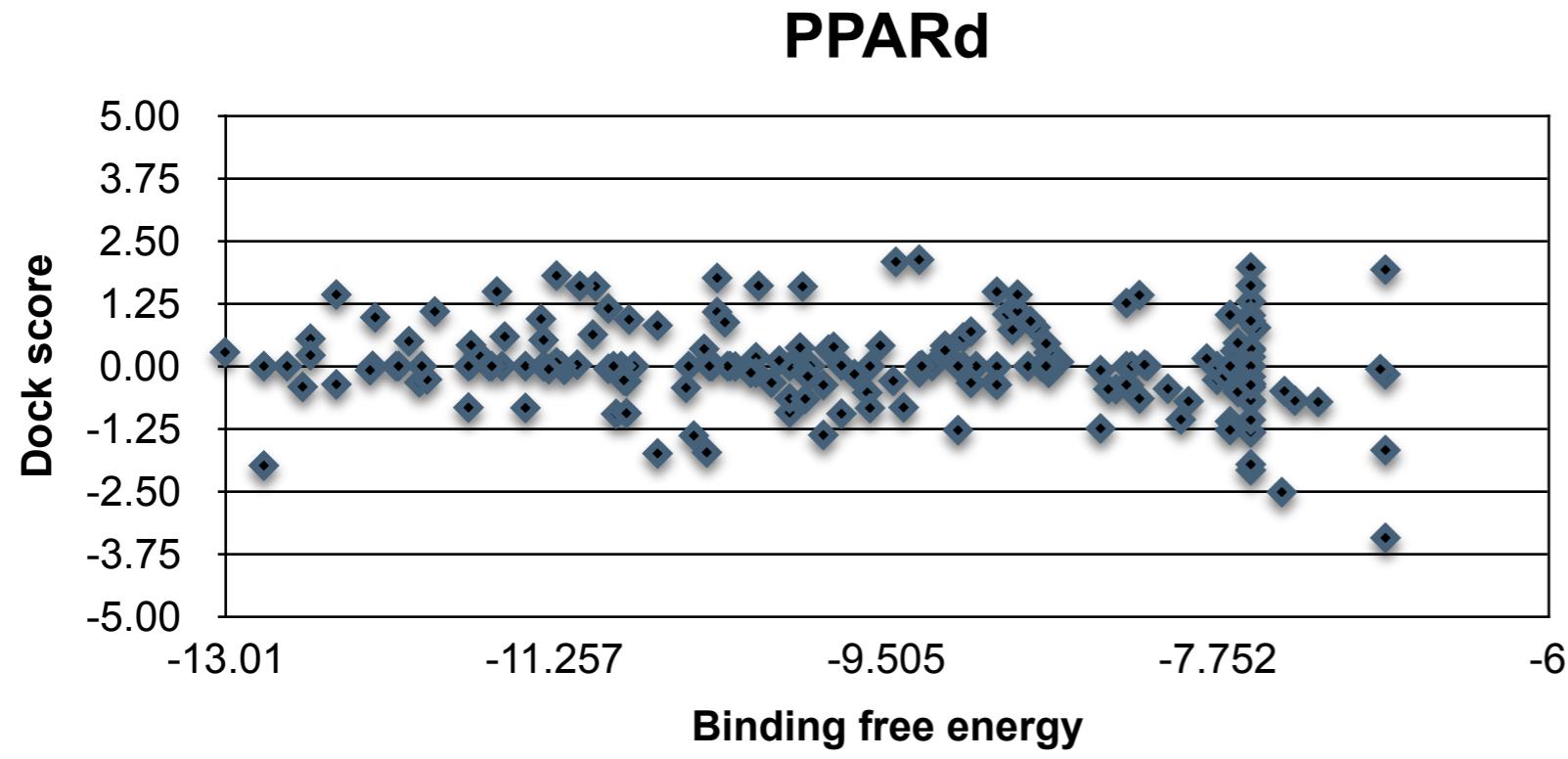
Huang et al. *J. Med. Chem.* 49:6789 (2006)



5F. InhA

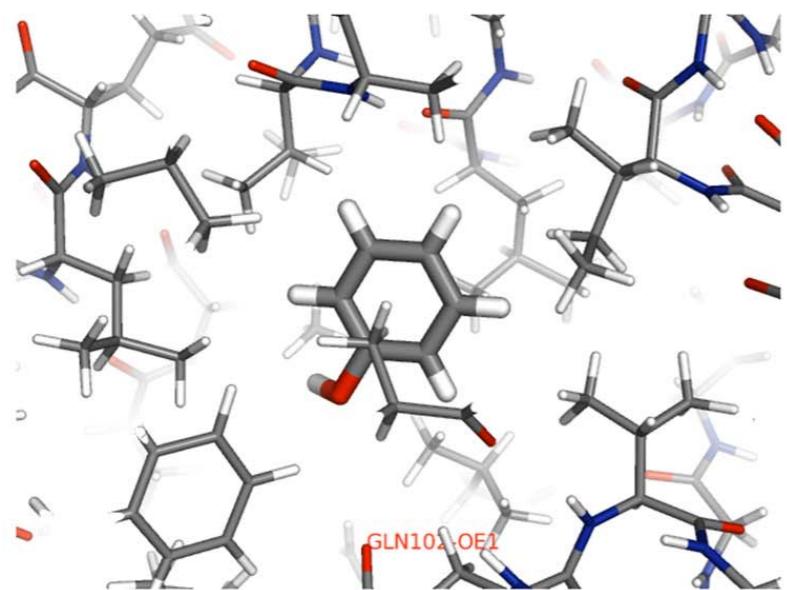


# Docking performs poorly at binding strength prediction

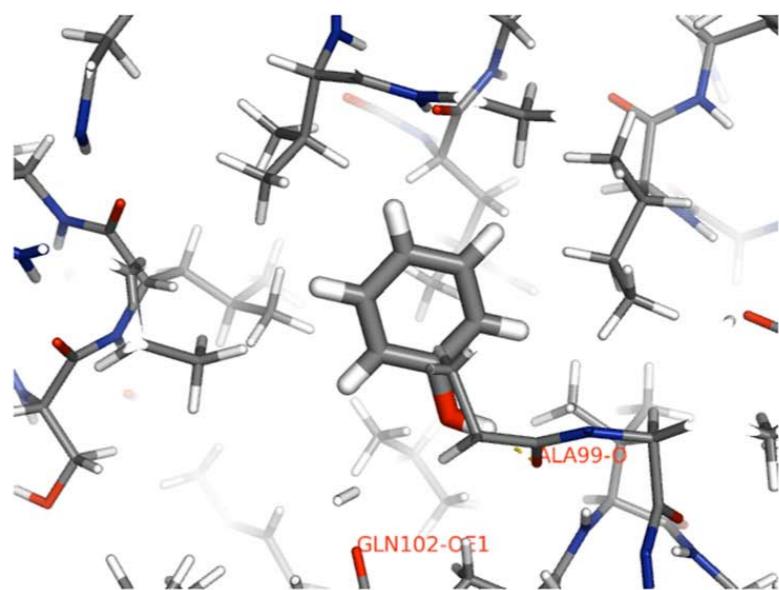


- Docking can't reliably calculate binding free energies, or even relative binding strengths

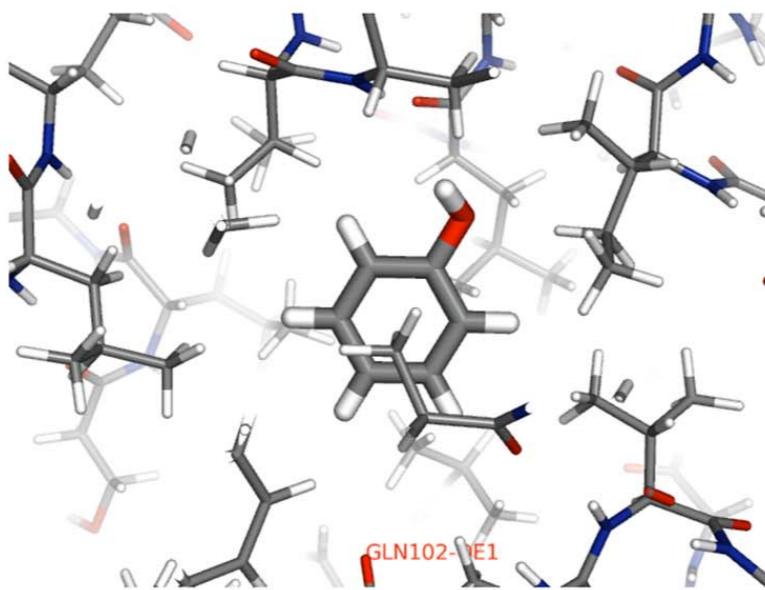
# Binding mode prediction is one part of the scoring problem



