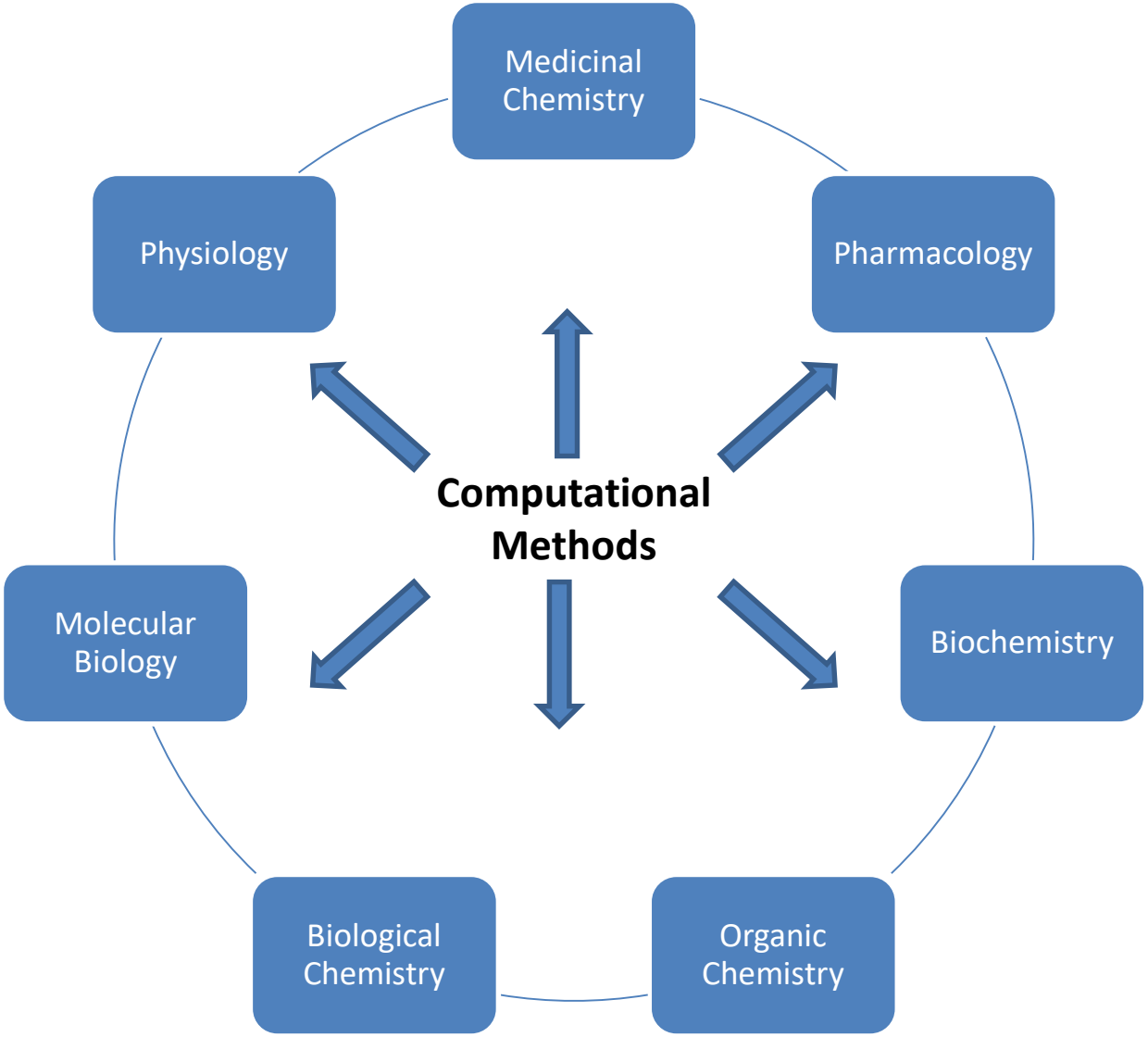
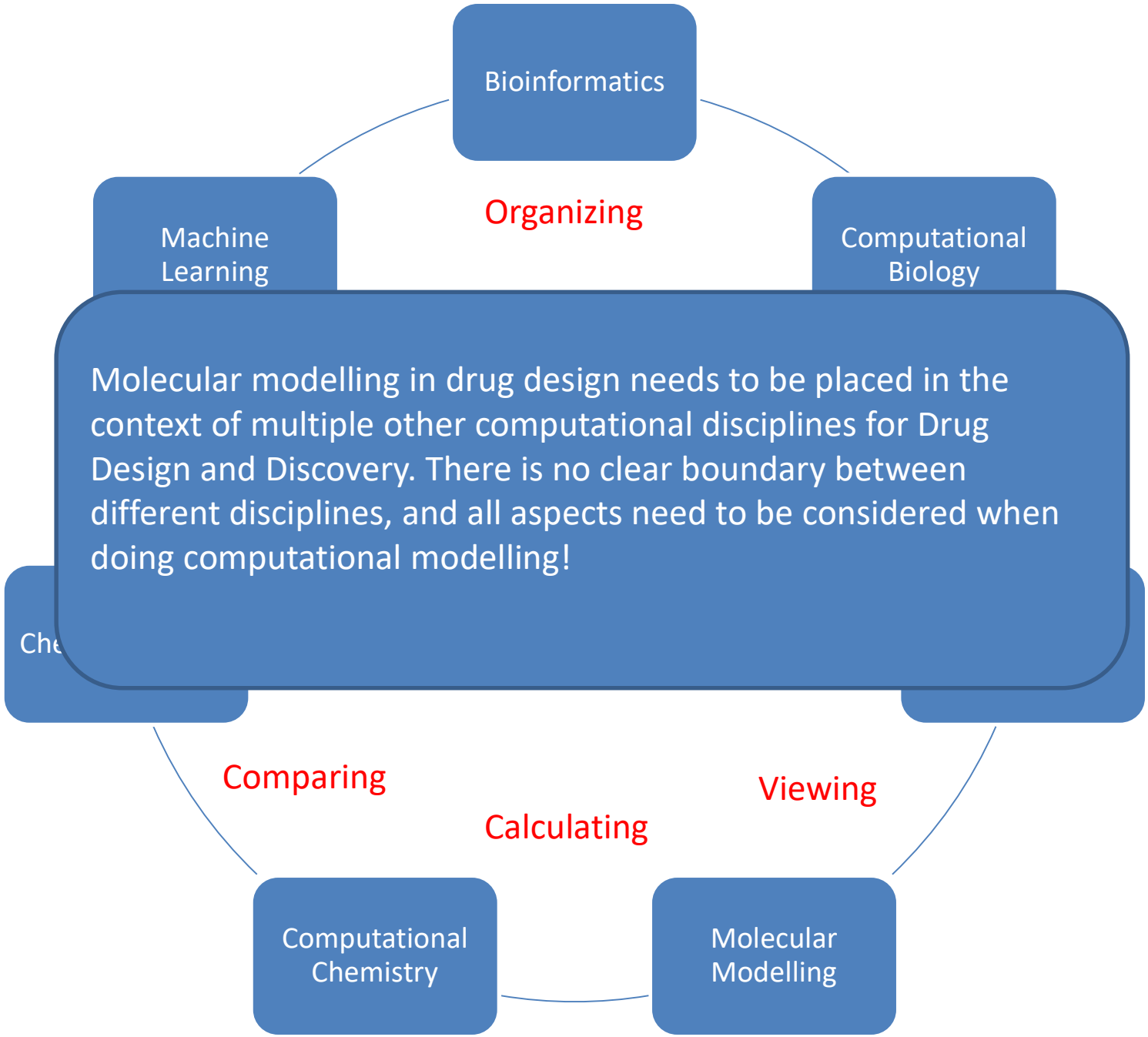
Computational Drug Design: what is it?

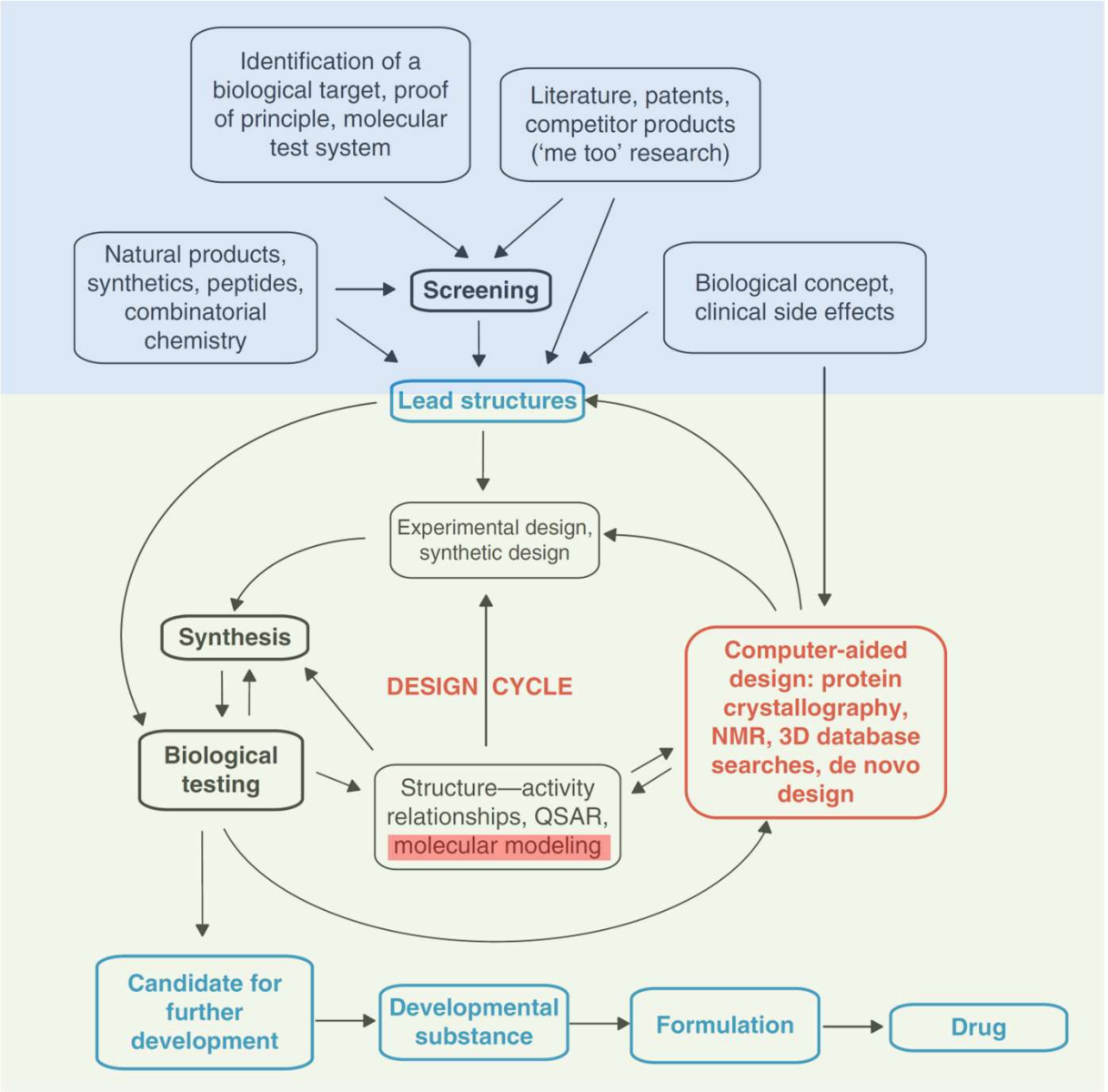
- Modern Drug Design arises from the convergence of multiple disciplines
- The chemical space is extraordinarily big and computational tools are required to fully explore it (too large for synthetic chemistry)
- Abstract and computational representation of small molecule structures
- Management of very large small molecule *virtual* databases
- Target are very large molecules (generally proteins) whose structure determination requires special methods where the computer is a necessary tool
- Analysis of target structures requires computational methods (very large structures with many thousands of atoms.
- Interaction between ligands and potential targets is a physicochemical process that can be modelled in a computer (docking)
- Computational techniques for molecular similarity can be used to identify new molecules sharing essential features with know ligands (pharmacophores, molecular fields, 3D QSAR)
- Sets of features (descriptors) can be used to classify and cluster molecules according to desired properties (Rule of 5, Golden Triangle, etc.)
- Automated machine learning methods can be used to classify molecules and predict potential activities, sites of metabolism or ADMET properties, and to genereate new structures for molecules with desired properties.

