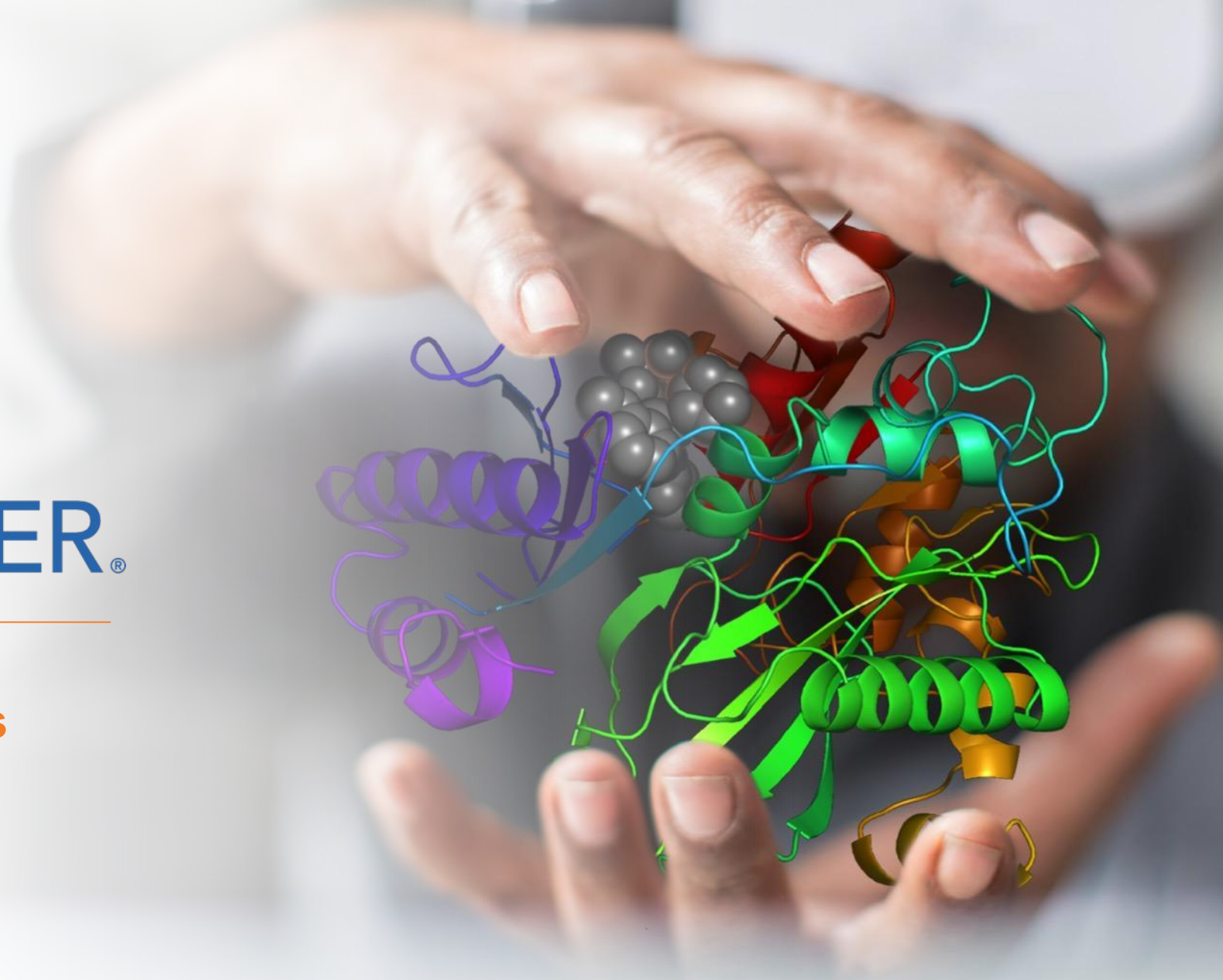
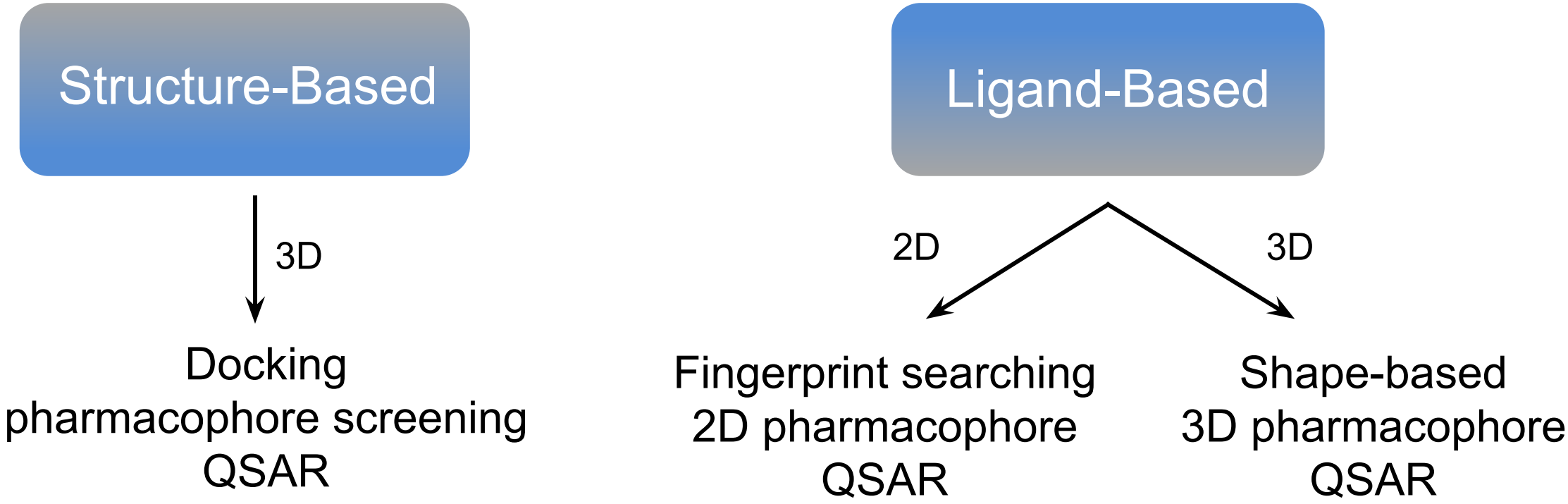
SCHRÖDINGER®

