CABANA

Capacity building for bioinformatics in Latin America

CABANA Workshop:

Chemoinformatics in Drug Discovery

Introduction to Structure-based drug discovery and molecular docking.

Edgar, López-López

edgar.lopez.593@hotmail.com

CINVESTAV

Unidad Irapuato

DIFACQUIM

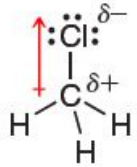
UNAM

DIFACQUIM

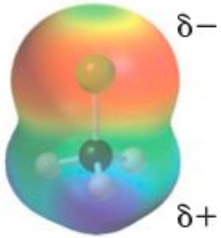
COMPUTER-AIDED DRUG DESIGN AT UNAM

Introduction to Structure-based drug discovery and molecular docking.

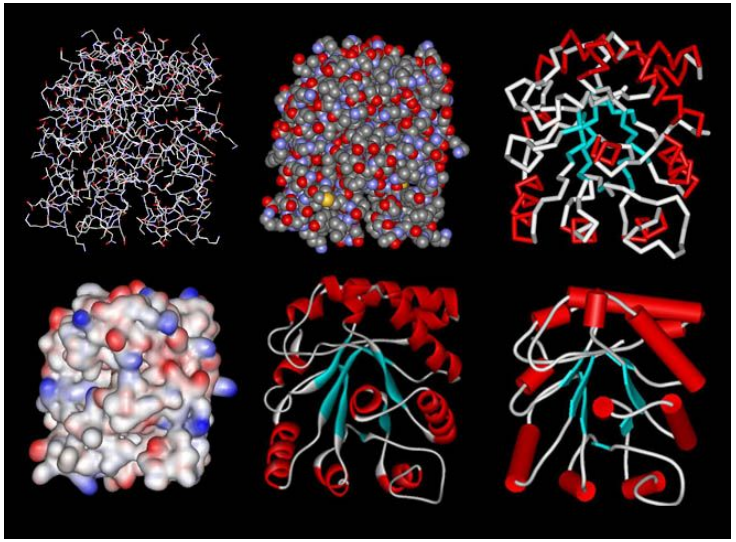
Molecular modeling.