The “classical” disciplines of Drug Design




Bioinformatics
Organizing
Machine Learning
Computational Biology
Molecular modelling in drug design needs to be placed in the context of multiple other computational disciplines for Drug Design and Discovery. There is no clear boundary between different disciplines, and all aspects need to be considered when doing computational modelling!
Comparing
Viewing
Calculating
Computational Chemistry
Molecular Modelling

How do computational techniques integrate into the Drug Discovery process?



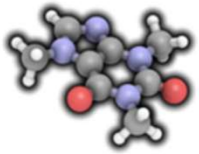
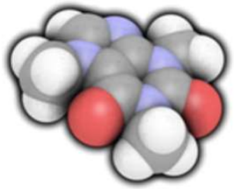
Klebe, G. *Drug Design.* Springer 2013, chap 1

Techniques in Molecular Modelling

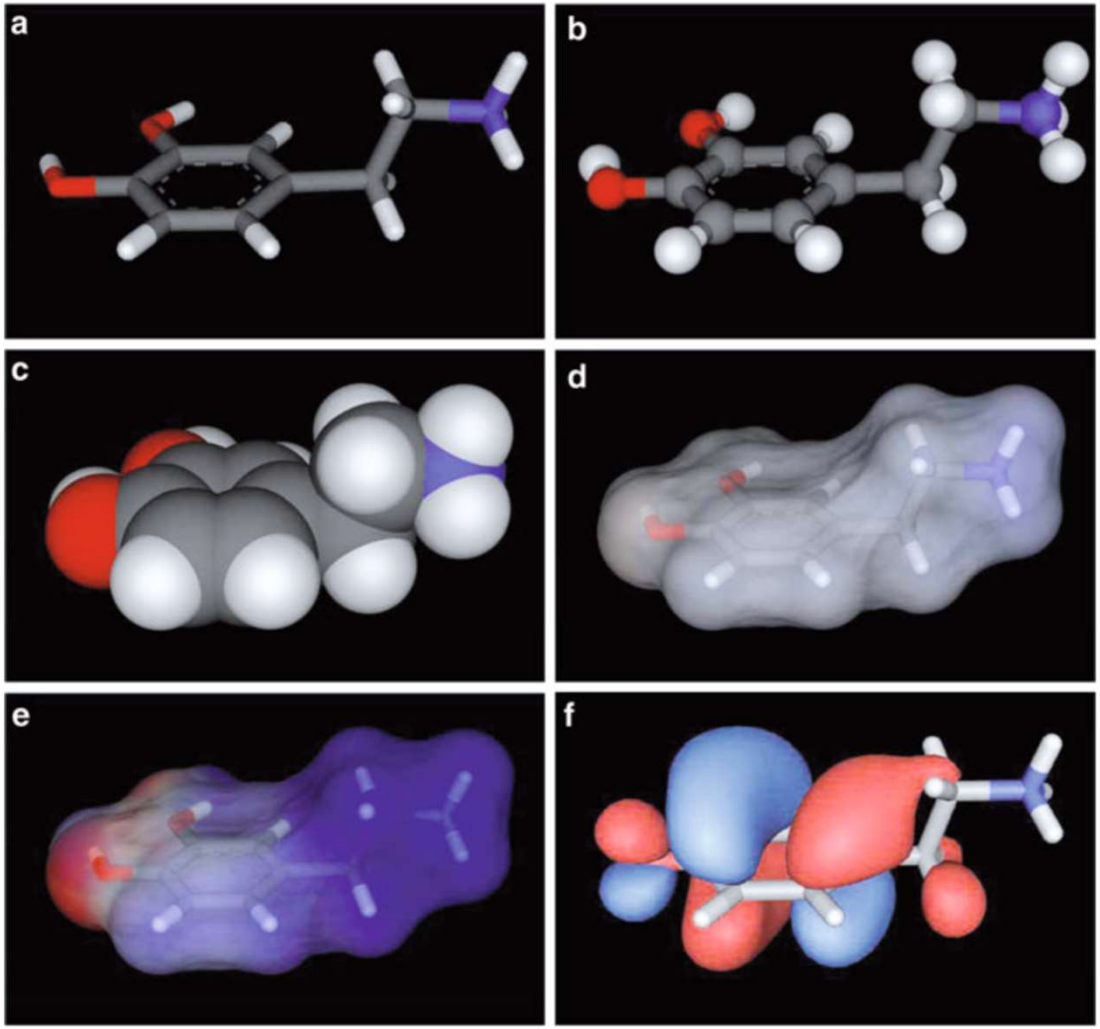
Technique	Objective
Interactive computer graphics	Display of 3D structures
Modeling small molecules	3D Structure generation (CONCORD, CORINA)
	Molecular mechanics—force fields
	Molecular dynamics
	Quantum mechanical techniques
	Conformational analysis
	Calculation of physicochemical properties
Comparing molecules	Superimposition of molecules according to their similarity
	Volume comparisons
	3D-QSAR (e.g., CoMFA methods)
Protein modeling	Sequence comparisons
	Protein homology modeling
	Protein-folding simulations
Modeling of protein–ligand interactions	Binding constant calculations
	Ligand docking
Ligand design	Searches in 3D databases
	Structure-based ligand design
	de novo design
	Virtual screening

Klebe, G. *Drug Design.* Springer 2013, chap 15

Representing chemical structures

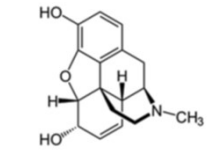
Representation Name	Representation of Caffeine
Common Name	Caffeine
Synonyms	Guaranine
	Methyltheobromine
	1,3,7-Trimethylxanthine
	Theine
Empirical Formula	$C_8H_{10}N_4O_2$
IUPAC Name	1,3,7-trimethylpurine-2,6-dione
CAS Registry Number	58-08-2
ChEMBL ID	CHEMBL113
Wiswesser Line Notation (WLN)	T56 BN DN FNVNVJ B F H
SMILES	CN1C=NC2=C1C(=O)N(C(=O)N2C)C
Aromatic SMILES	CN1C(=O)N(C)c2ncn(C)22C1=O
InChI	1S/C8H10N4O2/c1-10-4-9-6-5(10)7(13)12(3)8(14)11(6)2/h4H,1-3H3
InChIKey	RYYVLZVUVIJVGH-UHFFFAOYSA-N
Topography	
Surface	

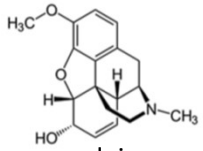
Vsizualizing chemical structures

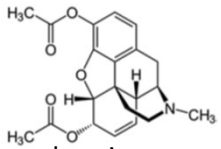


a – dreiding model
b – ball-and-stick
c – vdW (CPK)
d – molecular surface
e – surface potential
f – HOMO orbitals

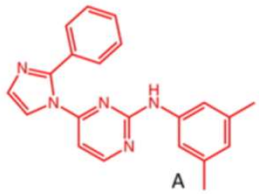
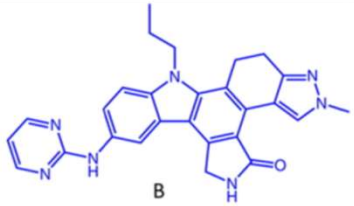
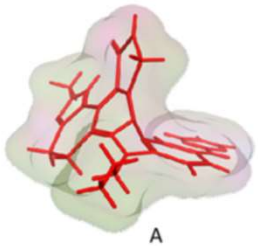
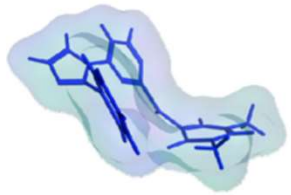
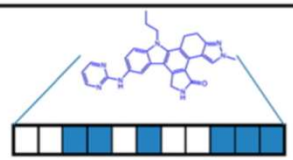
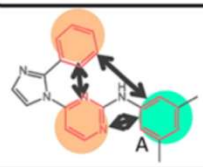
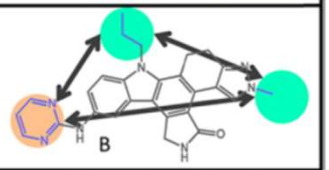
The Importance of molecular similarity