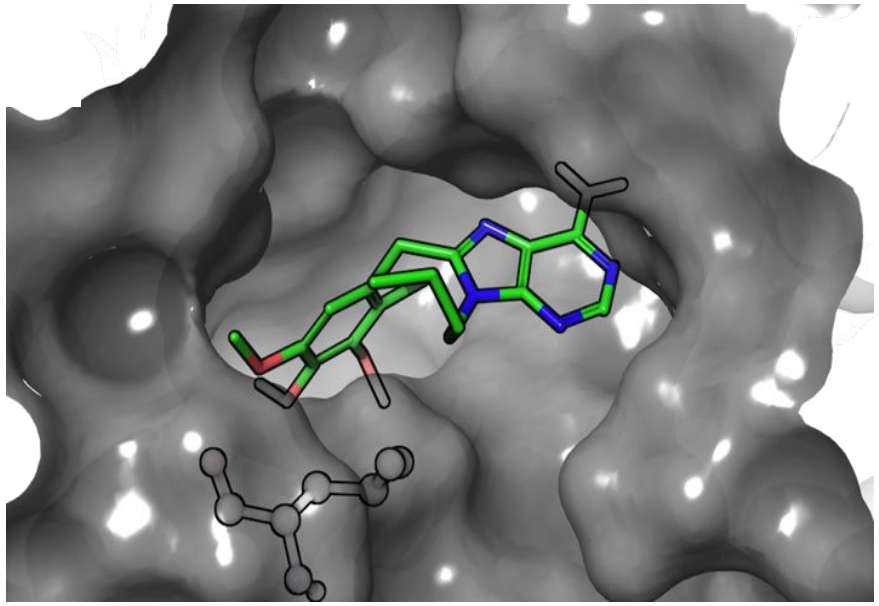
Ligand-based Virtual Screening Methods



Virtual screening can quickly evaluate ligand libraries

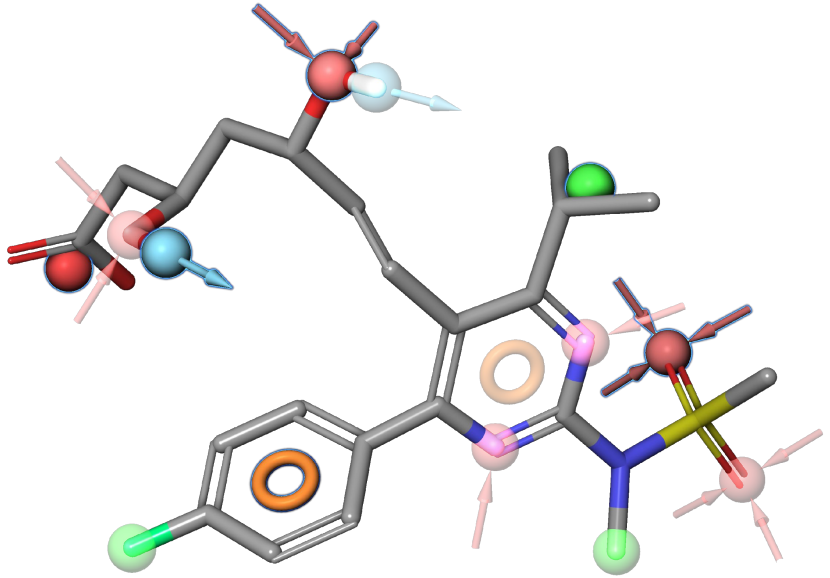


Use different screening methods for different available data



Structure-based virtual screens

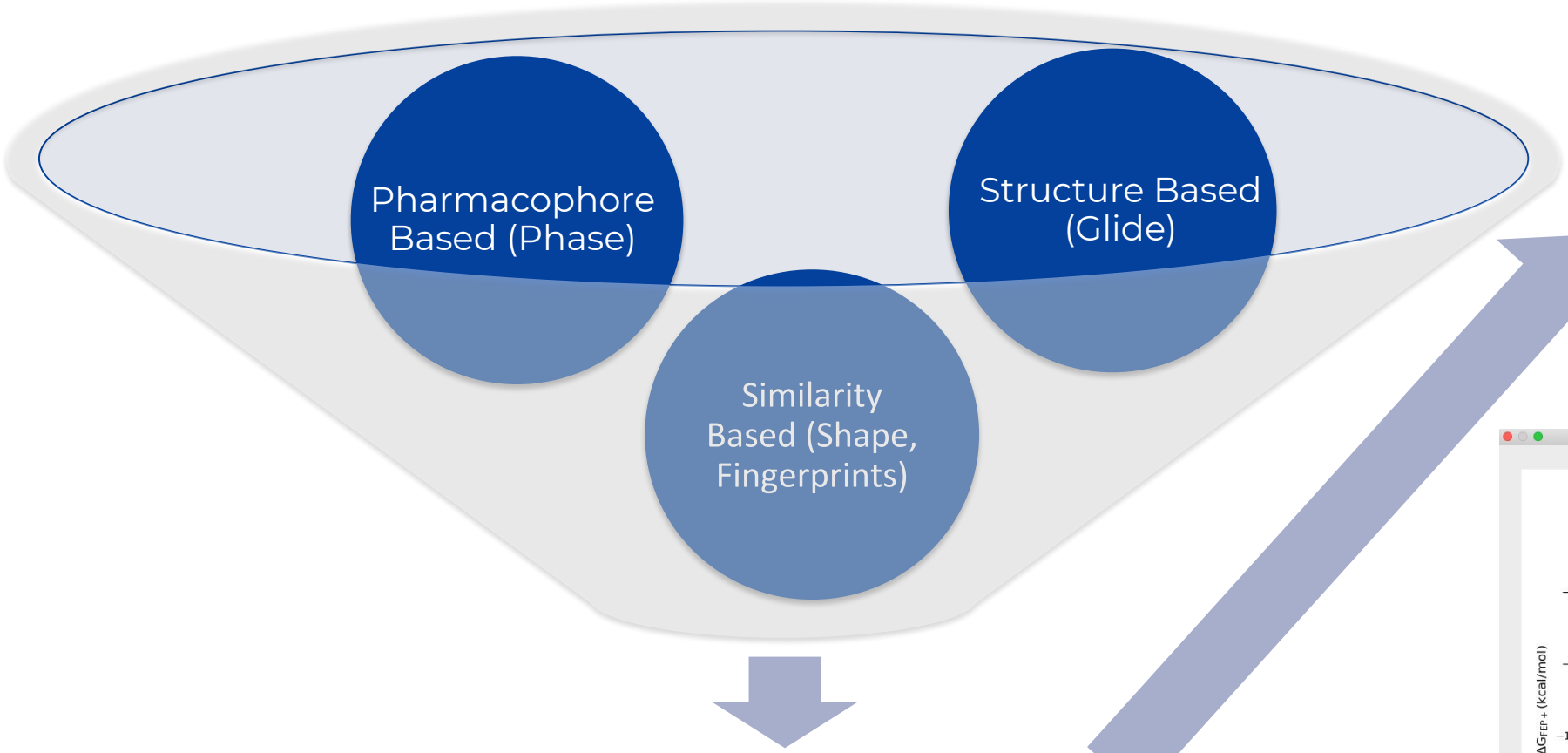
- Structure of Target
- Ligand-binding site known
- (optional) ligand/hit bound



