Ligand docking

CS/CME/BioE/Biophys/BMI 279

Nov. 9 and 14, 2017

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?

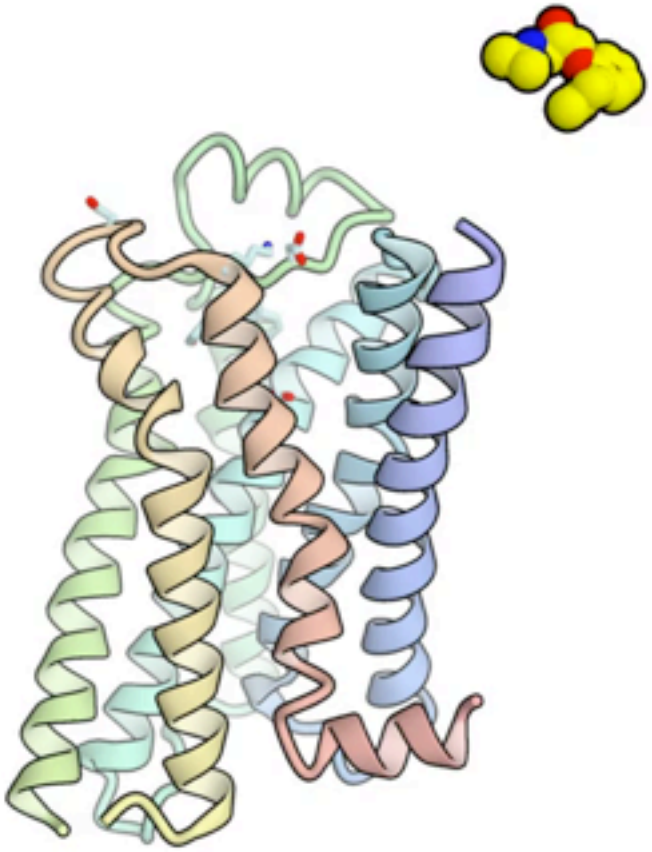
Goals of ligand docking

A drug binding to its target

(The great majority of drug targets are proteins)

0.00 us

There is constant binding and unbinding. It's not that the ligand goes straight to the binding pocket and sits there. It just *prefers* a bound conformation because that conformation has lower free energy.



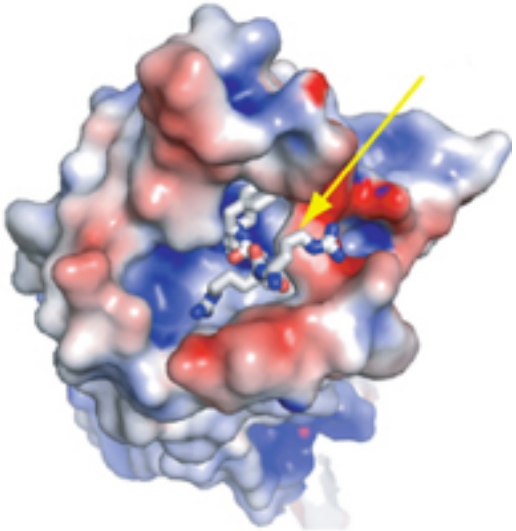
Beta-blocker alprenolol binding to an adrenaline receptor

Dror et al., *PNAS* 2011

Problem definition

- A *ligand* is any molecule that binds to a protein
  - We'll also use *ligand* to refer to any molecule that *might* bind to a protein (e.g., any candidate drug)
- *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?

Remember that many ligands are flexible molecules as well.



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

Why is docking useful?