morphine

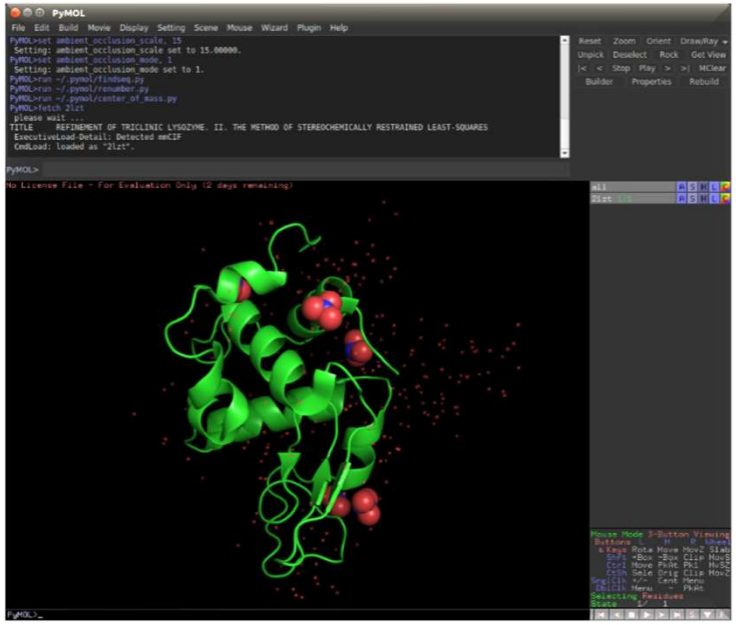

codeine

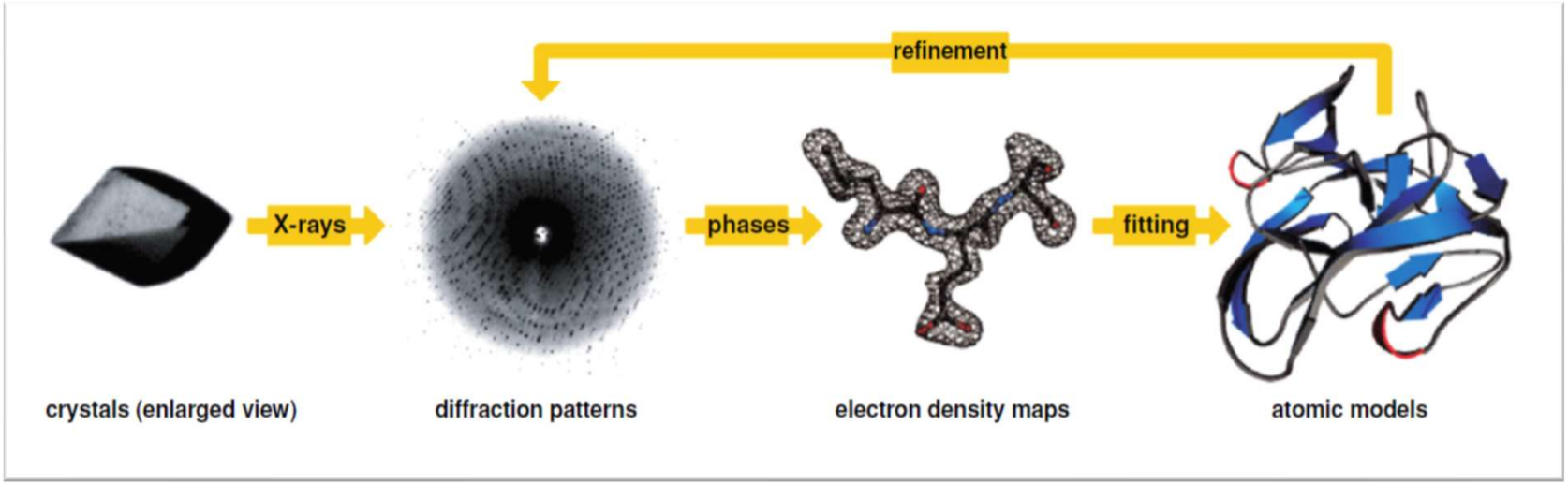

heroine

Similar structures similar functions

Chemical similarity		Mol. weight	LogP	Rotatable bonds	Aromatic rings	Heavy atoms
	A	341.4	5.23	4	4	26
	B	463.5	4.43	4	5	35
Molecular similarity / 2D similarity		 A		 B		
3D similarity		 A		 B		
Biological similarity		Vascular endothelial growth factor receptor 2	Tyrosine-protein kinase TIE-2			
	A	active	inactive			
	B	active	active			
Global similarity						
Local similarity		 A		 B		

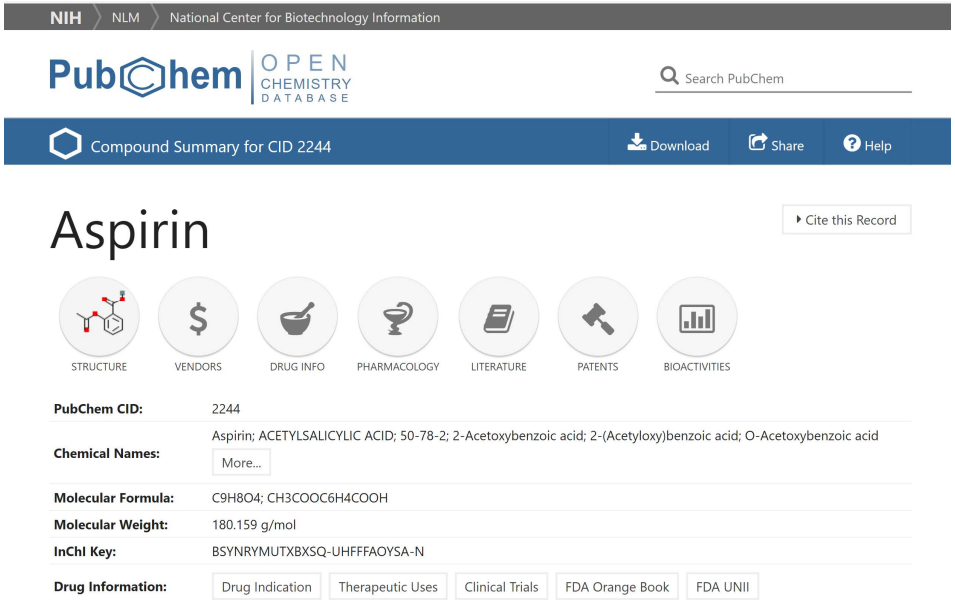
Target Structure





Protein structure determination by X-ray crystallography

Chemical Databases



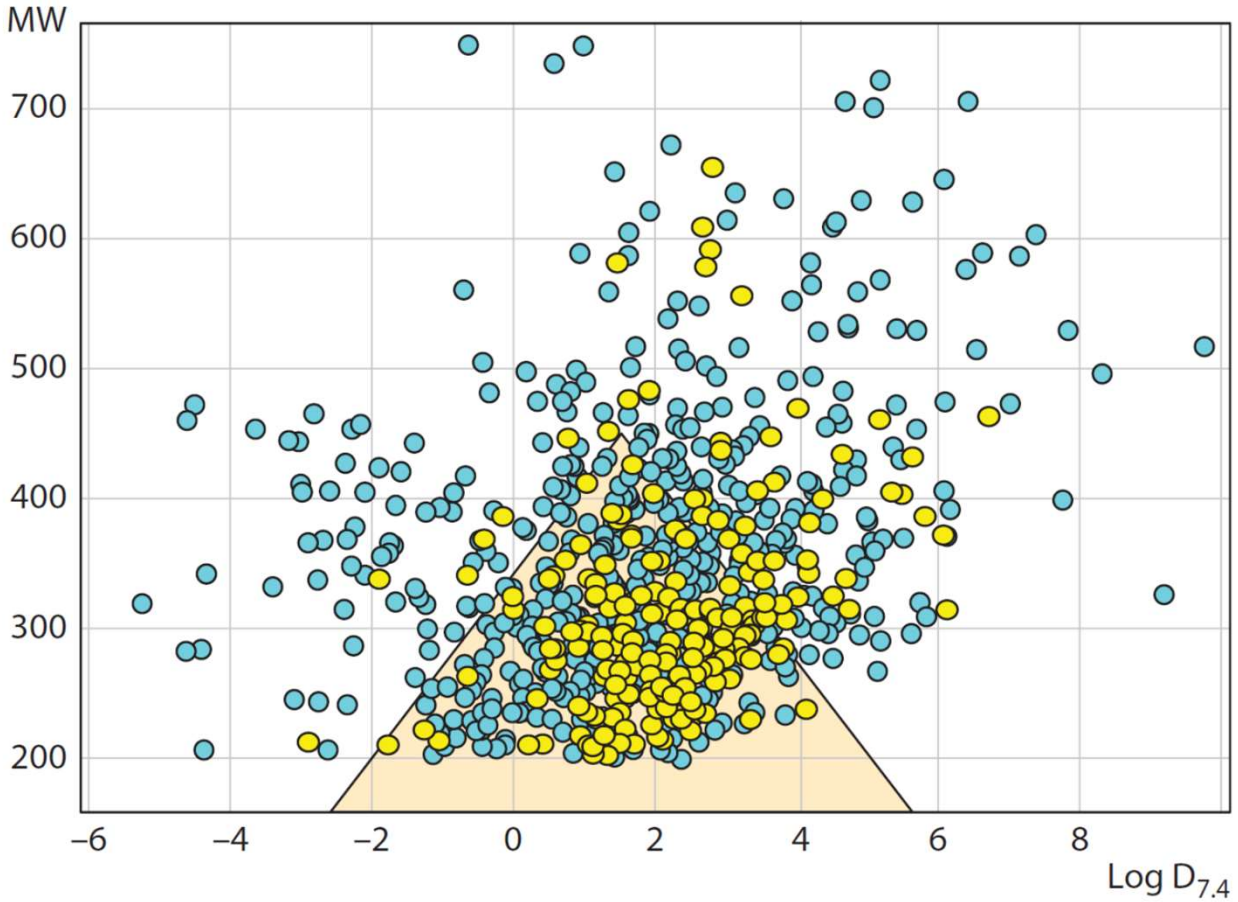
Database	Description	Size	web addresses
DrugBank[5]	Collection of approved and experimental drugs	7895	*https://www.drugbank.ca/*
CTD[6]	Toxicogenomics database	12 K	*http://ctdbase.org/about/dataStatus.go*
NCI[7]	National cancer institute chemical database	265 K	*https://cactus.nci.nih.gov/*
BindingDB[8]	Bioactive small molecules annotated with experimental data	600 K	*https://www.bindingdb.org/bind/index.jsp*
ChEMBL[9]	Bioactive small molecules annotated with experimental data	1.7 M	*https://www.ebi.ac.uk/chembldb*
SureChEMBL[10]	Collection of patented compounds	17 M	*https://www.surechembl.org/search/*
eMolecules	Commercial small molecules for screening	7 M	*https://www.emolecules.com/*
ChemSpider	Collection of compounds from various institutions and commercial companies	58 M	*http://www.chemspider.com/*
PubChem[11]	NIH repository of molecules	93 M	*http://pubchem.ncbi.nlm.nih.gov*
ZINC 15[12]	Commercial small molecules for screening	378 M	*http://zinc15.docking.org/*
GDB-11[13]	Possible small molecules up to 11 atoms of C, N, O, F	26 M	*http://gdb.unibe.ch*
GDB-13[14]	Possible small molecules up to 13 atoms of C, N, O, S, Cl	980 M	*http://gdb.unibe.ch*
GDB-13.FL[15]	Fragrance-like subset of GDB-13	59 M	*http://gdb.unibe.ch*
GDB-17[16]	Possible small molecules up to 17 atoms of C, N, O, S and halogens	166 B	*http://gdb.unibe.ch*
FDB-17[17]	Fragment like subset of GDB-17	10 M	*http://gdb.unibe.ch*

What Makes a Good Drug ?