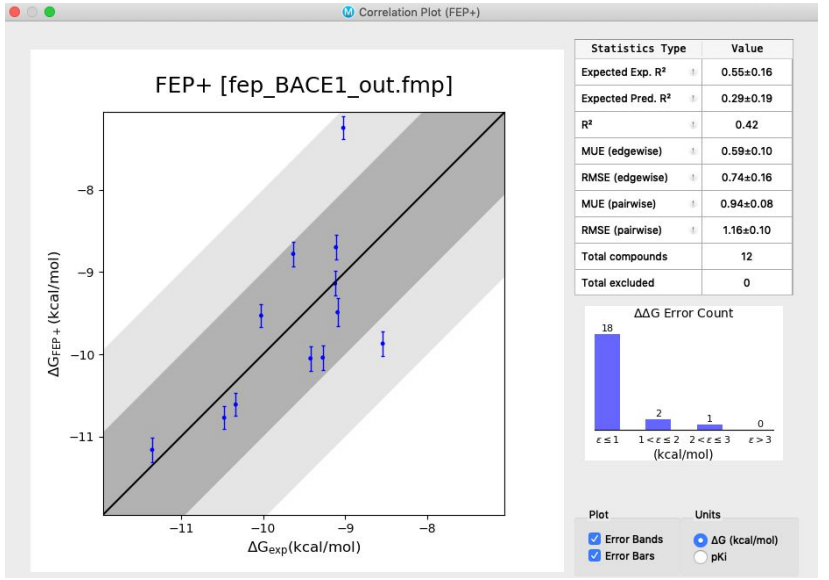
Ligand-based virtual screens

- a single known hit
- multiple known hits
- (optional) active conformation

Virtual screen with many tools for best results



- MM-GBSA
- AB-FEP
- FEP+

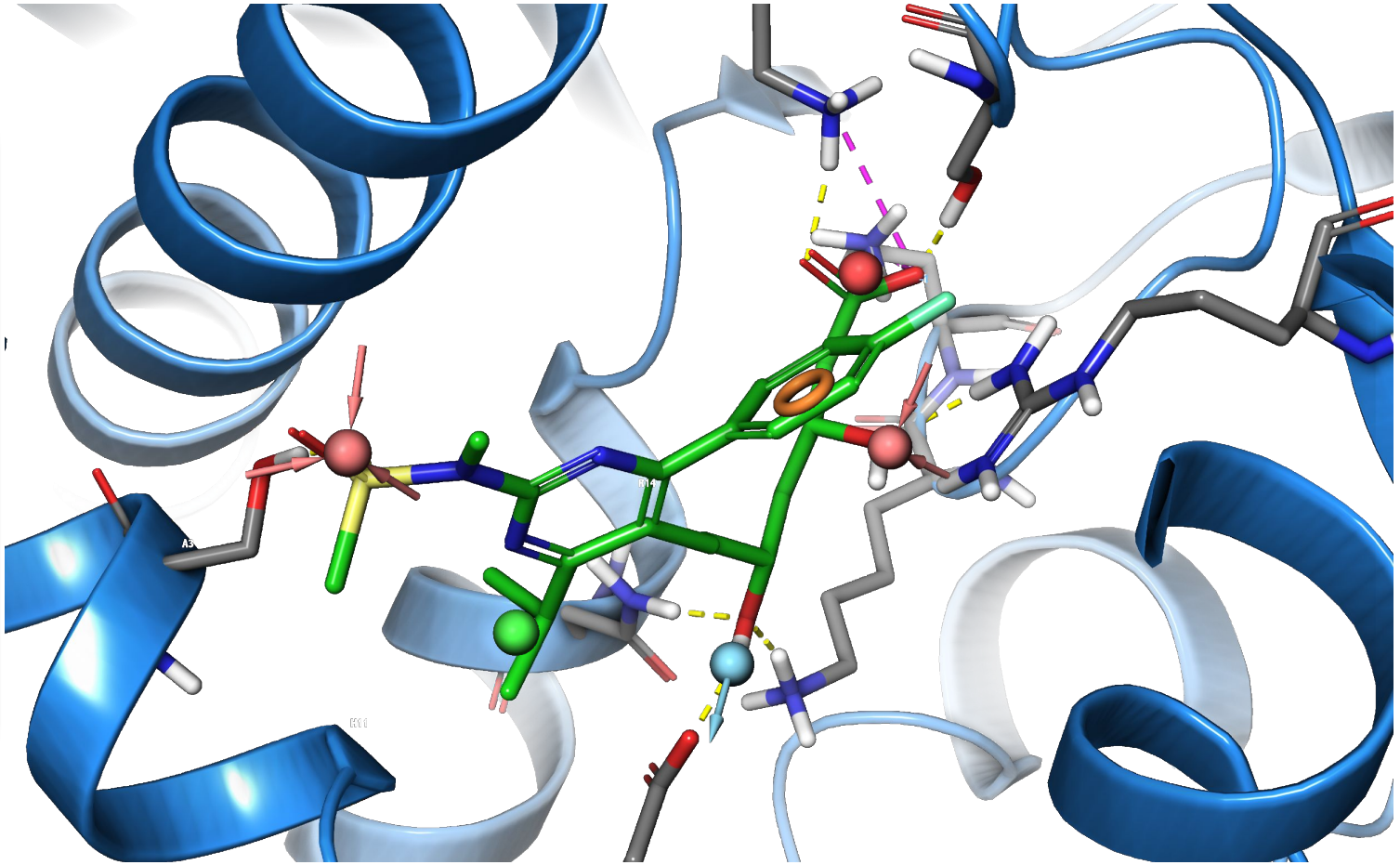


Screening Results

A pharmacophore is an abstract representation of interactions

