Introduction:
Improving virtual screening through physics-based methods

CS/CME/Biophys/BMI 371
Jan. 30, 2018
Ron Dror
Virtual screening

• Virtual screening: Identifying drug candidates by considering large numbers of possible ligands
  – A ligand is any molecule that might bind to a protein
• Virtual screening is an alternative to experimental high-throughput screening (done by robots)
• Once a candidate is identified, it undergoes an extensive optimization process in which it is modified chemically to improve its properties
  – This optimization is a big part of drug discovery
Ligand docking: standard approach to virtual screening

- Predicts...
  - The **pose** of the molecule in the binding site
  - The binding affinity or a **score** representing the strength of binding

Note that predicting binding **pose** (i.e., where each atom of the ligand ends up) is very important in its own right, particularly for the ligand optimization process
## Ligand docking software

### Most popular (based on citations 2001–2011):

- AutoDock
- GOLD
- DOCK
- FlexX
- Glide
- FTDOCK
- QXP

Sousa et al., Current Medicinal Chemistry 2013

So what’s the problem with ligand docking?

• Ligand docking is a physics-based heuristic approach with two key components
  – A scoring function that very roughly approximates the binding affinity (i.e., binding strength) of a ligand to a protein given a binding pose
  – A search method that searches for the best-scoring binding pose for a given ligand
• Accuracy is poor!
Why aren’t standard (physics-based) docking methods very accurate?

• Protein flexibility
  – The binding pocket may adopt different conformations when bound to different ligands
  – Most docking protocols treat it as rigid

• Both the protein and the ligand are continually wiggling around, both before and after binding
  – Most docking protocols don’t account correctly for entropic effects (“proteins and ligands like to be free”)
  – They also don’t account for some of the effects of water molecules
In theory, we could determine binding affinity by simply running molecular dynamics simulations.

- We would watch the ligand bind and unbind multiple times and determine what fraction of the time it was bound, on average.

- This isn’t practical—the simulations would need to be much, much too long.

Beta-blocker alprenolol binding to an adrenaline receptor

Dror et al., *PNAS* 2011
“Alchemical” simulation methods

- Binding affinity depends on the difference in energy between the bound and unbound states.
- It does not depend on the binding/unbinding pathways.
- However, one needs to a pathway to compute the difference in energy.
- Solution: use a fictitious unbinding pathway, in which the ligand gradually disappears from the binding pocket and rematerializes in water.
Another approach: exploit experimental information on protein flexibility

• If you have a very high-resolution crystal structure, you can extract information on different conformations the binding pocket can adopt in the absence of a ligand
• You can then dock to those different protein conformations
  – Include an energetic penalty for the protein conformations that are less populated in the absence of a ligand
Background material

• Ligand docking slides from CS/CME/BioE/Biophys/BMI 279:

• Slides on the relationship between probabilities and energy of a state (the Boltzmann distribution) from CS/CME/BioE/Biophys/BMI 279: