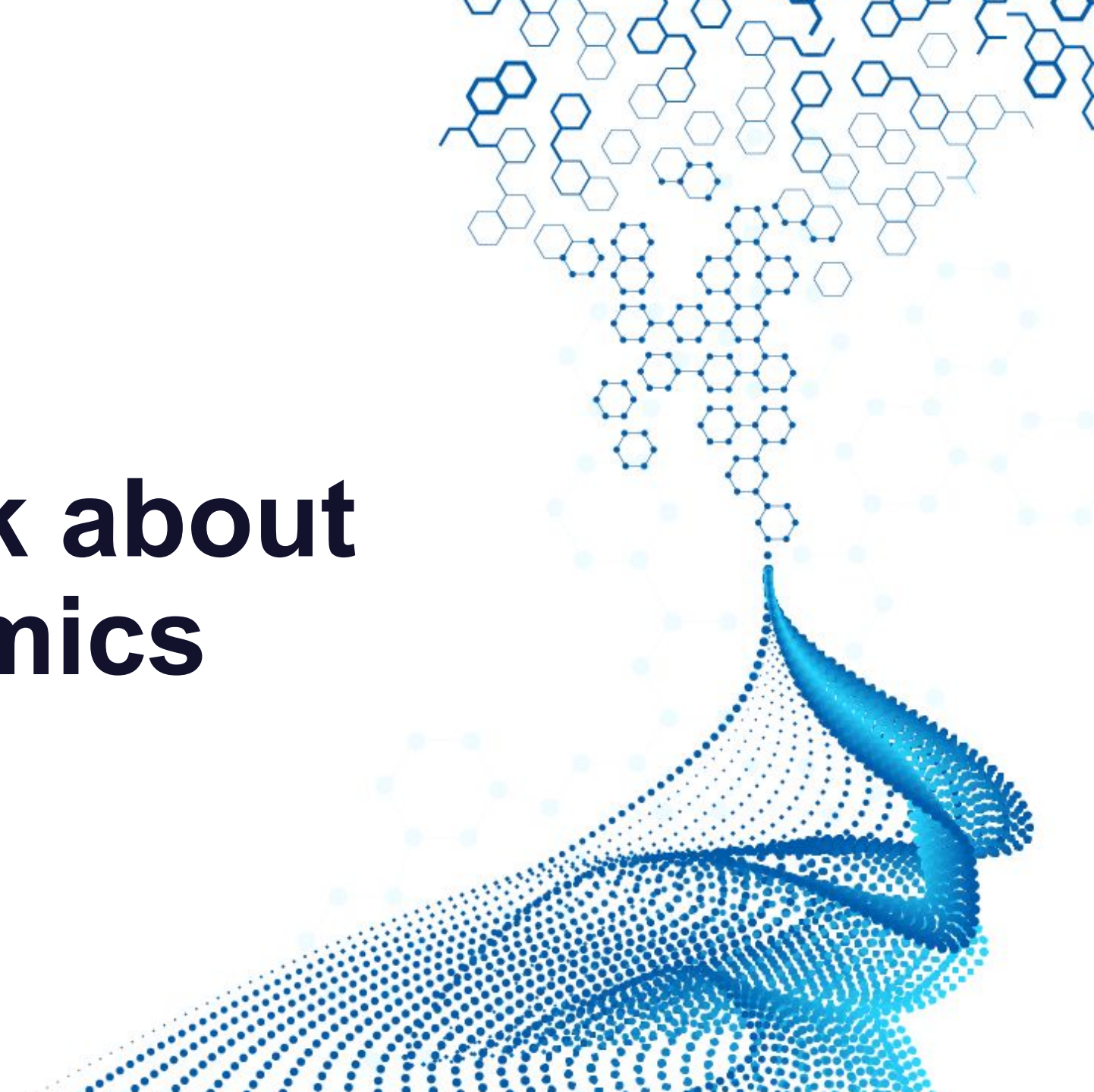
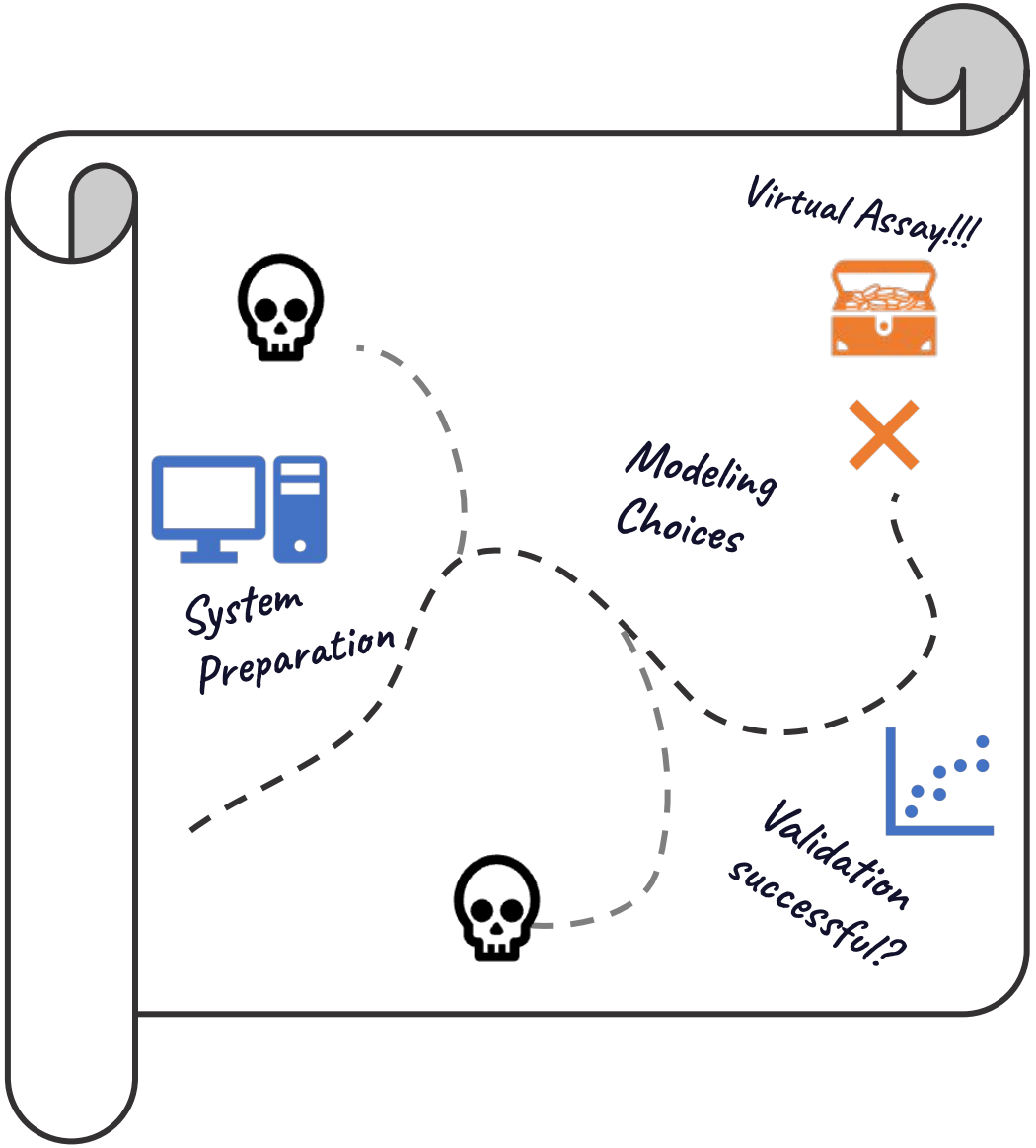
Schrödinger