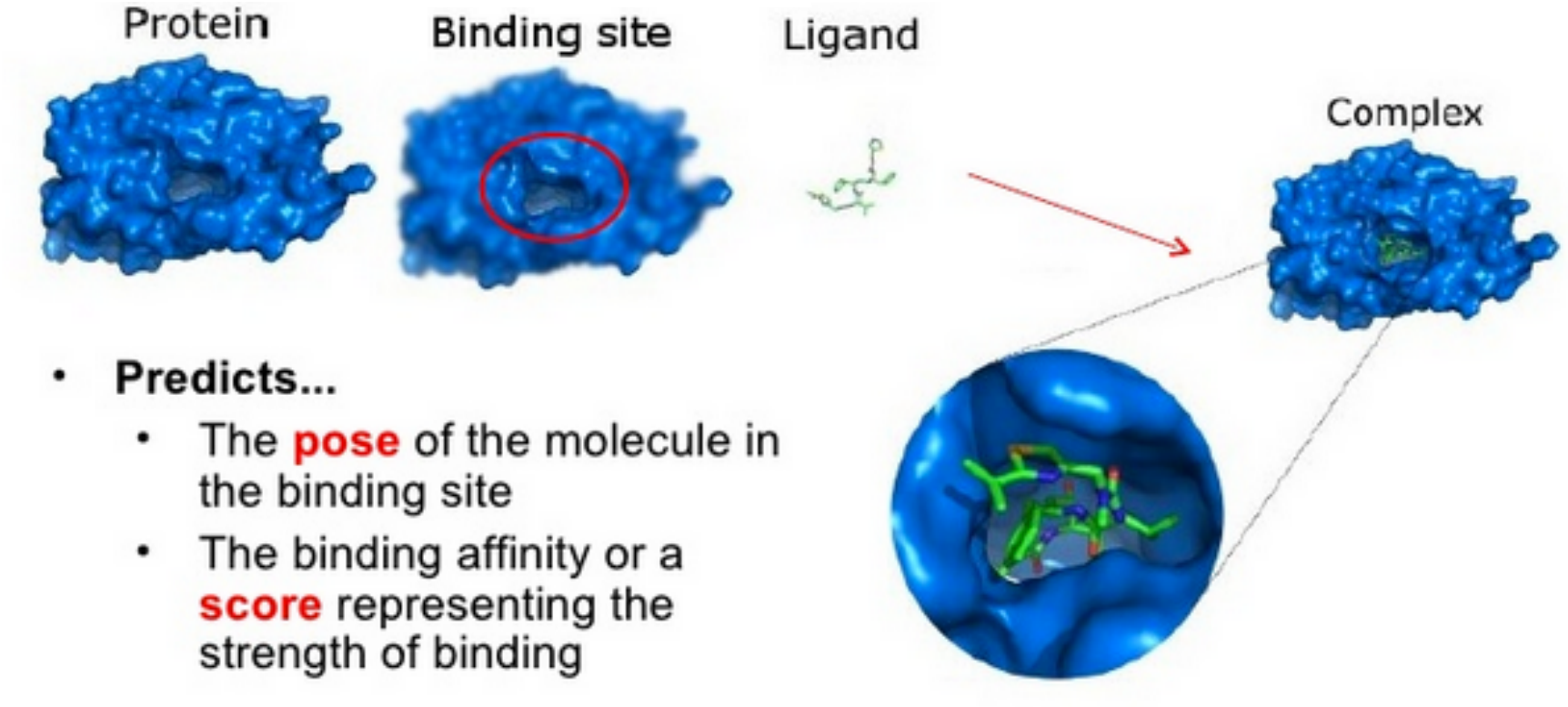
Though important, virtual screening is relatively a smaller part of the drug development process.

- *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

Initial drug candidates are often bad drugs. For example, although they might bind to the target, they often don't bind with high enough affinity.

Ligand docking: a graphical summary

Most software packages use a user-specified binding site.



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

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) Binding affinity is an equilibrium property, which means you need to give the proteins and ligands a lot of time to bind/unbind and find equilibrium.
  - 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? a higher affinity ligand will do some combination of binding faster and unbinding slower.
    - This depends on ligand concentration, so we assume that the ligand is present at some standard concentration.
  - Binding affinity is usually expressed as either:
    - 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

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

Consider the simplified reaction

RL <—> R + L

where R is the receptor, L is the free ligand, and RL is the ligand-receptor complex. Then the dissociation constant K_d is given by

$K_d = [R][L] / [RL]$

where [R] is the concentration of the receptor R, etc. In other words, K_d is the equilibrium constant of the reaction above.

Assume the ligand is in excess. When half the receptors are bound with the ligand, [R] = [RL], so [L] = K_d — in other words, K_d is the concentration of ligand at which half the receptors will be bound. People will often say they "have a 1pM drug", meaning that the drug has K_d = 1pM. If the drug is present at 1pM concentration, roughly half the receptors will be bound. If the drug is present at 1nM concentration (1000 times higher), then the vast majority of receptors will be bound.

Computing binding affinity: Simplifying the problem

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.

The direct approach doesn't work

- It is so computationally intensive that we usually cannot do it for even a single ligand
  - Drug molecules usually take seconds to hours to unbind from their targets.
  - Microsecond-timescale molecular dynamics simulations usually take days.

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
- Working with your neighbor, try to come up with at least two ways to do it
  - Don't look at the next couple slides. If you already have, try to come up with approaches that are not on those slides.

