Introduction to Computer-Aided Drug Design

http://www.glycopedia.eu/resources/article/computer-aided-drug-design

About the CADD course

Computer Aided Drug Design CADD represents an approach for the design of new bioactive molecules using molecular modelling methods.

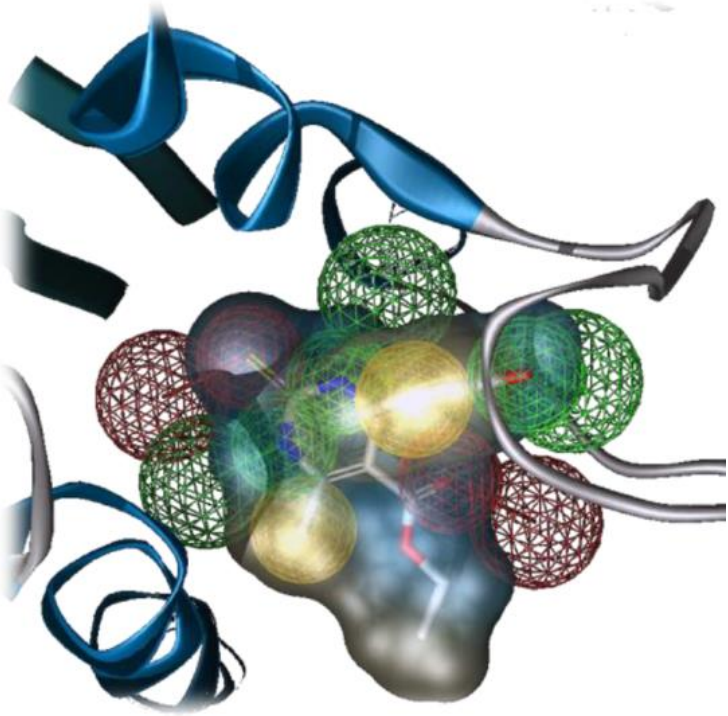
What you will learn….

- Molecular modelling methods used for drug design (methods & examples)
- Experimental methods complementary to molecular modelling (crystallography, NMR, Isothermal Calorimetry and Surface Plasmon Resonance)

Pharma Research

a few definitions

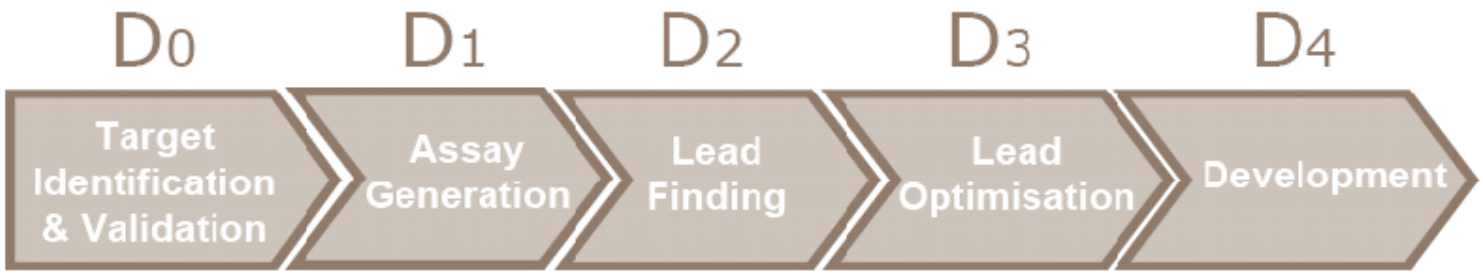
- **Target** : Pathopysiologically relevant biomolecule (enzyme, receptor, ion channel or transporter)
- **Assay** : Type of experiment. It measures the effect of a compound on the function of a target, cell, or living organism



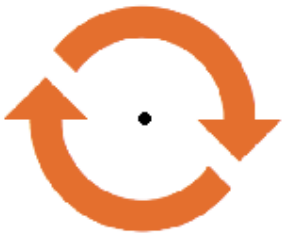
- **Hit** : Chemical compound active in assay. Does not mean necessarily that it acts on a specific target ... needs validation
- **Lead** : Chemical compound for which we demonstrate SAR, ipotential issues are identified (selectivity, physicochemical, ADME/Tox), novelty
- **LMW** : Low Molecular Weight compounds (< 900 Da)

Pharma Research

*The **D** notation of drug discovery phases*



Organon · novo nordisk · AMGEN · gsk GlaxoSmithKline · Pfizer · Boehringer Ingelheim · Astrazeneca · NOVARTIS · Roche · sanofi aventis · Eisai · Schering-Plough · Lilly · Wyeth · Johnson & Johnson · MERCK



Constant looping along the process

Pharma Research

***D0**: Target identification & validation*

- **Biology**

Understanding a disease pathway

  - What are the dysfunctional nodes that cause the disease?
  - Are they on the critical path?
  - How can we prove this at a clinical level?
  - How can we set up a surrogate *in vitro* assay?

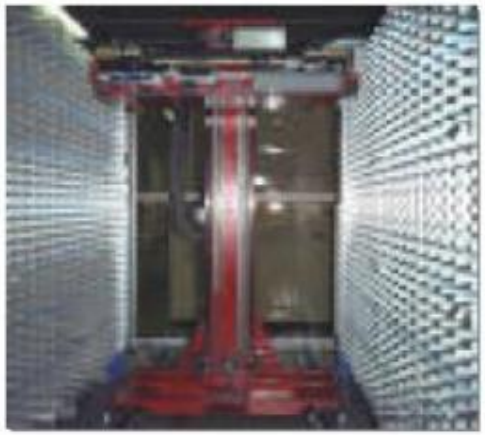
- **Chemistry**
  - Can the target be modulate by small molecules?
  - What do we know from the past about similar targets?

Target

pathopysiologically relevant biomolecule (enzyme, receptor, ion channel or transporter)

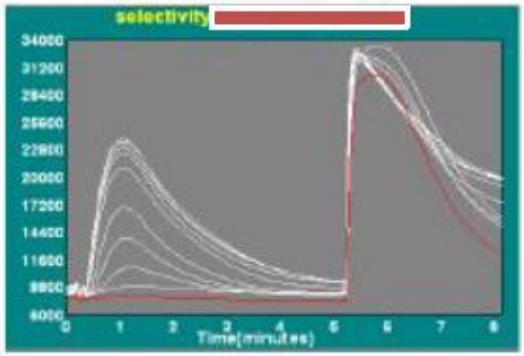
Pharma Research

***D1** : Assays generation for hit finding*



Compound libraires
(1,000,000 compounds)

Virtual screening and/or
Biological test system



Data analysis

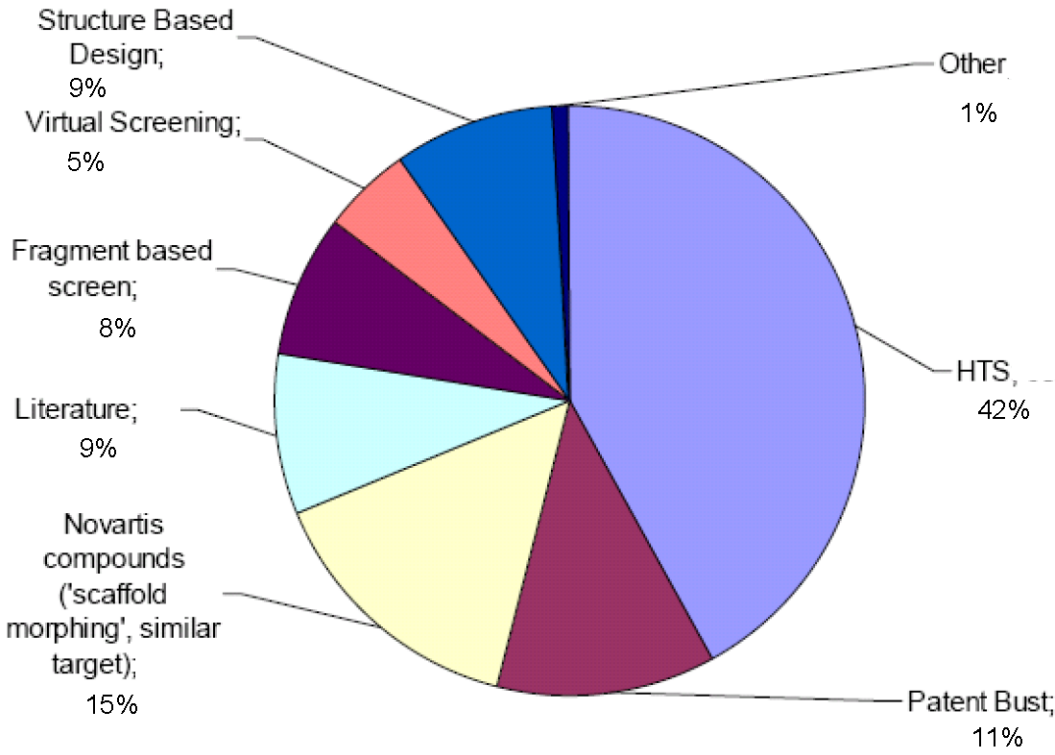
Primary screen: ca. 2 months
Validation: ca 1.5 months

Overall from start D1 to validated hit list (D2a): ca. 12 months

Pharma Research

***D2a** : Hit finding – the approaches*

- High-Throughput Screening 42%
- Virtual screening 5%
- Structure-Based Drug Design 9%
- Fragment-Based Screening 8%
- Literature 9%
- Scaffold morphing 15%
- Patent bust 11%
- Other 1%



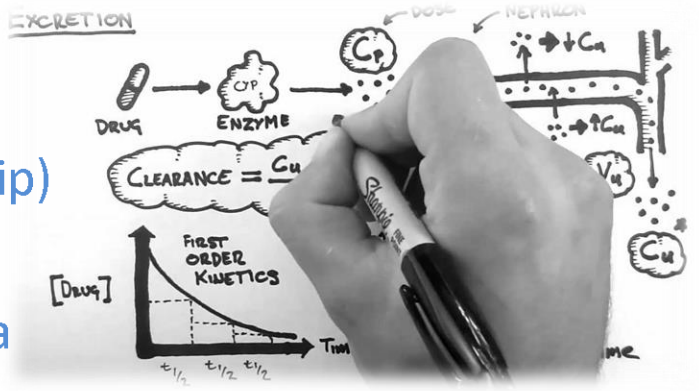
Novartis® data (2014)
with courtesy of Dr. Richard Lewis

Hit

Chemical compound active in assay. Does not mean necessarily that it acts on a specific target ... needs validation

Pharma Research

***D2b** : Hit to Lead*



- Improved activity on target
- Demonstrated SAR (Structure-Activity Relationship)
- Selectivity issues determination
- *In vitro* PK (PharmacoKinetics) : solubility and pKa
- *In vivo* PK : bioavailability (BAV) and half-life ($t_{1/2}$)
- P450, hERG, PLP, Plasma/whole blood stability
- Novelty!

Lead

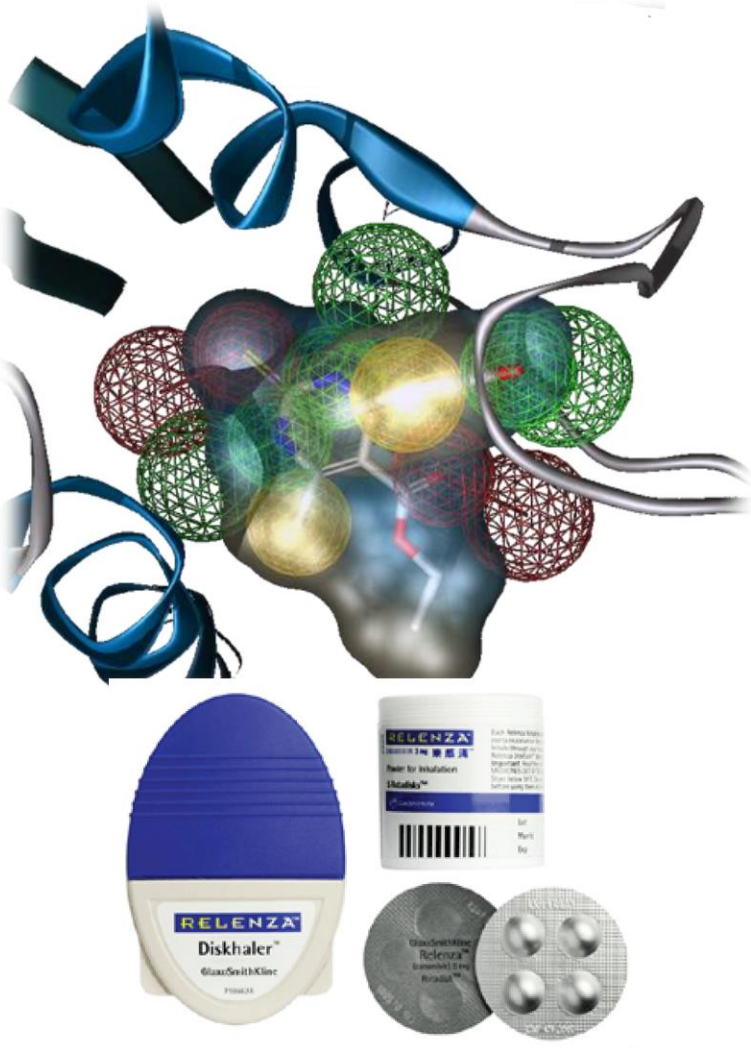
Chemical compound for which we demonstrate SAR, ipotential issues are identified (selectivity, physicochemical, ADME/Tox), novelty

Pharma Research

***D3** : Lead optimisation*

A successful drug candidate will:

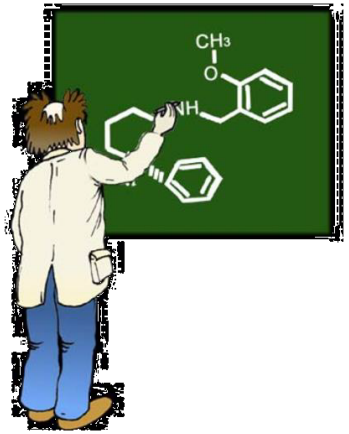
- Have a clear mechanistics identified
- Address an unmet medical need
- Have a competitive advantage *versus* current and emerging therapies
- Be bioavailable with appropriate PK and PD
- Be efficacious and safe (therapeutic window)
- Have a synthetic route that can be operated on a large scale
- Be amenable to convenient formulation
- Have intellectual property rights secured



Pharma Research

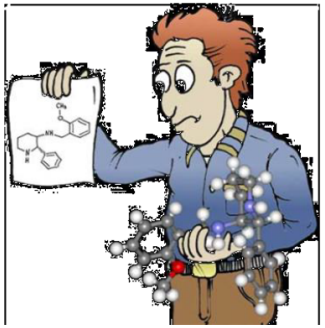
D4 : Development

Molecular perception in the 1970s… 2D!