Finding the essential chemical descriptors (dimensionality reduction), classifying, filtering, selecting.
Machine learning-methods

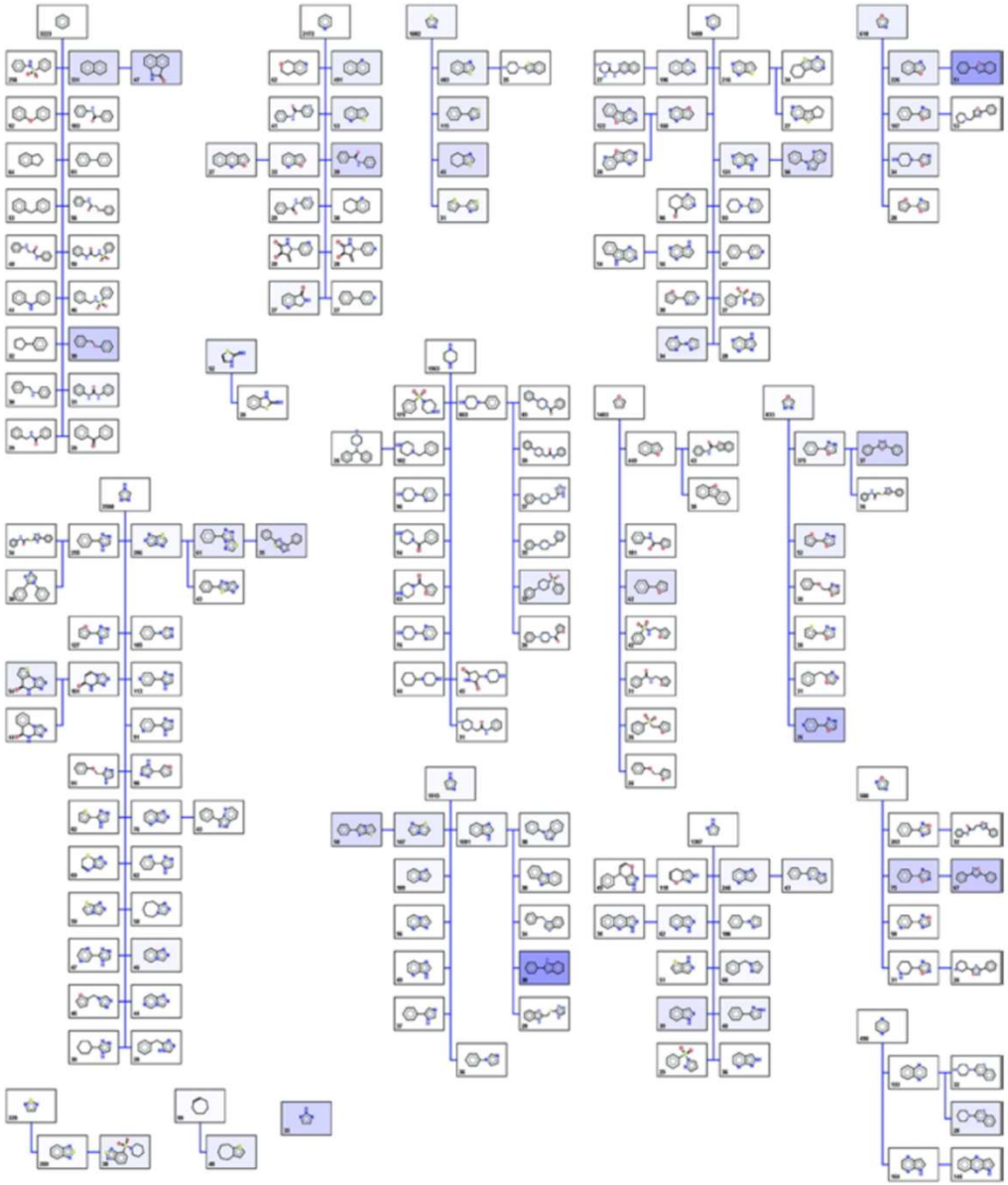
Lipinski's rule of 5

Peripheral drugs
84% Ro5 compliant 53% inside the Golden Triangle 70% have CNS MPO score > 4
CNS drugs
92% Ro5 compliant 77% inside the Golden Triangle 70% have CNS MPO score > 4



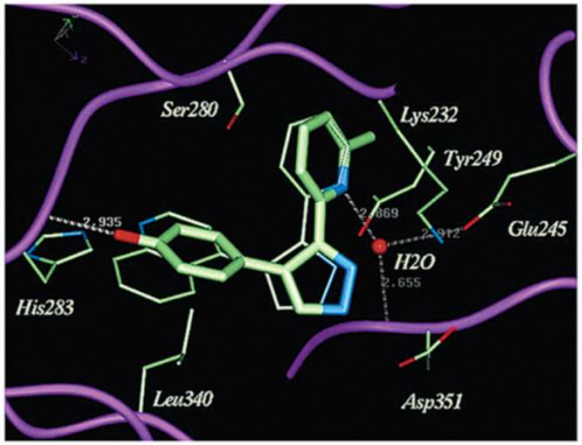
Scaffold Trees

Guided search through chemical space.