We need to talk about (Thermo)dynamics



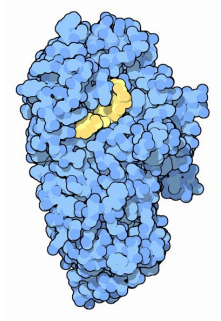
Mila Krämer, Rita Podžuna
LRZ, 2022

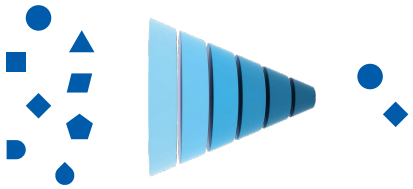
Path to Virtual Assay Is Not Necessarily Straightforward



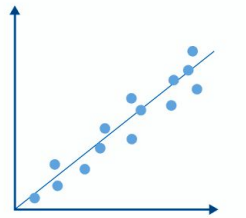
Virtual screening in HitID and LeadOpt

Target selection | Hit identification | Lead optimisation | Candidate selection | Clinical trials

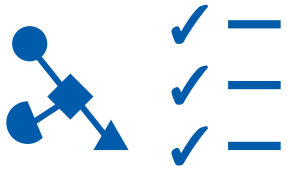




- Aim of virtual screening is to filter down large libraries of diverse compounds
- Requirement is enrichment wrt random screening, and to find diverse hits
- Additional properties are nice to have but less important at this stage

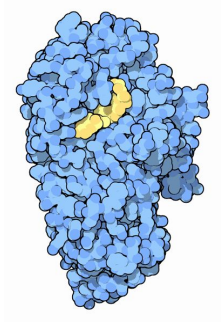


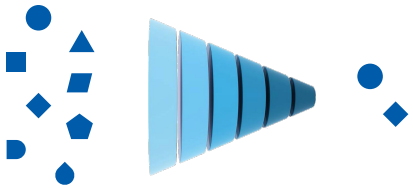
- Aim of virtual screening is to rank order congeneric compound ideas
- Requirement is to accurately predict binding and rank similar compounds
- Compounds have to optimally balance activity and other properties

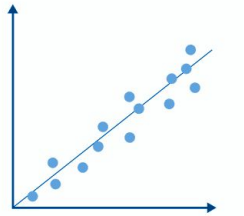


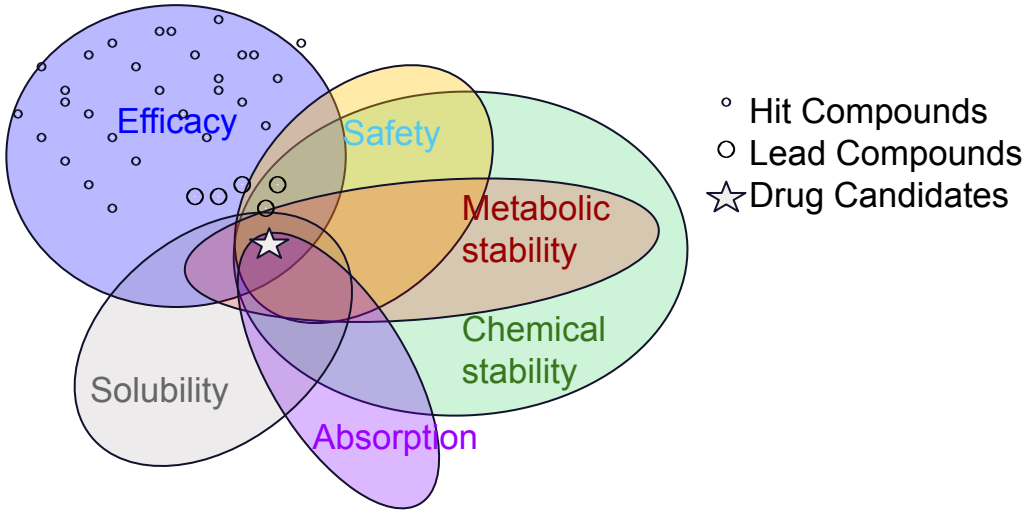
Where does Lead Optimization fit in a project?

Target selection | Hit identification | Lead optimisation | Candidate selection | Clinical trials

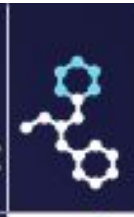
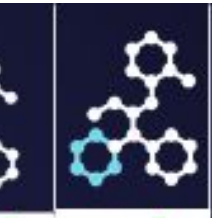