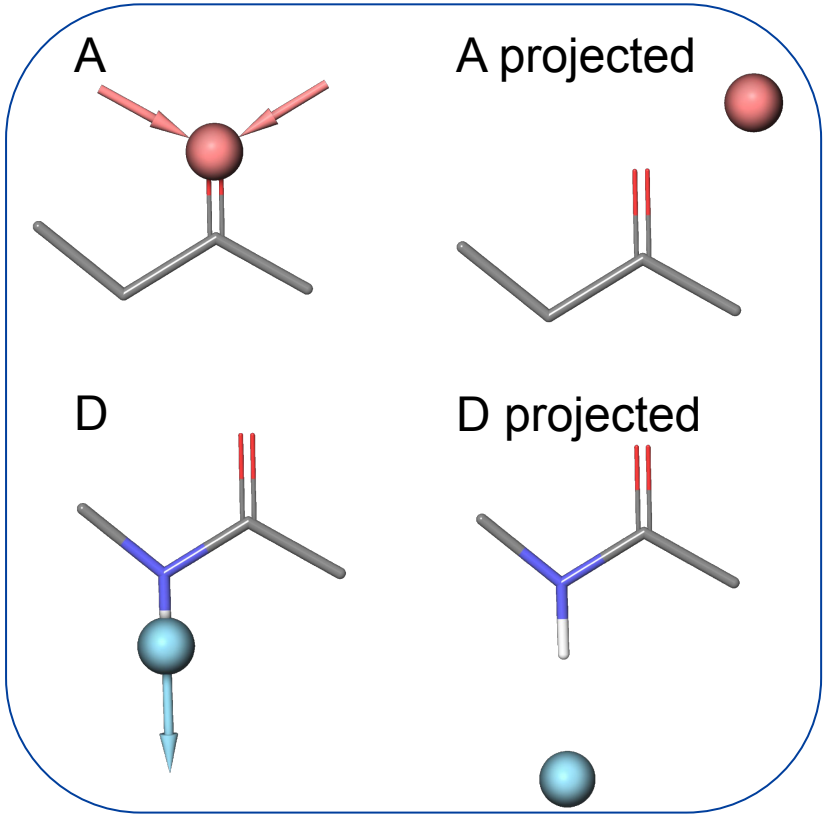
A pharmacophore is the ensemble of steric and electronic features that is necessary to ensure the optimal supramolecular interactions with a specific biological target and to trigger (or block) its biological response

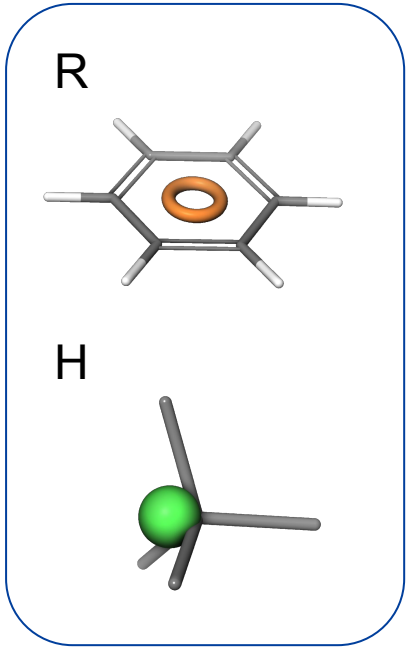
-IUPAC

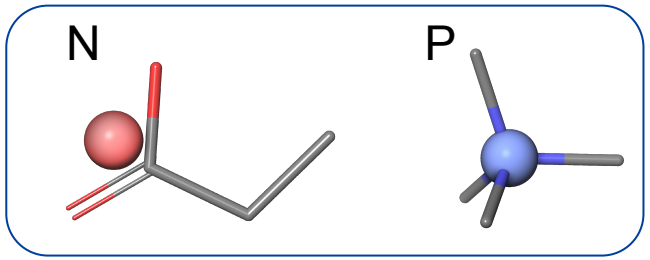


Pure Appl. Chem. 1998; 70: 1129-1143.