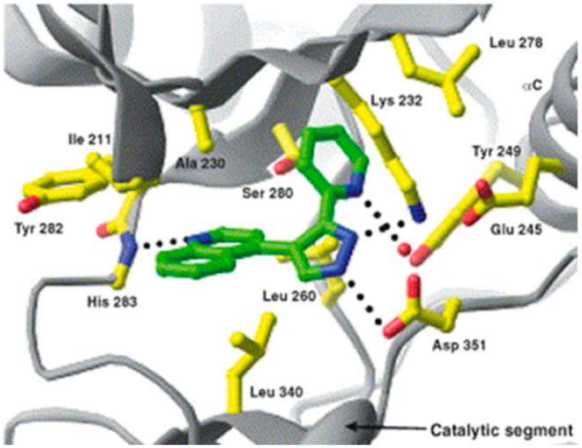


Color intensity represents potency

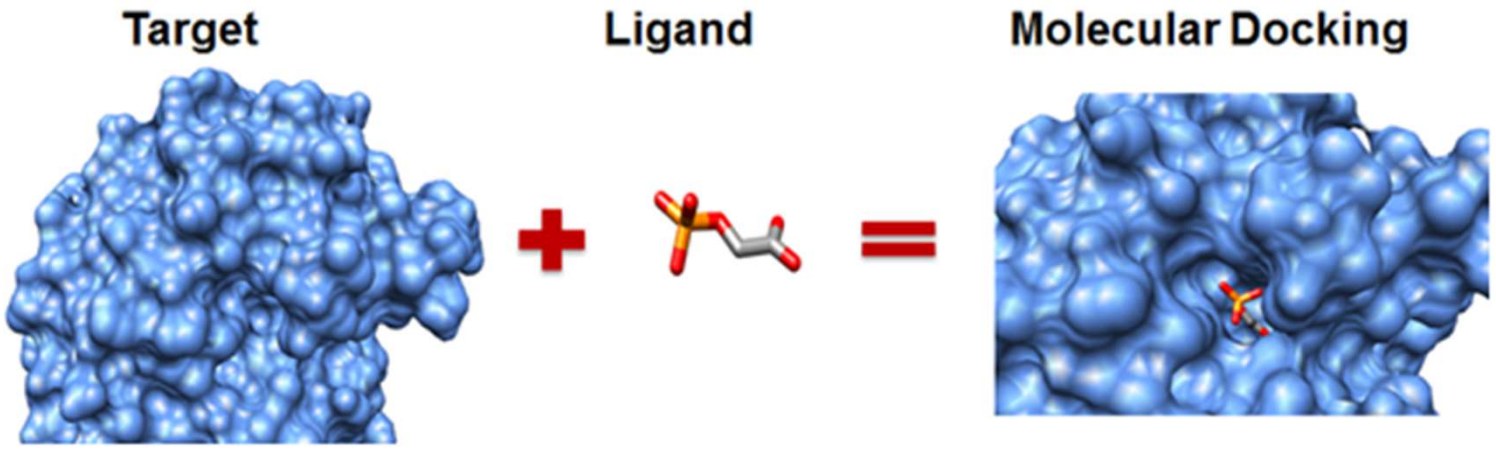
Target-Ligand Docking



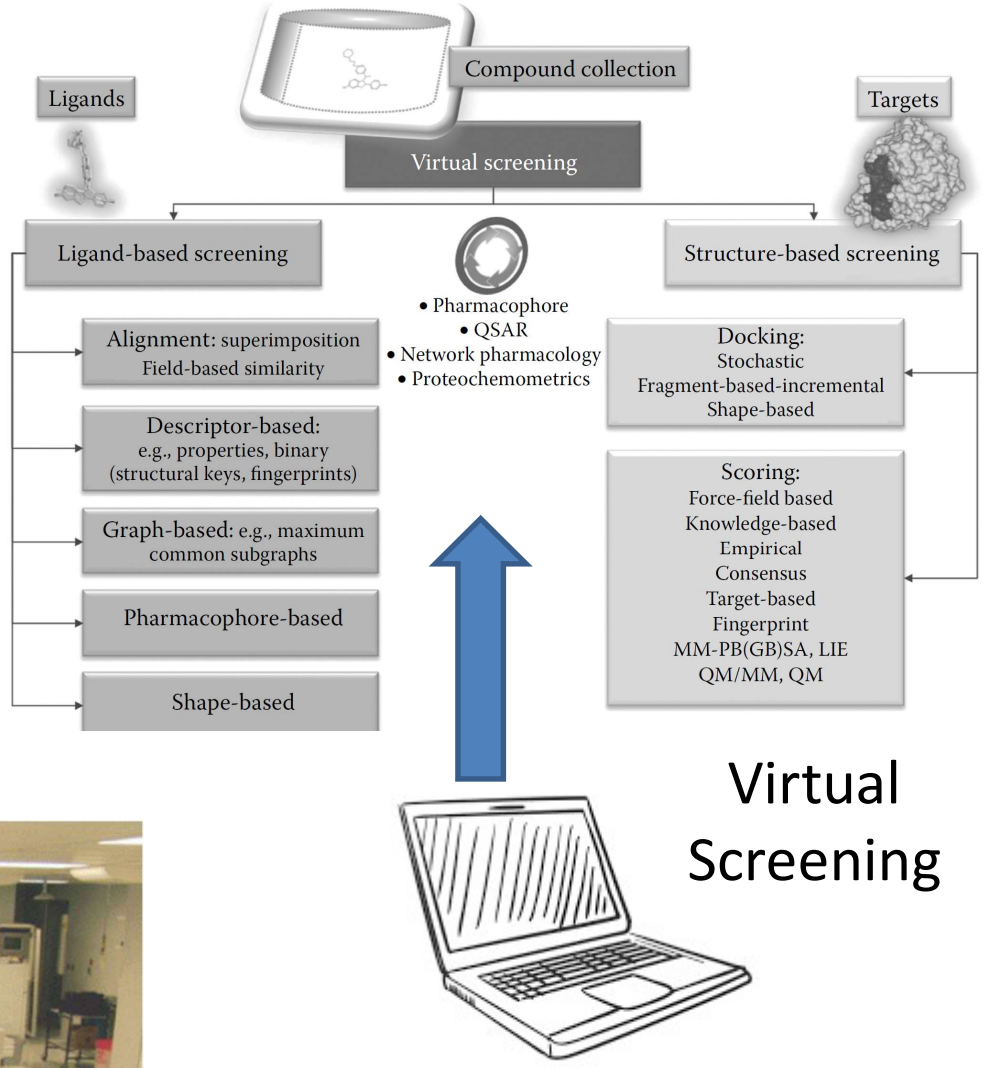
HTS



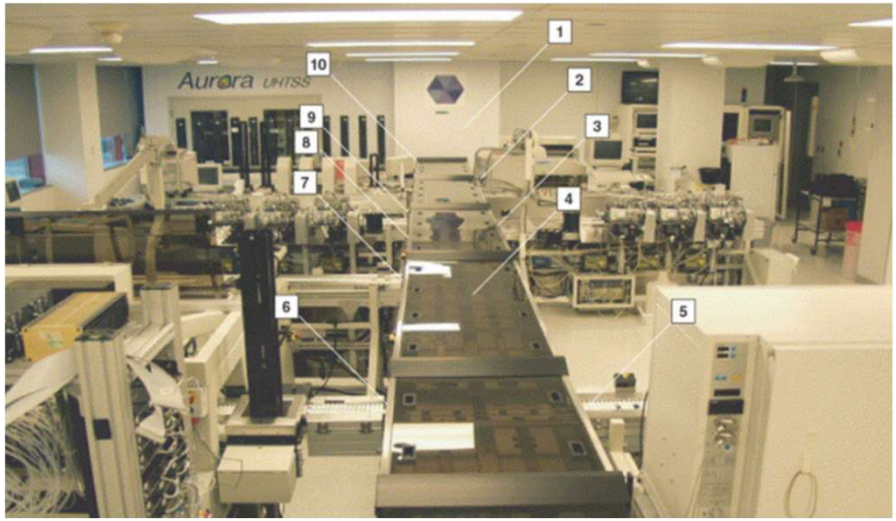
Docking



Virtual Screening *versus* Real Screening

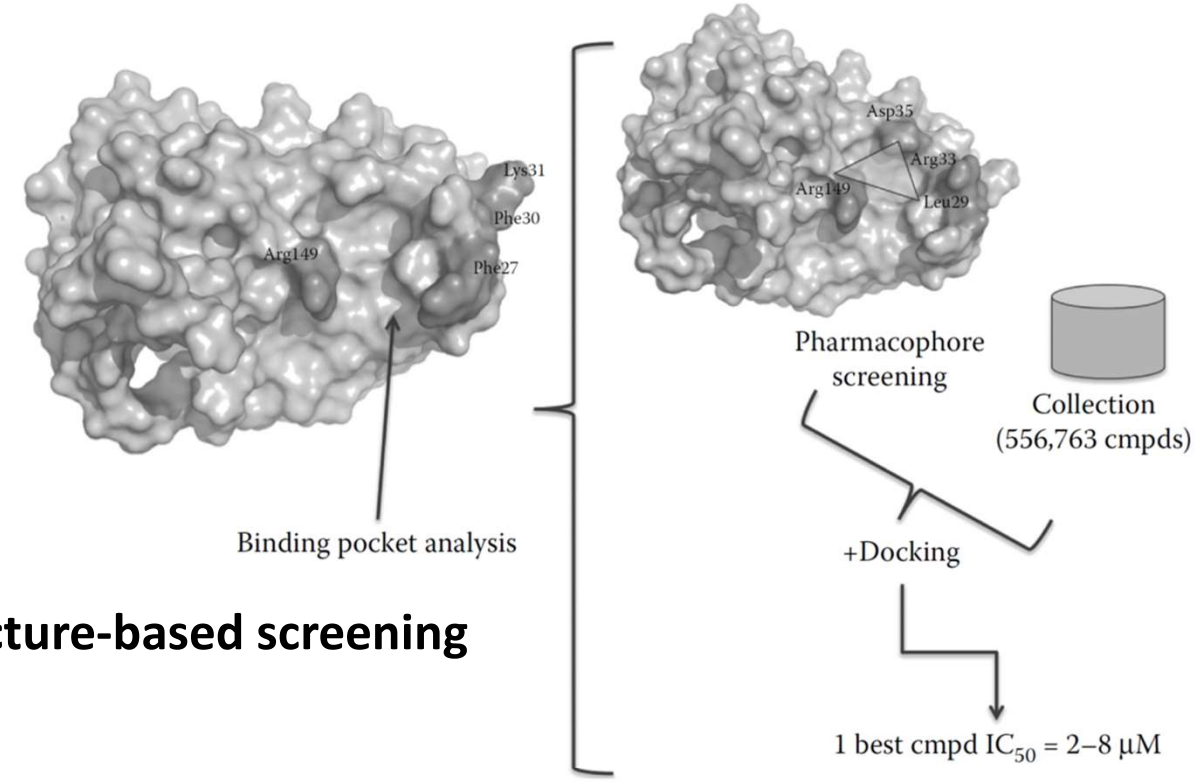


Virtual Screening

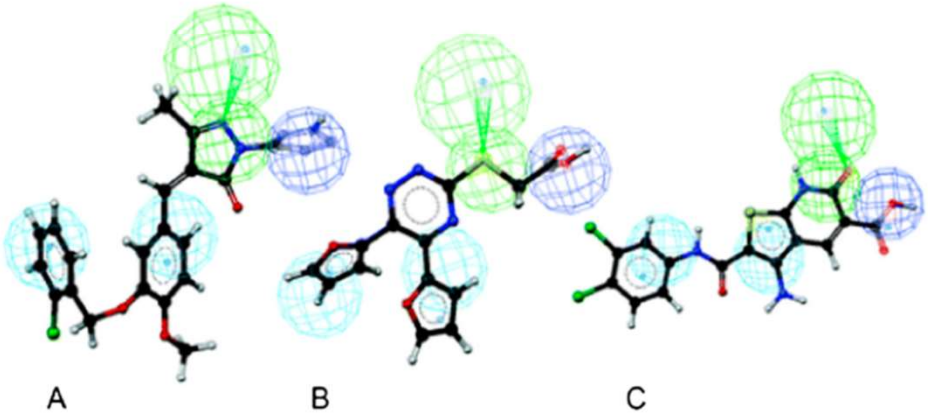


Hight-Throughput Screen Laboratory

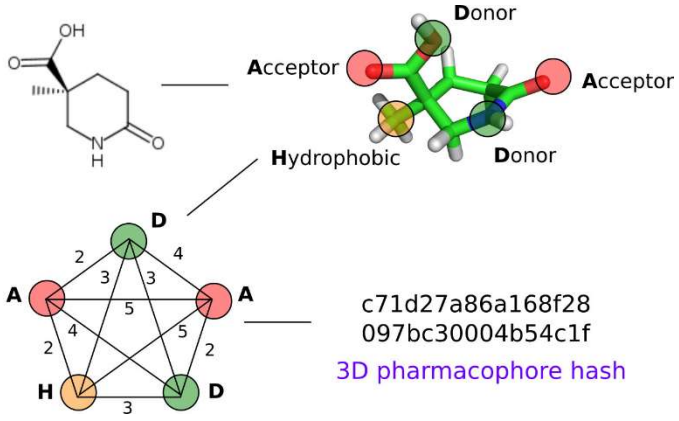
Pharmacophore screening



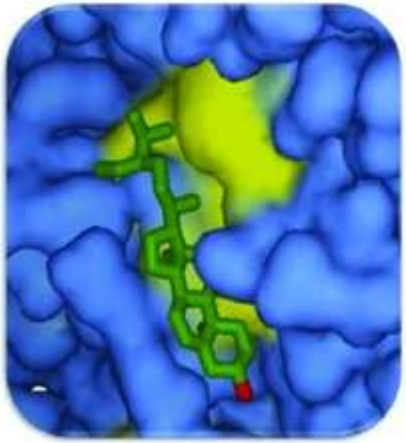
Structure-based screening



Ligand-based screening



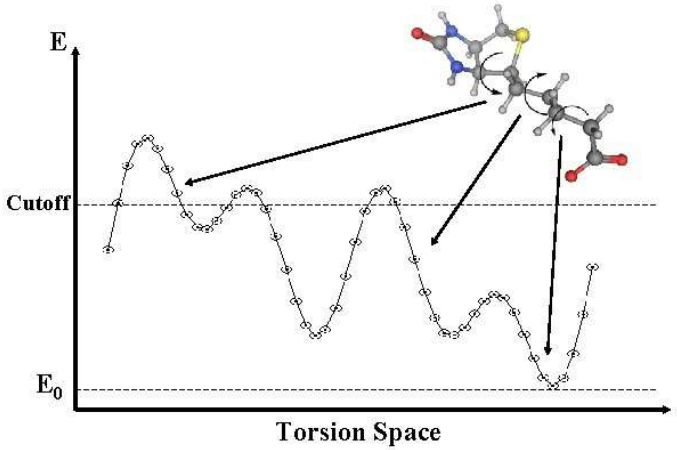
Importance of Conformational Search

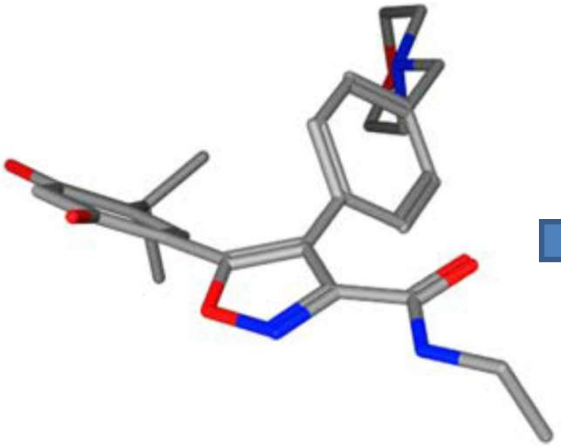


Molecule docks in the "right" conformation

