Ligand docking and virtual screening

CS/BioE/CME/Biophys/BMI 279

Oct. 27, 2020

Ron Dror

Outline

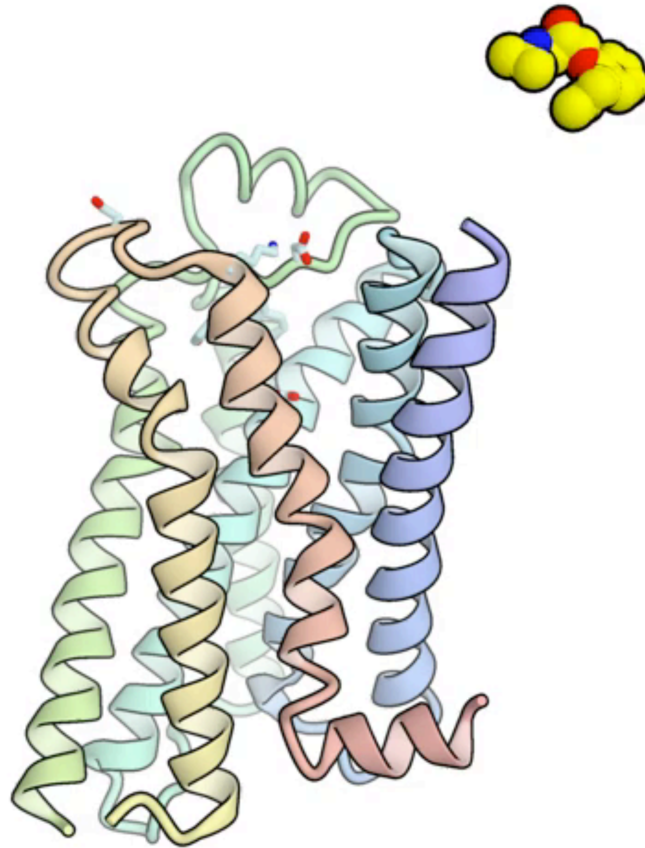
- Goals of ligand docking
- Defining binding *affinity* (strength)
- Computing binding affinity: Simplifying the problem
- Ligand docking methodology
- How well does docking work?

Goals of ligand docking

A drug binding to its target

(The great majority of drug targets are proteins)

0.00 us

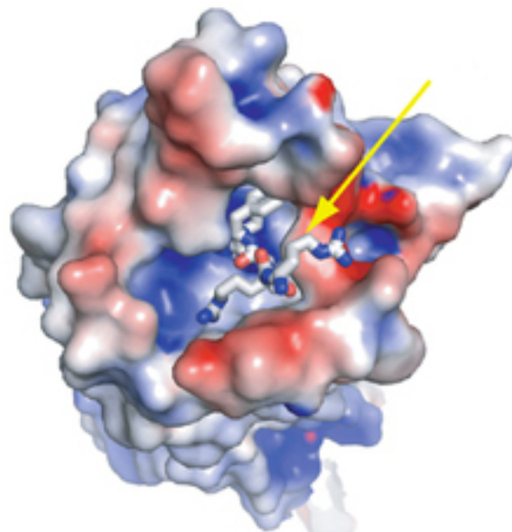


Beta-blocker alprenolol binding to an adrenaline receptor

Dror et al., *PNAS* 2011

Problem definition

- A *ligand* is any molecule that binds to a macromolecule (e.g, protein or RNA)
  - We'll also use *ligand* to refer to any molecule (e.g., any candidate drug) that *might* bind to a given macromolecule (e.g., a drug target)
- *Ligand docking* addresses two problems:
  - Given a ligand known to bind a particular protein, what is its binding *pose* (that is, the location, orientation, and internal conformation of the bound ligand—basically, the position of each ligand atom when bound)
  - How *tightly* does a ligand bind a given protein (or other macromolecule)?