Chloromethane

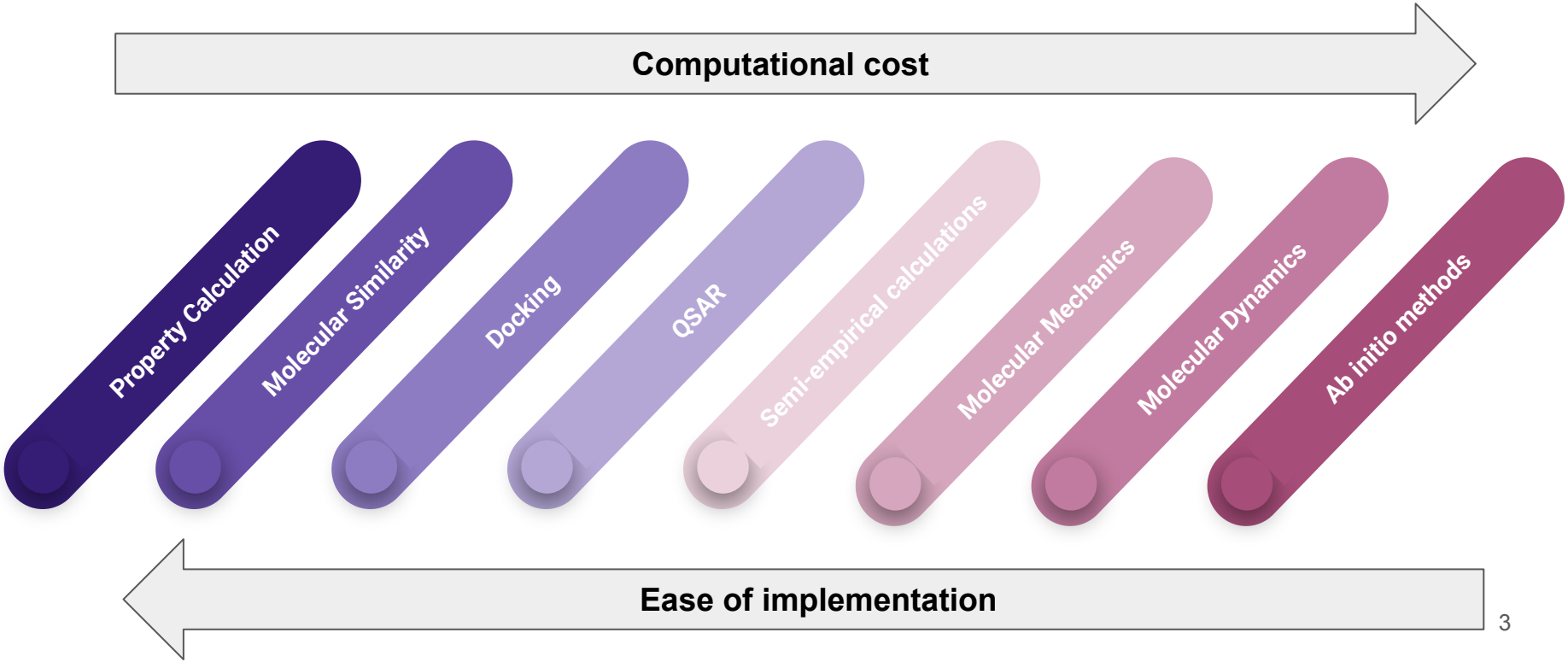


Electrostatic potential map of chloromethane



Introduction to Structure-based drug discovery and molecular docking.

Computational Chemistry.



Introduction to Structure-based drug discovery and molecular docking.

Docking.

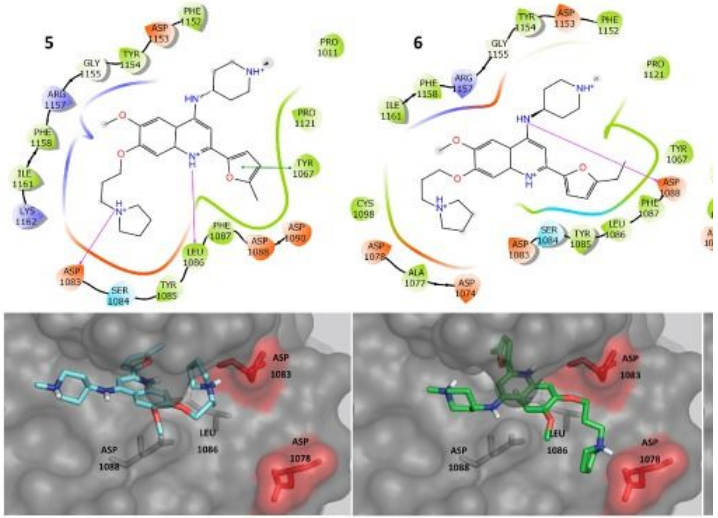
1. Generalities
2. Search algorithms
3. Scoring function
4. Post Processing

Introduction to Structure-based drug discovery and molecular docking.

Docking.

- **Haptoforic region:**
  Allows ligand binding and orientation.

- **Pharmacophoric region:**
  Allows specific recognition / catalytic activity.

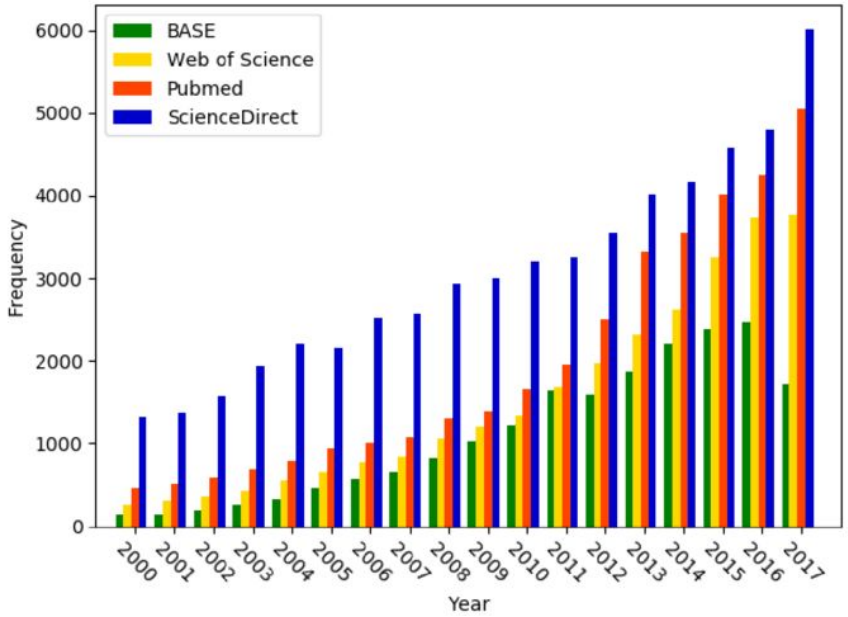


Molecules **2018**, *23*(12), 3282

Introduction to Structure-based drug discovery and molecular docking.