(a) docking orientation 1

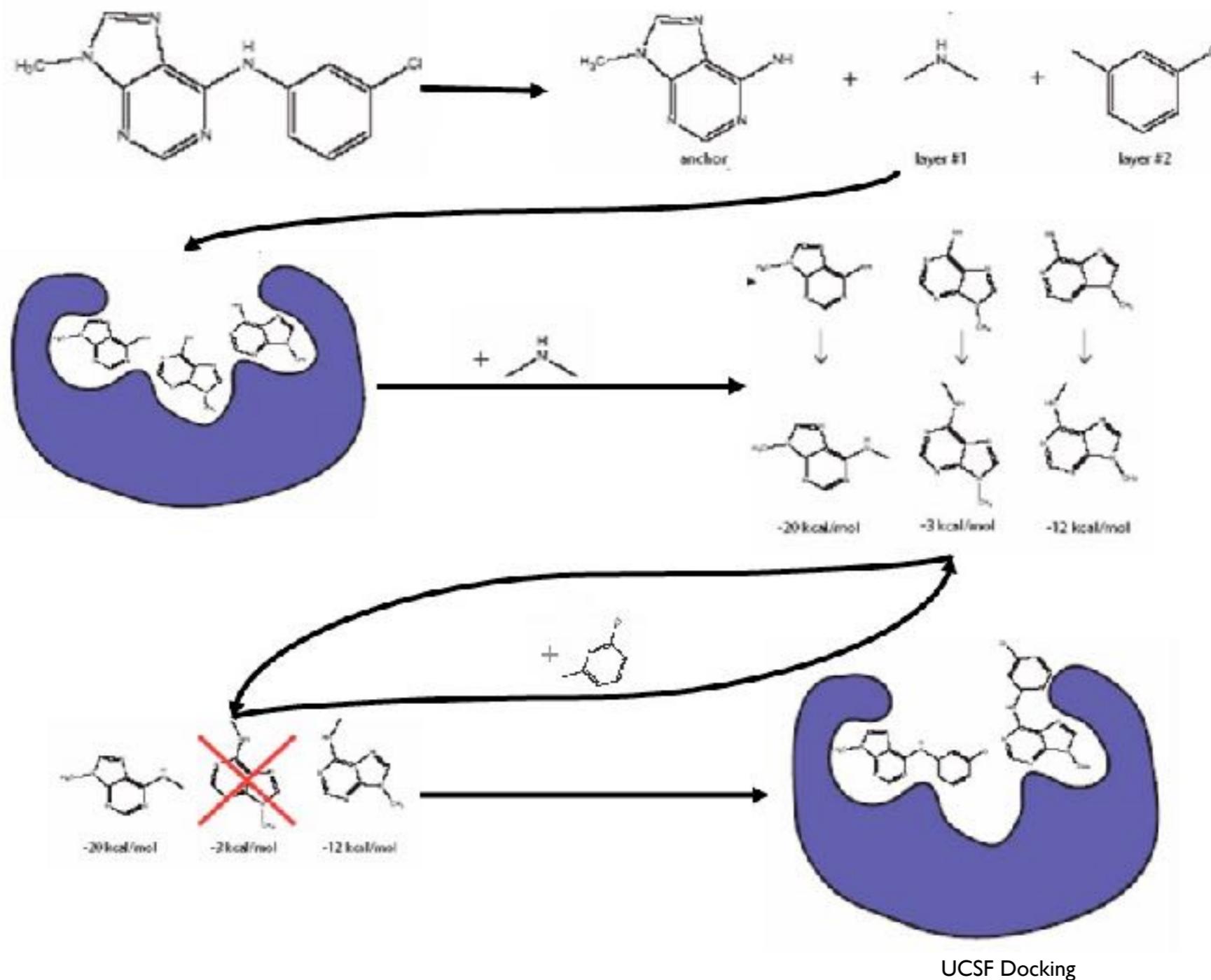


(b) docking orientation 2

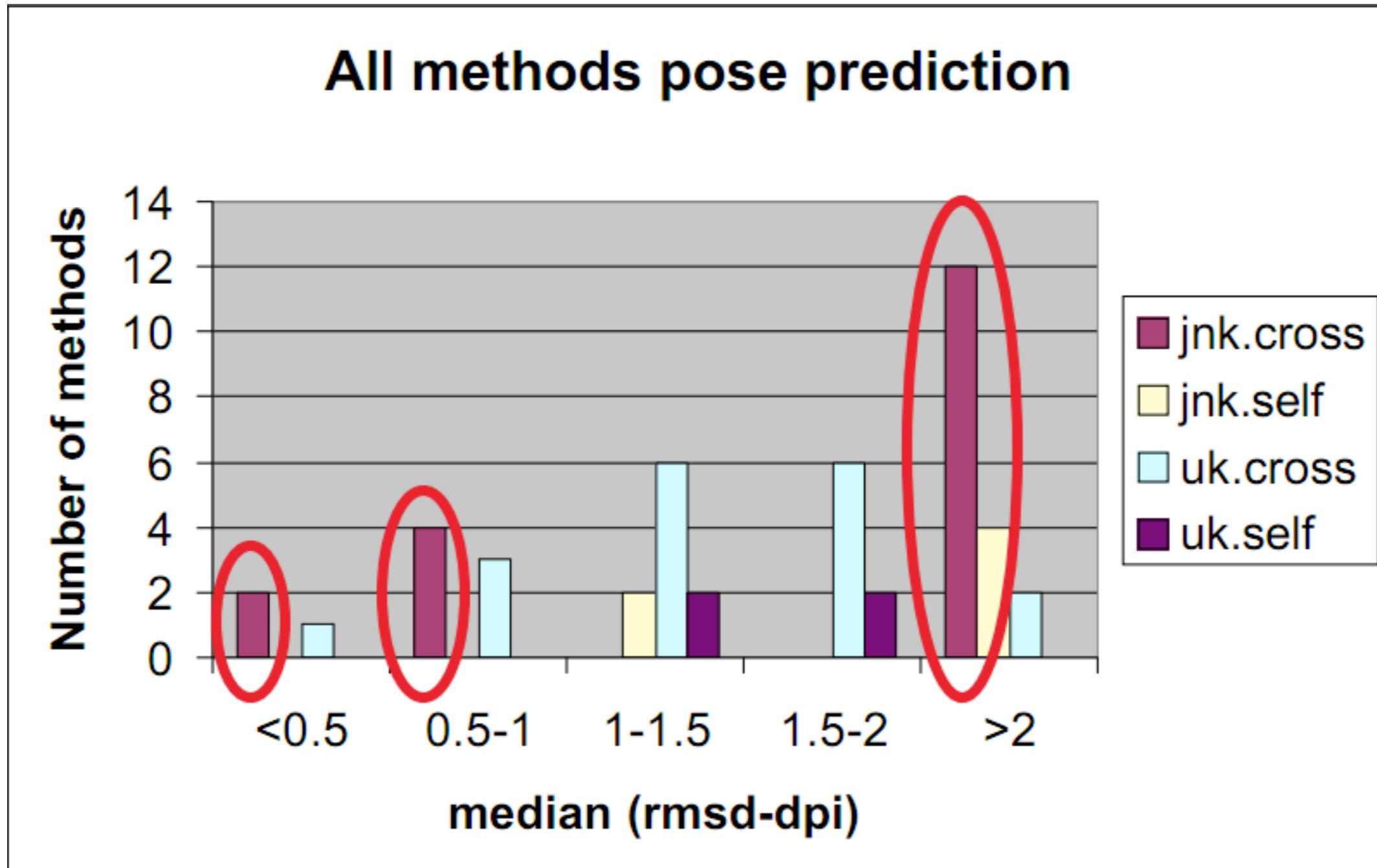


(c) docking orientation 3

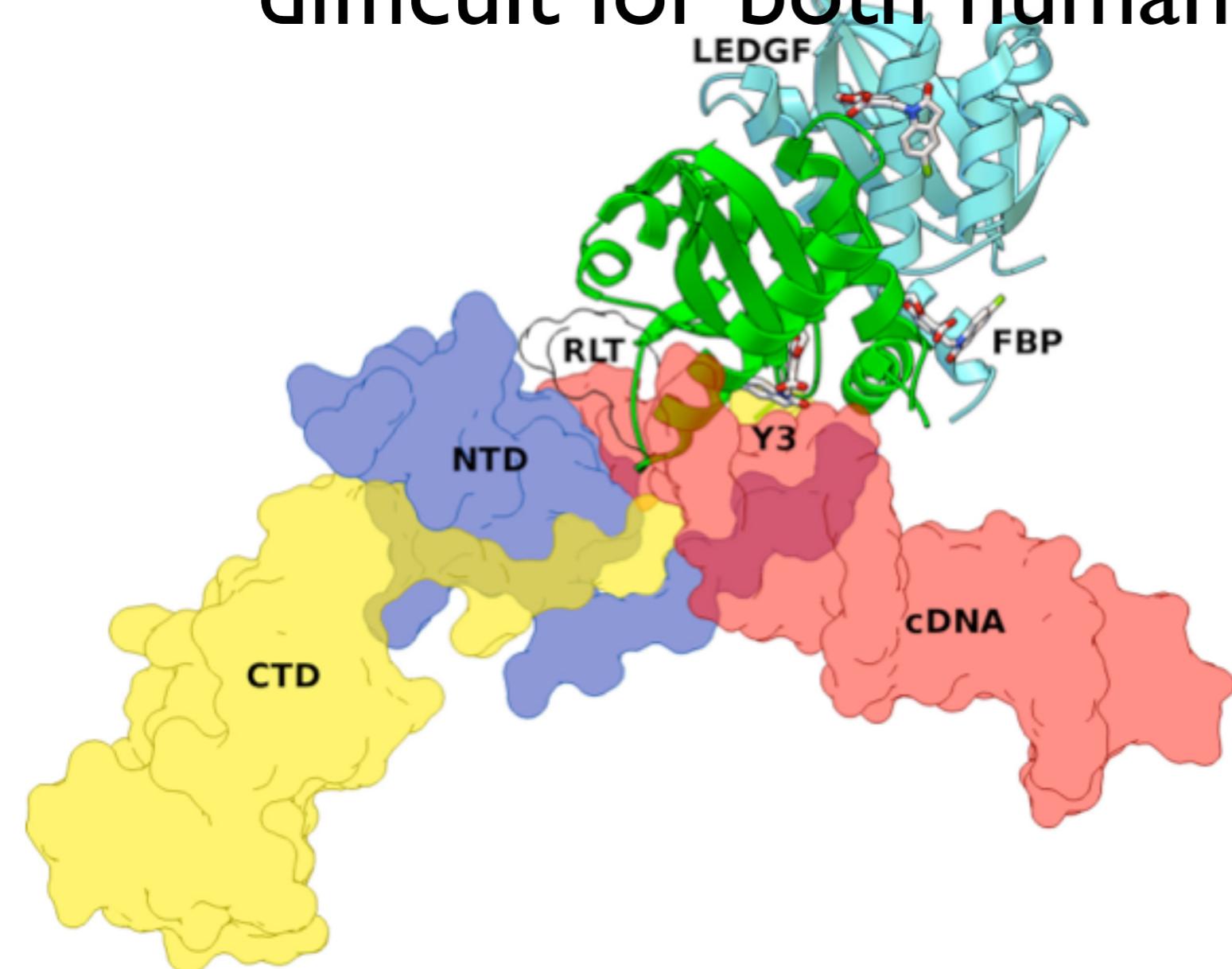
# Docking typically places flexible ligands into “rigid” binding sites