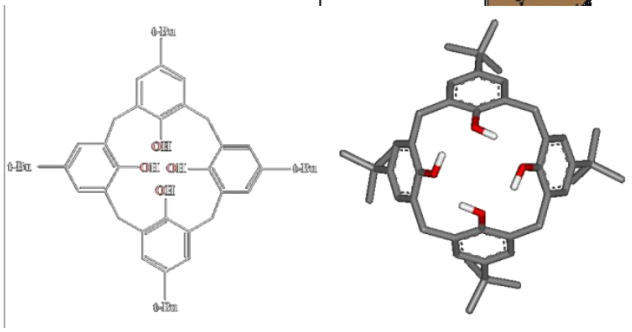


- The way molecules have been perceived and defined has changed over the years. **In the early 1970s, medicinal chemists considered molecules as topological two dimensional (2D) entities**.

- **The formula of a molecule can be drawn in two dimension as a piece of paper. However, it really exists in three-dimensional (3D) space with precise geometrical features.** The 3D geometries are of great importance because they represent the very molecular determinants that control molecular interactions.

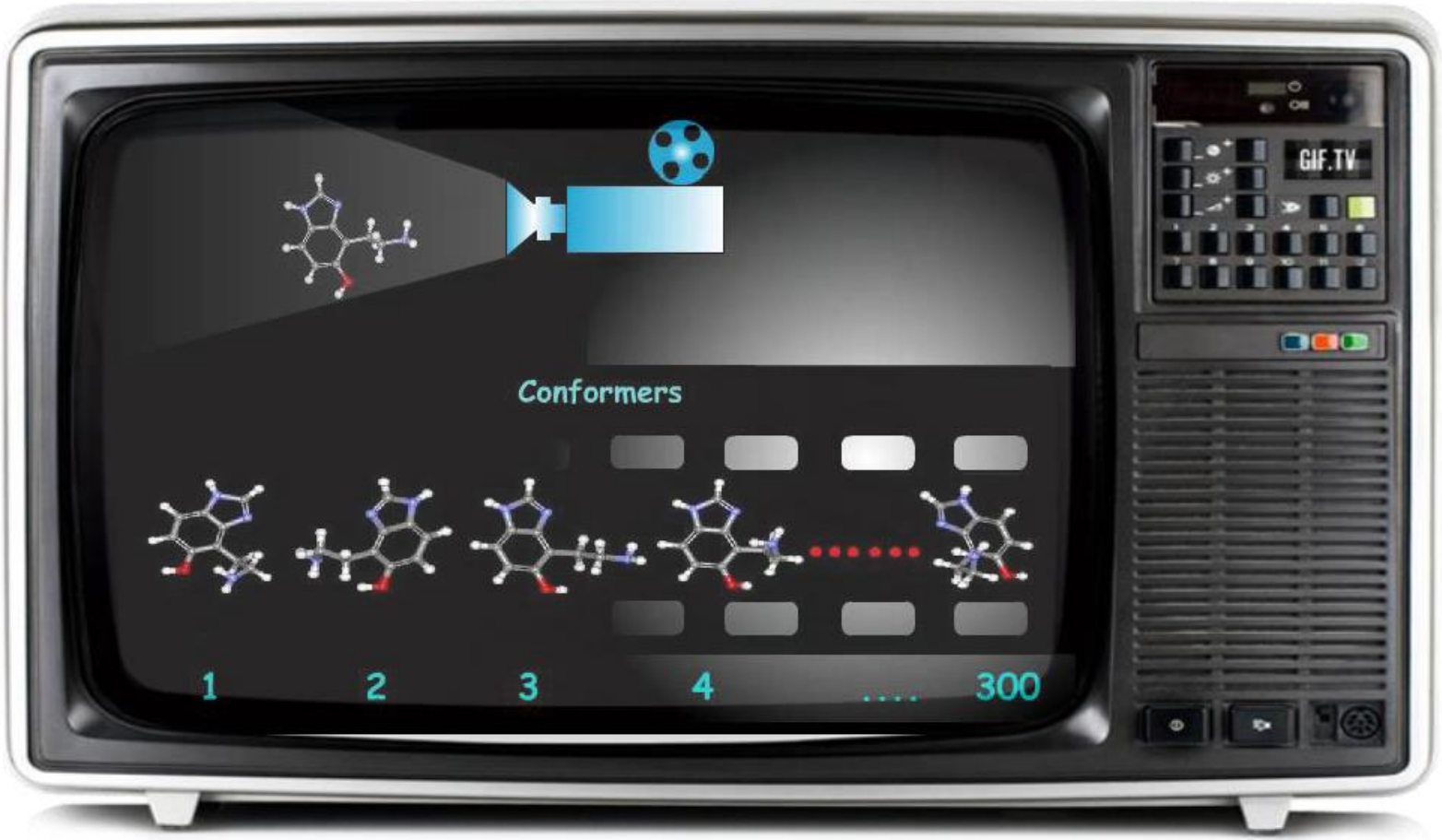


calix[4]arene



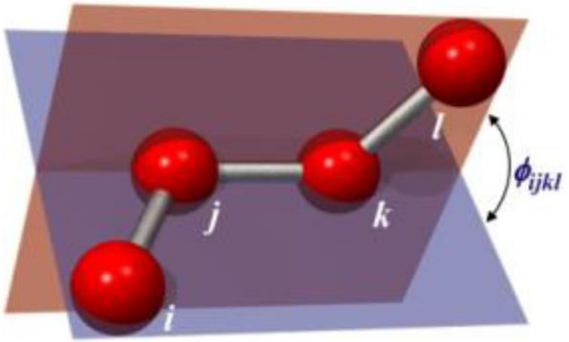
Molecular conformations

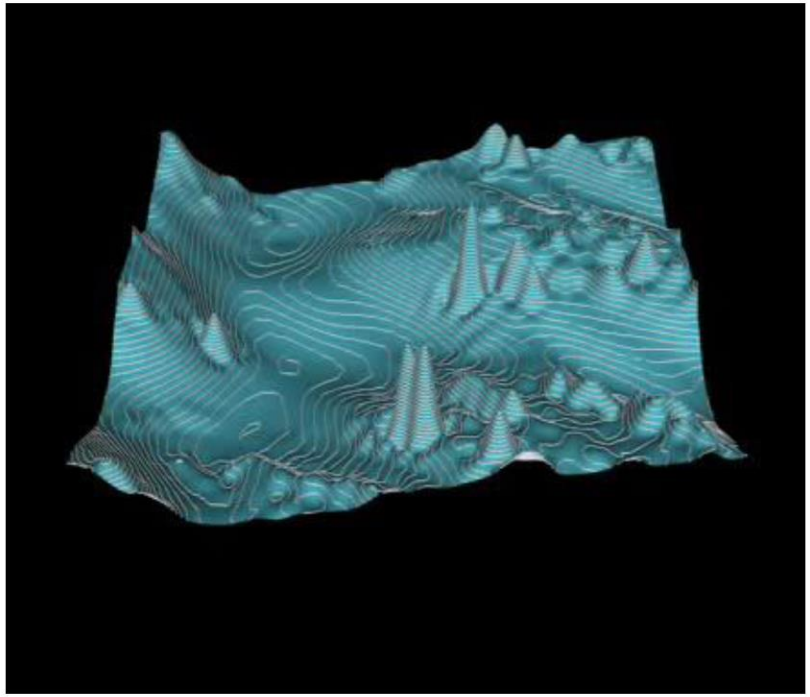
- **Molecules are dynamic structures that can change their geometries. An individual geometry is called a conformation.** A molecule consists of a set of 3D conformers in solution. Single bonds can be rotated, increasing the flexibility of a molecule. Double, triple bonds and rings reduce the flexibility.

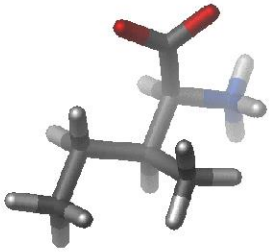


Molecular conformations

- A system of 4 consecutive atoms and 3 consecutive bonds defines two half planes. **The torsion angle is the angle between these two half planes**.







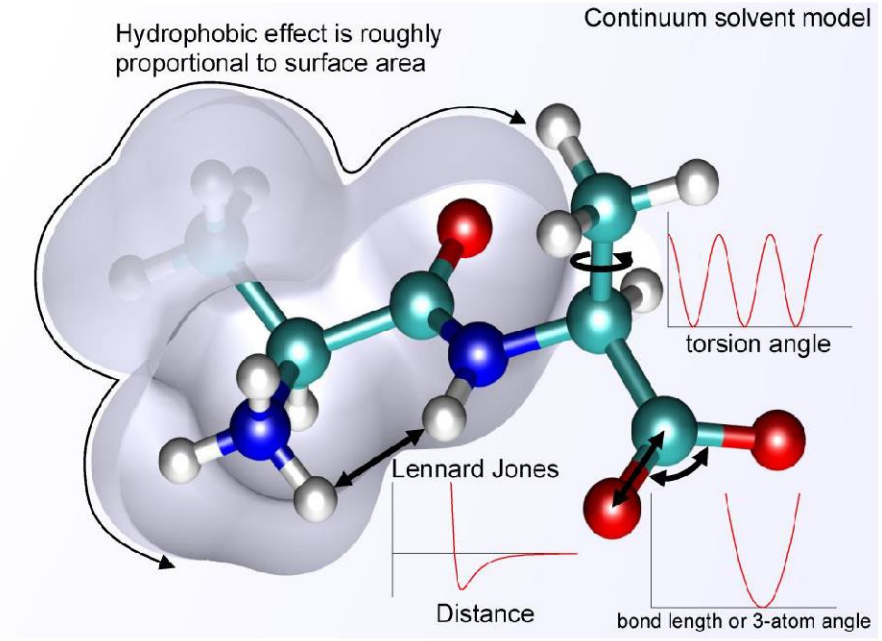
- **A molecule is a mixture of conformers in equilibrium**. All possible conformations of a molecular entity and their corresponding energy are enclosed in an **energy map**.

Energy in CADD

- A mathematical (molecular mechanics-based) equation called **force field** allows the computational simulation of forces regulating conformational changes and estimates **energy**.

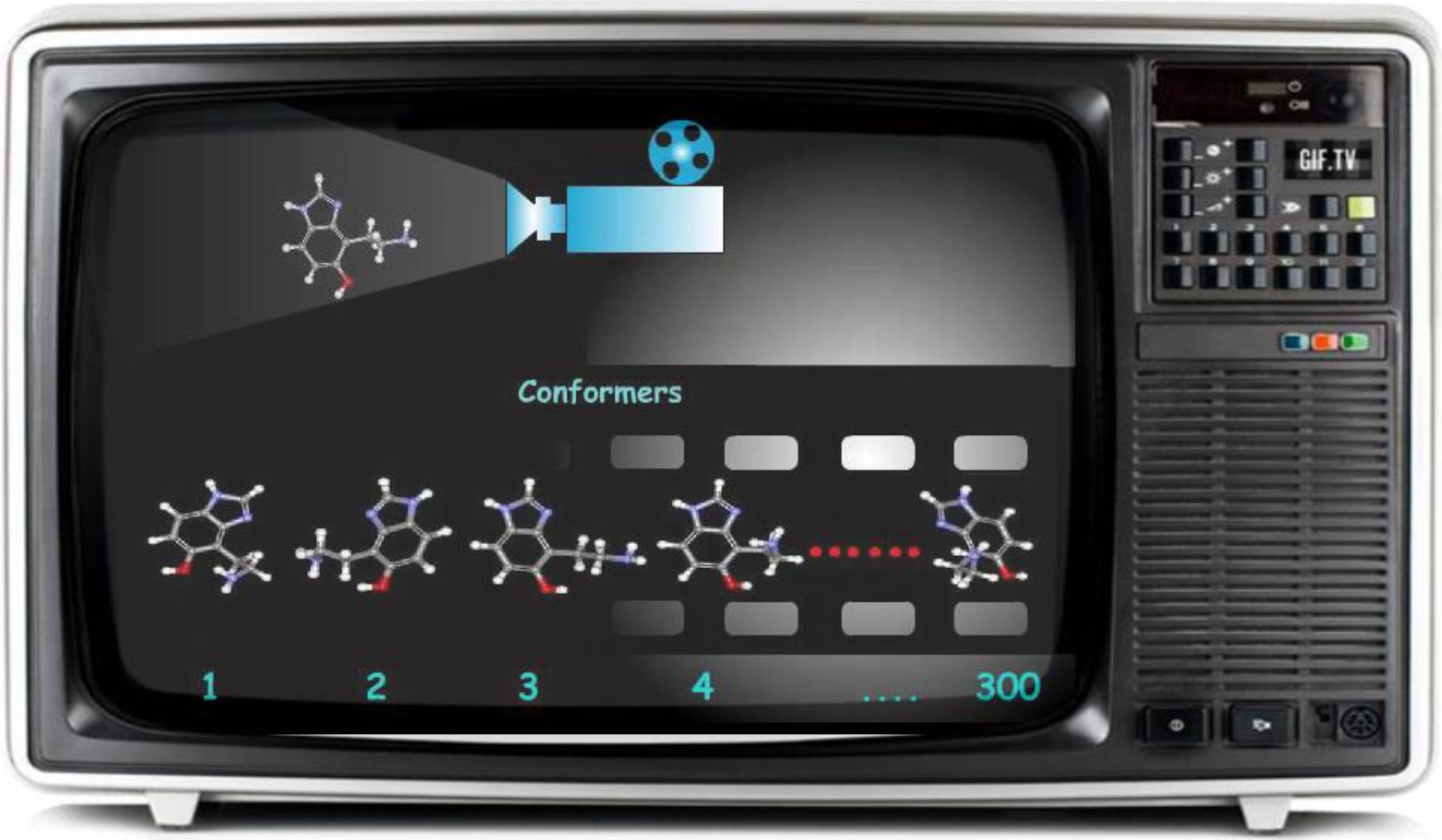
$$E_{pair} = \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 +$$

$$\sum_{dihedrals} \frac{V_n}{2}\left[1 + \cos(n\phi - \gamma)\right] + \sum_{i<j}\left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\varepsilon R_{ij}}\right]$$