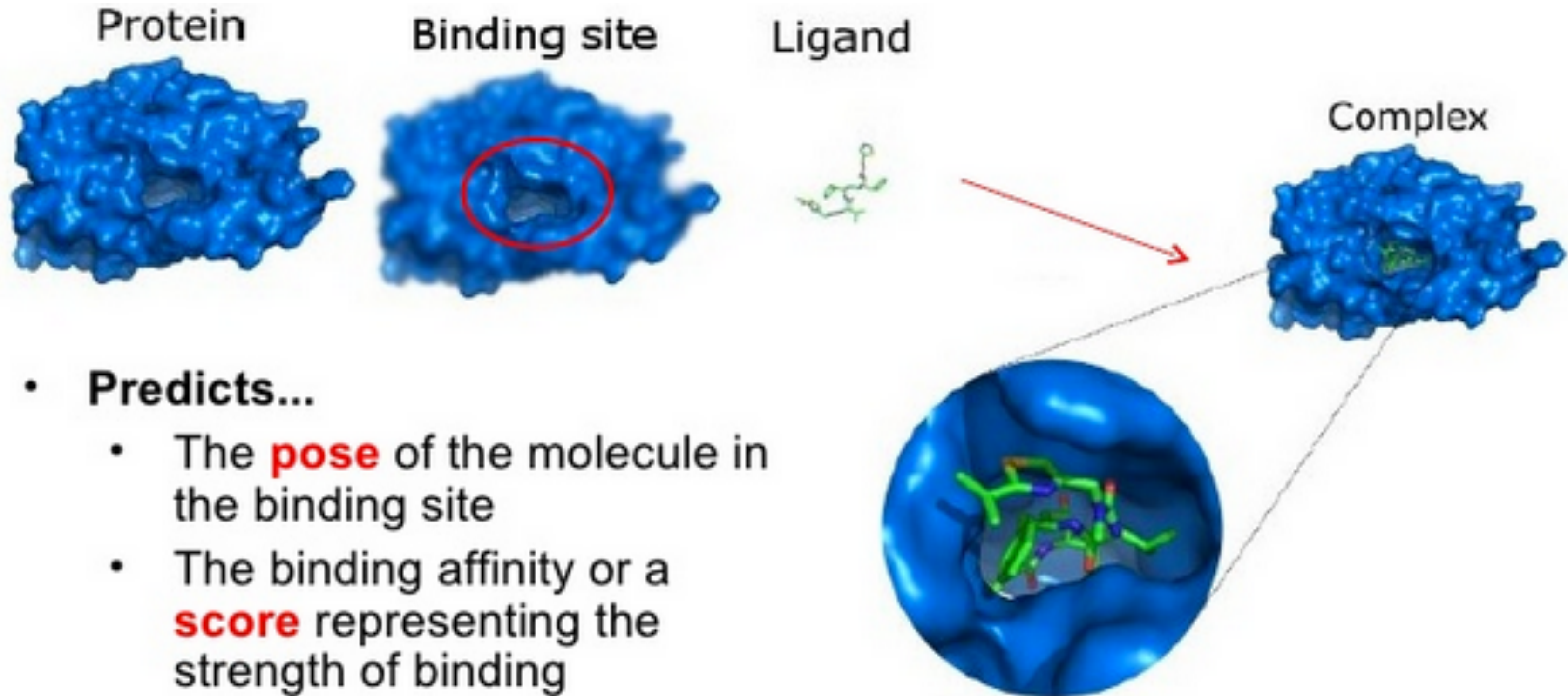


http://www.nih.gov/researchmatters/october2012/images/structure_l.jpg

Why is docking useful?

- *Virtual screening*: Identifying drug candidates by considering large numbers of possible ligands
- *Lead optimization:* Modifying a drug candidate to improve its properties
  - If the binding pose of the candidate is unknown, docking can help identify it (which helps envision how modifying the ligand would affect its binding)
  - Docking can predict binding strengths of related compounds

Ligand docking: a graphical summary



http://www.slideshare.net/baoilleach/proteinligand-docking-13581869

Defining binding *affinity* (strength)

How do we measure how tightly a ligand binds to a protein?