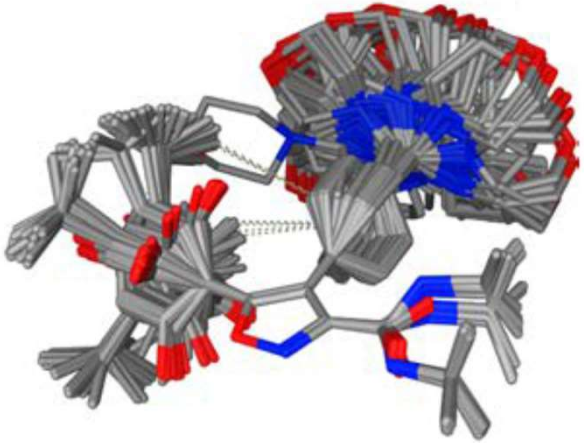
Systematic Conformational Search

- Exhaustive incremental dihedral rotation search





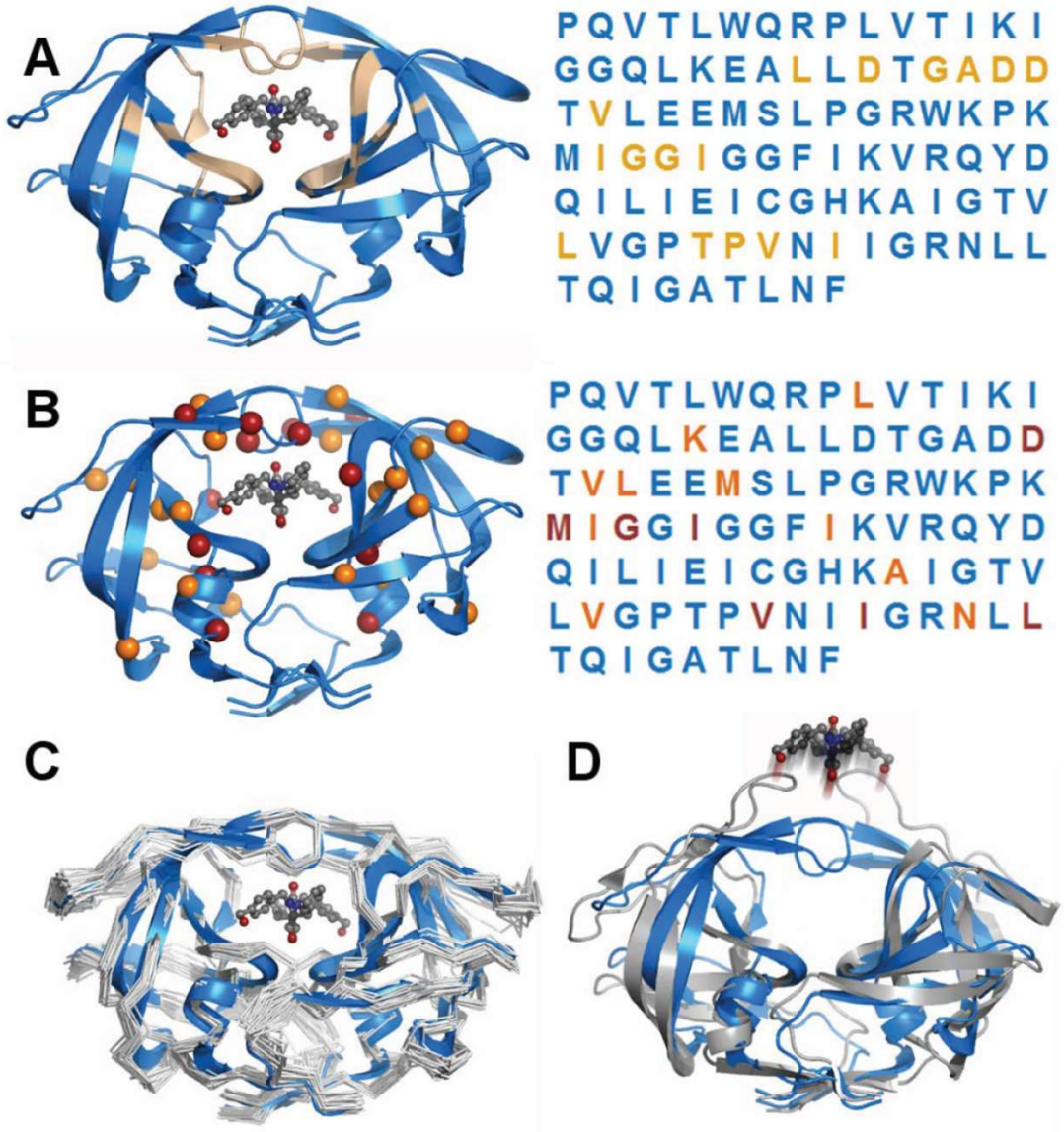
Scanning Conformational Space



Single conformation

Mutiple conformations

Sequence and Structure Analysis of Protein Targets



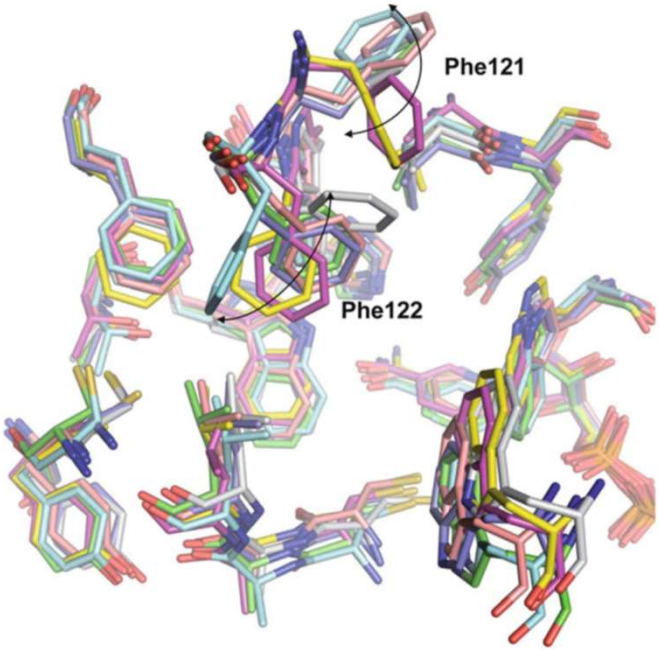
HIV protease

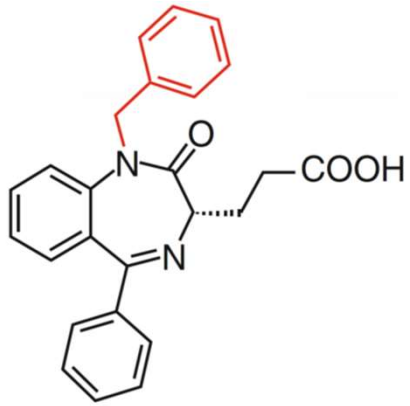
A – Residues near bound inhibitor
B – Mutations leading to resistance
C – Mutations can affect flexibility
D – Dynamics of ligand free protein (studied by *MD simualtions*)

Urrutia (2016). F1000, 5:766

Importance of Molecular Dynamics Simulations

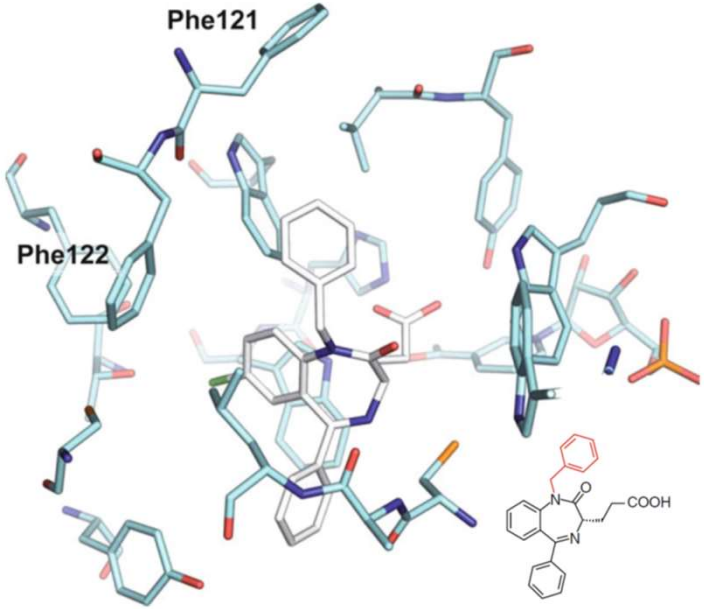
MD simulation shows wide movement of Phe121 residue, enlarging the binding pocket of the receptor