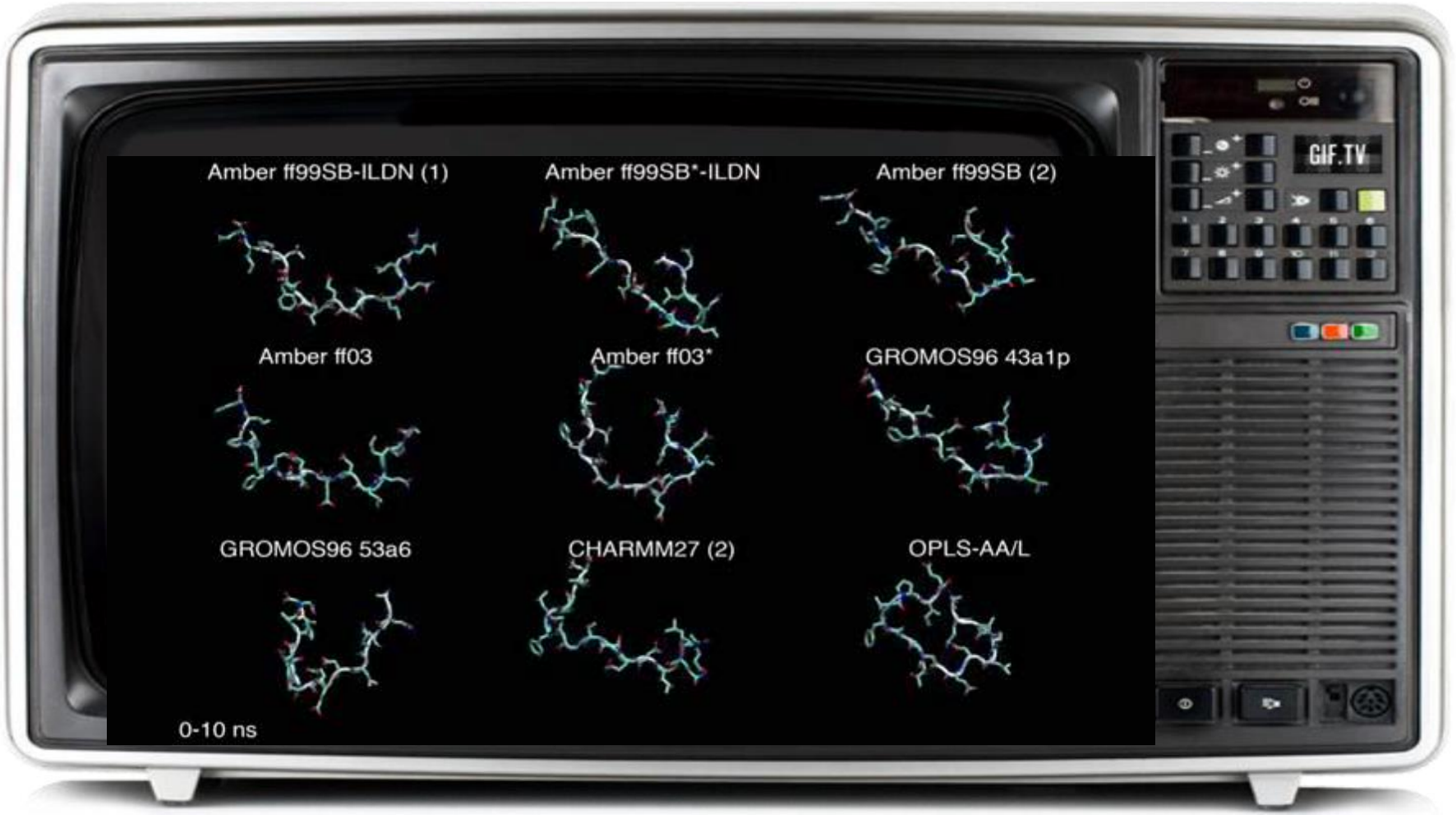


Molecular conformations

- **Molecules are dynamic structures that can change their geometries. An individual geometry is called a conformation.** A molecule consists of a set of 3D conformers in solution. Single bonds can be rotated, increasing the flexibility of a molecule. Double, triple bonds and rings reduce the flexibility.

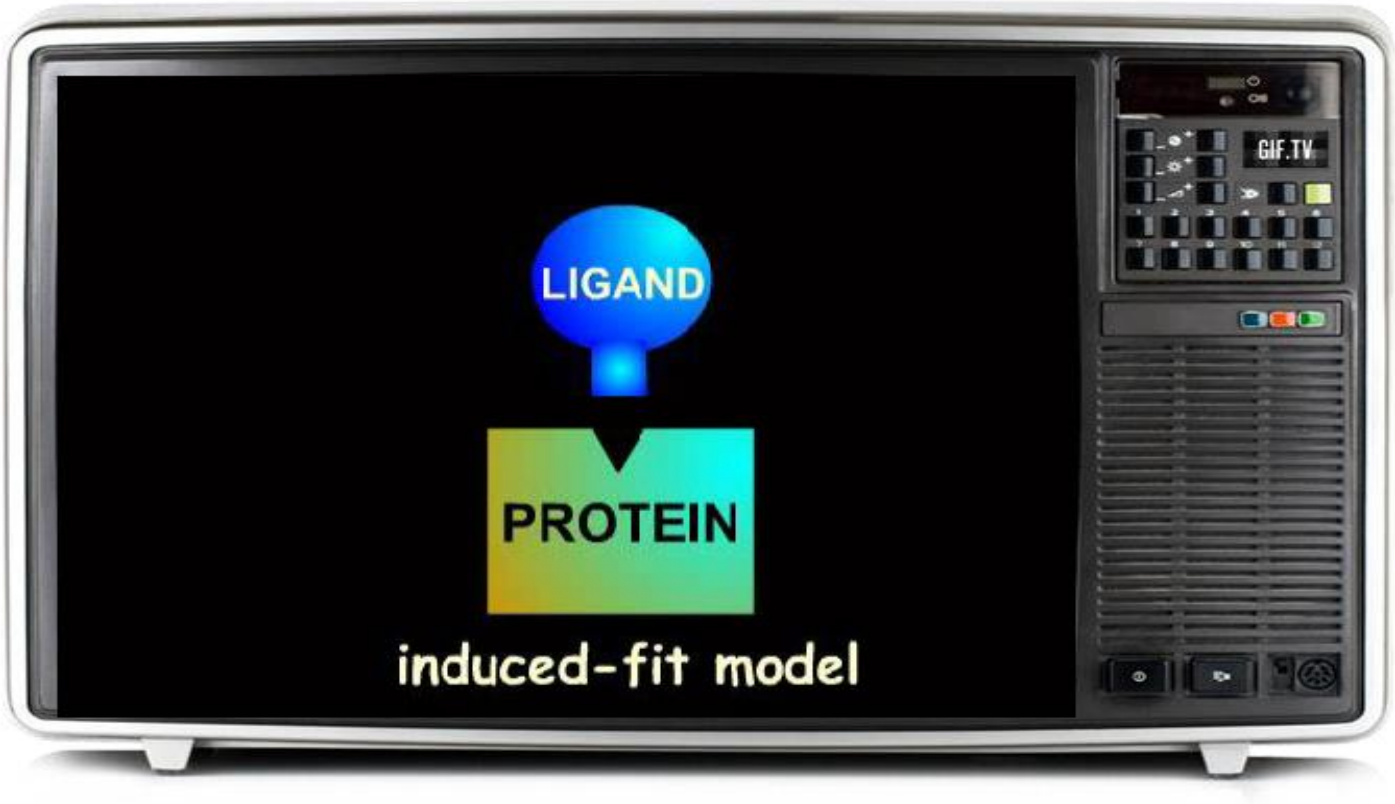


Conformational Flexibility



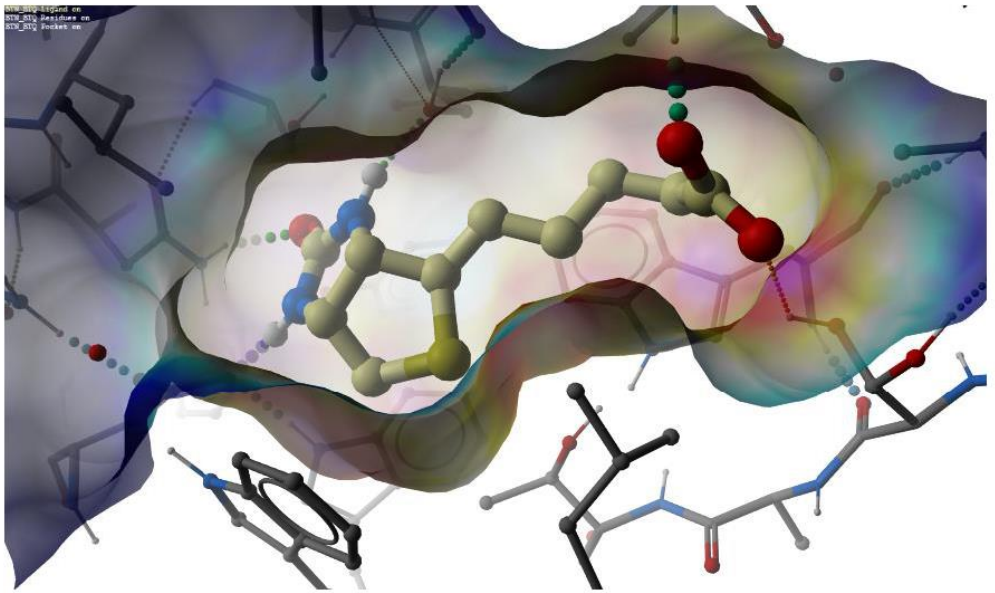
Bioactive conformation

- Daniel Koshland in 1958 introduced the "**induced fit theory**". This theory proposed that in the recognition process both ligand and target mutually adapt themselves by small conformational changes until an optimal fit is achieved.



Free energy of binding

- $\Delta G = \Delta H - T\Delta S$



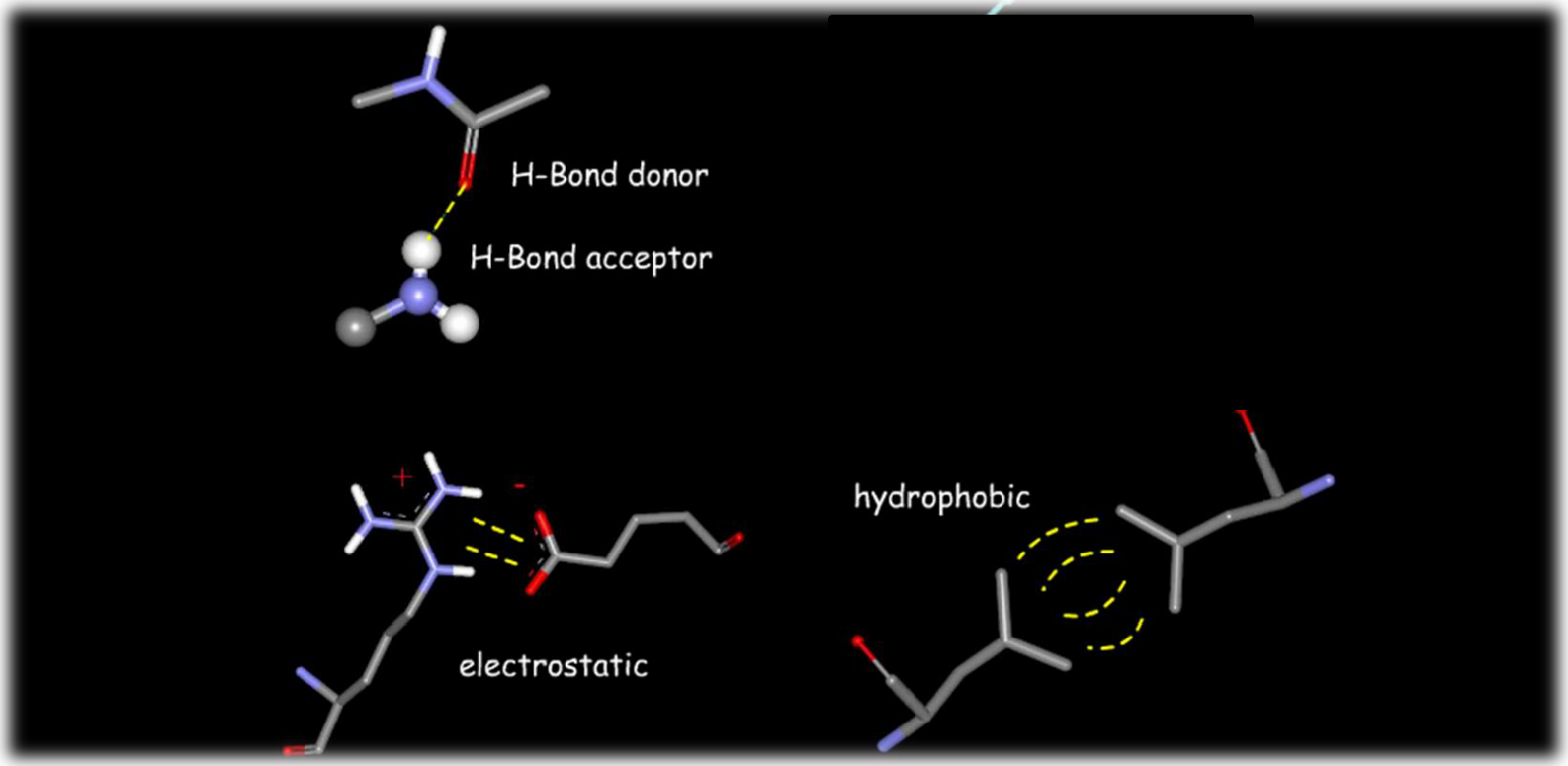
- **ΔH Entalpy:** measures the strength of the intermolecular interactions
- **TΔS Entropy**: increases disorder & costs you energy