Pharmacophores are represented using features

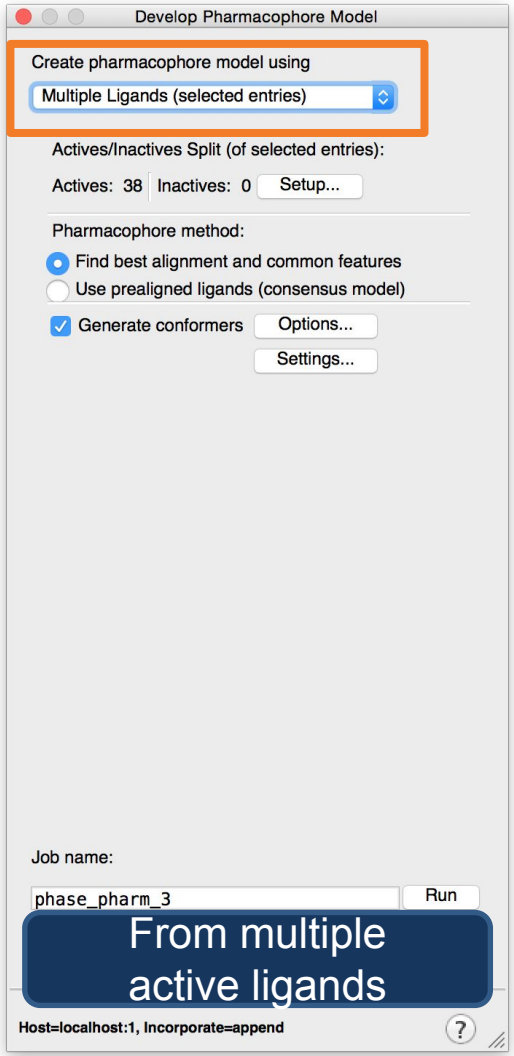




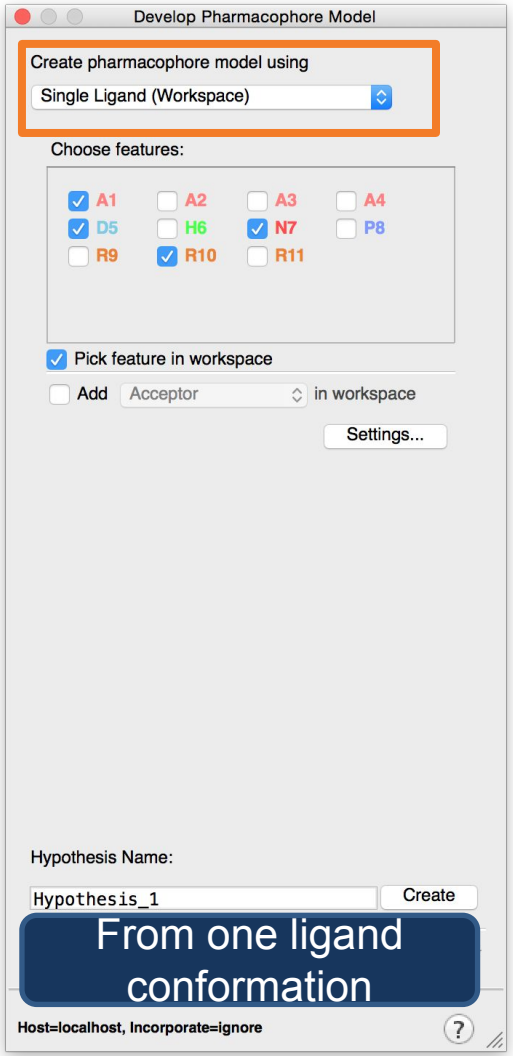


- Common features
  - D for h-bond donor
  - A for h-bond acceptor
  - R for aromatic ring
  - H for hydrophobic
  - N for negative ionic
  - P for positive ionic
- D/A/R features have vector characteristics
- D/A features can be treated as projected points or vectors
- Can customize features

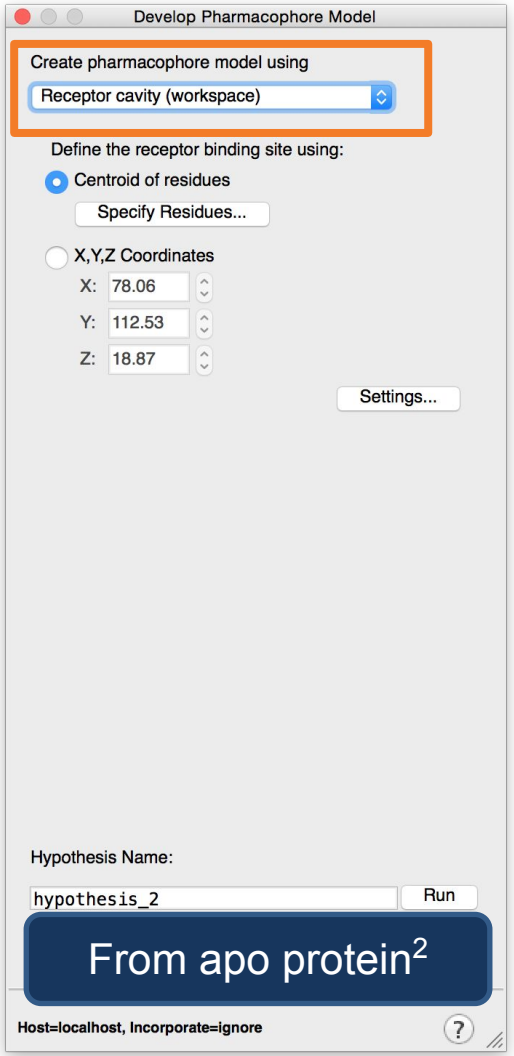
There are several ways to create pharmacophore hypotheses



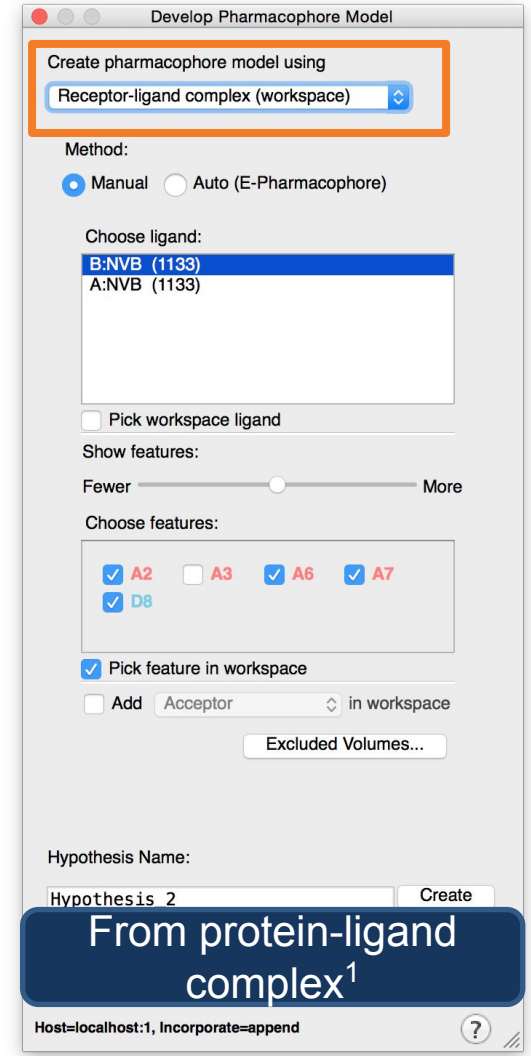
From multiple active ligands



From one ligand conformation



From apo protein[2]