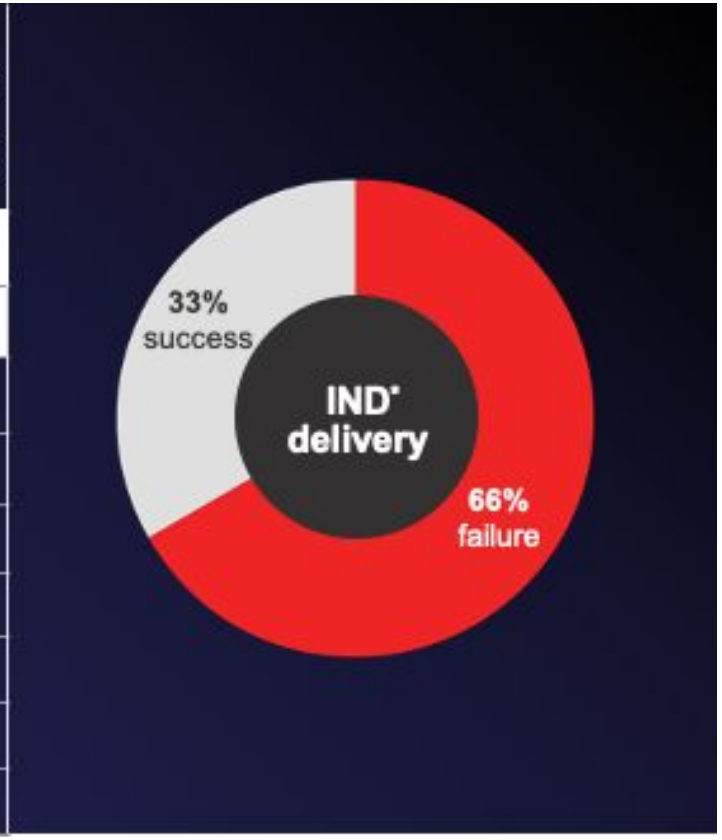






Designing drugs is an extremely hard multi-dimensional optimization problem

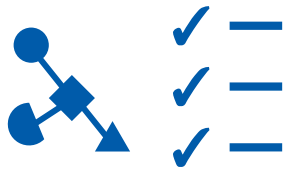
Need to identify a molecule that balances a large number of anti-correlated properties:						
Potency	✓	×	✓	×	✓	✓
Selectivity	×	✓	✓	✓	×	✓
Solubility	×	×	×	✓	✓	×
Bioavailability	×	×	×	×	×	×
Clearance / half-life	×	×	×	×	×	×
Permeability	×	×	×	×	×	×
Drug-drug interactions	×	×	×	×	×	×
Synthesizability	×	×	×	×	×	×
Toxicity	×	×	×	×	×	×



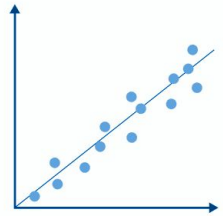
** Based on average, industry-wide success rates*

Lead Optimization is Complex, Slow and Expensive

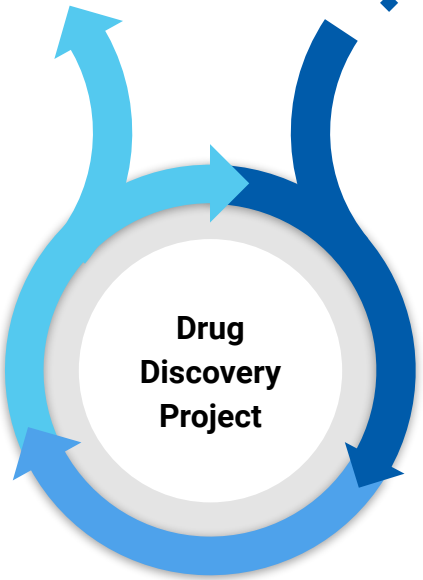
Drug candidate
Compound with optimal property combination



Initial data
Identify hits and/or reference compounds for target



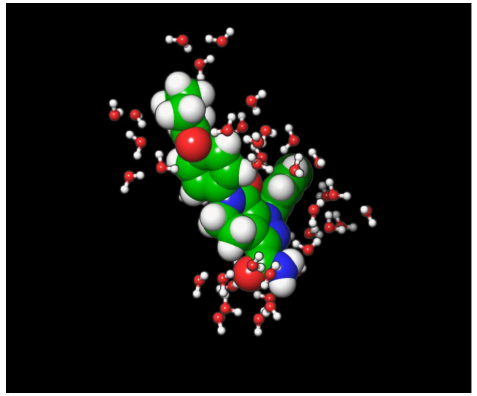
Test
Obtain bioactivity data and SAR for new compounds



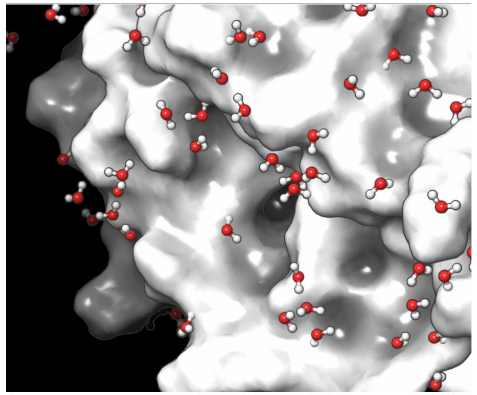
Design
Utilize the currently available data to generate new ideas