- Water in bulk solvent is often **more disordered** than when bound to a receptor
- A ligand is often **more ordered** when bound to a receptor than in bulk solvent

Isothermal Calorimetry
Surface Plasmon Resonance

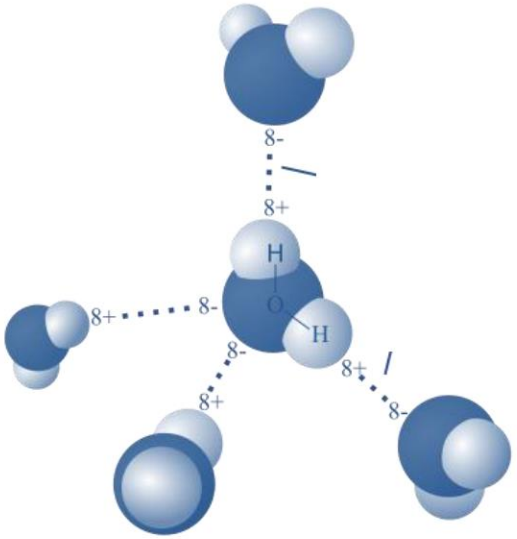
Intermolecular interactions

- Molecular interactions are responsible for the assembly of biological structures. These forces include hydrogen bonds, electrostatic interactions and hydrophobic interactions.

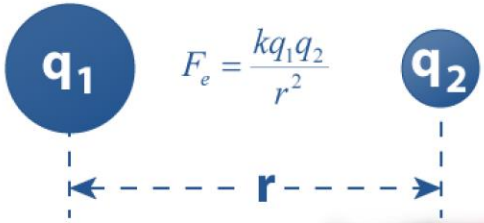


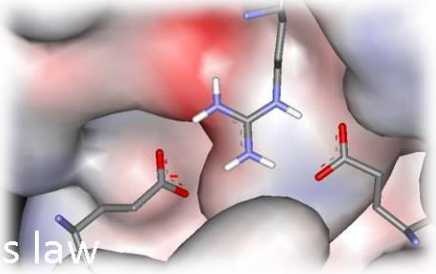
Intermolecular interactions

- A **hydrogen bond** is the electrostatic attraction that occurs when a hydrogen (H) atom, bound to a highly electronegative atom such as nitrogen (N) or oxygen (O), experiences attraction to another nearby highly electronegative atom.



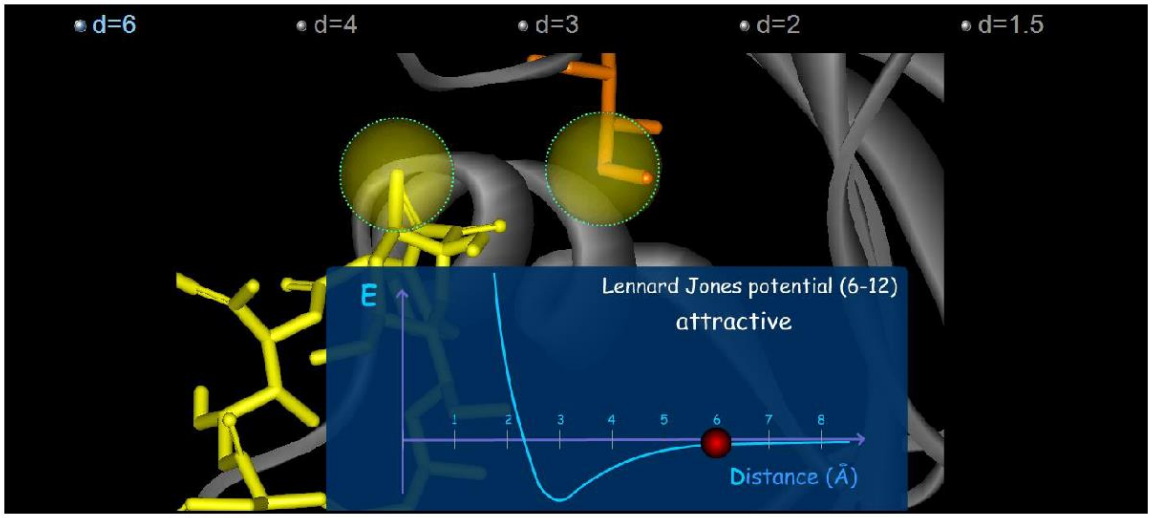
- **Electrostatic interactions** are strong forces acting between charged particles. They can be attractive (if the two charges are unlike) or repulsive (if the charges are like). The electrostatic potential can be simply described by the **Coulomb's law**.





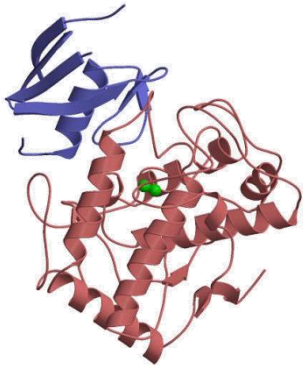
Intermolecular interactions

- Van der Waals (hydrophobic) forces between atoms are due to temporary atomic induced polarization between interacting atoms (induced dipoles). At long distances, these forces are very weak. At short distances, the vdW become strongly repulsive because of the steric clash between electron clouds.

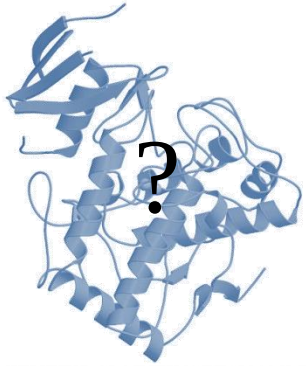


Homology Modelling

How to build an homolgy model



DSITYRVRKGDSLSSIAKRHGVNIK
DVMRWNSDTANLQPGDKLTLFVK



ANITYTIKLGDNYFIVSTTSYQNLTN
YVEMENFNPNLSPNLLPPEIKVVVP

DSITYRVRKGD--SLSSIAKRHGV-N I KDVMR-WNSDTANLQPGDKLTLFVK
ANITYT I KLGDNYFIVSTTSYQNLTNYVEMEN-FNPNLSPNLLPPE I KVVVP

Target identification

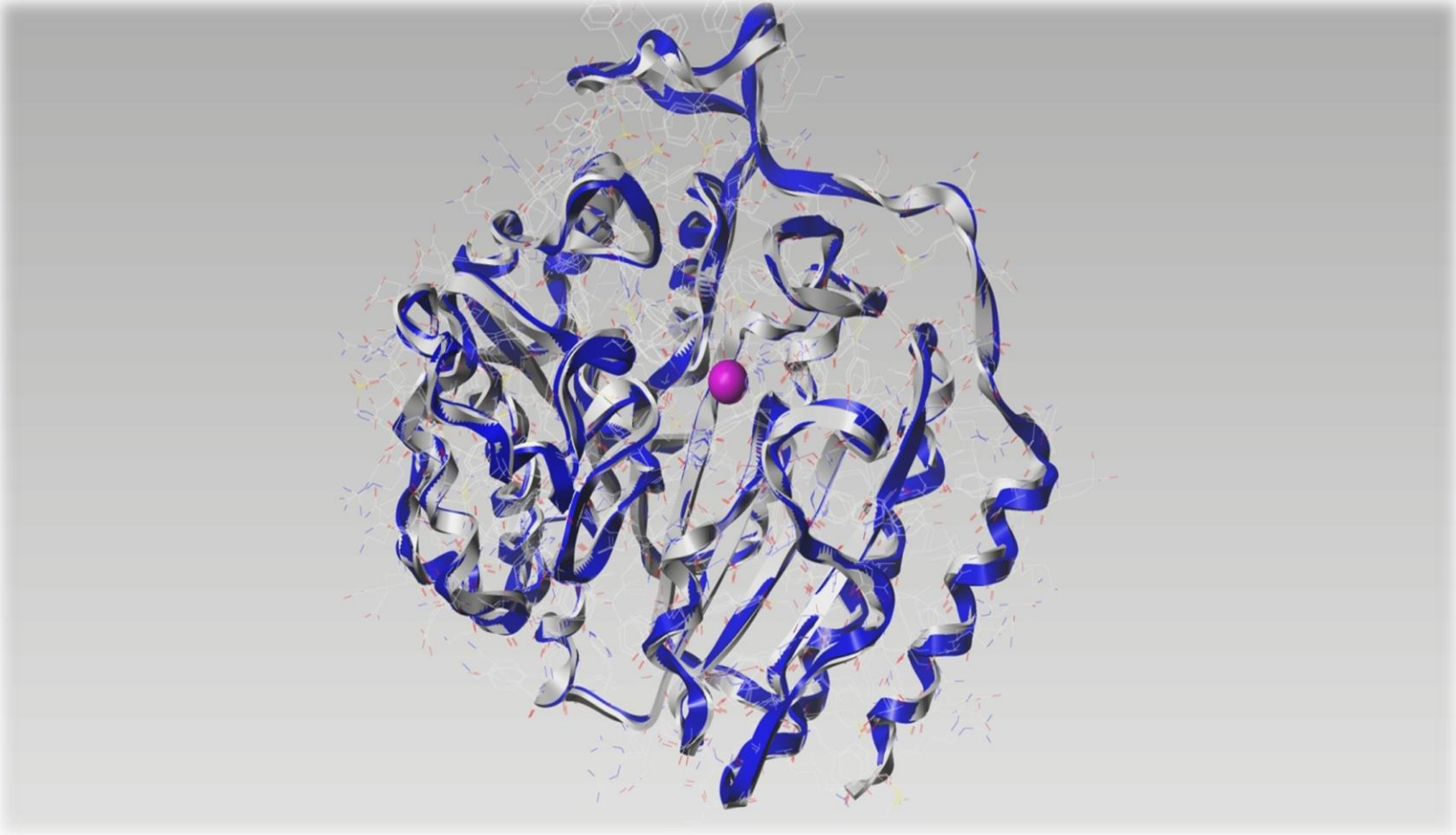
Template Selection

Target-Template Alignement

Model Construction and Refinement

Homology Modelling

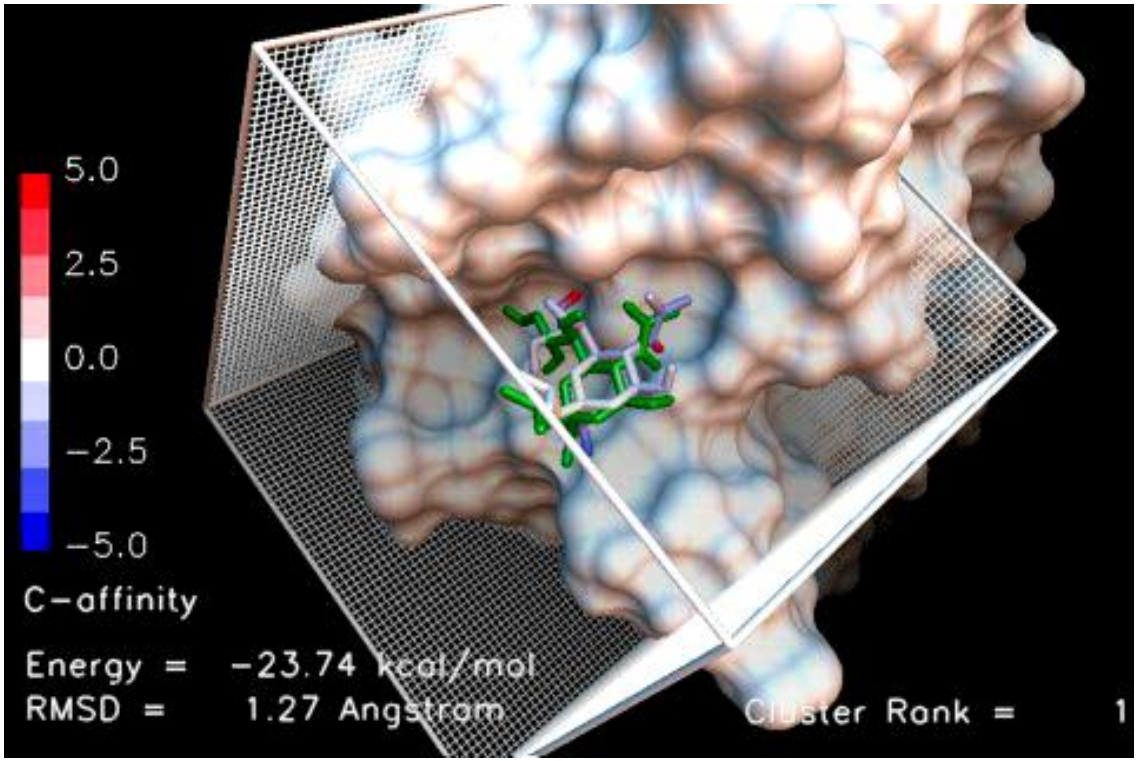
- **HDAC1 homology model**
- HDAC1 X-ray structure (PDB id 4BK,)



Molecular Docking

Search algorithm:
moves the ligand (**flexible**) into the protein pocket (**rigid**) and generates different ligand conformations.