- *Binding affinity* quantifies the binding strength of a ligand to a protein (or other target)
  - Conceptual definition: if we mix the protein and the ligand (with no other ligands around), what fraction of the time will the protein have a ligand bound?
    - This depends on ligand concentration, so we assume that the ligand is present at some standard concentration.
  - Binding affinity can be expressed in two ways:
    - The difference ΔG in free energy of the bound state (all atomic arrangements where the protein is ligand-bound) and the unbound state (all atomic arrangements where the protein is not ligand-bound)
      - Again, assume standard concentration of ligand
      - From ΔG, one can compute the fraction of time the ligand will be bound
    - A dissociation constant (K_d), which is (roughly) the ligand concentration at which half the protein molecules will have a ligand bound

Binding affinity: Clarifications

- Binding affinity is different from “how long the ligand remains bound” (the off-rate) or “how quickly the ligand binds” (the on-rate)
  - Binding affinity is a ratio of the on-rate and off-rate; you can’t calculate it from either one alone
  - These rates are also of interest in drug discovery, and predicting them is a different (and challenging) computational problem

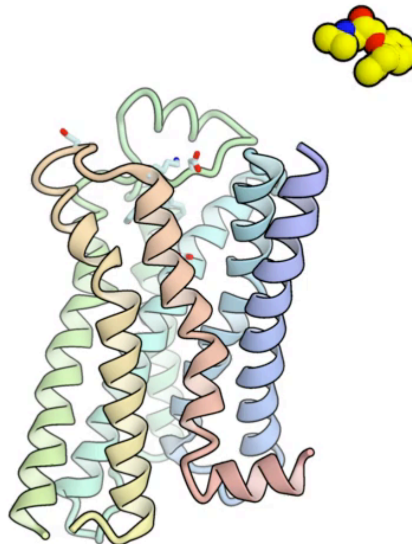
- Binding affinity is different from “how strong are the inter-atomic forces between the ligand and the target when the ligand is bound”
  - Binding affinity also depends a great deal on what happens when the ligand isn’t bound—for example, how favorable are the interactions of the ligand and the binding pocket with water

Computing binding affinity: Simplifying the problem

A hypothetical direct approach to computing binding affinity

- Run a really long molecular dynamics (MD) simulation in which a ligand binds to *and* unbinds from a protein many times.
- Directly observe the fraction of time the ligand is bound.

0.00 us



This direct approach doesn't work

- It is so computationally intensive that we usually cannot do it for even a single ligand, let alone millions
  - The toughest part is the unbinding (dissociation): Drug molecules usually take seconds to hours to unbind from their targets.
  - Microsecond-timescale molecular dynamics simulations usually take days.
  - We'd have to simulate many cycles of binding and unbinding.

How would you compute a binding affinity?

- Suppose you're given the structure of a target protein, and you want to compute the affinity of a particular ligand
  - To simplify the problem a bit, assume that you're given the binding pose

What can we do instead?
Option 1: Use alternative MD-based approaches

- It turns out that one can compute binding affinities by MD in more efficient ways
  - In these methods, called free energy perturbation (FEP) and thermodynamic integration (TI), the ligand gradually dematerializes from its bound position and materializes in an unbound position, rather than following a realistic path between bound and unbound positions. *This works because binding affinity does not depend on the binding pathway.*
  - These methods currently represent the most accurate way to determine binding affinities computationally
  - They are very expensive computationally and thus cannot be used on large numbers of ligands
  - They assume that one knows the binding pose
- There are also methods based on implicit solvent MD simulation (where water molecules are not represented explicitly)
  - For example, MM-PB/SA or MM-GS/SA (Molecular Mechanics - Poisson Boltzman (or Generalized Born)/Surface Area))
  - These methods are faster, but still computationally intensive
  - They are somewhat less accurate
  - They again assume that one knows the binding pose