Make
Purchase or synthesise new compounds

One way to look at protein-ligand binding



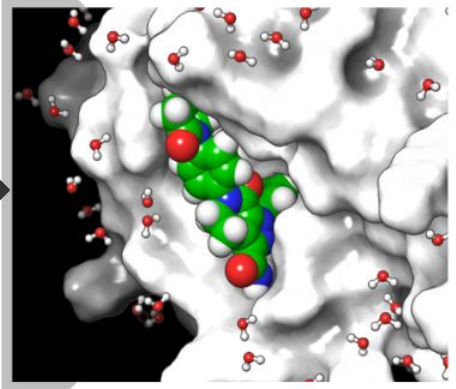
Flexible molecule in water



Flexible protein in water

ΔG_{bind}

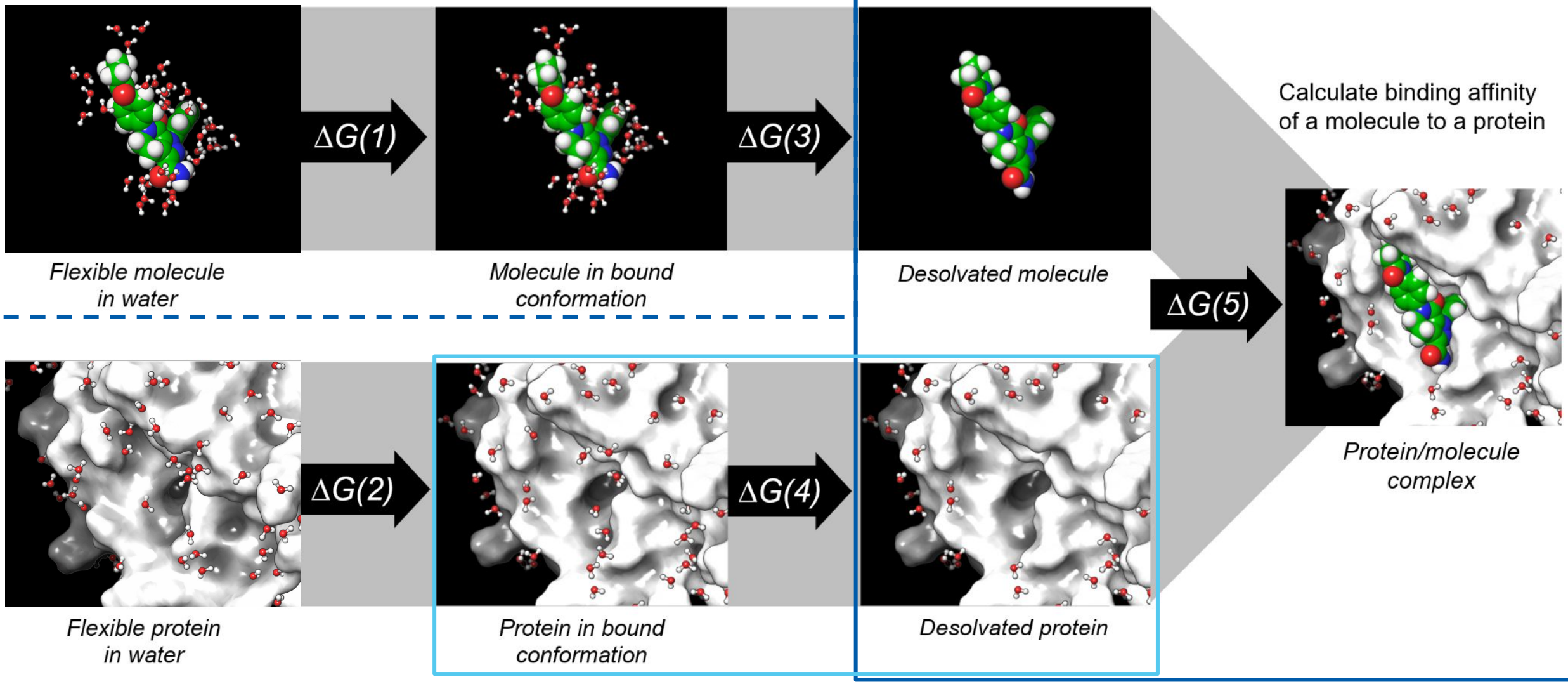
Calculate binding affinity of a molecule to a protein



Protein/molecule complex

One way to look at protein-ligand binding

Docking (Glide)



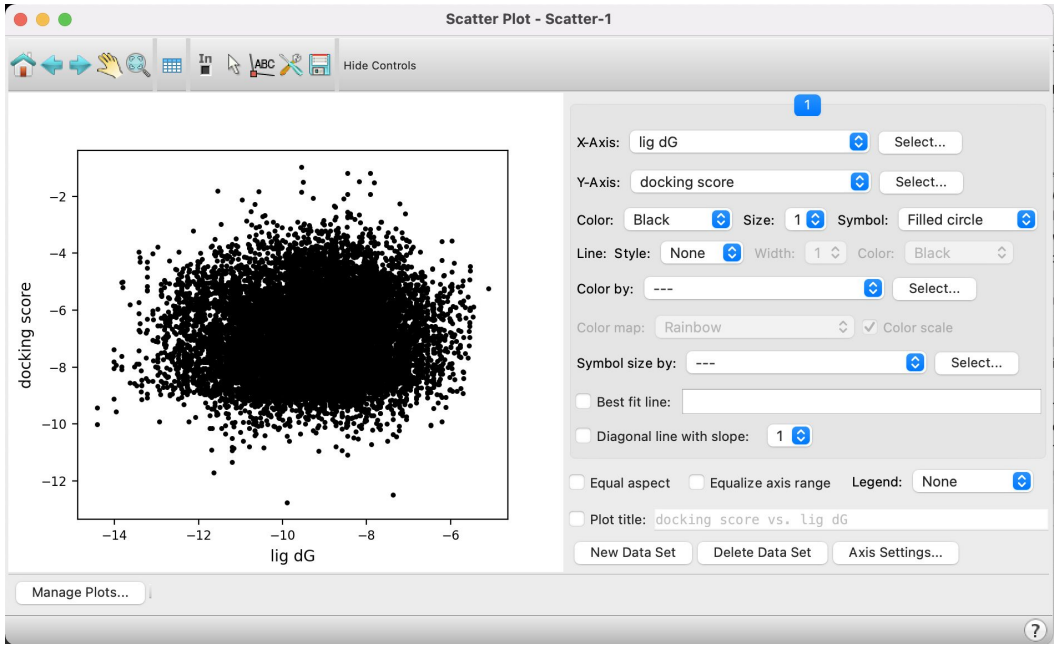
WaterMap

Schrödinger

$\Delta G_{bind} = \Delta G(1) + \Delta G(2) + \Delta G(3) + \Delta G(4) + \Delta G(5)$

Correlation of Docking Results to Binding Affinity

- In many compound sets, there is only a very weak correlation between docking score and experimental binding affinity or none at all.
- Generally docking score can not be used to distinguish between less and more active compounds

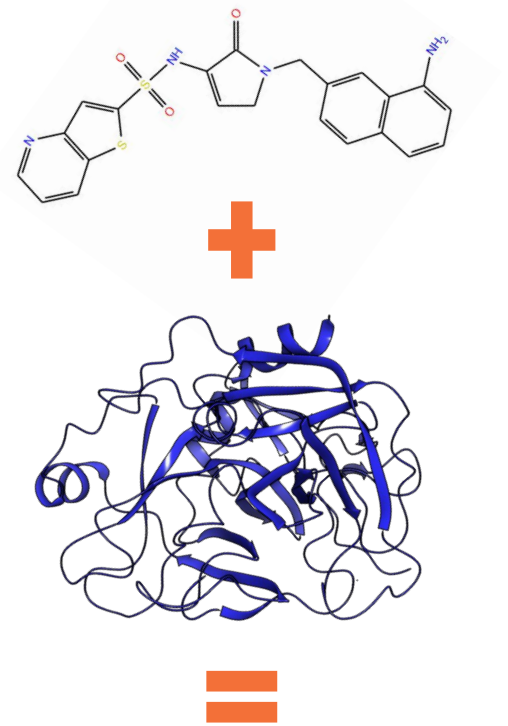


⇒ Reminder: the docking score is parametrized to efficiently distinguish binders from non-binders, not as a proxy for binding affinity