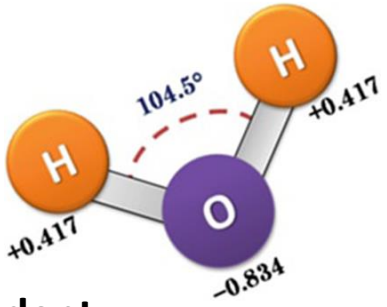
Scoring function:
valuates the quality of interactions



- Quick structural estimation of interactions
- No full flexibility
- No solvent effects
- Qualitative energy of binding

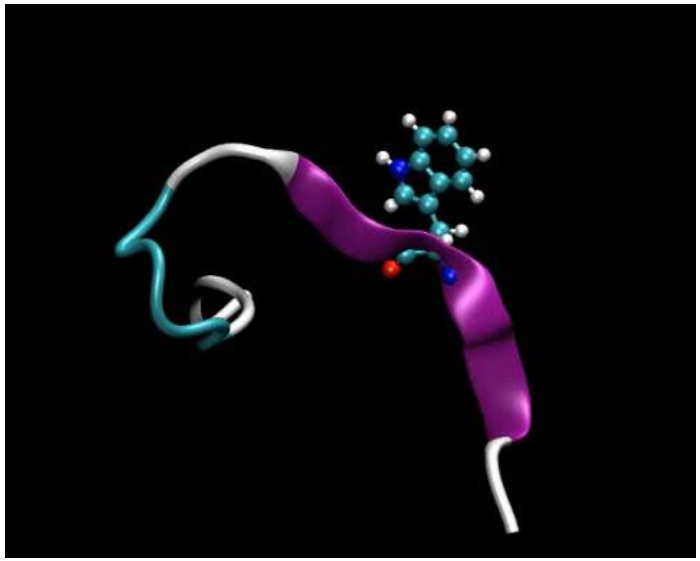
Molecular Dynamics

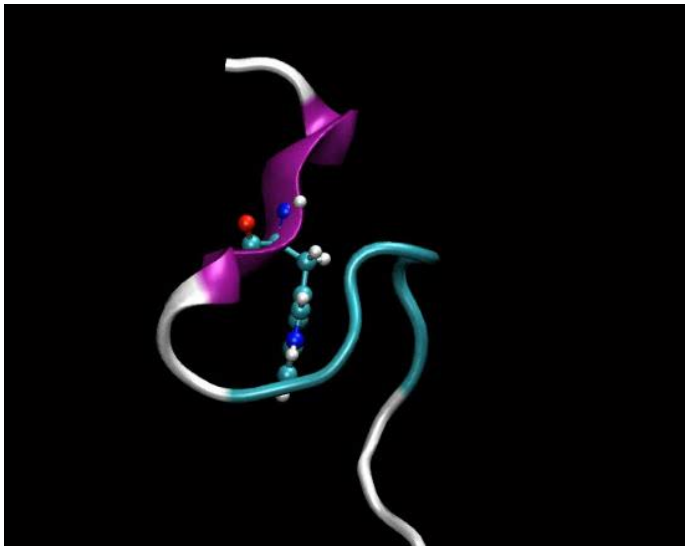
A realistic biological system is always expected to be located in a solvated environment. Systems are embedded in box of explicit solvent molecules



Generation of representative time-dependent molecular conformations (trajectories)

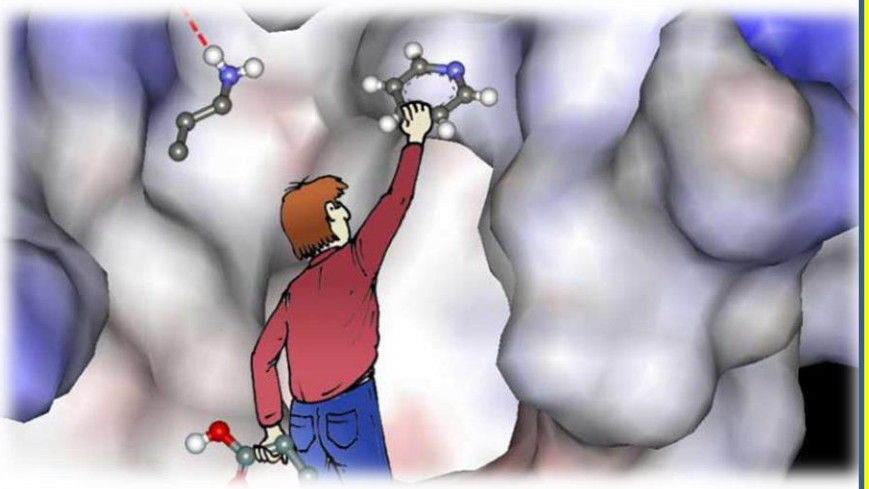
Properties calculations as a function of time. eg. The root-mean-square deviation (RMSD), the measure of the average distance between the atoms of superimposed proteins





Structure-based drug design

- **Structure-based design** allows one to use detailed 3D features of the active site by introducing appropriate functionalities in the designed ligand. The modeller can rapidly assess the validity of a possible solution and can measure the progress achieved in the course of successive design attempts.



- Crystallography
- NMR
- Homology modelling

Molecular Docking

Homology Modelling)

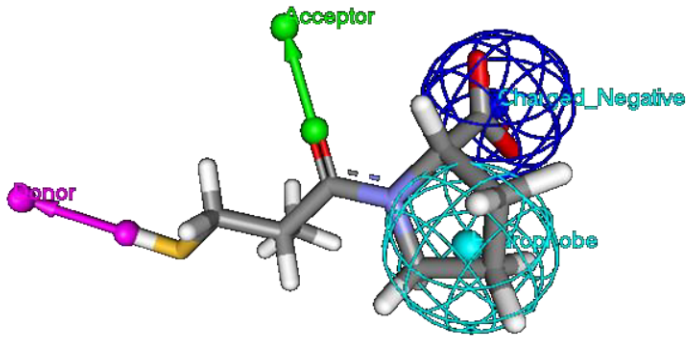
Crystallography:

Ligand-based drug design

- When the 3D structure of the target protein is not available, one can exploit the information provided by known biologically active molecules. This approach is called ligand-based, pharmacophore-based or indirect drug design.

"Similar compounds should have similar biological profiles"

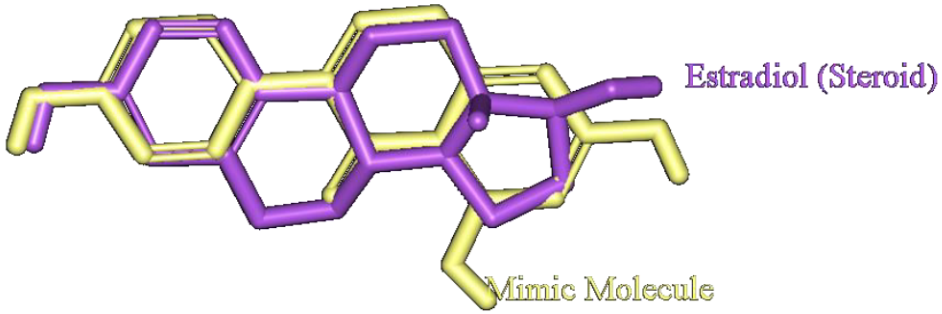
Ligand-based drug design 10 Nov. 15:30-17:30



-Ligand-based drug design → 12.11 (MC)

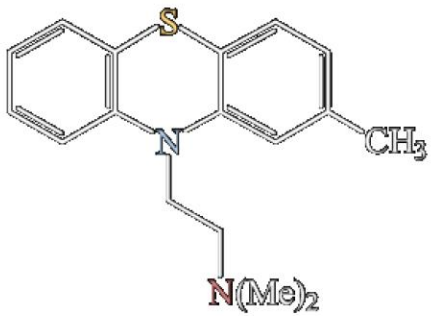
Ligand-based drug design

- The analysis of the active and inactive molecules gives a feeling of how structural variations can change biological properties and allows to generate hypotheses about the interactions of the ligands with the receptor. This strategy consists on the preservation of the structural elements recognized as necessary for the biological activity in the structure of new chemical entities.

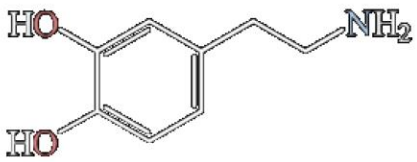


Ligand-based drug design

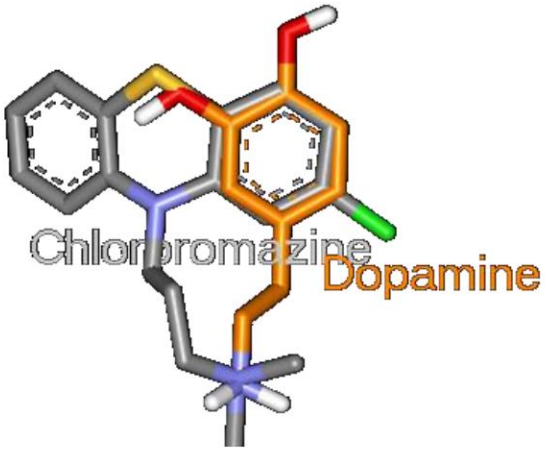
- Superimposition techniques have been very successful in elucidating the mechanism of action of drugs like the antipsychotic drug chlorpromazine. It was observed that the X-ray structures of chlorpromazine and dopamine were perfectly superimposed. On the basis of this discovery, it was suggested that anti-psychotic activity of chlorpromazine may be controlled by modulating central dopaminergic activity.