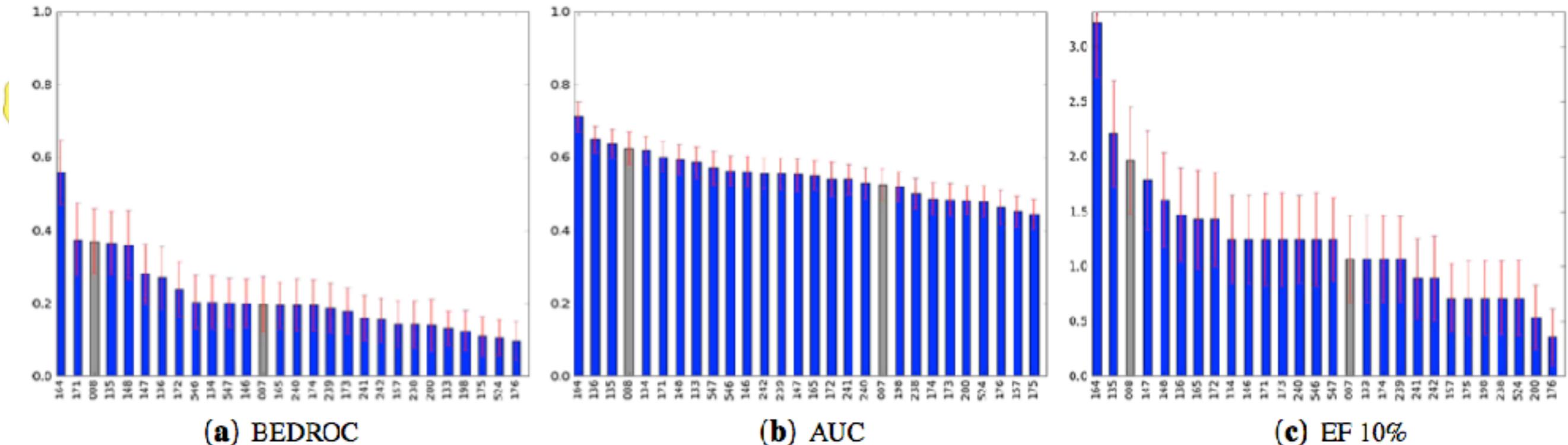
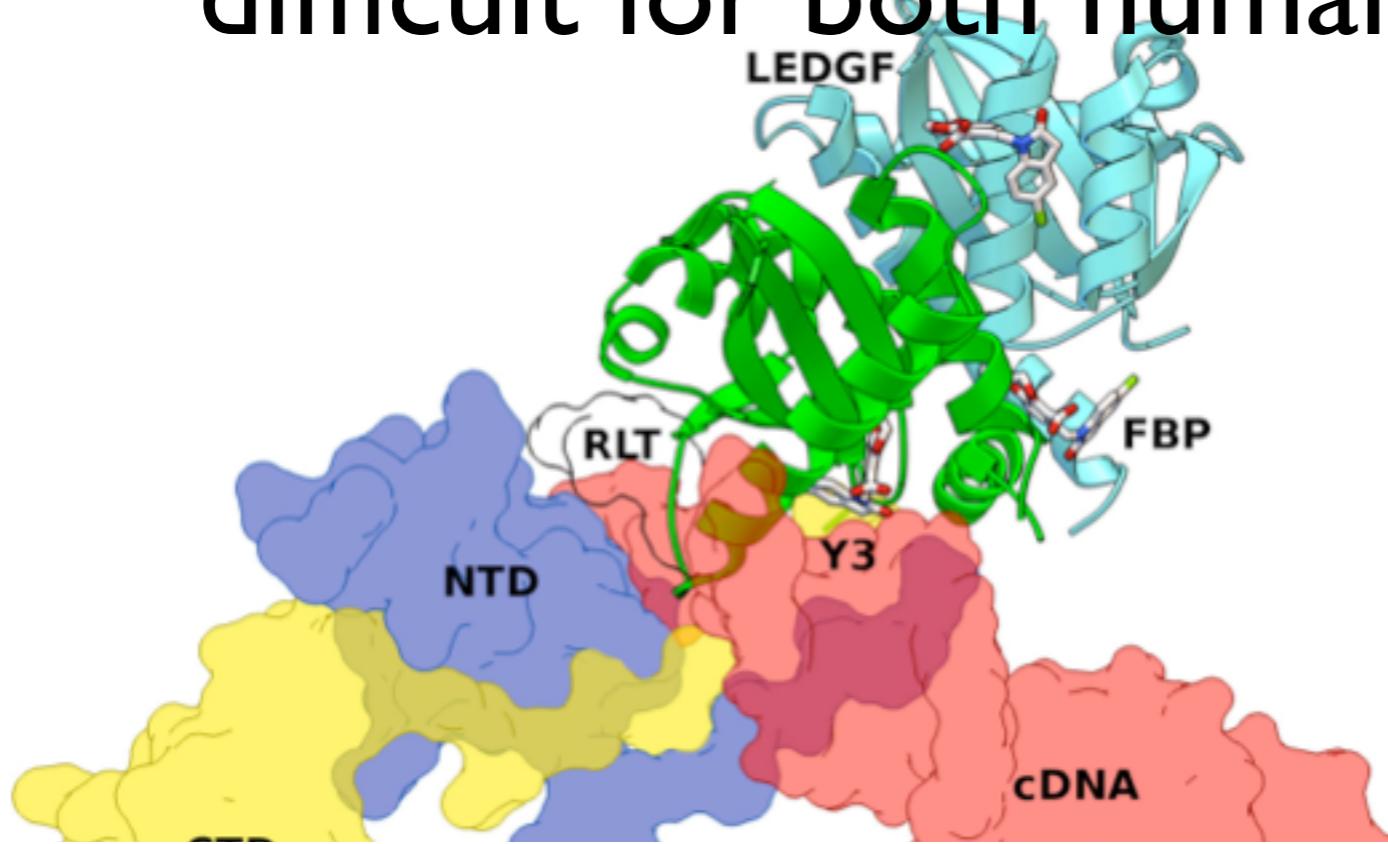
# Binding mode prediction is extremely difficult for both humans and computers



Binding mode prediction is extremely difficult for both humans and computers



# Binding mode prediction is extremely difficult for both humans and computers

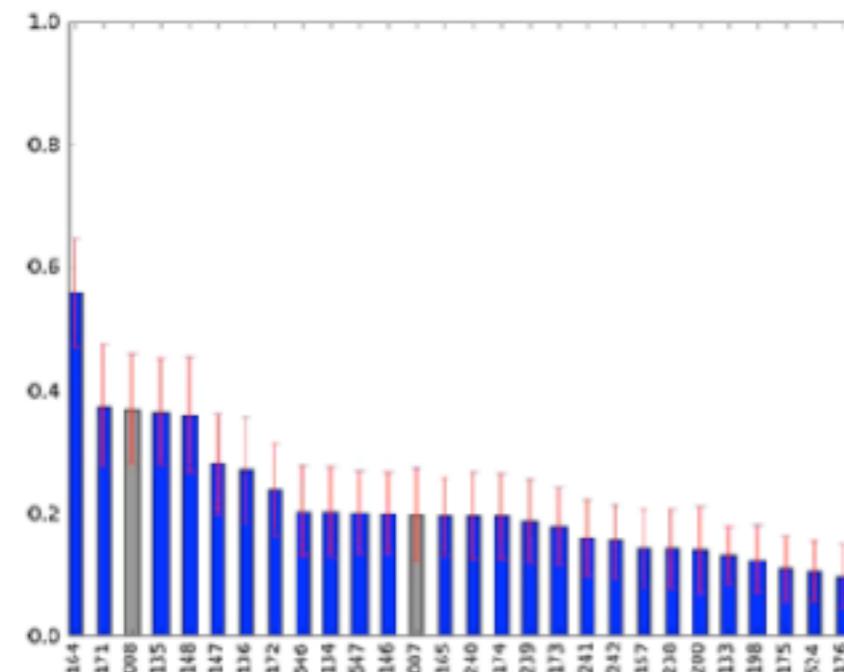
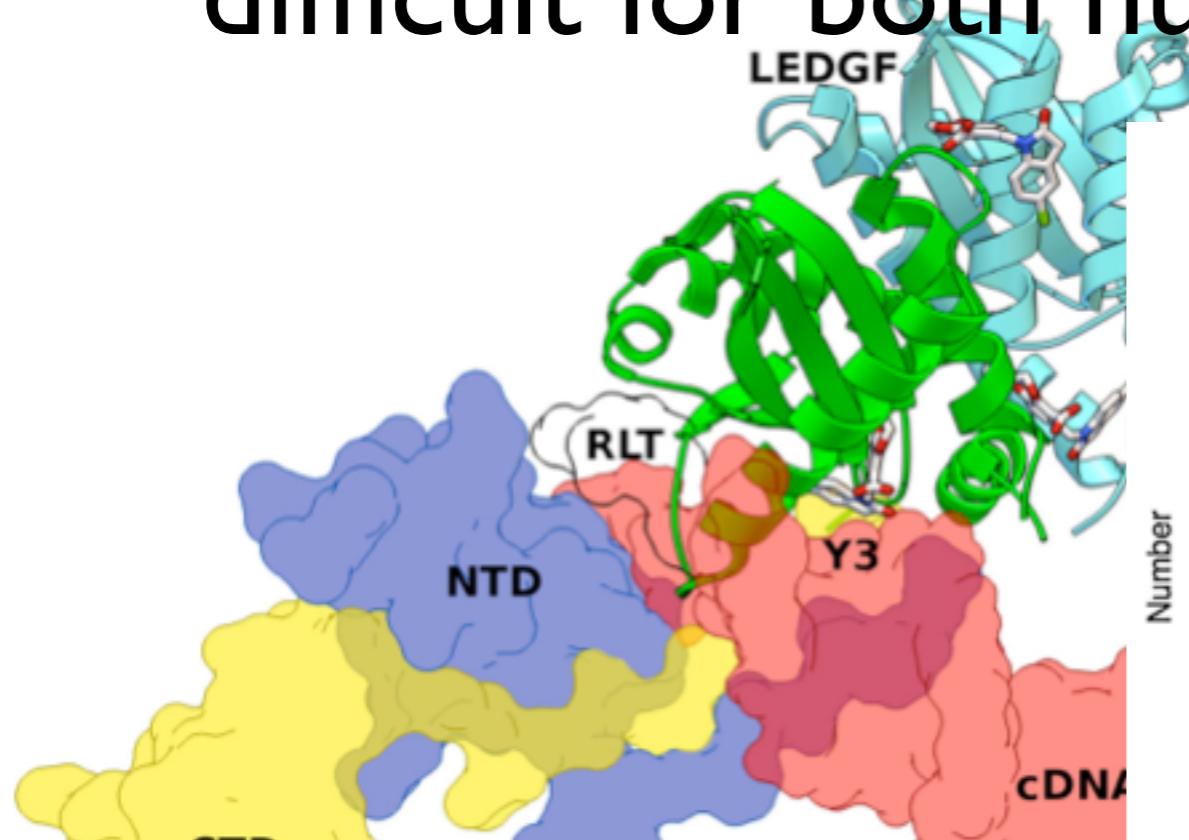