Docking.

Is it still worth doing docking?



TIP **2018**, *21(1)*, 1-23

Why do docking?

1. It helps optimize time and resources.
2. It is a good approximation (if done correctly).
3. Rational design aid.
4. Provide preliminary information.

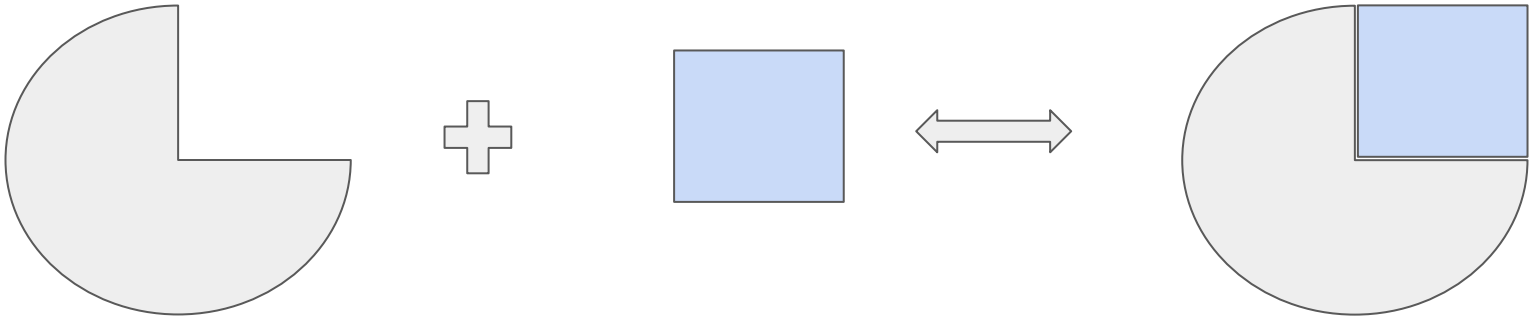
What can't you do the docking?

1. Isimulate covalent reactions (QM / MM).
2. Model complex systems.
3. Consider a solvated environment (MM PBSA / GBSA).
4. **Generate real inhibition values.**
5. **Distinguish between a pair of enantiomers.**
6. **Discover a drug.**

Introduction to Structure-based drug discovery and molecular docking.

Docking.

The cornerstone of docking!



$[E]_{aq}$ $[I]_{aq}$ $[E+I]_{aq}$

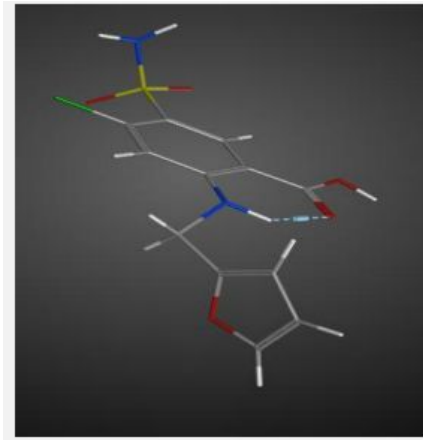
$$K_A = \frac{1}{K_i} = \frac{[EI]}{[E][I]} \Rightarrow -RTlnK_A = \Delta G \Rightarrow \Delta G = \boxed{\Delta H} - T\boxed{\Delta S}$$

$$\Delta G = E_{inter} + E_{intra} + \Delta S_{conf}$$

The essence of the conformational search!!

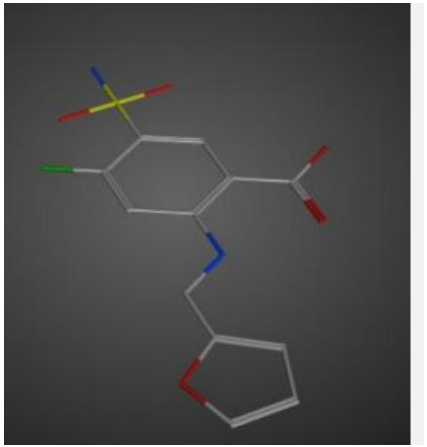
How to perform the conformational search?