Option 2: Ligand docking
(most common in practice)

- Ligand docking is a fast, heuristic approach with two key components
  - A *scoring function* that very roughly approximates the binding affinity 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
- Most ligand docking methods assume that
  - The target protein is rigid
  - The approximate binding site is known
    - That is, one is looking for ligands that will bind to a particular site on the target
- In reality, ligand mobility, protein mobility, and water molecules all play a major role in determining binding affinity
  - Docking is approximate but useful
  - The term *scoring function* is used instead of *energy function* to emphasize the highly approximate nature of the scoring function

Docking software

Program	Country of Origin	Year Published
AADS	India	2011
ADAM	Japan	1994
AutoDock	USA	1990
AutoDock Vina	USA	2010
BetaDock	South Korea	2011
DARWIN	USA	2000
DIVALI	USA	1995
DOCK	USA	1988
DockVision	Canada	1992
EADock	Switzerland	2007
eHiTS	UK	2006
EUDOC	USA	2001
FDS	UK	2003
FlexE	Germany	2001
FlexX	Germany	1996
FLIPDock	USA	2007
FLOG	USA	1994
FRED	UK	2003
FTDOCK	UK	1997
GEMDOCK	Taiwan	2004
Glide	USA	2004
GOLD	UK	1995
Hammerhead	USA	1996
ICM-Dock	USA	1997
Lead finder	Canada	2008
LigandFit	USA	2003
LigDockCSA	South Korea	2011
LIGIN	Germany	1996
LUDI	Germany	1992
MADAMM	Portugal	2009
MCDOCK	USA	1999
MDock	USA	2007
MolDock	Denmark	2006
MS-DOCK	France	2008
ParDOCK	India	2007
PhDOCK	USA	2003
PLANTS	Germany	2006
PRO_LEADS	UK	1998
PRODOCK	USA	1999
ProPose	Germany	2004
PSI-DOCK	China	2006
PSO@AUTODOCK	Germany	2007
PythDock	South Korea	2011
Q-Dock	USA	2008
QXP	USA	1997
rDock	UK	2013
SANDOCK	UK	1998
SFDOCK	China	1999
SODOCK	Taiwan	2007
SOFTDocking	USA	1991
Surflex	USA	2003
SYSDOC	USA	1994
VoteDock	Poland	2011
YUCCA	USA	2005

Most popular (based on citations 2001–2011):

AutoDock
GOLD
DOCK
FlexX
Glide
FTDOCK
QXP

Sousa et al., Current Medicinical Chemistry 2013

http://en.wikipedia.org/wiki/Docking_(molecular)

You are not responsible for the details on this slide

Ligand docking methodology

Scoring functions

- Scoring functions used for docking tend to be empirical
  - Capture chemists' intuition about what makes a ligand–receptor interaction energetically favorable (e.g., hydrogen bonding, or displacement of water from a hydrophobic binding pocket)
  - Parameters are often optimized based on known binding affinities of many ligands for many receptors
  - Some scoring functions borrow terms from molecular mechanics force fields, but a molecular mechanics force field is rarely used directly as a scoring function for docking
    - The scoring function is an (extremely rough) attempt to approximate the binding *free energy*. By contrast, molecular mechanics force fields give *potential energy* associated with a particular arrangement of atoms.