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
  - These methods, called free energy perturbation (FEP) and thermodynamic integration (TI), are very clever
  - They 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

This method is based on the observation that affinity is determined by the difference in free energy between the bound and unbound pose and does not depend on how the ligand got there. So you can simulate the ligand slowly disappearing from the bound pose and use that to calculate the free energy.

- There are also methods based on implicit solvent MD simulation (water molecules 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

Water molecules are very important for defining binding affinity. For example, water molecules in the binding pocket need to be displaced by the ligand.

You are not responsible for any of the methods on this slide

Option 2: Ligand docking

- 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 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

If you look under the hood, these algorithms are different in the details but use similar broad approaches.

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

You are not responsible for the details on this slide

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

Ligand docking methodology

Scoring functions

- Scoring functions used for docking tend to be empirical
  - Capture chemists' intuition about what makes a 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*. Molecular mechanics force fields give *potential energy* associated with a particular arrangement of atoms.

Remember that the bound macrostate free energy includes the potential energy of all possible arrangements of atoms in which the ligand is bound.

Example: Glide scoring function

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

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

"lipo" = "lipophilic" ≈ "hydrophobic"

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

solvation terms account for specific water effects

  - 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

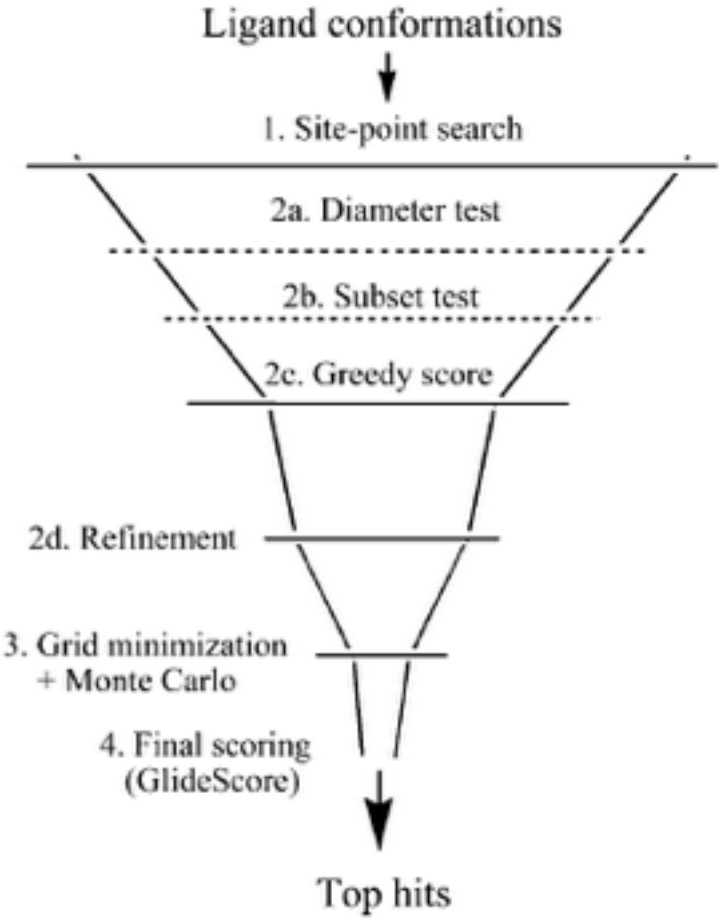
You are not responsible for the details on this slide

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:
  - Heuristic assumptions about what poses will/won't work
  - Monte Carlo methods
  - Hierarchical methods in which one uses approximate measures to identify promising groups of poses, then evaluates them in more detail

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?

Basically, it works so-so, but it's still useful and there isn't a better alternative, so this kind of software is very widely used.

- 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

e.g. suppose you have 1 million possible ligands but can only test 1000. People will perform "virtual screening" and evaluate its enrichment: namely, how many more ligands bind experimentally when you test the top 1000 ranked ligands versus 1000 randomly chosen ligands? typical enrichment values are 5-fold or 10-fold, so this method is broadly useful but far from perfect for individual cases.