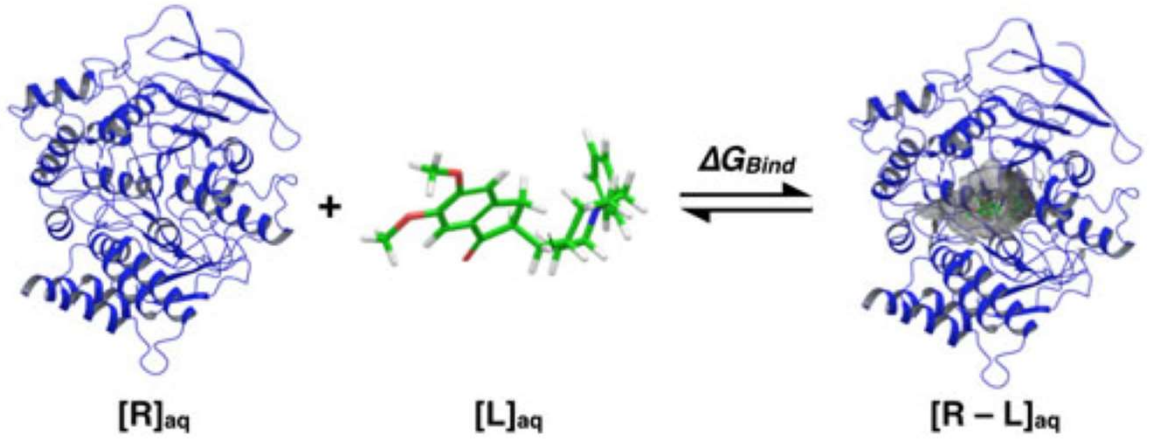


Benzodiazepine-like inhibitor

The open conformation can accommodate ligands with extended functional groups, like the red group of the benzodiazepine-like inhibitor,

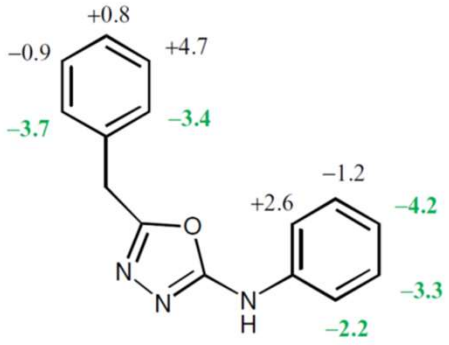


Binding free energies of ligands by Molecular Dynamics

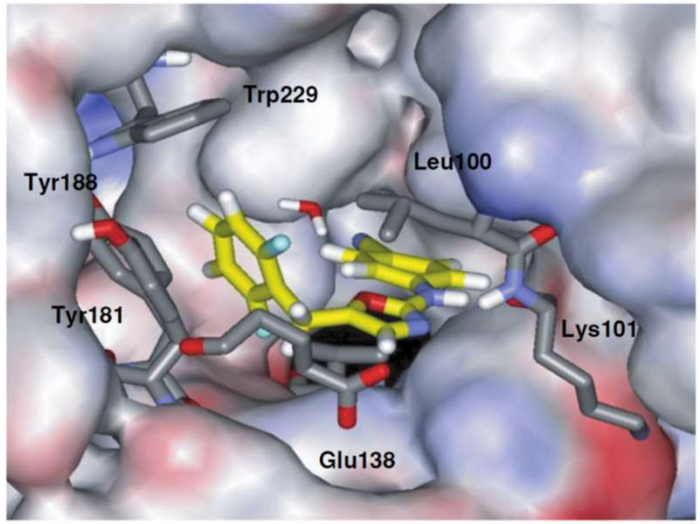


$$K_d = \frac{[R][L]}{[RL]}$$

$$\Delta G = RT \ln K_d$$

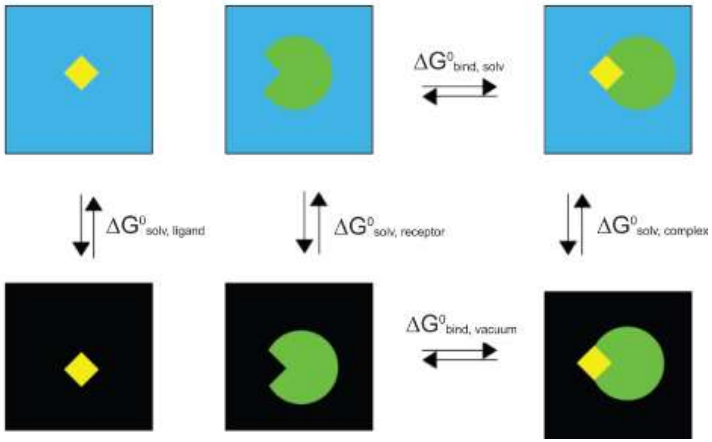


a)



b)

Prediction of binding affinities



MM-PBSA binding energy calculation

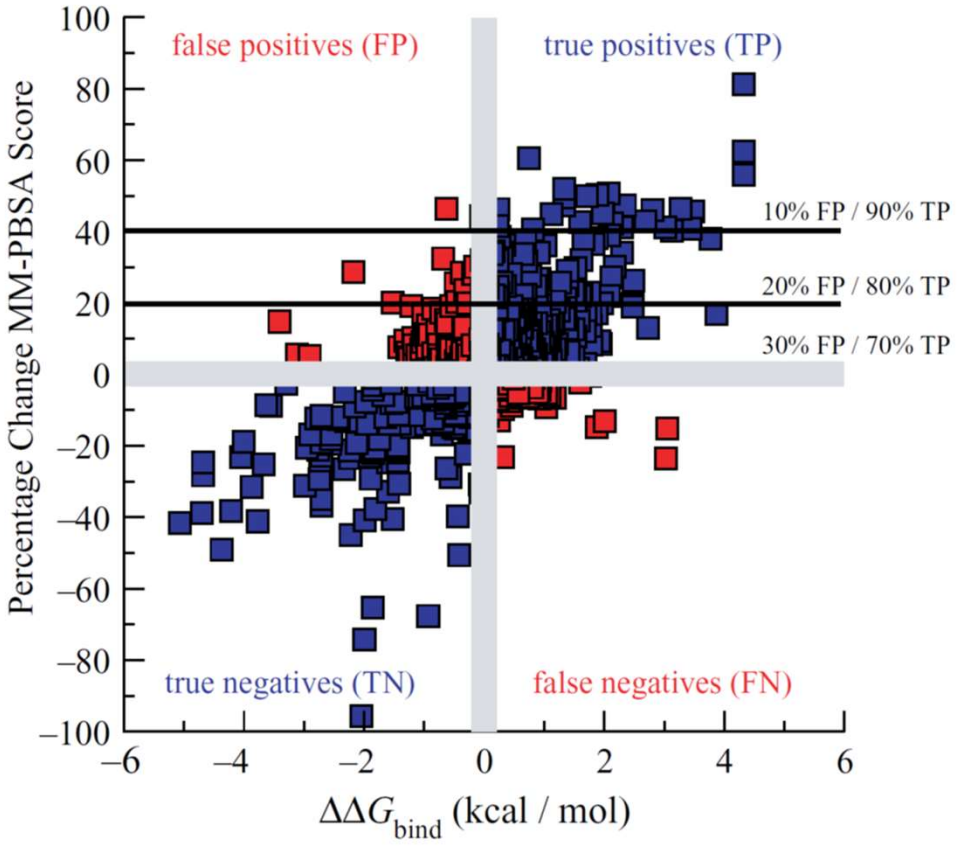
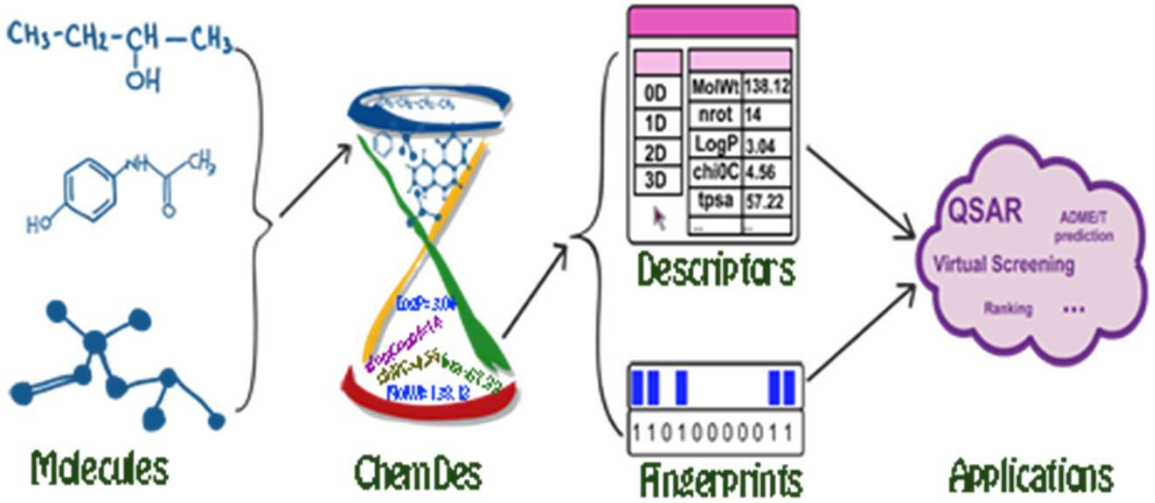


Figure 5.3. Data showing change in compound potency (relative to a reference compound) versus percentage change in MM-PBSA score (relative to same reference compound) for 480 compounds across eight targets, which span 292 x-ray crystallographic complexes.