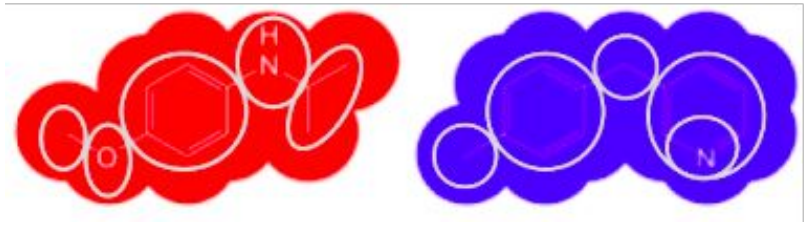


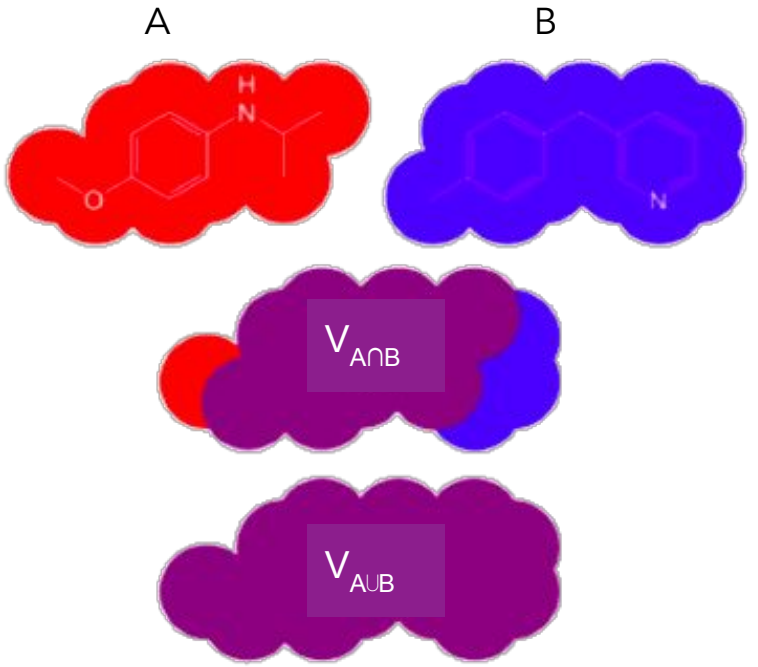
From protein-ligand complex[1]

[1]Salam, N., Nuti, R., Sherman, W., *J Chem Inf Model*, 2009, 49(10), 2356.
[2]Loving, K., Salam, N., Sherman, W., *J Comput-Aid Design*, 2009, 23(8), 541.

Shape screening makes use of known binders

- Hard sphere volume overlaps are used for similarity assessment
- Incorporate property information by allowing only atoms with similar types to match
- Sphere types
  - Heavy atom (shape only)
  - Heavy atom by element/atomtype (color)
  - Pharmacophore feature-located (color)
    - Hydrophobic Aromatic
    - Acceptor Donor
    - Negative Positive

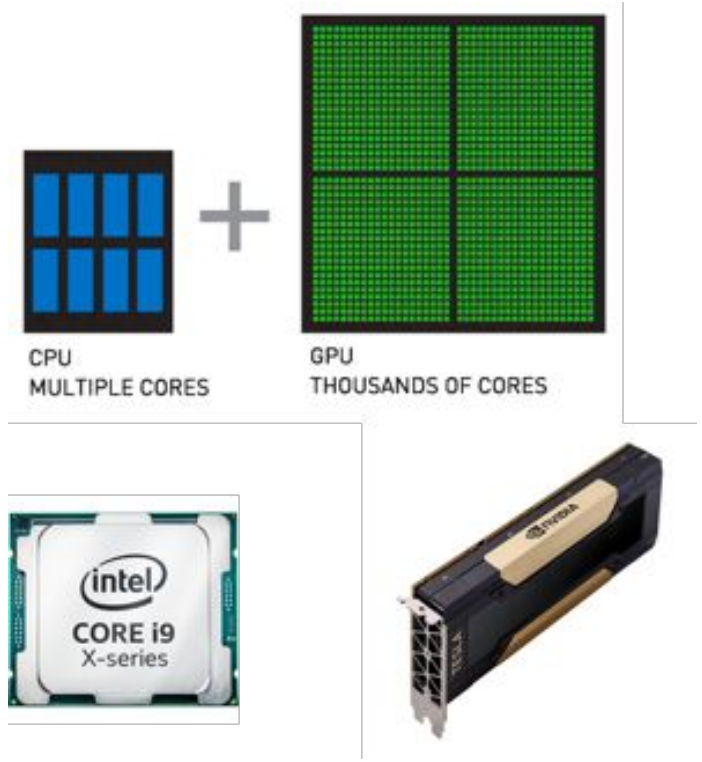




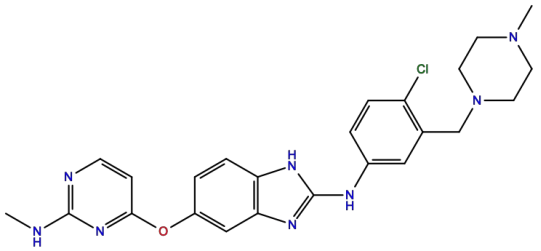
CPU and GPU Shape screening are both useful

- Shape algorithm is primarily linear algebra
  - Same mathematical operations performed many times across different datasets
  - Small amount of data required per conformer compared

CPU	GPU
Small number of faster cores	Large number of slower cores
Works on a variety of tasks simultaneously	Optimized for taking huge batches of data and performing the same operation over and over very quickly



Generalized shape screening workflow



Identify which probe(s)/active compound(s) you would like to use to build the shape

Are the active conformation and tautomeric states of your hit molecule(s) known?

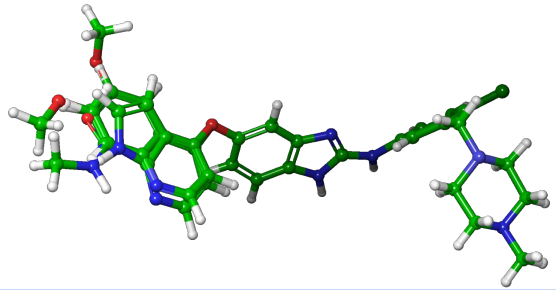
Yes

Align the conformations of the active/probe molecule(s).

No

Use LigPrep and Epik to generate the lowest energy ligand states.

Choose a ligand library to screen against the shape(s) created from the probe(s).



CPU or GPU-based screen?

CPU

GPU

Create shape data file.

Perform Shape Screen!

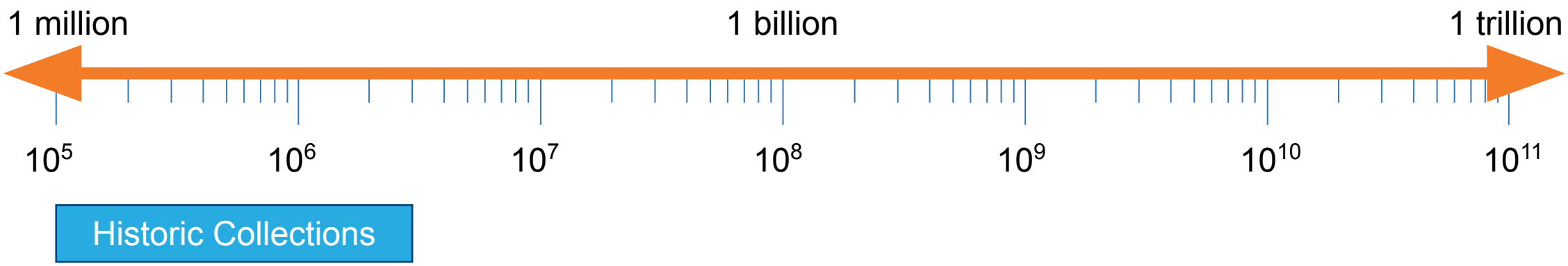
Align the conformations of the active/probe molecule(s).