Binding Affinity Prediction from Static Structures

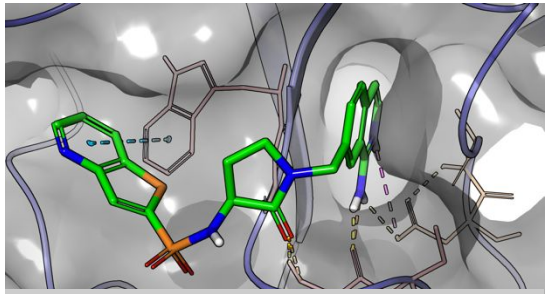
Limitations:

- 3D structure(s):
  - Homology or X-ray (but which co-crystal ligand?)
  - Flexibility of the receptor and ligand often marginally included (sometimes via multiple structures)
  - Experimental conditions might not be reflected by the structure
- Computational model:
  - Implicit solvent models: continuum dielectric models do not reflect the complex effects of microsolvation
  - Force fields: need to reproduce the energy gain upon complex formation but also the relative energy gain upon solvation



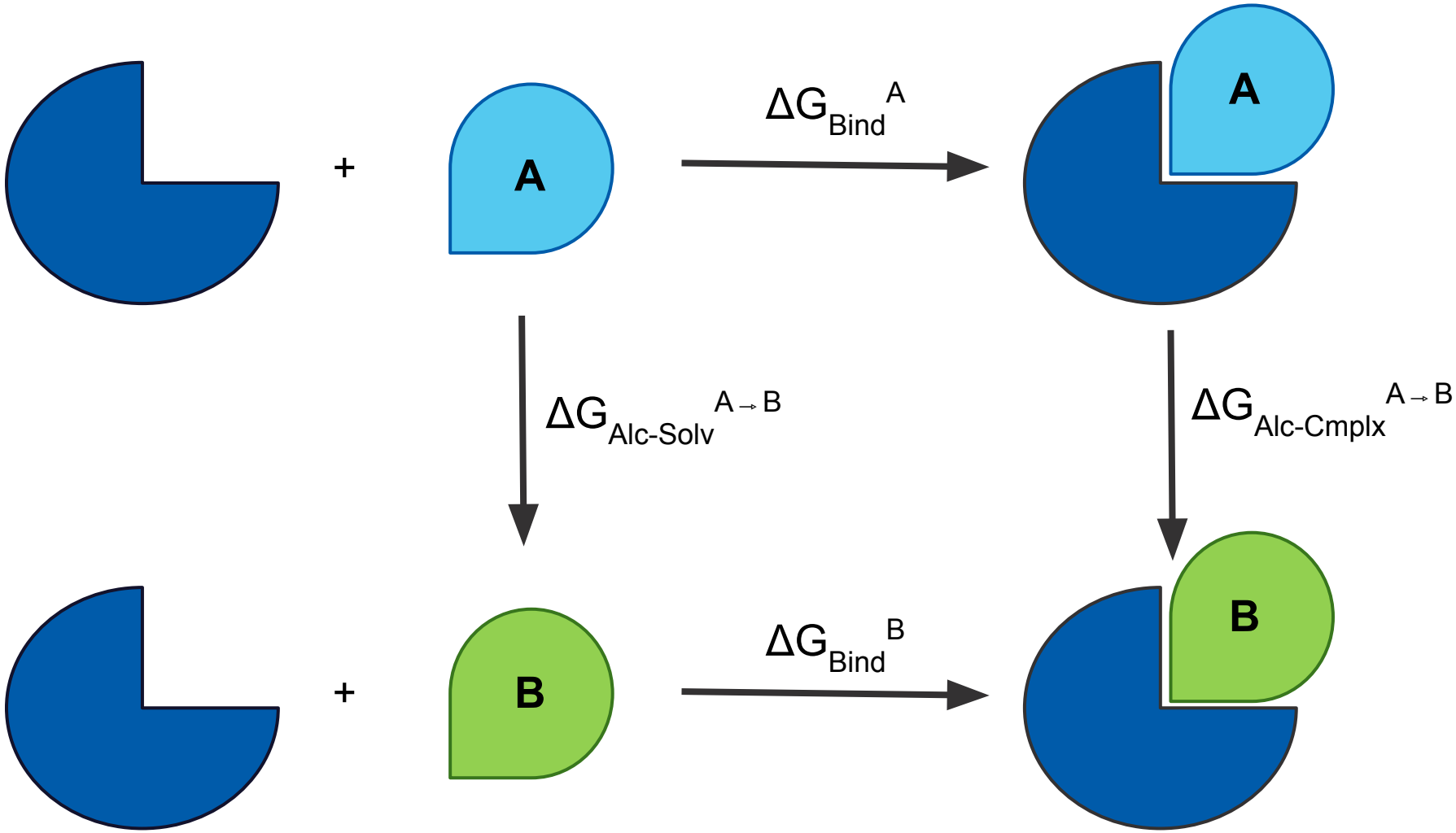
Results:

- Out-of-the-box correlation with experimental data can be poor
- Experience plays a crucial role in correcting these limitations