Chlorpromazine

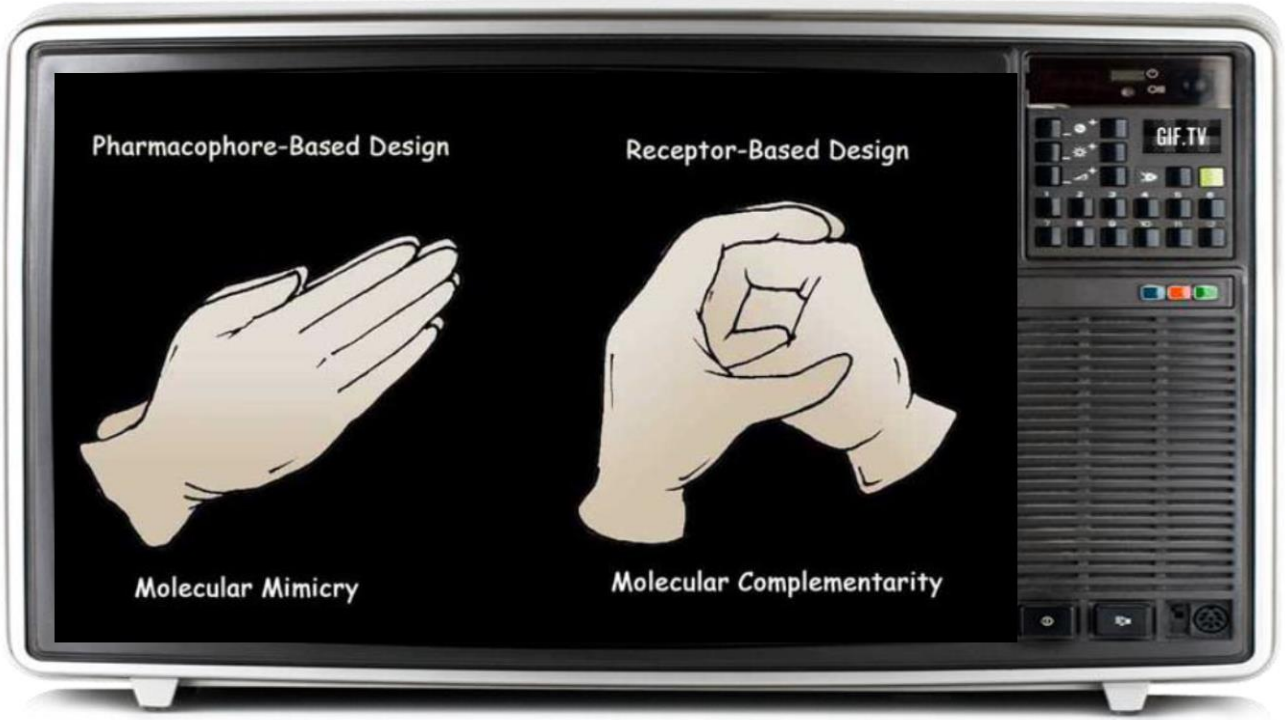


Dopamine



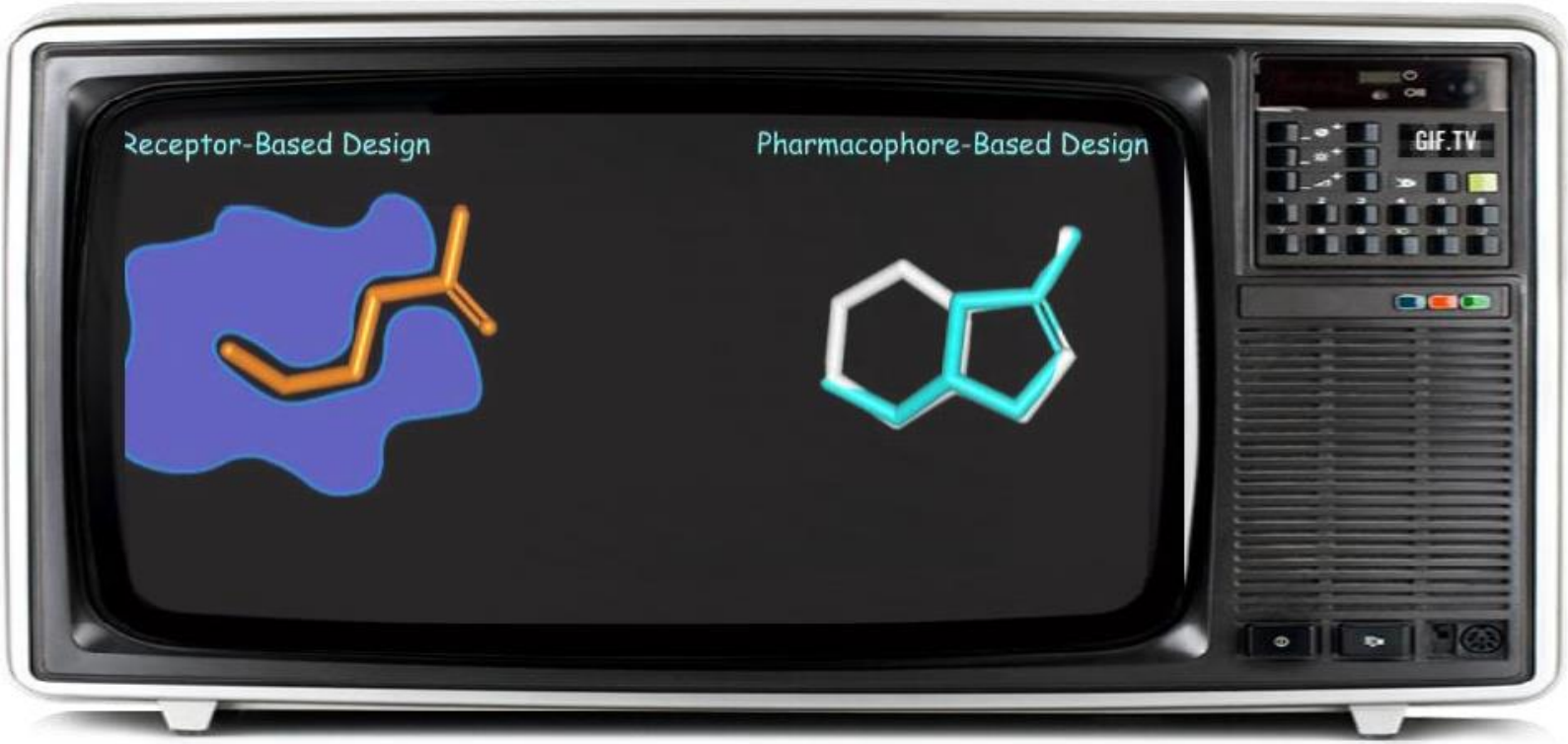
Synergy between the two methods

- When information is available for both the target protein and active molecules, the two approaches can be developed independently. In the first case, the design will concentrate on the binding to the 3D structure of the protein, and in the second case it will be based on the structures of the reference active molecules.

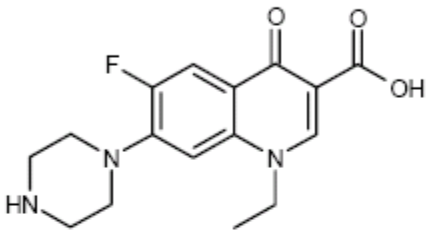


Synergy between the two methods

- When a promising docked molecule is designed and modeled with favorable interactions with the target protein, it is compared to the available active structures. Likewise, when a mimic of an active compound is considered, it is docked into the protein to see if the two approaches lead to convergent conclusions (synergy).

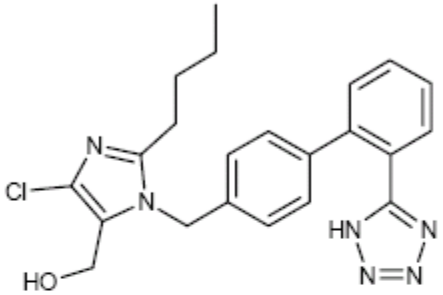


Examples of drugs designed with CADD



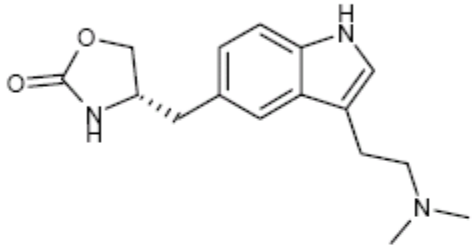
Antibiotics / Merck

Two-dimensional structure of norfloxacin.



Angiotensin II receptor antagonist / DuPont & Merck

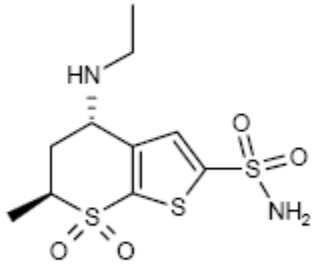
Two-dimensional structure of losartan.



5-HT1B/1D receptor agonist / Wellcome & AZ

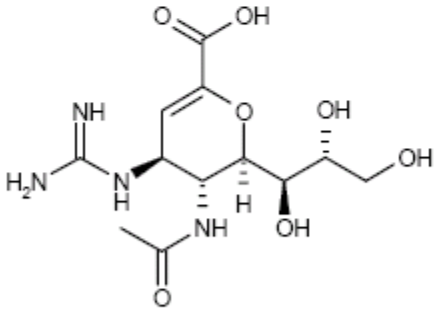
Two-dimensional structure of zolmitriptan.

L
B
D
D



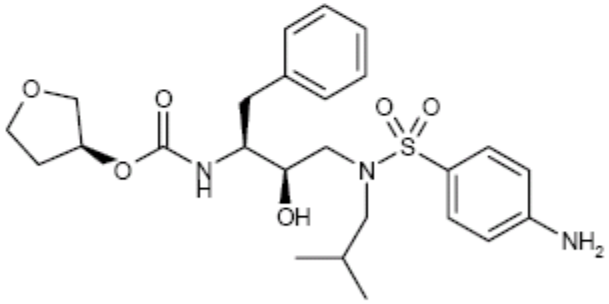
Carbonic anhydrase II / Merck

Two-dimensional structure of dorzolamide.



Neuraminidase inhibitor/ GSK

Two-dimensional structure of zanamivir.



HIV protease inhibitor/ GSK

Two-dimensional structure of amprenavir.

S
B
D
D

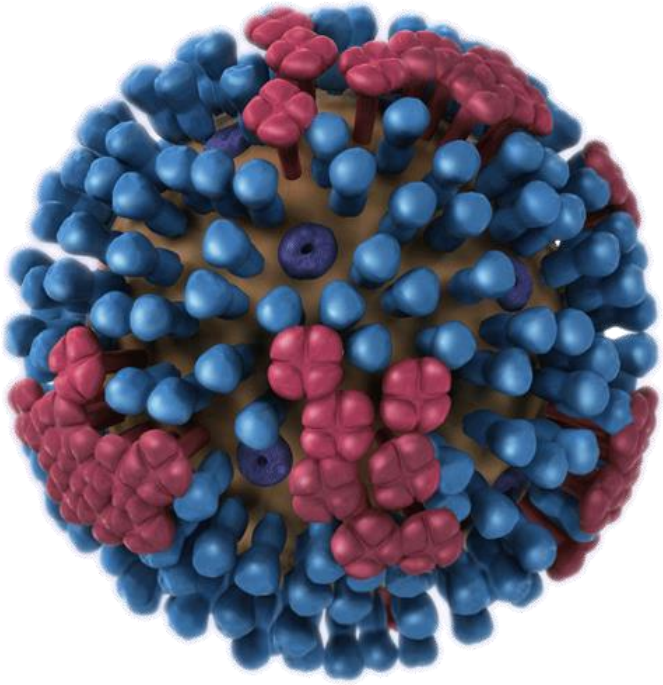
with courtesy of Dr. Richard Lewis (Novartis®)

Antibiotics design by CADD:

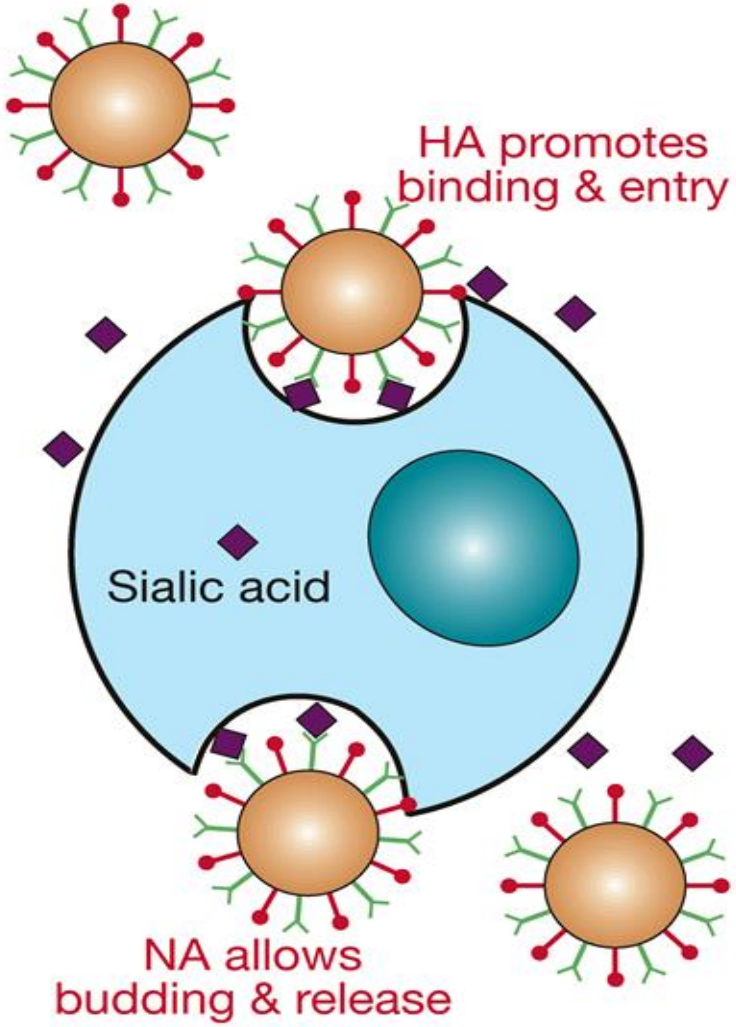
Seasonal *versus* Avian Flu

The viral strain are called according to the Surface Proteins: H1N1, H7N9…..