Example: Glide scoring function

- Glide (considered one of the most accurate docking software packages) uses the following "GlideScore" function in SP ("standard precision") mode:

$$
\begin{aligned}
\Delta G_{\mathrm{bind}} = {} & C_{\mathrm{lipo-lipo}} \sum f(r_{\mathrm{lr}}) + \\
& C_{\mathrm{hbond-neut-neut}} \sum g(\Delta r)\ h(\Delta \alpha) + \\
& C_{\mathrm{hbond-neut-charged}} \sum g(\Delta r)\ h(\Delta \alpha) + \\
& C_{\mathrm{hbond-charged-charged}} \sum g(\Delta r)\ h(\Delta \alpha) + \\
& C_{\mathrm{max-metal-ion}} \sum f(r_{\mathrm{lm}}) + C_{\mathrm{rotb}} H_{\mathrm{rotb}} + \\
& C_{\mathrm{polar-phob}} V_{\mathrm{polar-phob}} + C_{\mathrm{coul}} E_{\mathrm{coul}} + \\
& C_{\mathrm{vdW}} E_{\mathrm{vdW}} + \text{solvation terms}
\end{aligned}
$$

Friesner et al., Journal of Medicinal Chemistry 47:1739-49 (2004)

  - The first term rewards contacts between hydrophobic atoms of the ligand and protein, and is a function of the distance between them
  - The next three terms reward specific kinds of hydrogen bonds, and are a function of both angle and distance
- The final ranking of ligands in Glide SP is determined by a combination of the GlideScore, an interaction energy computed using a molecular mechanics force field (OPLS-AA), and an estimate of the internal strain of the ligand in the bound pose
- Glide's XP ("extra precision") mode uses an even more complicated scoring function

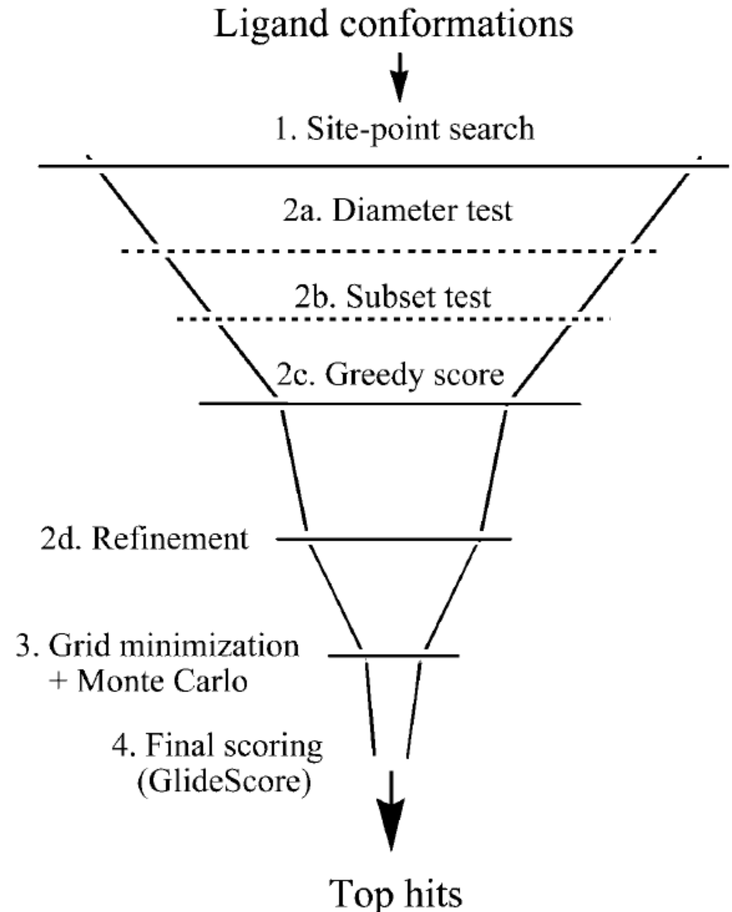
You are not responsible for the details on this slide

Search methods

- Docking software needs to search for the best-scoring pose for each ligand
- The search space is huge, because one needs to consider all possible ligand positions and orientations, and the ligand’s internal degrees of freedom
- To search this space efficiently, docking software typically employs some combination of:
  - Heirarchical methods in which one uses approximate measures to identify promising groups of poses, then evaluates them in more detail
  - Monte Carlo methods

Example: Glide search

- Glide SP uses a hierarchical search method
- It first identifies a discrete set of "reasonable" conformations for each ligand, by varying internal torsion angles
- For each ligand, it scans possible positions and orientations, using a rough metric of fit
- The most promising approximate poses undergo further "refinement" and evaluation



Friesner et al., *J Med Chem* 47:1739, 2004

You are not responsible for the details on this slide

How well does docking work?