GlideScore = - 7.42
MMGBSA Score = - 9.01

Free Energy Perturbation - Calculating Relative Binding Energies



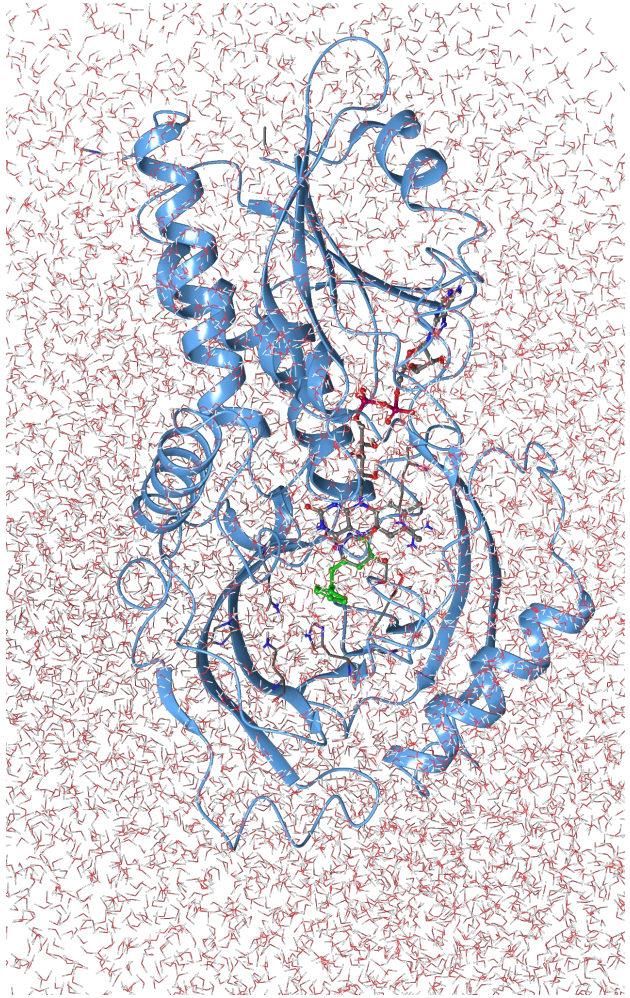
$$\Delta\Delta G_{Bind}^{A \to B} = \underbrace{\Delta G_{Bind}^{B} - \Delta G_{Bind}^{A}}_{\text{hard but interesting}} = \underbrace{\Delta G_{Alc\text{-}Cmplx}^{A \to B} - \Delta G_{Alc\text{-}Solv}^{A \to B}}_{\text{simpler to calculate now}}$$

Understanding Hydration

using WaterMap

Reminder: Why is Water Important?

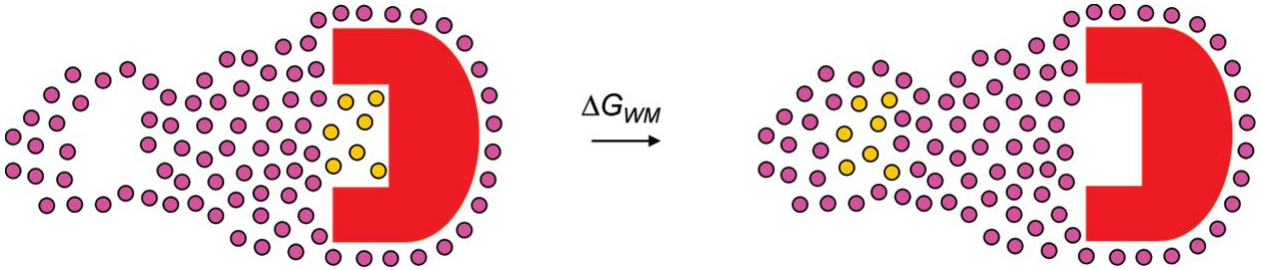
- Water is everywhere in biology
- “Empty” binding sites are mostly filled with water
- Ligands must displace that water to bind



⇒ Water energetics can drive potency, but can’t be obtained from static structures

What do ΔG, ΔH and -TΔS correspond to?

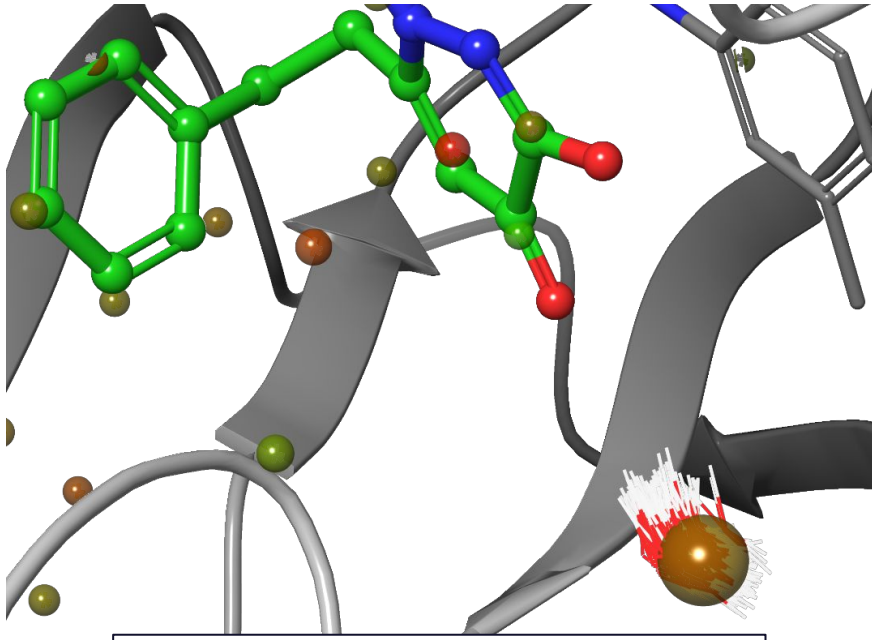
- The values calculated by WaterMap correspond to the average excess enthalpy, entropy and free-energy that a water molecule, located at the hydration site, would possess relative to bulk water



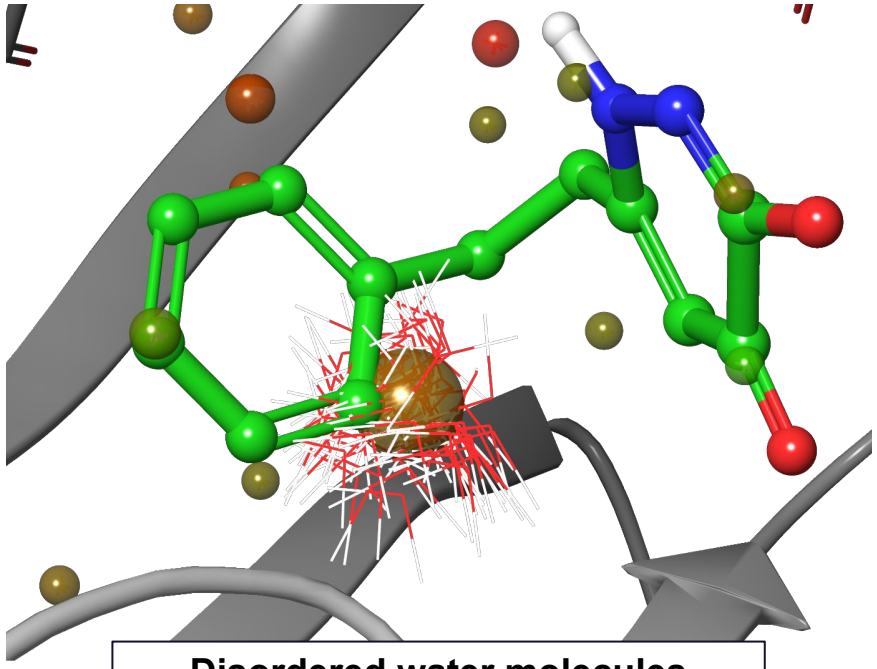
- This means that:
  - A hydration site with a negative ΔH-value is making stronger interactions with the surrounding protein than it would with surrounding water molecules in solution e.g. near a charged group
  - A hydration site with a positive ΔH-value is making weaker interactions with the surrounding protein than it would with surrounding water molecules in solution e.g. near a hydrophobic residue

What do ΔG, ΔH and -TΔS correspond to?