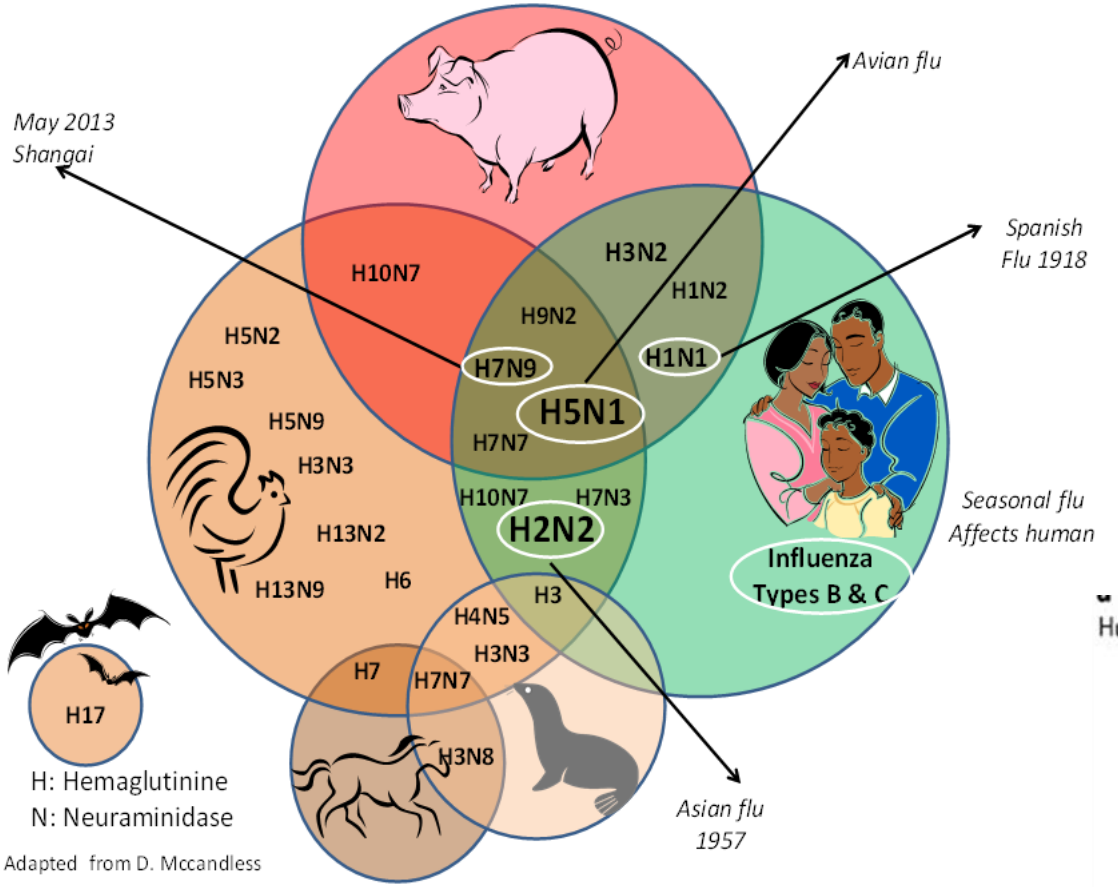
Hemagglutinin HA = H

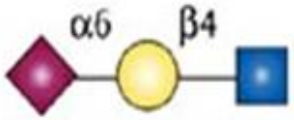


Neuraminidase = Sialidase: NA = N

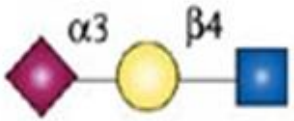


The Viral Reservoirs

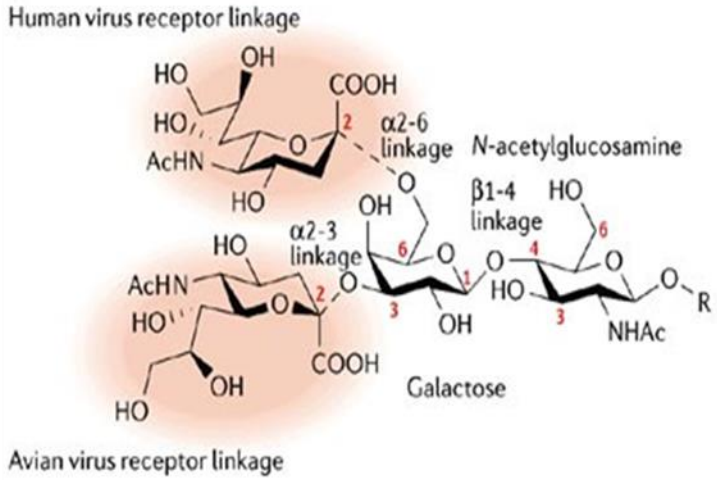




Human virus receptor linkage

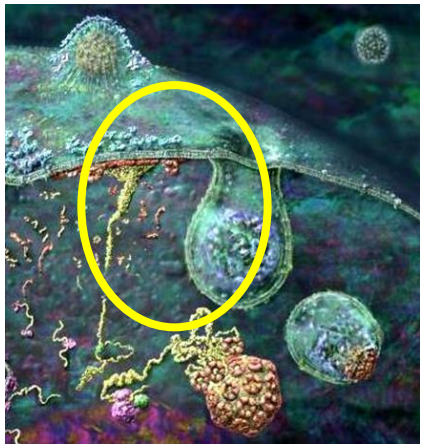


Avian virus receptor linkage

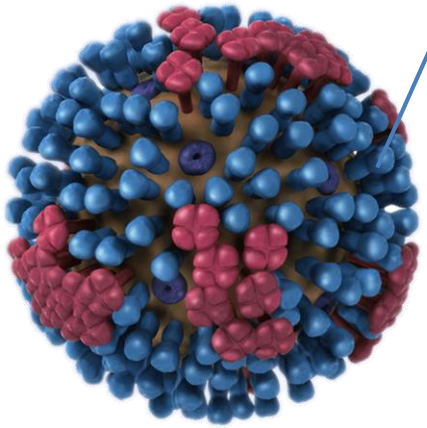


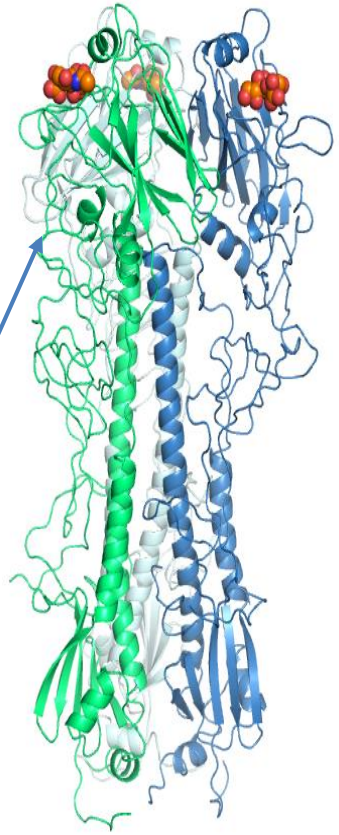
Hemagglutinin

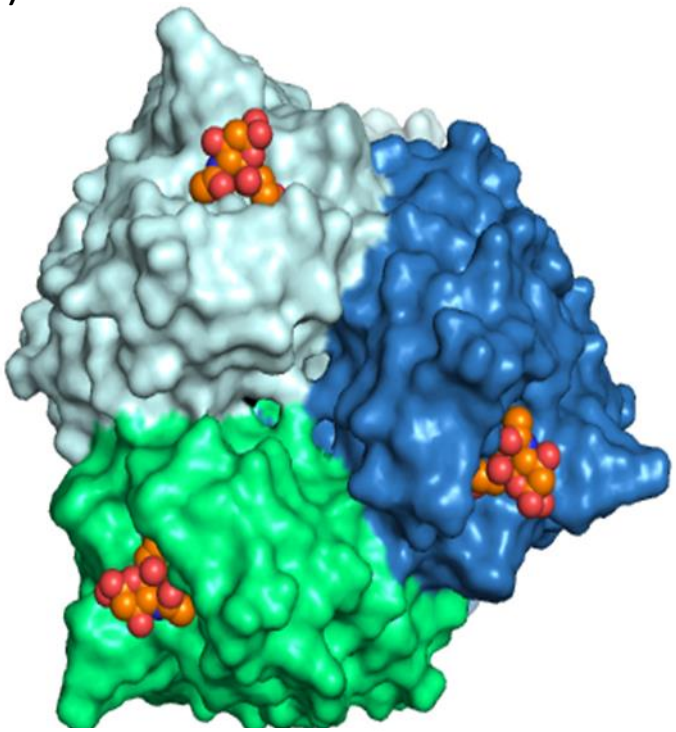
Lectin Binding to Sialic acid (N-Acetyl Neuraminic acid) on the Membrane

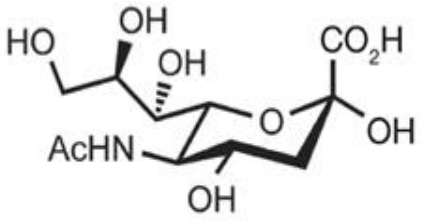


Phagocytosis of virus





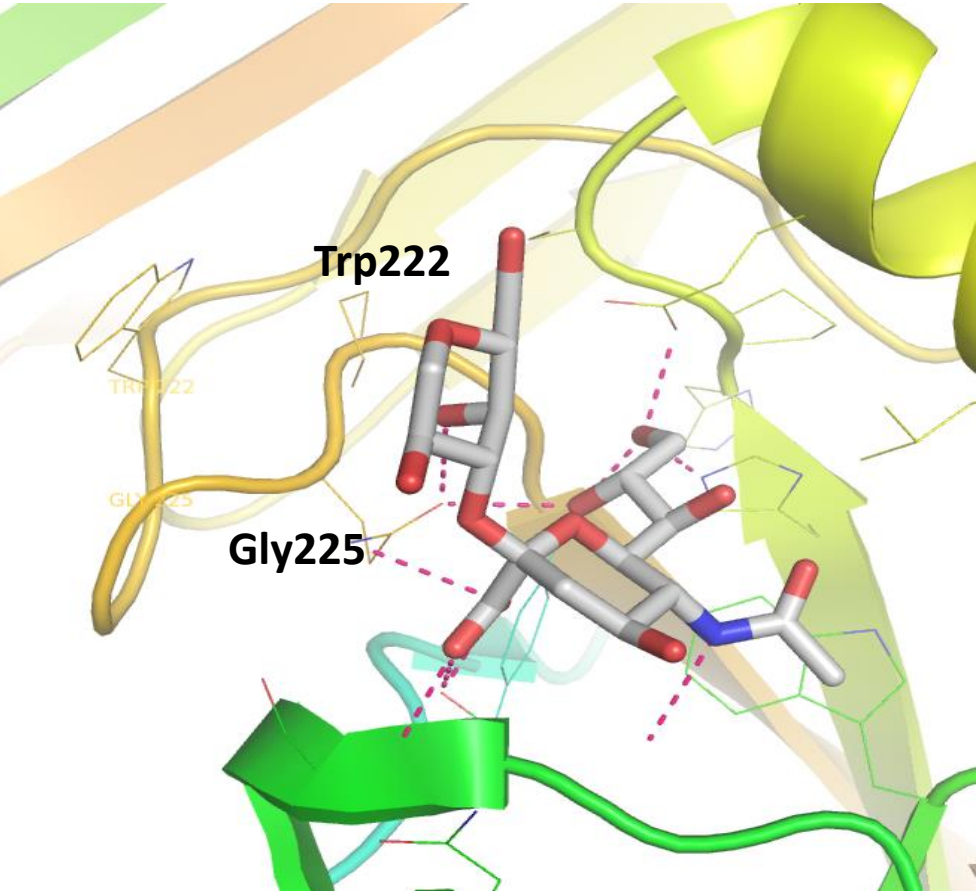


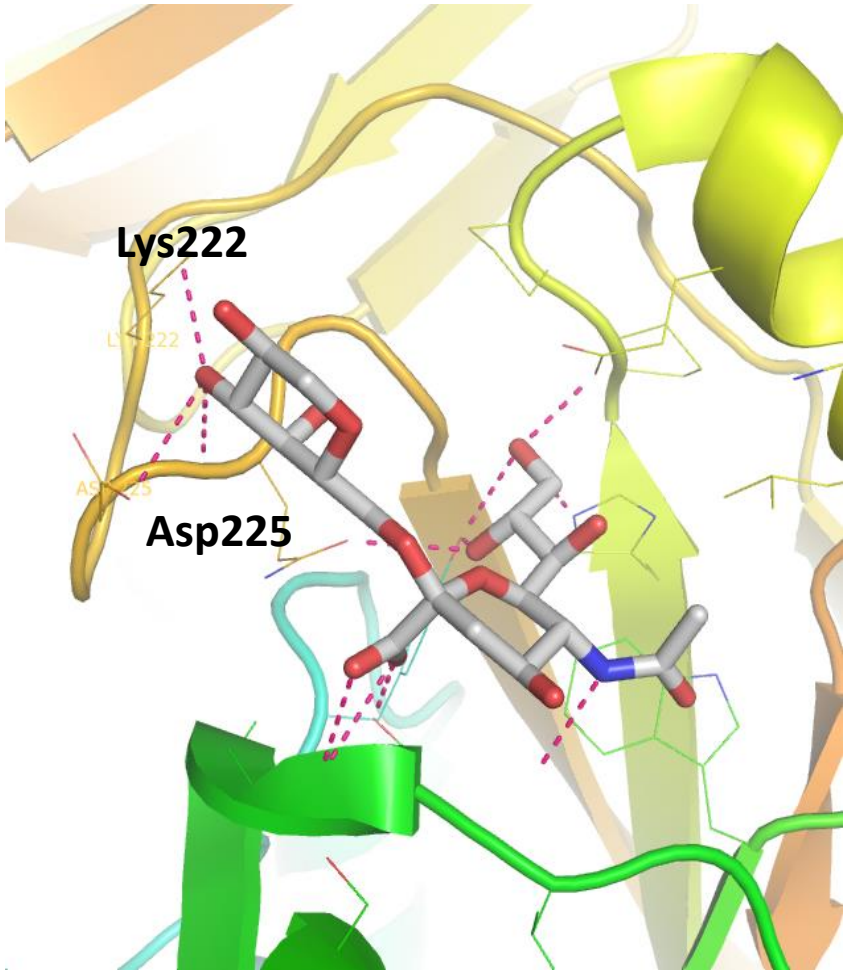


N-Acetylneuraminic acid
(NeuAc)