(a) BEDROC

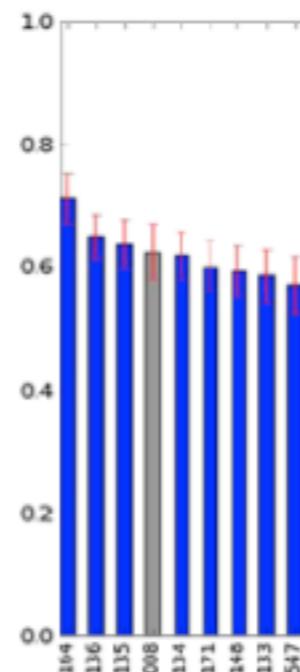
(b) AUC

(c) EF 10%

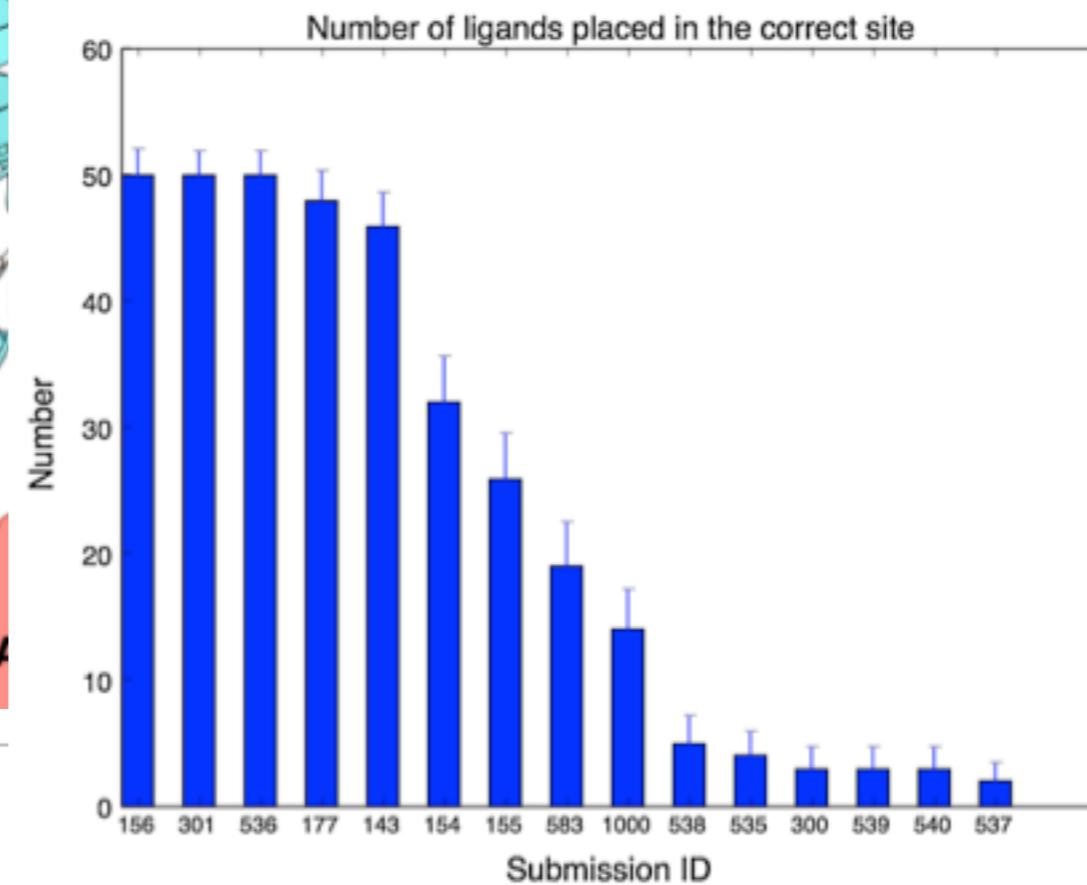
# Binding mode prediction is extremely difficult for both humans and computers