- The offensive mathematics is just quantifying the ‘randomness’ of the water molecules at each hydration site



Highly ordered water molecules
Unfavourable entropy
-TΔS=5.29 kcal/mol

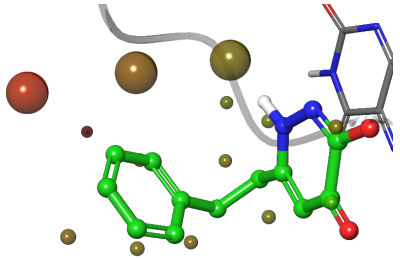


Disordered water molecules
Favourable entropy
-TΔS=0.71 kcal/mol

Analyzing WaterMap Results

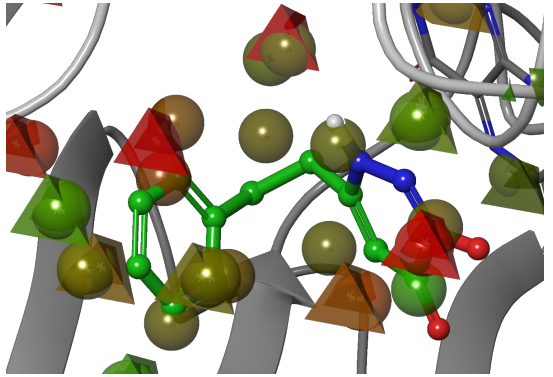
show waters occupying hydration sites

select sites in workspace to highlight them and select corresponding rows in table



filter sites by properties

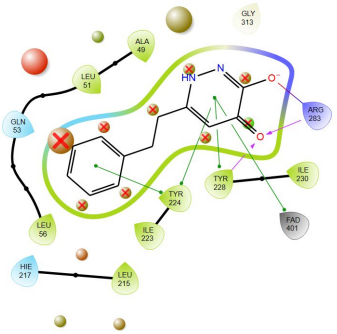
useful for comparing multiple WaterMaps:



Comparing apo (spheres) vs holo (pyramids) WaterMap highlights waters displaced by ligand

visualize water density, cavities and continuous WaterMap

2D view



import raw MD simulation data (rarely necessary)

reset panel to return to blank slate

WaterMap - Examine Results

Load results: Import Files... | Analyze Workspace | ☐ Adjust the view when analyzing

Display: ☑ Receptor ☑ Ligand ☑ H-bonds ☐ Markers ☑ Shape by entry ☐ Dipoles

☐ Waters ☐ Ligand surface ☑ Water density ☐ Cavity map ☐ Free energy density

Site label: None

Color sites by: ΔH and -TΔS | Color scale: ○ Relative ◉ Absolute ☐ Show color in table

Site	Occupancy	Overlap	ΔH	-TΔS	ΔG	#HB(WW)	#HB(PW)	#HB(LW)
47	0.29	0.00	-0.52	0.59	0.07	2.70	0.00	0.00
45	0.31	0.00	1.05	0.70	1.75	2.48	0.00	0.00
42	0.35	0.00	0.19	0.71	0.90	2.43	0.00	0.00
46	0.31	1.00	2.56	0.71	3.27	2.19	0.00	0.00
43	0.34	0.00	0.19	0.75	0.94	2.44	0.02	0.00

☐ Pick to select sites | Select All | ☐ Show Only Selected Rows | Export to CSV...

Show only water sites with:

☐ Enthalpy (ΔH) -4.43 3.55 kcal/mol ☐ Invert range

☐ Entropy (-TΔS) 0.59 5.29 kcal/mol ☐ Invert range

☐ Free energy (ΔG) -0.22 8.08 kcal/mol ☐ Invert range

☐ Overlap factor 0.00 1.00 ☐ Invert range

☐ Distance 10.0 Å of entry.id 13 Select...