There are speed limitations to using 3D-based screening at scale

Method	Timings	Time to process 500M compounds on a single CPU or GPU	CPUs required to process 500M compounds in 7 days
Docking*	~10 sec/ligand	~158 years	~8200
Pharmacophore*	~5 ligands/sec @ 50 confs each	~3 years	~165
CPU Shape	~200 ligands/sec @ 10 confs each	~1 month	~4
GPU Shape	**~10,500 ligands/sec @ 10 confs each**	**13 hours**	**NA**

*Necessary Ionization/tautomeric expansion typically expands number of structures to be screened by 2.5X

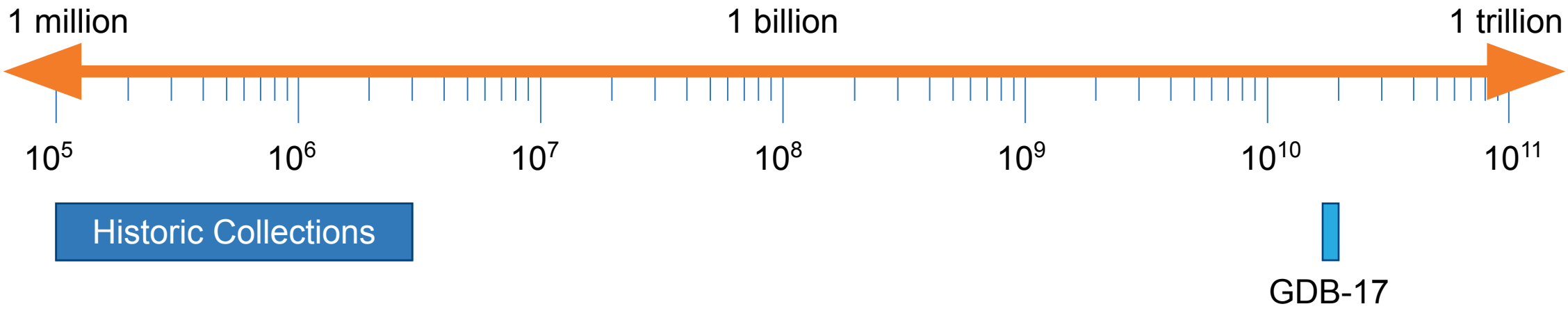
Real & Virtual chemical screening libraries are growing substantially



Historically compound collections were comprised of already synthesized molecules with < 20 million compounds

J. Chem. Inf. Model., 2012, 52(11), 2864-2875.

Real & Virtual chemical screening libraries are growing substantially

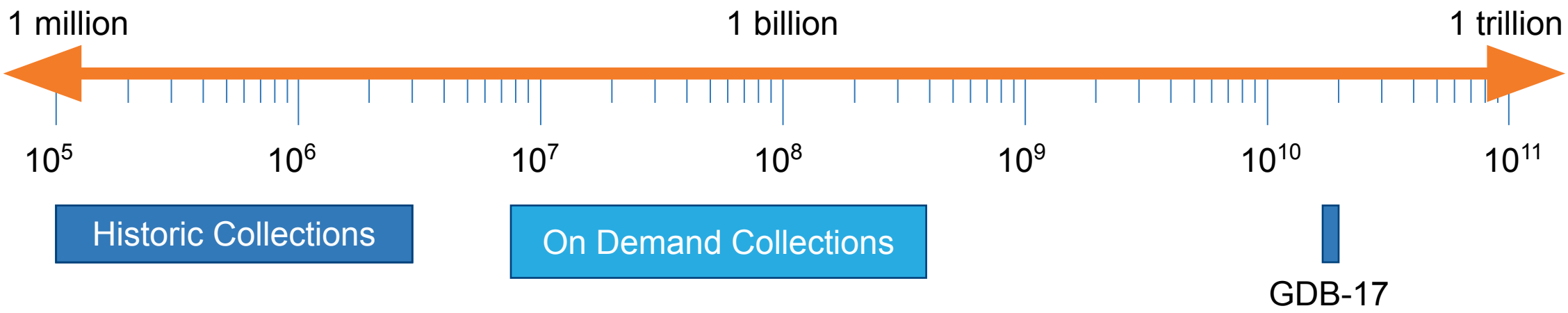


GDB-17* with 166.4 billion compounds

All molecules with up to 17 atoms containing C, N, O, S, and halogen atoms

J. Chem. Inf. Model., 2012, 52(11), 2864-2875.

Real & Virtual chemical screening libraries are growing substantially

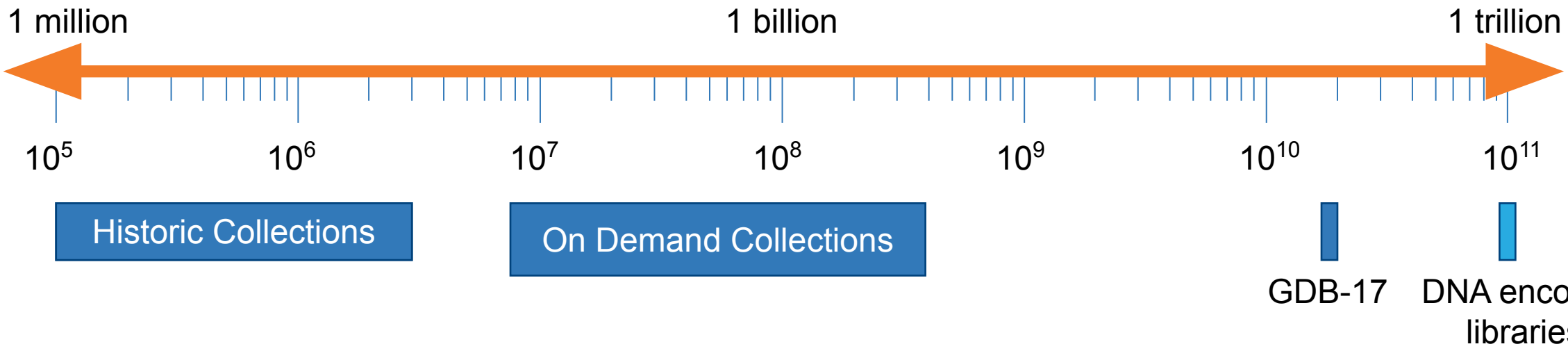


Diverse, synthetically-tractable libraries are becoming common

- Multiple vendors with >100M synthesizable libraries
- 720M - Enamine REAL set (readily accessible)
  - Synthesis success rates of 85% or higher reported