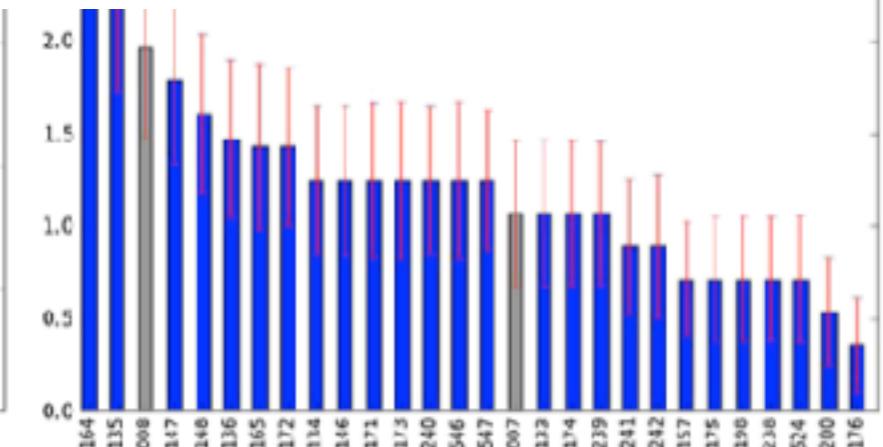
(a) BEDROC



(b) AUC

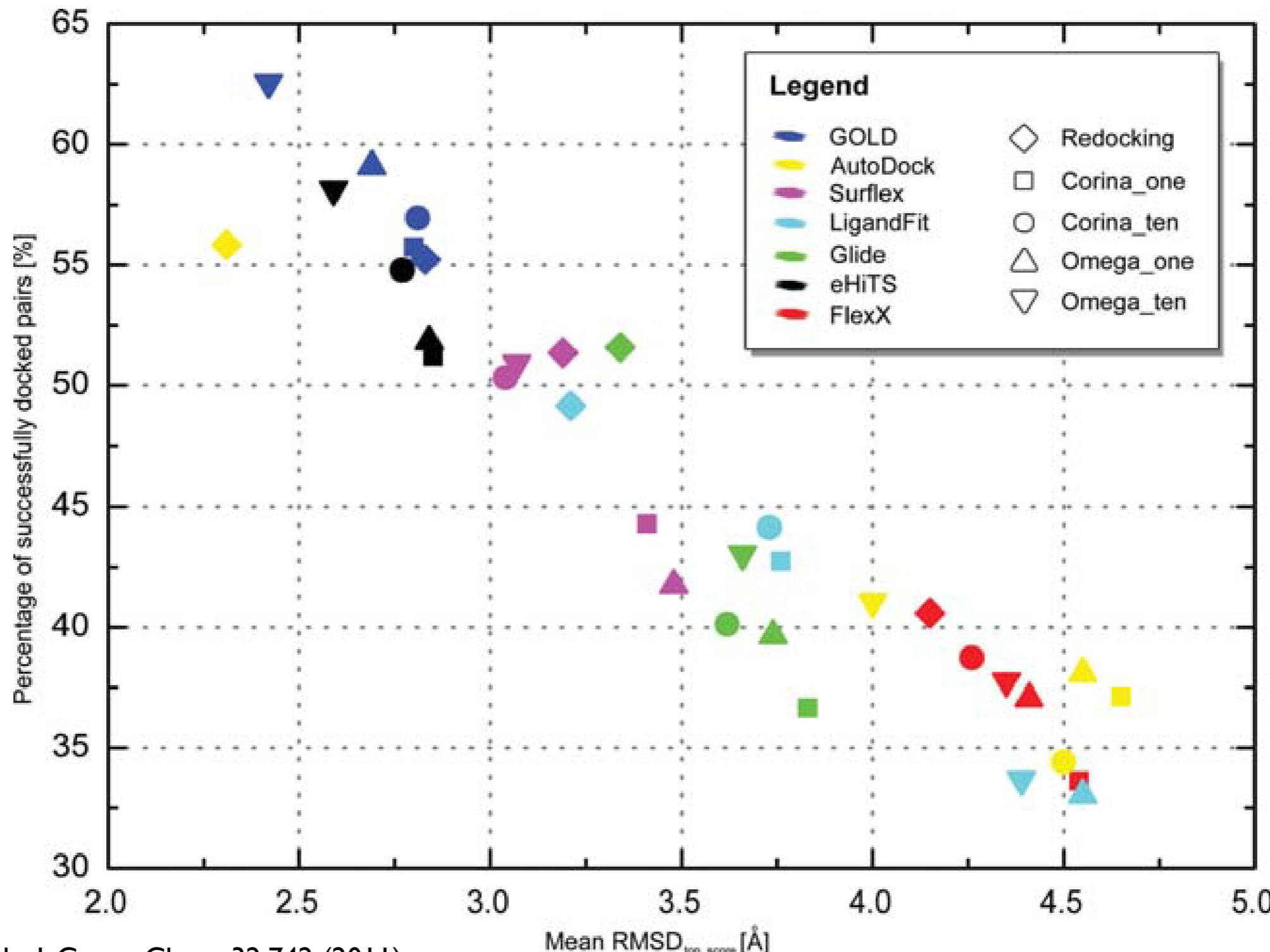


(e) Site identification

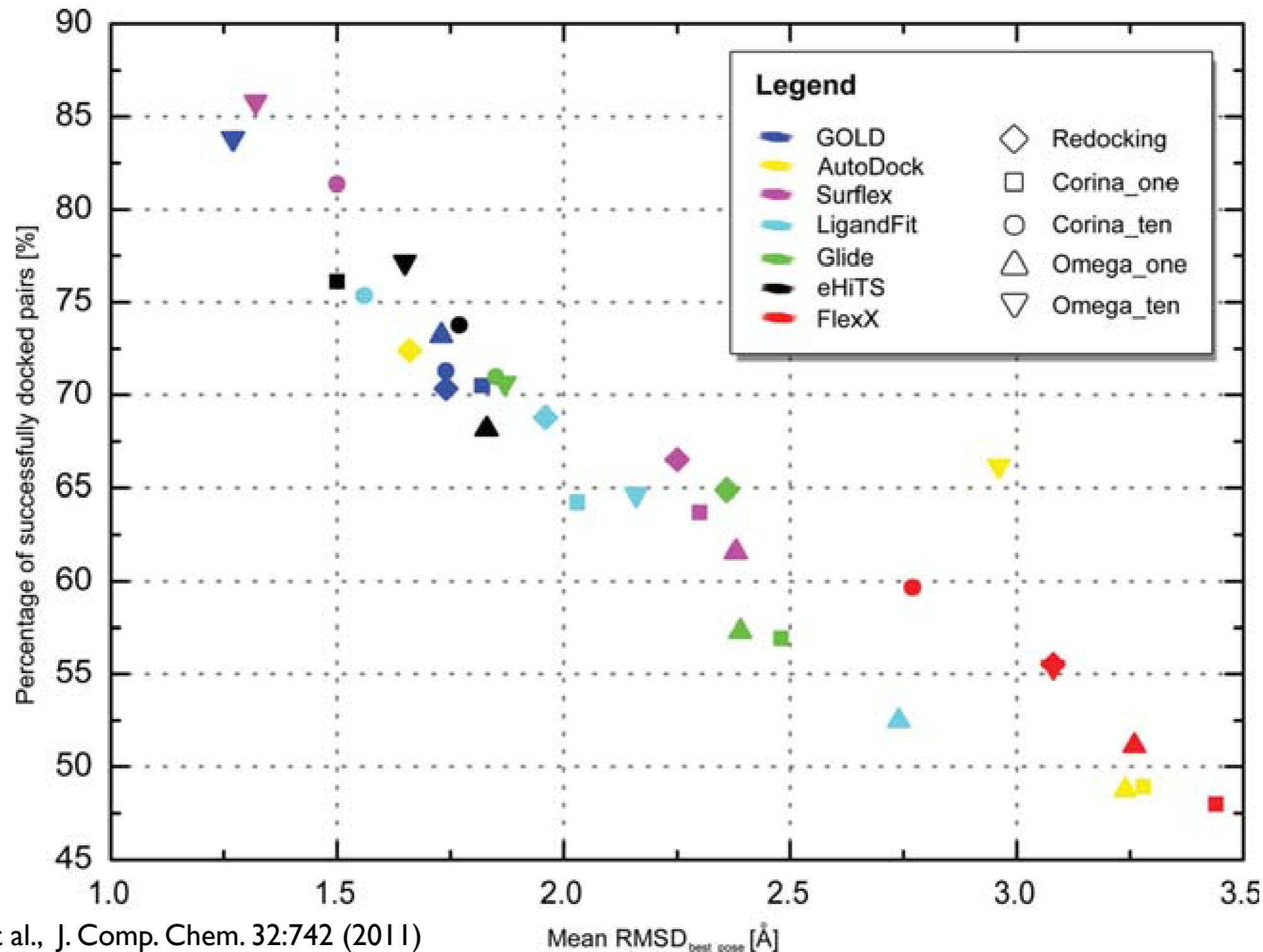


(c) EF 10%

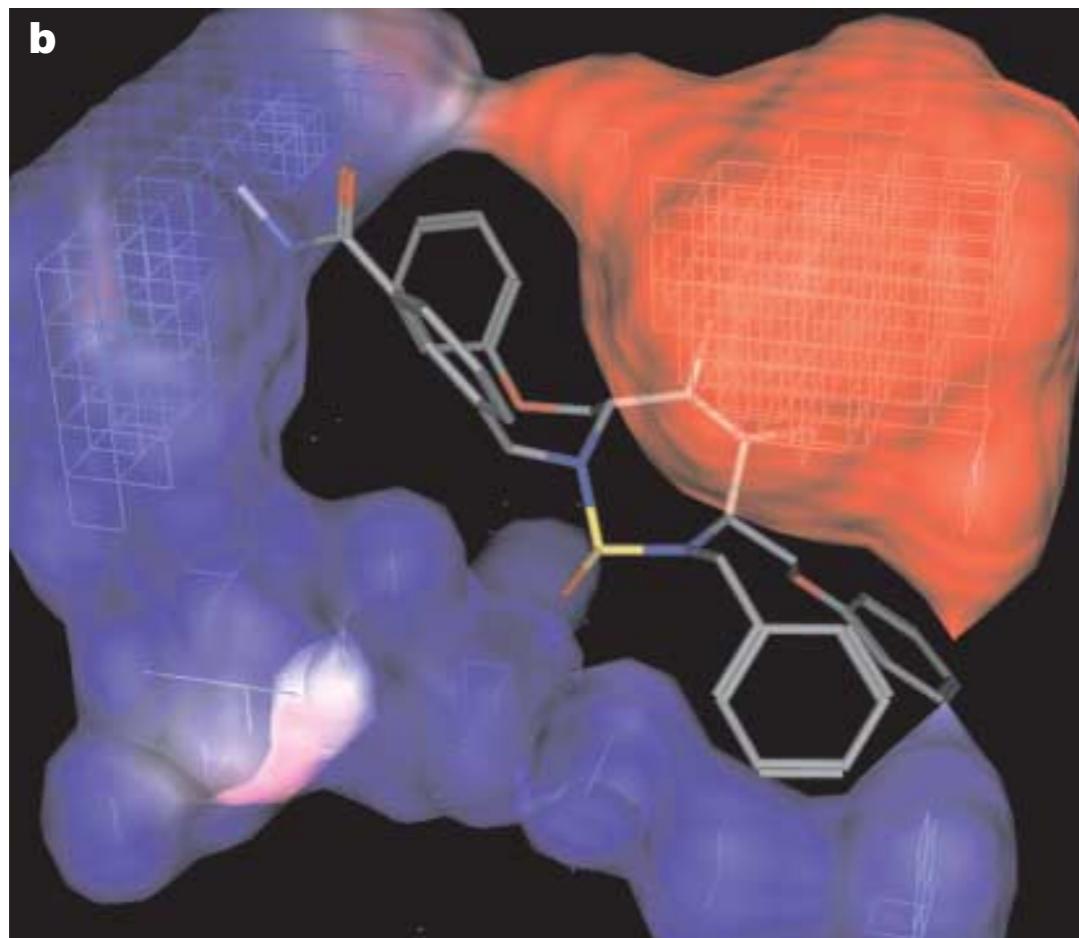
# Binding mode prediction is extremely difficult for both humans and computers



# This is partly a problem of “scoring”



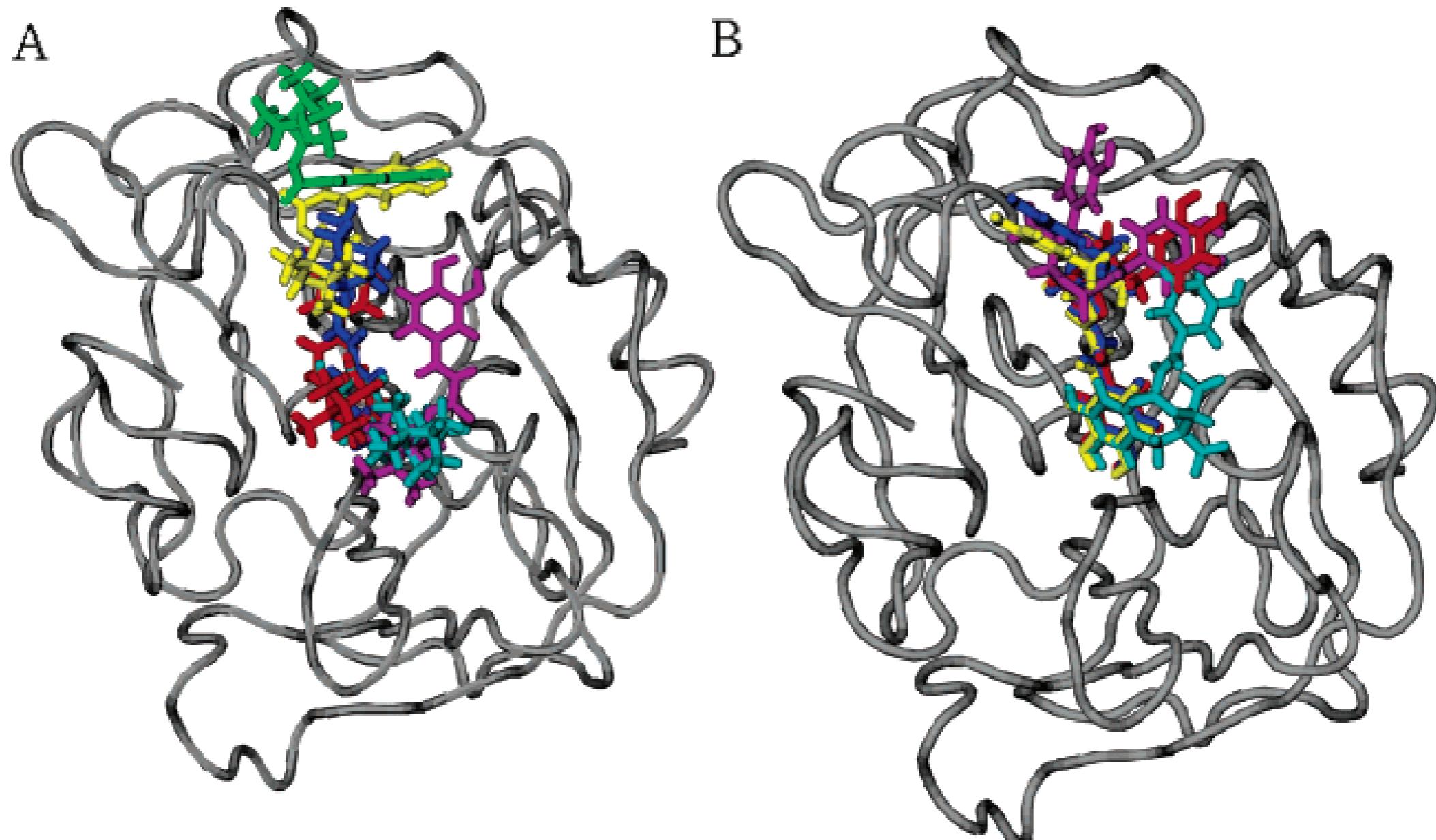
# Rescoring can be done using higher accuracy methods



Kitchen et al., NRDD 3:935 (2004)

- Initial docking done quickly:
  - Using grid
  - Using rapid scoring function
- After filtering, rescore top pose(s) with more detail:
  - No grid
  - More detailed energy calculation

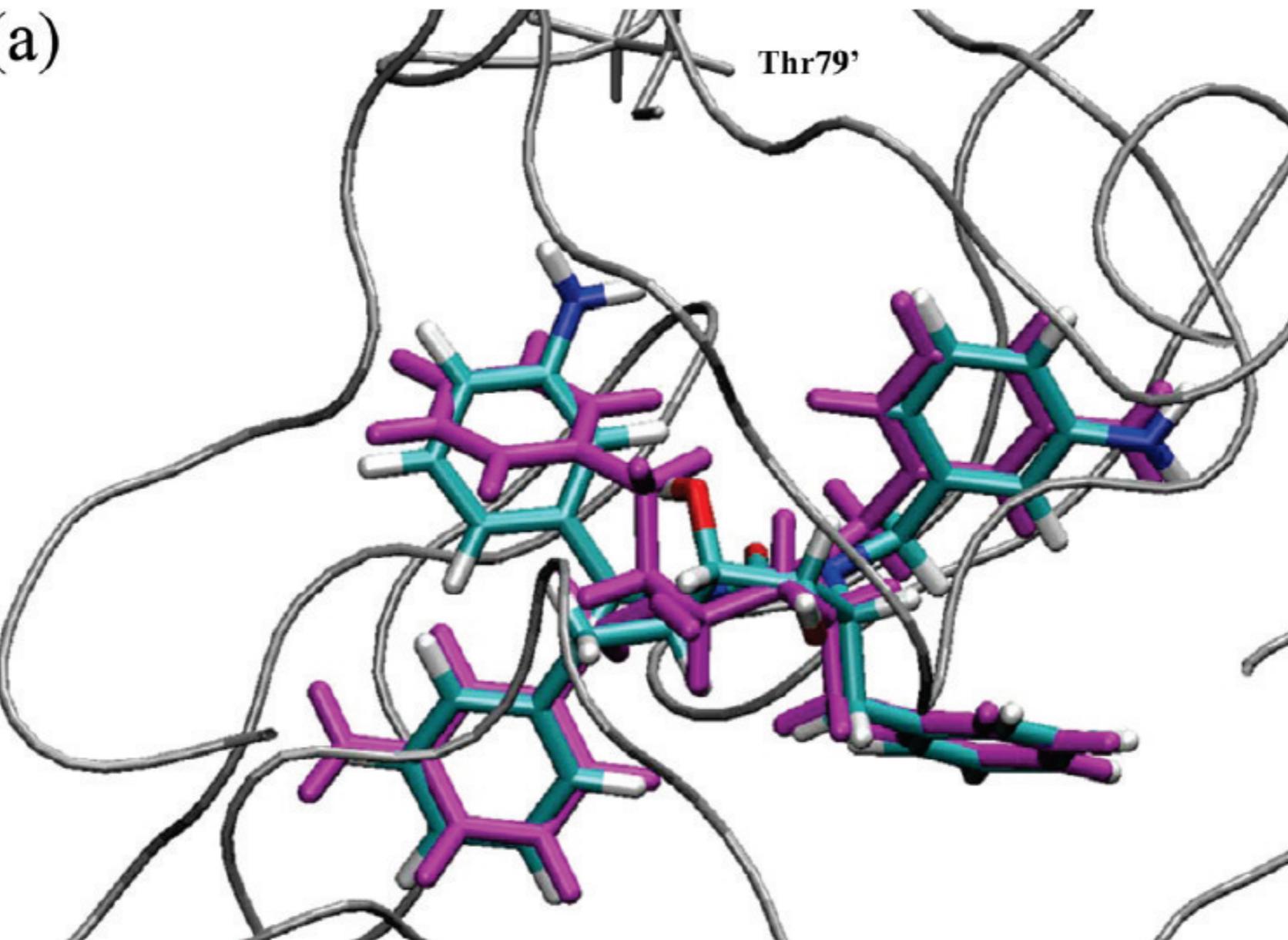
# Rescoring can be done with energy minimization and/or solvent effects