    - 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)

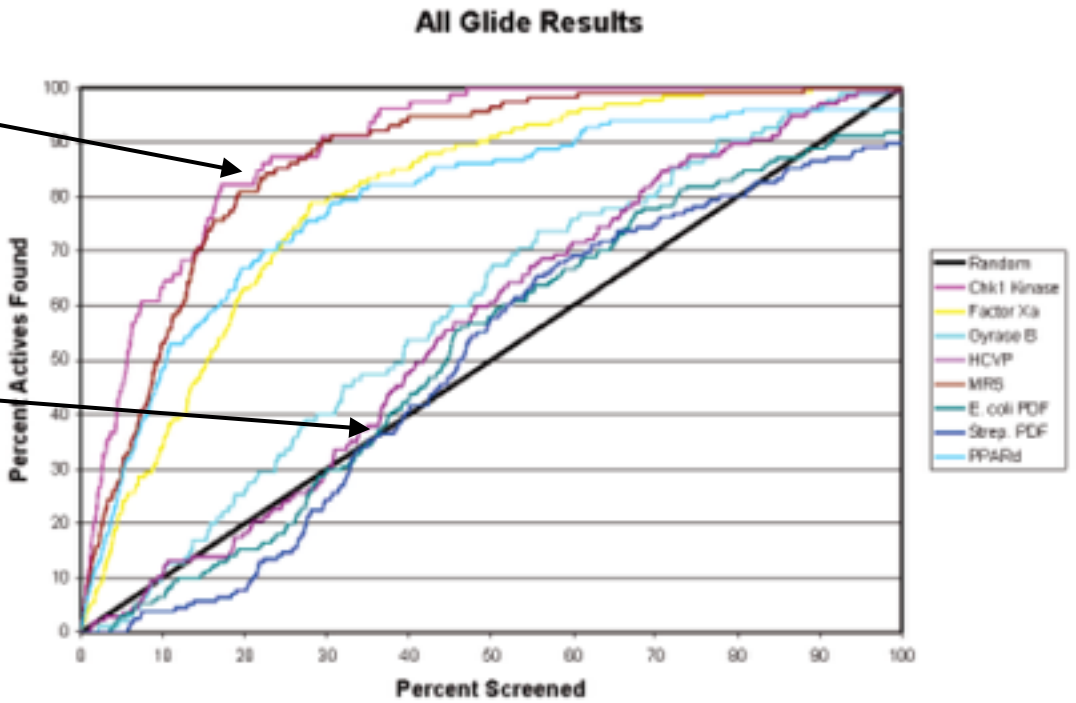
Warren et al., J Med Chem 49:5912 (2006)

How well does docking work?

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

Good performance on these targets

Poor (near-random) performance on these



Different target proteins

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.

For some targets, Glide gives results that are just as good as randomly choosing ligands! Performance is highly target-dependent.

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

How well does docking work?

Example: Correlation between docking scores and affinity for one target

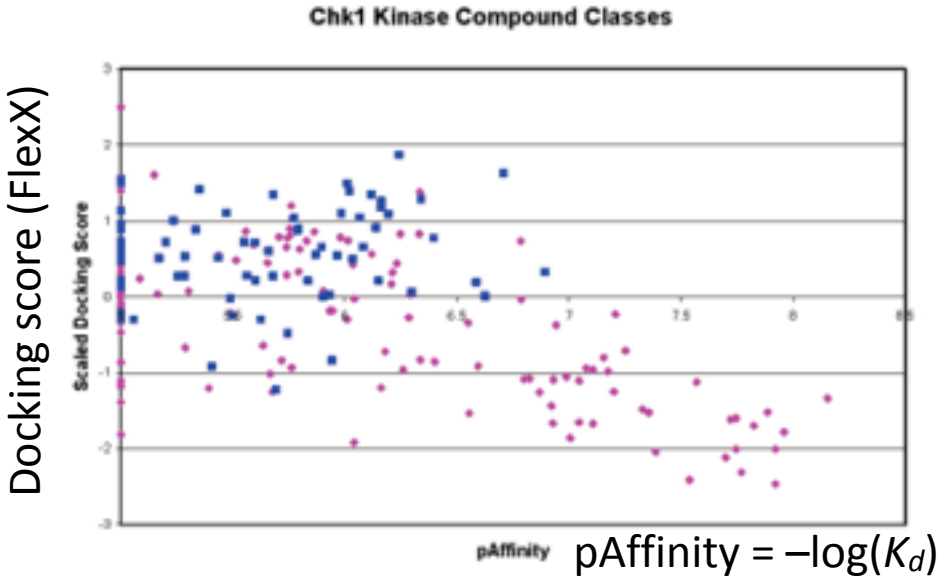


Figure 10. Plot of scaled score vs pAffinity where the two Chk1 kinase chemical classes are plotted in magenta (class **1**) and blue (class **2**). It is readily apparent that all of the correlation observed between the scaled docking score and affinity is found in the class **1** molecules and that 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)