J. Chem. Inf. Model., 2012, 52(11), 2864-2875.

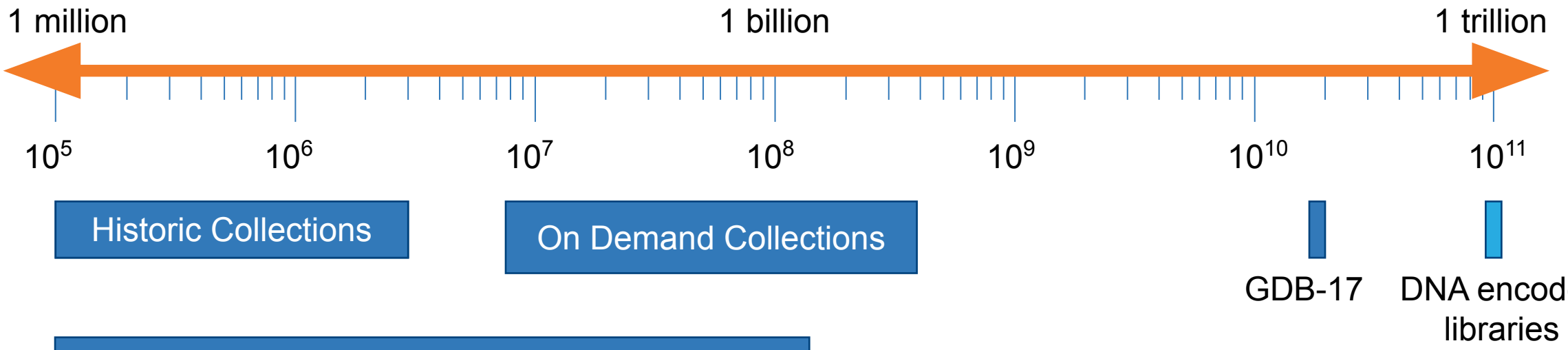
Real & Virtual chemical screening libraries are growing substantially



DNA encoded libraries of potentially trillions of compounds created combinatorially in one flask

J. Chem. Inf. Model., 2012, 52(11), 2864-2875.

Chemical screening libraries are growing substantially



DNA encoded libraries of potentially trillions of compounds created combinatorially in one flask

J. Chem. Inf. Model., 2012, 52(11), 2864-2875.

CADD can explore large chemical space, which can lead to new hits

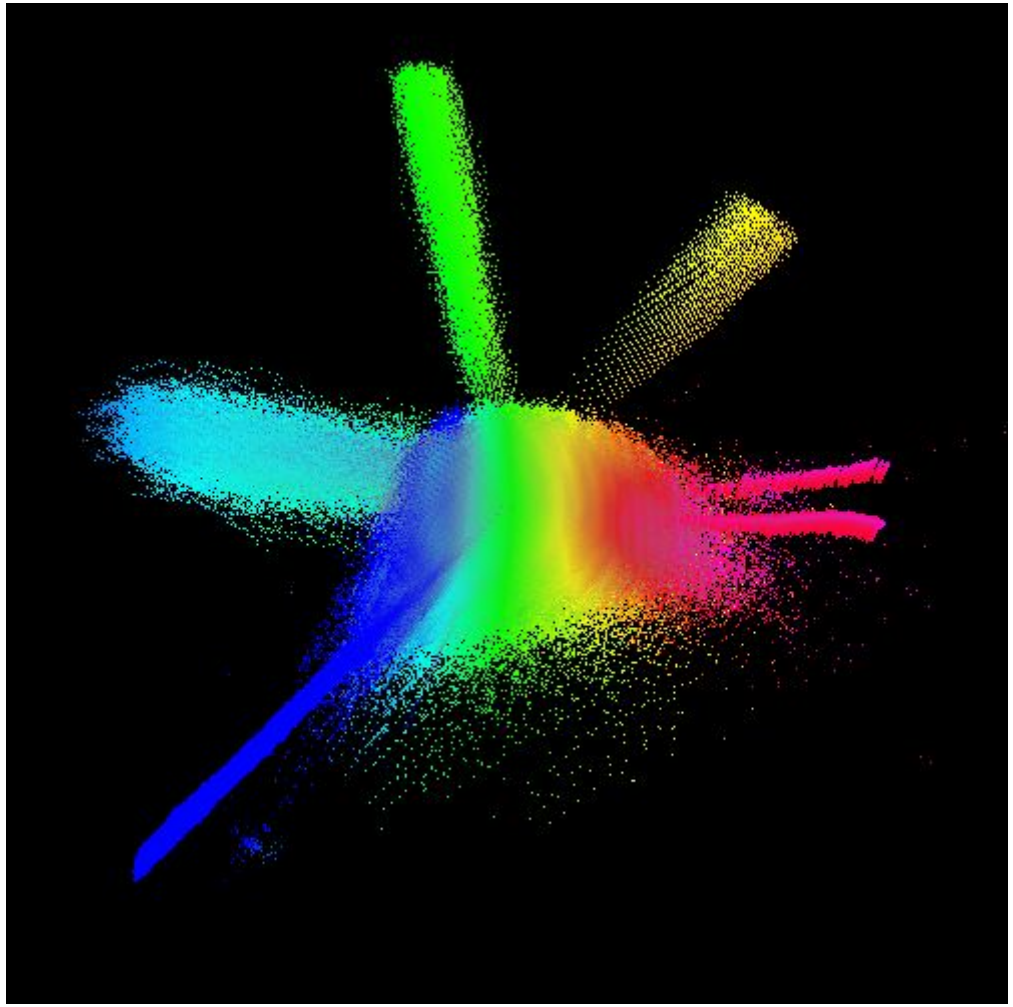


Image from: https://en.wikipedia.org/wiki/Chemical_space

- 19.2 million structures
- the further away to points are, the more chemically distinct they are

Shoichet Hypothesis

- Chemical diversity increases with larger chemical space searching
- Shoichet hypothesis: As screening decks expand, more tight binders and tighter binders found

Reymond, A.L. and Awale M. *ACS Chem. Neurosci.* 2012, *3*(9), 649–657.
Shoichet, et al. *Nature* 2019, *566*, 224–229.

Questions? Email us at online-learning@schrodinger.com