Steinbrecher et al., J. Med. Chem. 49(6): 1837 (2006)

# Rescoring can even be done with QM

(a)



Fong et al., JCLM 49:913 (2009)

In our docking assignment, (if chosen), we will use the FRED toolkit from OpenEye

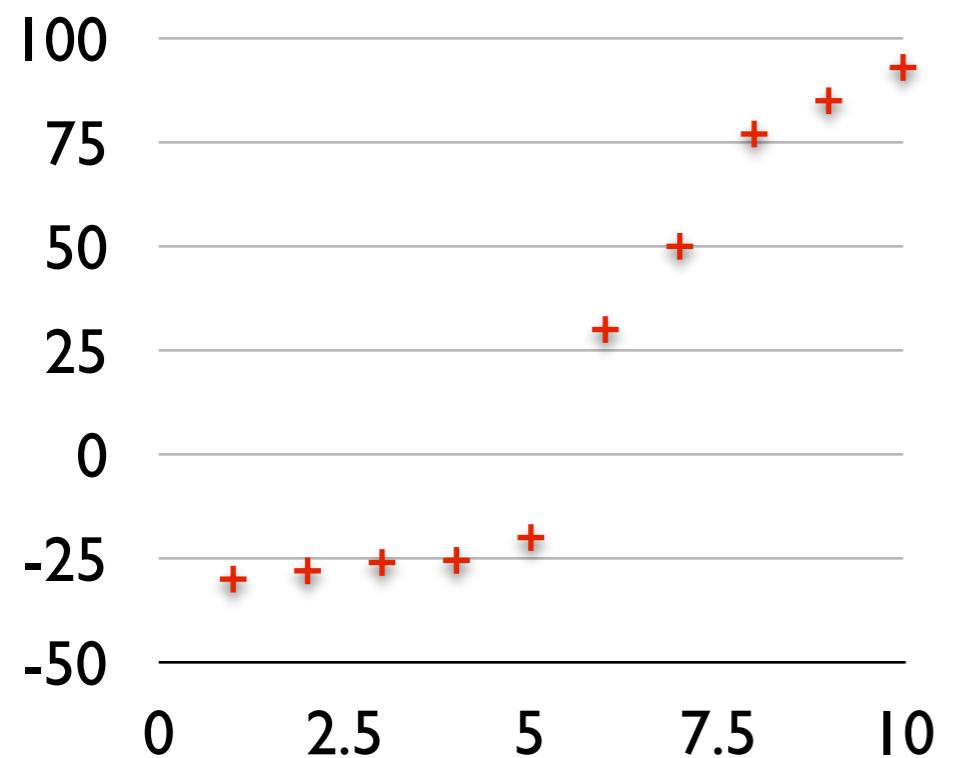
FRED works by:

- A) First prepare ligands by generating multiple conformations of each
- B) Load the receptor
- C) Define a “box” or identify the binding site -- the region to dock into
- D) Rigidly dock ligand conformers into binding site

# Docking outcomes depend a lot on the details

- Sometimes scores clearly separate “good” and “bad”
- A particular approach may work well in some sites, not others
- Difficult to predict when one scoring function will work well and when not
- No method works well across all targets

# The “hit rate” provides one way to measure success in virtual screening



## Hit rate:

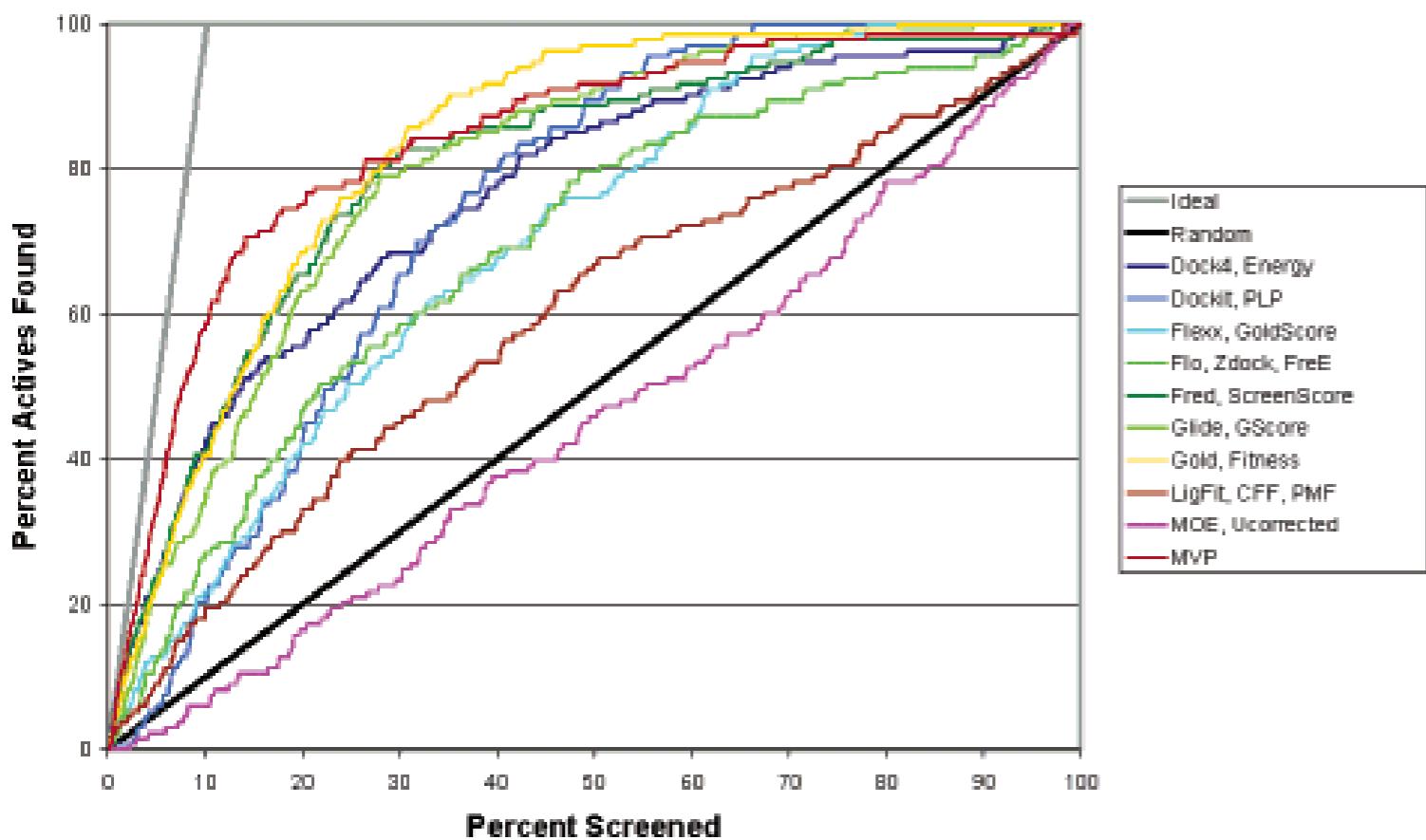
- What percentage of the actives (in a test set) are at the top few percent of the database
- Related: What percentage of compounds are active in the top few percent of the database

Problem: Be careful of confidence level on hit rate -- 22% may not be better than 19%

# An “enrichment plot” provides more information than a hit rate in *tests of docking methods*

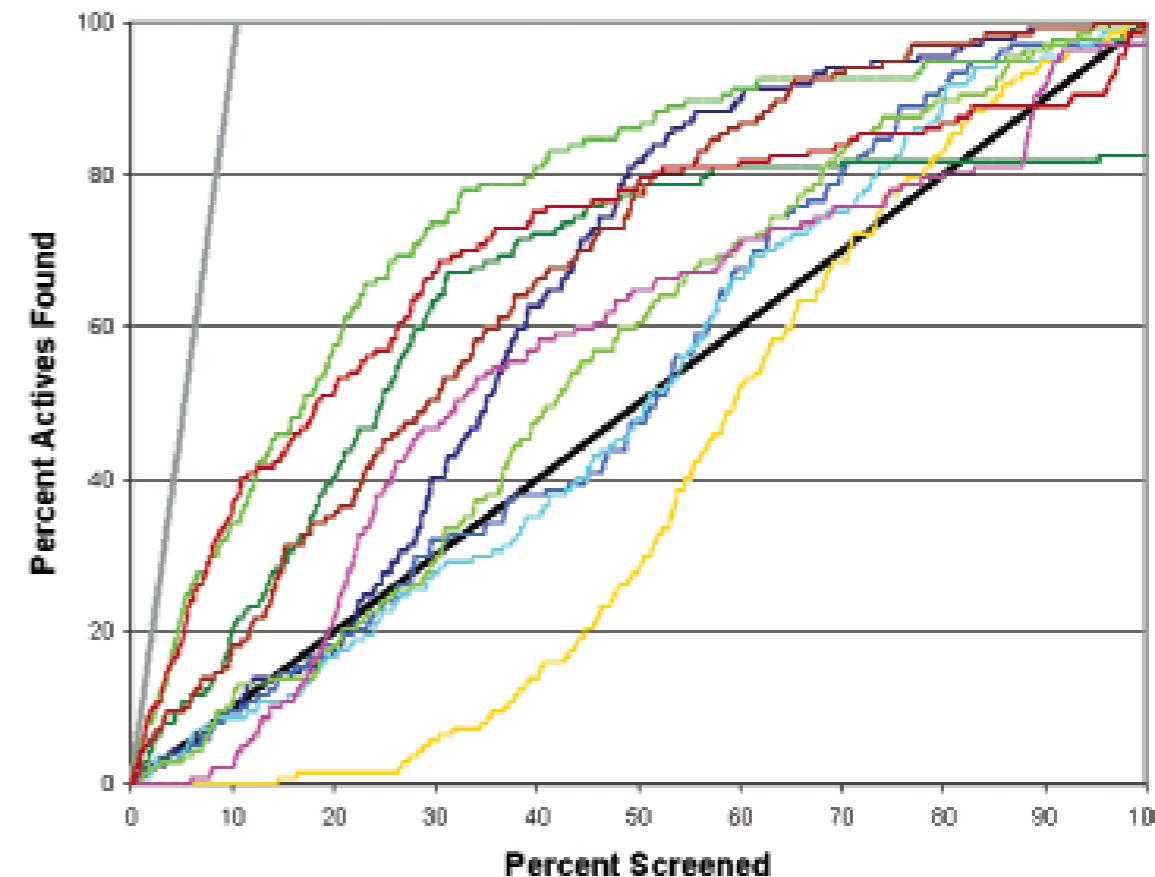
D.