Avian *versus* Human Hemagglutinin

Comparison of the Binding sites



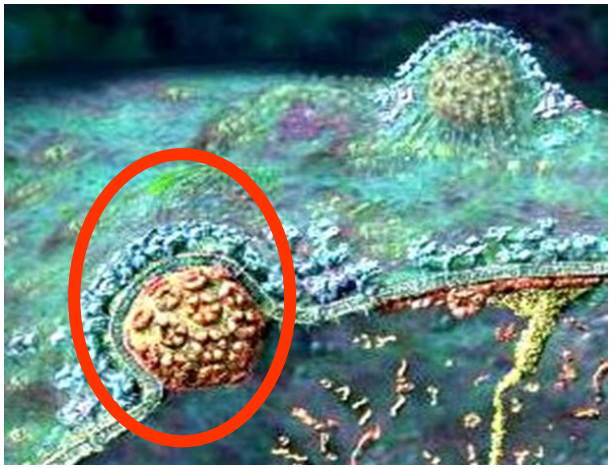


H3 avian – precursor Hong-Kong 1968

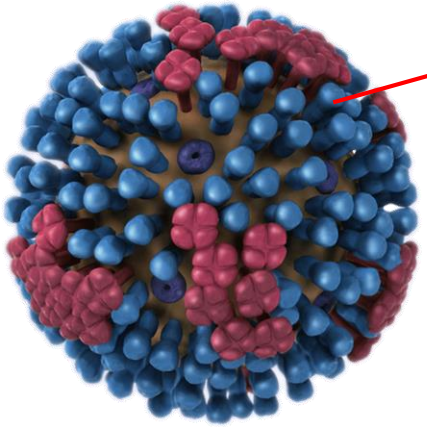
H1 human – 1918

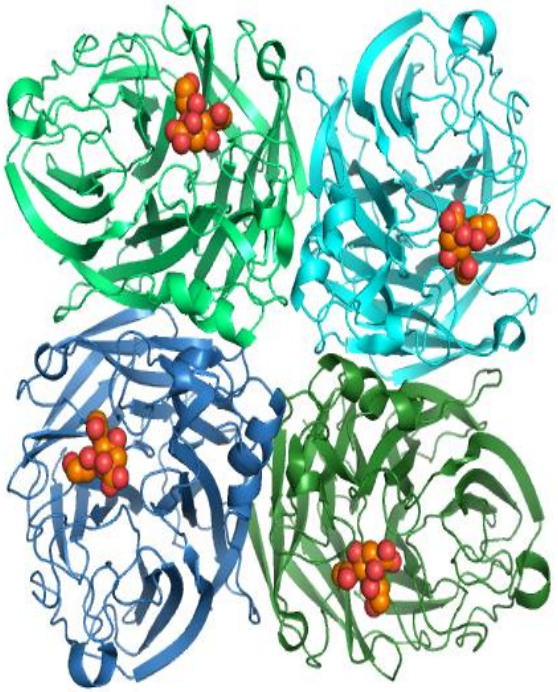
Neuraminidase

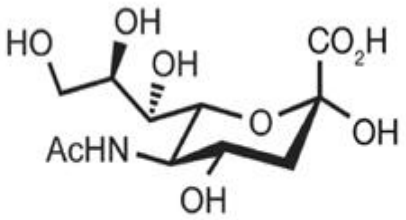
Cleaves Sialic acid on our Cells



Exocytosis of virus



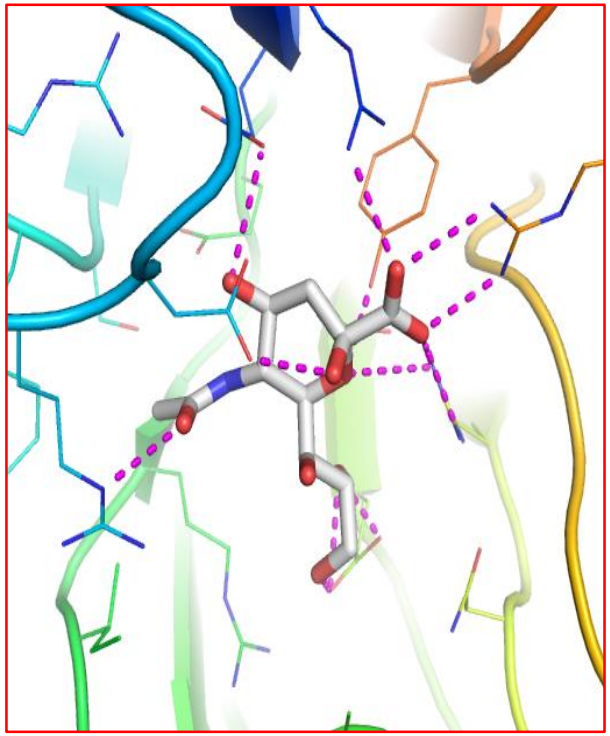




N-Acetylneuraminic acid
(NeuAc)

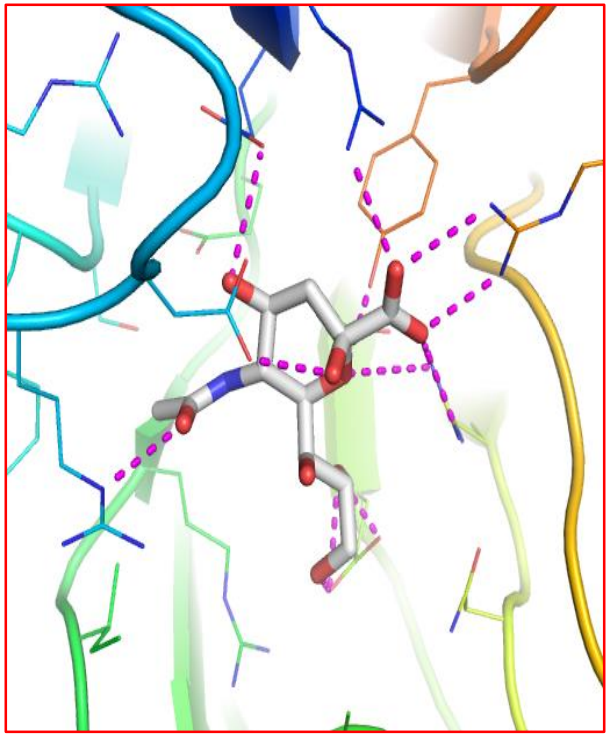
Conception of Neuraminidase Inhibitors

N1 with sialic Acid



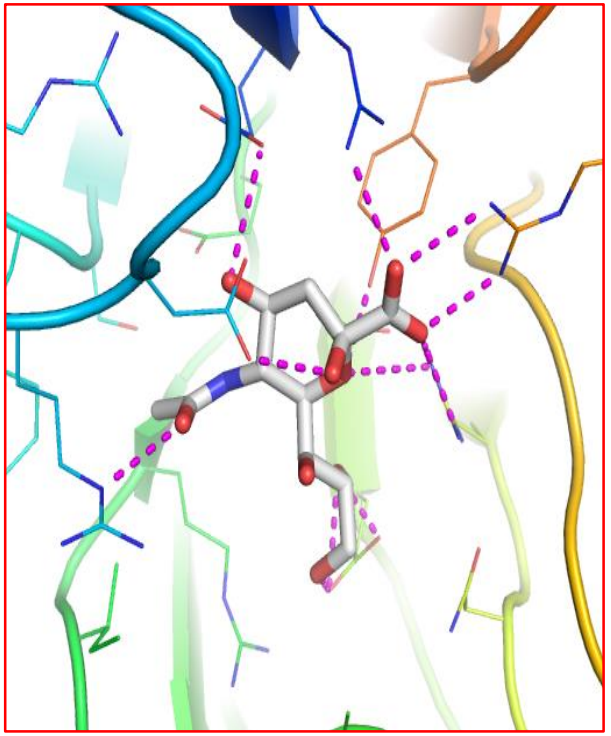
Varghese *et al*., Proteins 1992

N1 / oseltamivir (Tamiflu)



Russel *et al*., Nature 2006

N1 / anamivir (Relenza)



Wu *et al*., J. Virol. 2008

Leading to Drugs

Big challenges in CADD

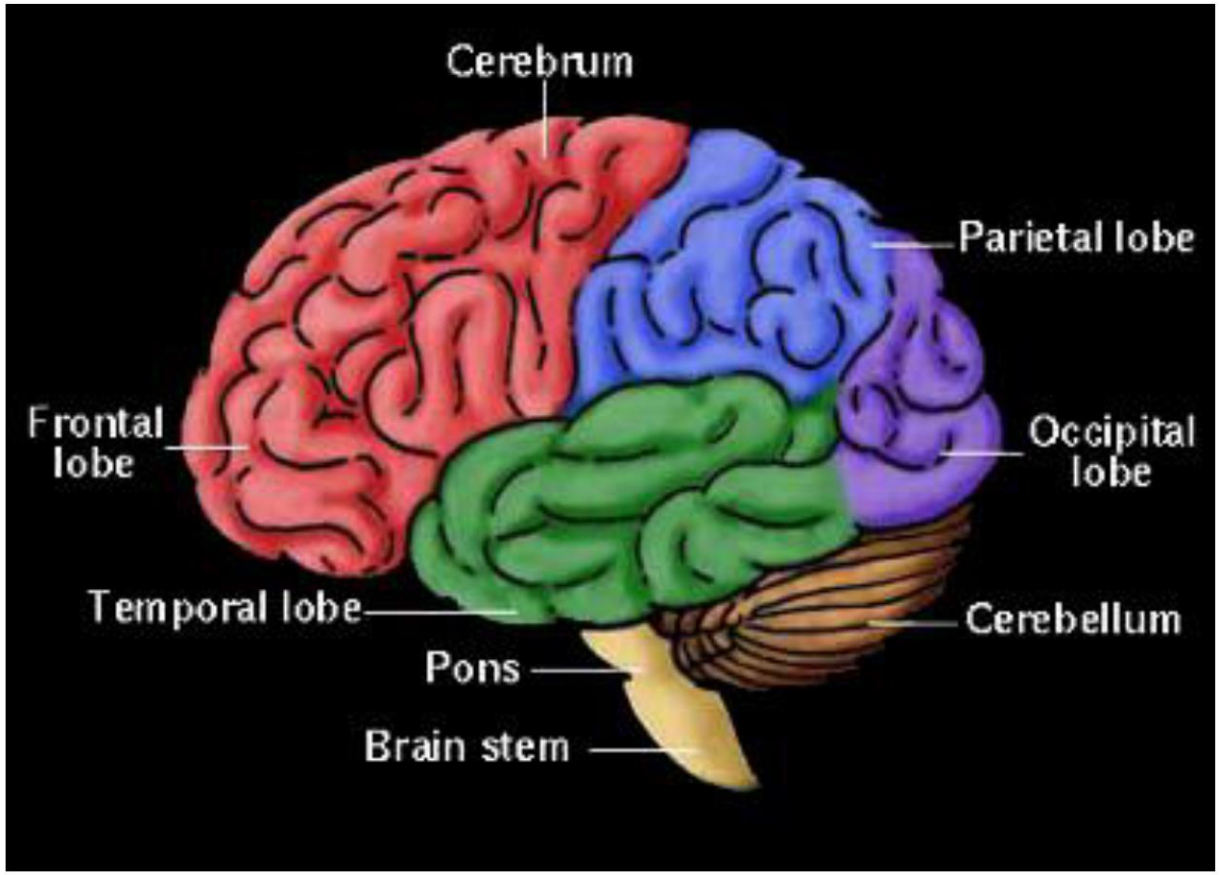
- **Thermodynamics**
  - Accurate binding free energy calculation (…and then, do it fast!)
  - Prediction of melting point (hence solubility)
  - Prediction of pKa
  - Energy distribution of molecular conformations in solution
  - Understanding entropy
- **Solvation & modeling water**
- Force-field: comprehensiveness, accuracy, speed
- **Predicting protein fold & conformation from sequence**
  - More accurate protein homology modeling
- **Incorporation of protein dynamics into design of modulators**
- **Targeting membrane proteins:** membrane models, lack of structure…

Take home message

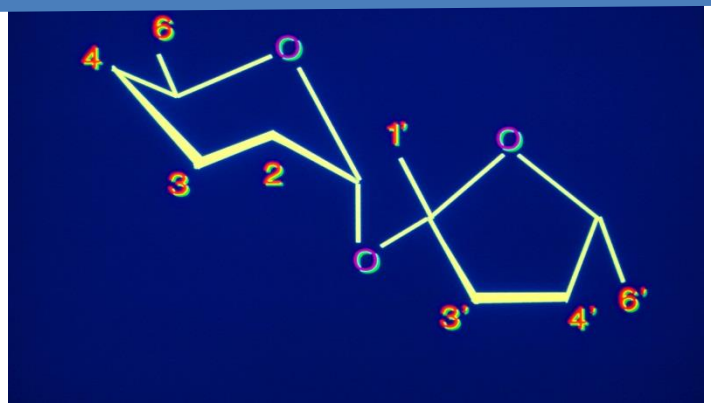
- Drug discovery is most effective when it is a partnership across several disciplines.
- There are still several significant challenges we need to solve to make our understanding more rigorous, and hopefully more useful.
- It is not just about making more compounds, or screening them faster.
- It is about doing better science on live projects with real data that test our assumptions to destruction, which one day may lead to a treatment for the patient.
- This is why it is so much fun to work in Pharma research!

Take home message

The most potent drug design tool …

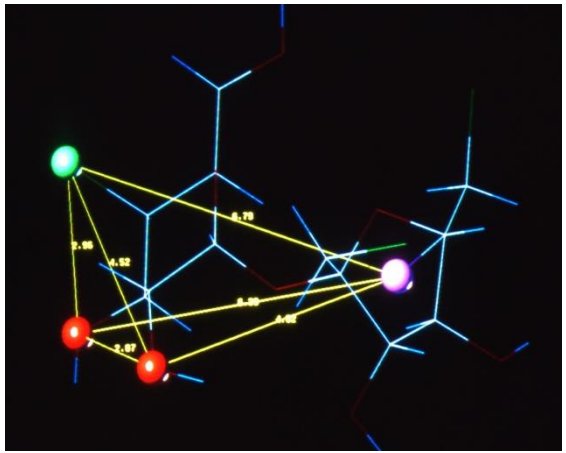


Sweet Taste Perception



2	3	4	6	1'	3'	4'	6'	Sw
-	-	-	-	-	-	-	-	1
-	-	-	-	Cl	-	-	-	20
-	-	-	-	-	-	-	Cl	20
-	-	-	Cl	-	-	-	-	bitter
-	-	Cl	-	-	-	-	-	5
-	-	-	-	Cl	-	Cl	-	30
-	-	-	-	Cl	-	-	Cl	76
-	-	-	Cl	-	-	-	Cl	bitter
-	-	Cl	-	Cl	-	-	-	120
-	-	-	-	Cl	-	Cl	Cl	100
-	-	-	Cl	Cl	-	-	Cl	25
-	-	Cl	Cl	-	-	-	Cl	4
-	-	Cl	-	Cl	-	-	Cl	650
-	-	Cl	-	Cl	-	Cl	-	220
-	-	Cl	-	-	-	Cl	Cl	160
-	-	F	-	F	-	-	F	40
-	-	Br	-	Br	-	-	Br	800
-	-	I	-	I	-	-	I	120

2	3	4	6	1'	3'	4'	6'	Sw
-	-	Cl	Cl	Cl	-	-	Cl	100
-	-	Cl	Cl	Cl	-	-	Cl	200
-	-	Cl	-	Cl	-	Cl	Cl	2200
-	-	Cl	-	Cl	-	Cl	Cl	200
Cl	-	-	Cl	Cl	-	-	Cl	bitter
-	-	Br	-	Br	-	Br	Br	7500
-	-	Cl	H	Cl	-	-	Cl	400
-	-	Cl	OMe	Cl	-	-	Cl	500
-	-	Cl	OiPr	Cl	-	-	Cl	0
-	-	Cl	-	Cl	-	F	Cl	1000
-	-	Cl	-	Cl	-	Br	Cl	3000
-	-	Cl	-	Cl	-	I	Cl	3500
-	-	F	-	Cl	-	Cl	Cl	200
-	-	Br	-	Br	-	Br	Br	7000



Sweet Taste Perception