How well does docking work?

- The best docking protocols:
  - Predict a reasonably accurate pose (for ligands that do in fact bind the target protein) about half the time for rigid targets (the “easy” cases)
    - In these cases, one of the highly-ranked poses is usually close to the correct one
  - Provide useful, but far from perfect results, when ranking ligands
    - Tend to work best when comparing closely related ligands
  - Are not particularly useful when it comes to quantitatively estimating binding free energies

Leach et al., J Med Chem 49:5851 (2006)

Warren et al., J Med Chem 49:5912 (2006)

How well does docking work?

Example: Performance of Glide on ligand-ranking tests for multiple targets.

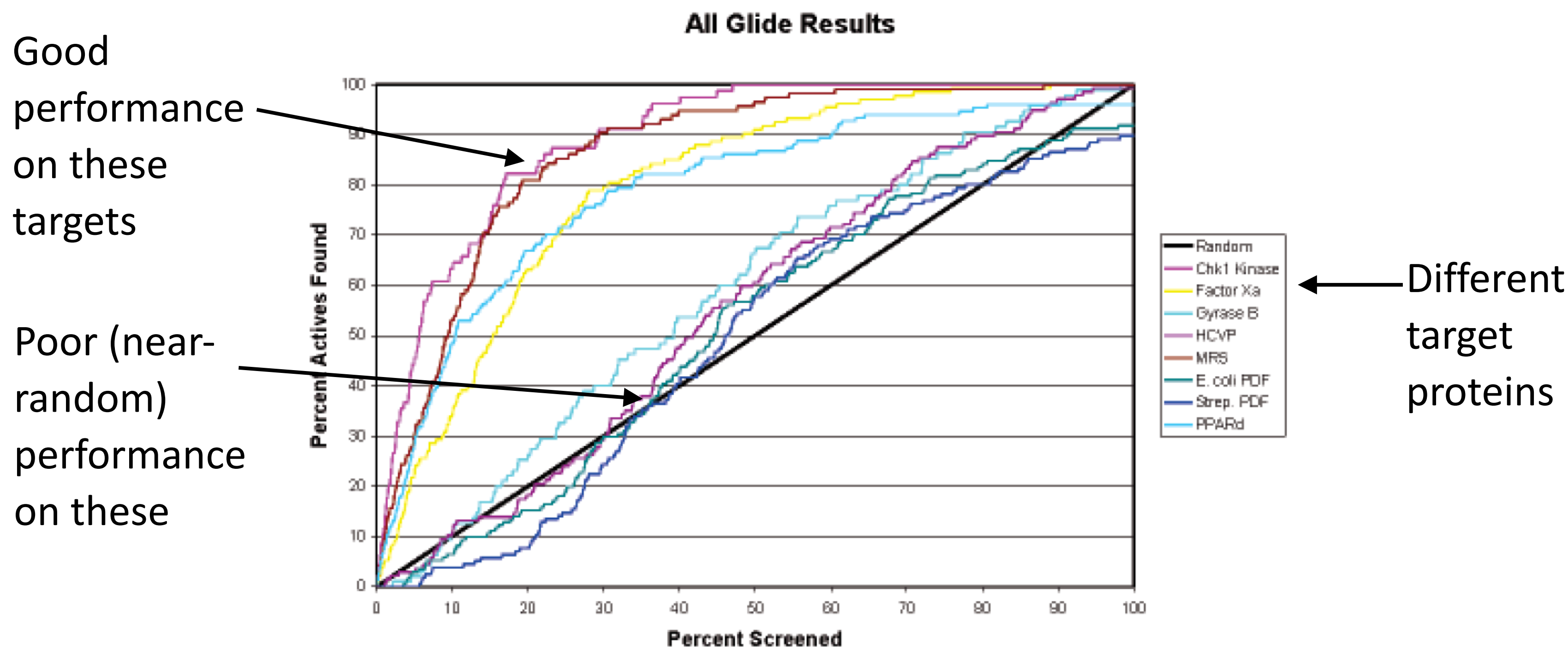


Figure 7. Illustrative example of how enrichment by a single program varied across the targets evaluated using data from the program Glide. Similar variation in performance was observed in all docking programs evaluated.

Warren et al., J Med Chem 49:5912 (2006)

How well does docking work?

Example: Correlation between docking scores and affinity for one target

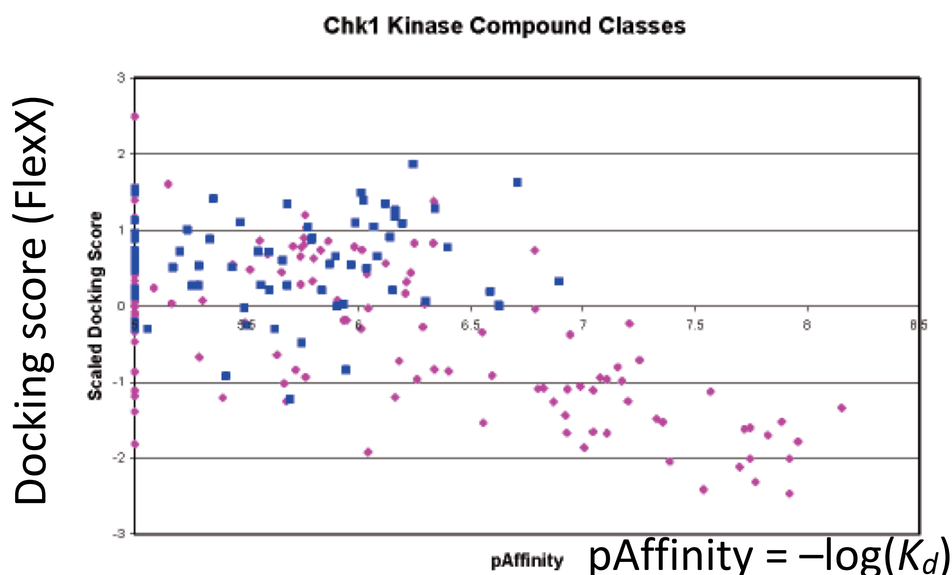


Figure 10. Plot of scaled score vs pAffinity where the two Chk1 kinase chemical classes are plotted in magenta (class **1**) and blue (class **2**). It is readily apparent that all of the correlation observed between the scaled docking score and affinity is found in the class **1** molecules and that no correlation exists between the docking score and class **2** compound affinities.

Magenta points correspond to ligands from one chemical family. Blue points correspond to a second chemical family.

Magenta points: decent correlation between docking score and affinity.

Blue points: no correlation.

Warren et al., J Med Chem 49:5912 (2006)

Despite this inaccuracy, docking has proven very useful

- Allows screening of 10^8–10^9 ligands — more than could be tested experimentally by “high-throughput screening” robots
- Typically used in combination with:
  - experimental validation of top “hits”
  - human intuition to choose which of the top-ranked ligands to test experimentally (“hit picking”)
  - optimization of experimentally validated binders by testing related ligands
- For example, see:
  - “Ultra-large library docking for discovering new chemotypes,” *Nature* 2019
    https://www.nature.com/articles/s41586-019-0917-9?fbclid=IwAR1HDXx0kEsNIRQZXVtPkmX7hU_gDoT2aqVEiBZj04qhz_6x1WCbNkj75IE