**Factor Xa**

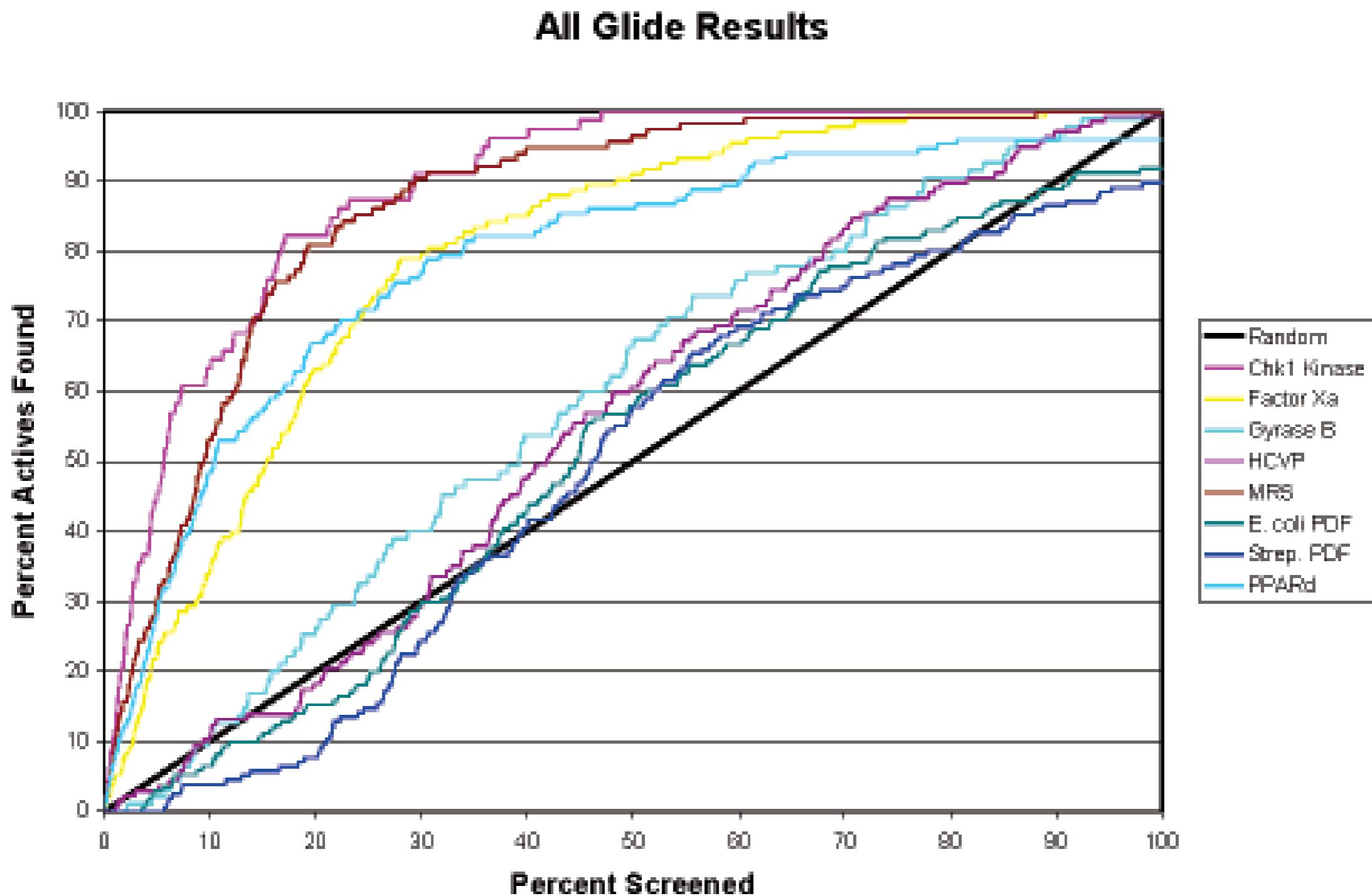


H.

**HCV Polymerase**

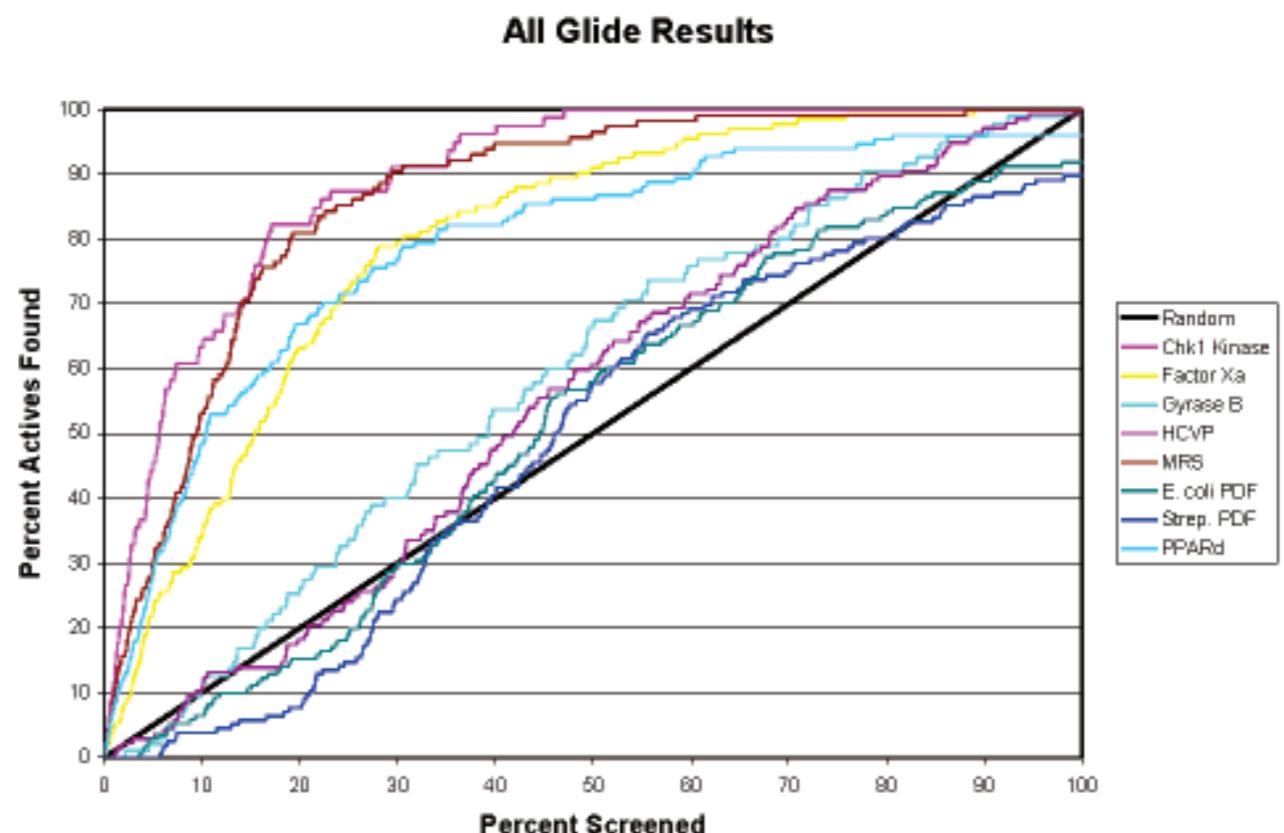


# Enrichment of course depends on the difficulty of the particular docking problem



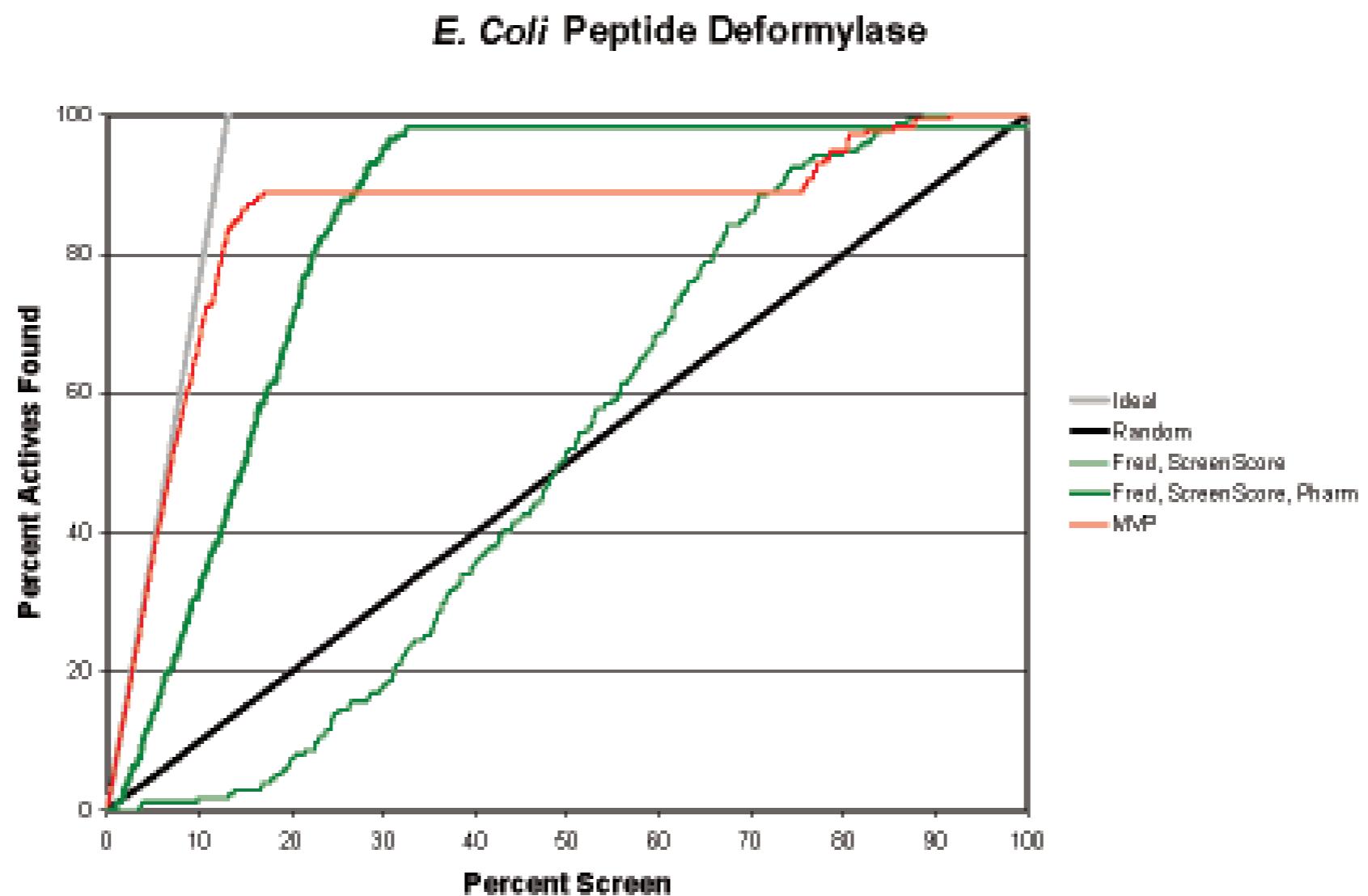
There are other metrics people use

- ROC and variants
  - Pose prediction quality
    - Such as RMSD
    - (Takes lots of structures!)