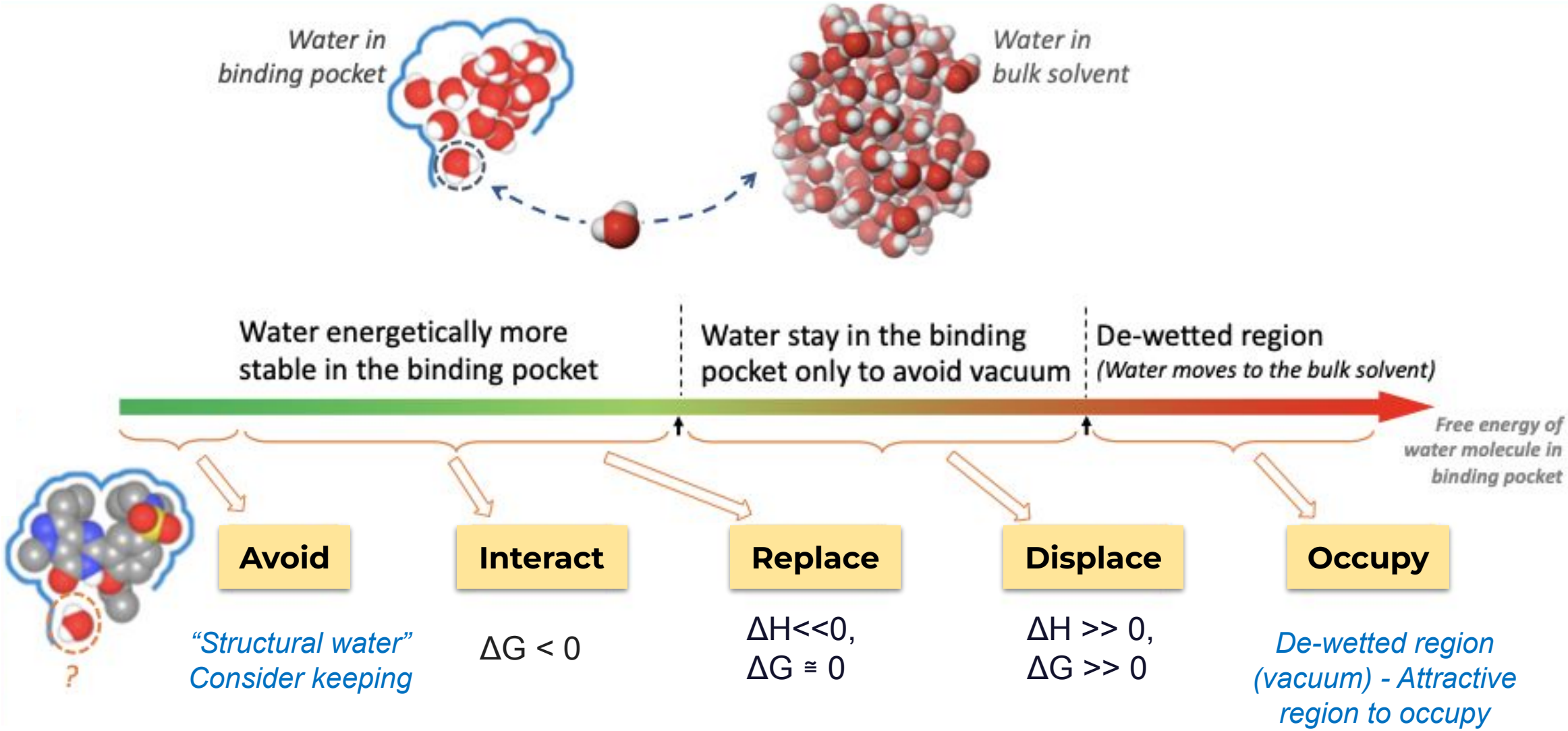
Score Ligands | ☐ Create properties for selected sites | ☐ Score only selected sites

Perform WM/MM Scoring...

Import Trajectory | Import Water Frames | Archive... | Export to LiveDesign... | Open Ligand Interaction Diagram...

Reset

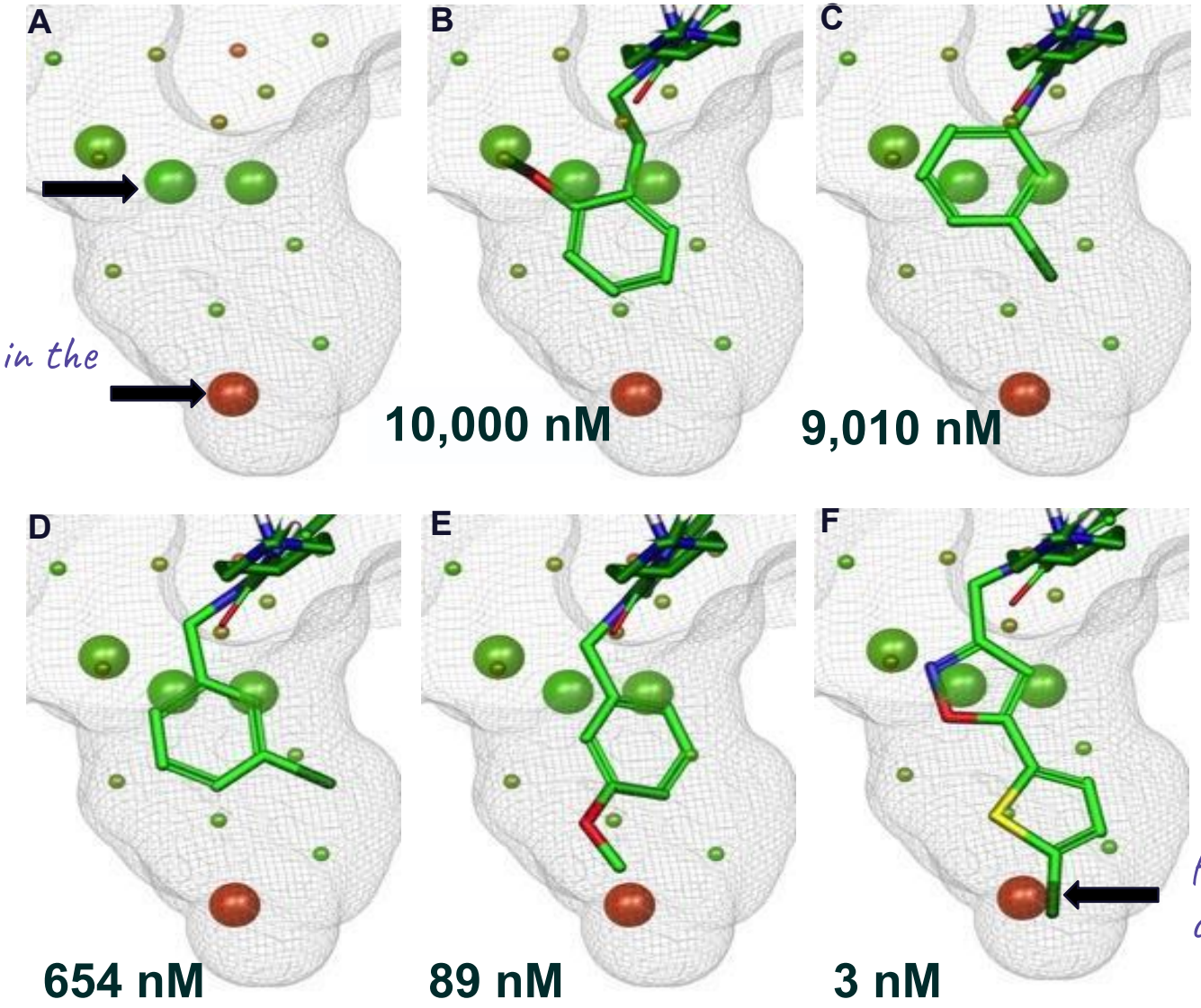
How can Understanding Hydration Guide Strategy?



Mapping out where water molecules are can guide and explain SAR

3 stable waters with favorable enthalpy

most unstable water in the binding site



full displacement of unstable water corresponds to tightest binder

Thank You!

mila.kraemer@schrodinger.com