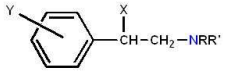
Drug Descriptors and QSAR



New compound generation

Hansch Equation

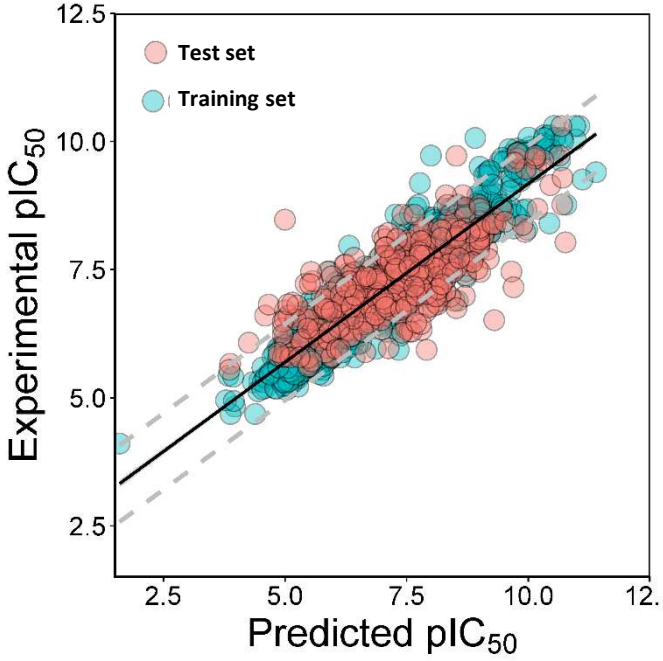
Example: Adrenergic blocking activity of β-halo-β-arylamines



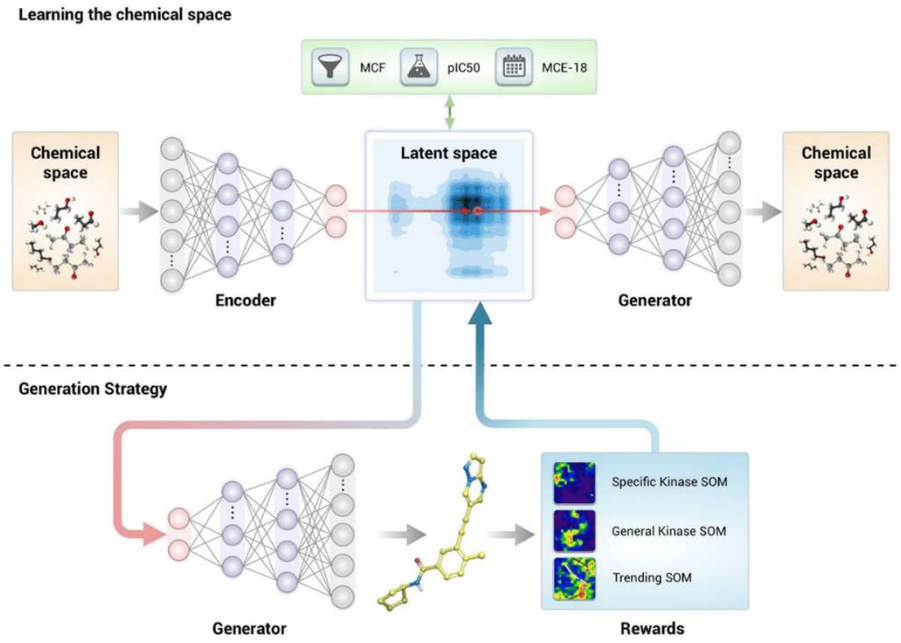
$$\text{Log}\left(1/C\right) = 1.22\,\pi - 1.59\,\sigma + 7.89$$

Conclusions:
- Activity increases if π is +ve (i.e. hydrophobic substituents)
- Activity increases if σ is negative (i.e. e-donating substituents)

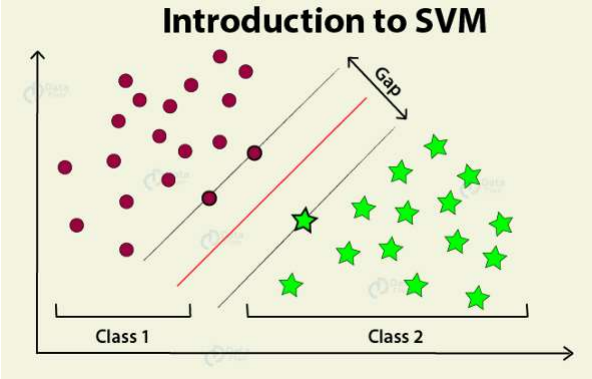
Model equation



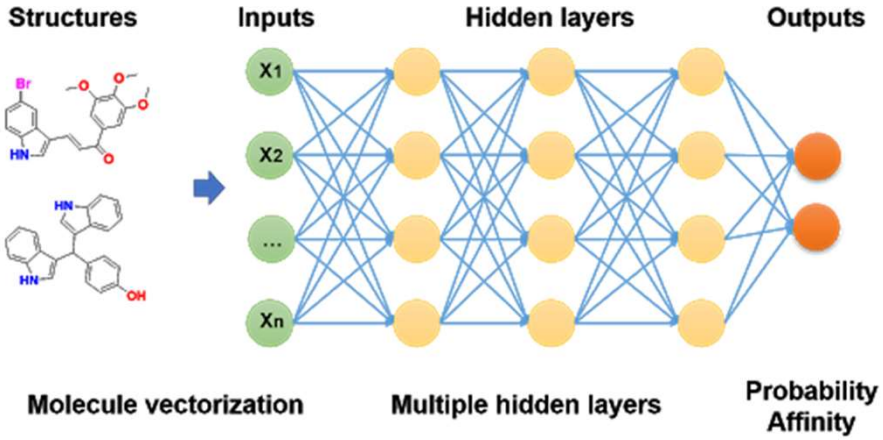
Machine Leaning and AI



New compound generation

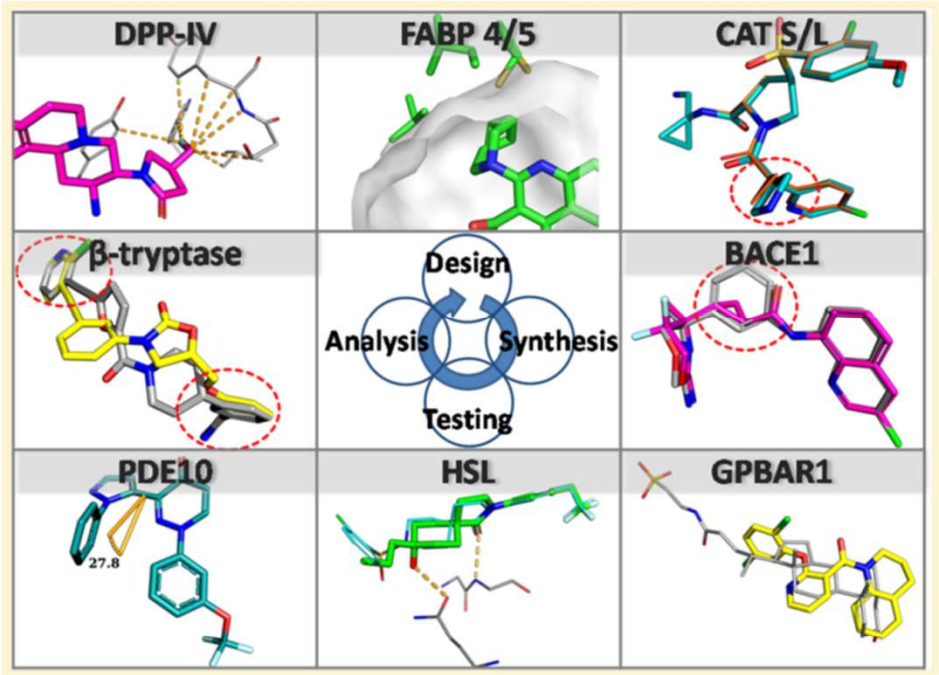


Automatic classification



Property prediction

CADD works in the "real world"



Journal of Medicinal Chemistry

Perspective

pubs.acs.org/jmc

A Real-World Perspective on Molecular Design

Miniperspective

Bernd Kuhn, Wolfgang Guba, Jérôme Hert, David Banner, Caterina Bissantz, Simona Ceccarelli, Wolfgang Haap, Matthias Körner, Andreas Kuglstatter, Christian Lerner, Patrizio Mattei, Werner Neidhart, Emmanuel Pinard, Markus G. Rudolph, Tanja Schulz-Gasch, Thomas Woltering, and Martin Stahl*

Roche Pharmaceutical Research and Early Development, Roche Innovation Center Basel, F. Hoffmann-La Roche Ltd., Grenzacherstrasse 124, 4070 Basel, Switzerland

Roche

Kuhn (2016). *J.Med.Chem.* 59:4087

CADD and diseases

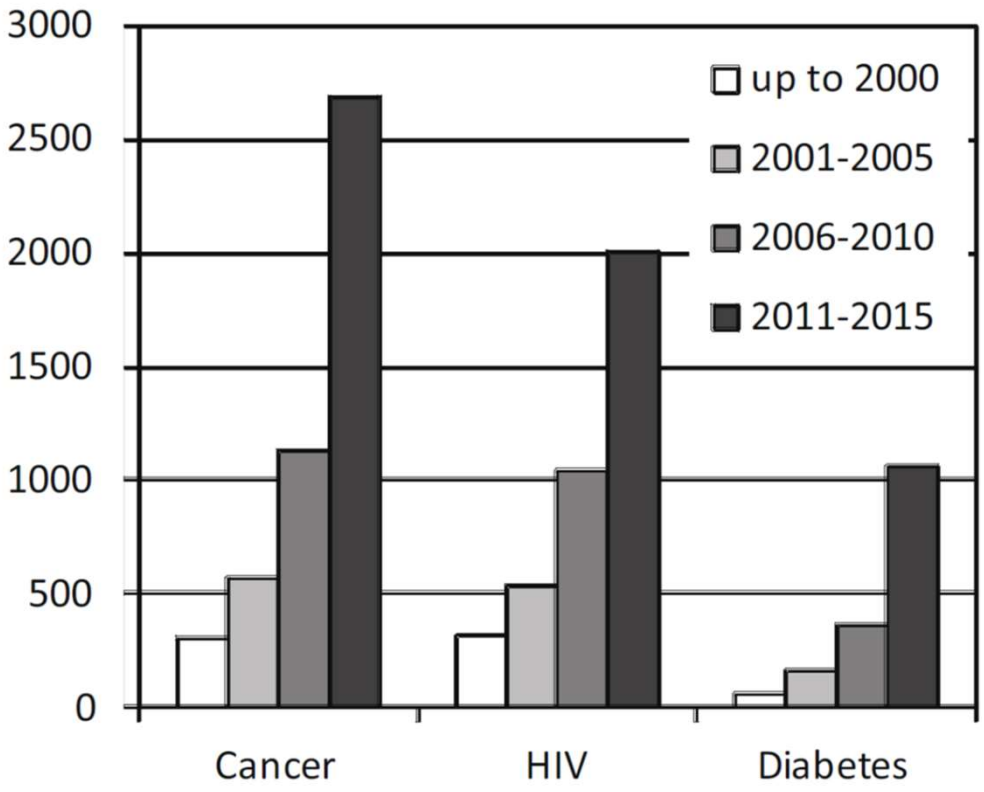


Fig. (1). The number of publications related to computer-aided drug design and diseases. Key words used in the Google Scholar search [16] were as follows: computer-aided drug design and disease; *e.g.* diabetes.