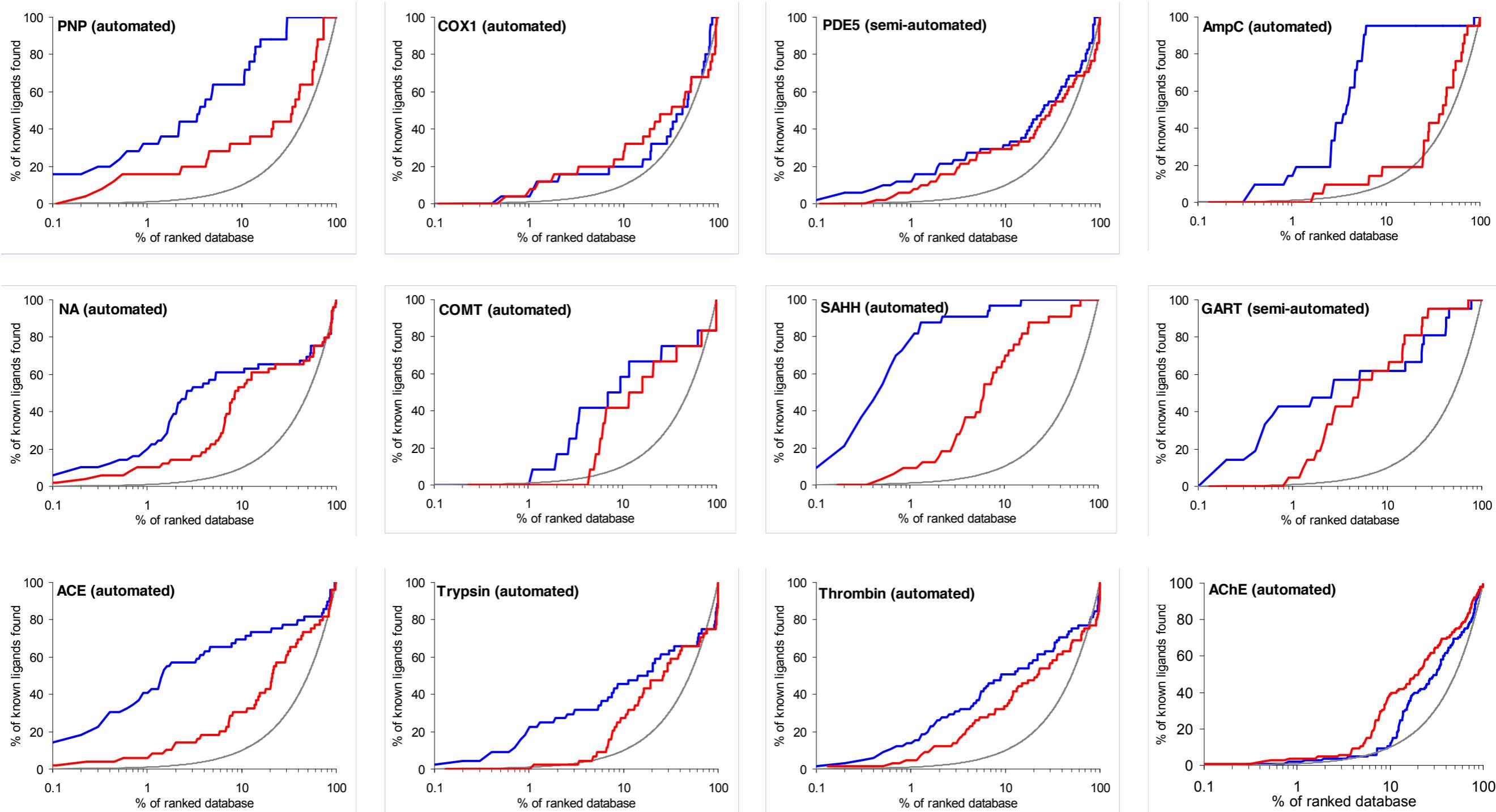
# To do well at screening using docking, use what you know

If you're targeting a protein that has a metal in the binding site (i.e. a metalloprotease), your ligands should be able to interact well with metals

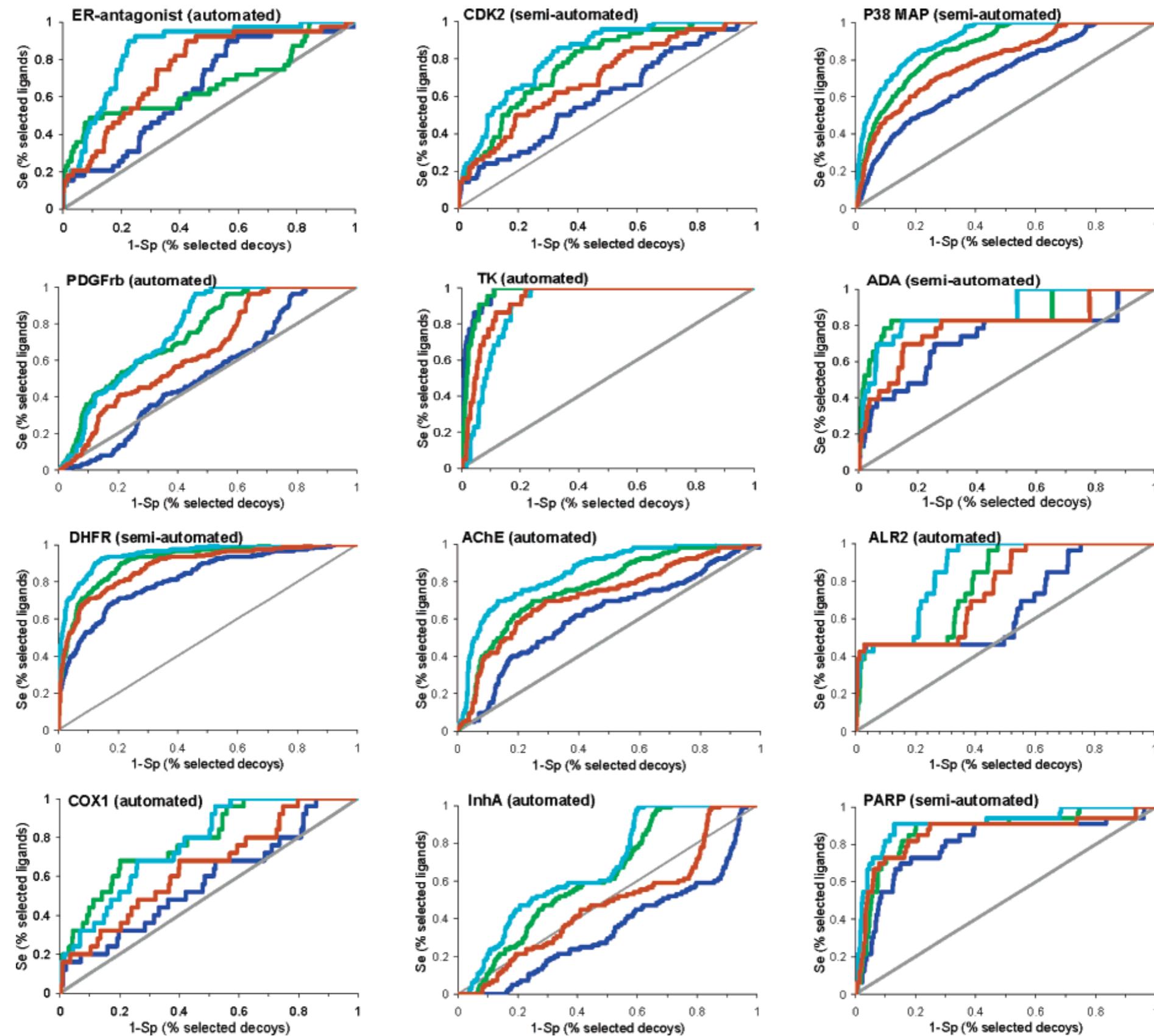


Looking for ligands with metal-binding groups leads to 26 times better enrichment

# Thus, it's often harder to find actives when looking at *similar* compounds than when looking elsewhere



# Thus, enrichment depends on the dataset



# What to screen in virtual screening, then?

- If you're *testing* or *developing* a method, use decoys like your actives (see DUD)
- Otherwise, start with what you've got or what you know
  - Ligands you have on hand
  - Ligands you can easily buy (ZINC)
  - Molecules you could easily make which are interesting for some reason
  - NOT stuff no one can ever synthesize

The screenshot shows the ZINC12 website. At the top, there's a dark header bar with the UCSF logo and links for "About UCSF", "Search UCSF", and "UCSF Medical Center". Below the header is a large blue banner with the "ZINC" logo and the number "12". Underneath the banner is a navigation menu with links for "About", "Search", "Subsets", "Help", and "Social". The main content area has a light blue background and contains a welcome message about the ZINC database.

Welcome to ZINC, a free database of commercially-available compounds for virtual screening. ZINC contains over 35 million purchasable compounds in ready-to-dock, 3D formats. ZINC is provided by the [Shoichet Laboratory](#) in the Department of Pharmaceutical Chemistry at the University of California, San Francisco (UCSF). To cite ZINC, please reference: Irwin, Sterling, Mysinger, Bolstad and Coleman, *J. Chem. Inf. Model.* 2012 DOI: [10.1021/ci3001277](https://doi.org/10.1021/ci3001277). The original publication is Irwin and Shoichet, *J. Chem. Inf. Model.* 2005;45(1):177-82 [PDF](#), [DOI](#). We thank [NIGMS](#) for financial support (GM71896).