Systematic Search



DOCK, FlexX or LeDock

Stochastic Search



GOLD, AutoDock or PLANTS

Mol. Bio. **2008**, 365-382

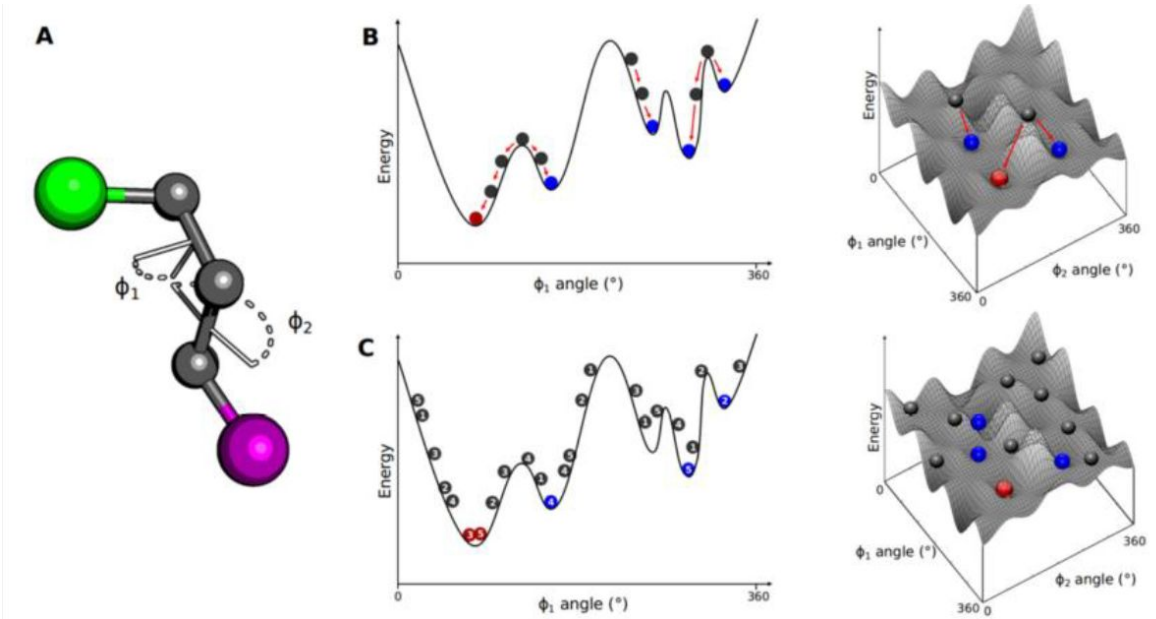
Introduction to Structure-based drug discovery and molecular docking.

Docking.

How to perform the conformational search?

Systematic Search

Stochastic Search



Molecules .20(7): 13384–13421.

Systematic Search

- ★ Obtains significant results in short times.
- ★ Their algorithms can be further optimized (hybrid approaches).
  - ○ It is prone to be "trapped" in local minimums.

Stochastic Search

- ★ Usually gives more effective results.
- ★ Allows you to explore the energy landscape more broadly.
- ★ Computationally is more expensive.
  - ○ Strictly, it must be bounded.

How to simulate ligand flexibility?

- Growing anchor
- Simulated annealing.

Systematic

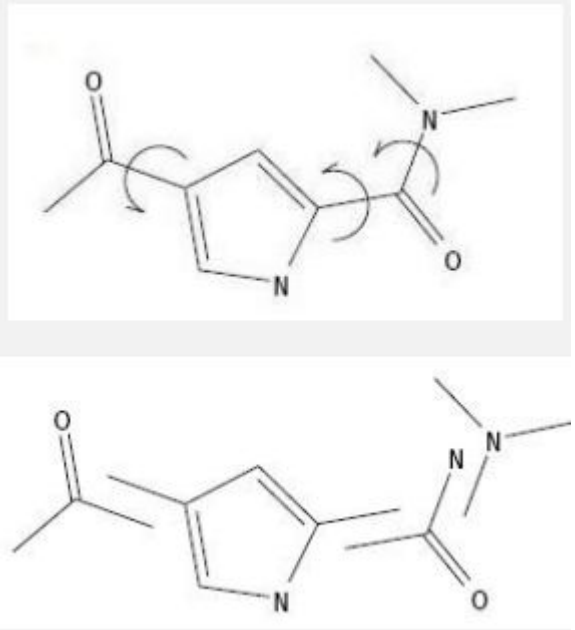
- Metaheuristic

Stochastic

- Local Search (Vina)
- Hierarchical filters

Hybrid

Growing anchor.



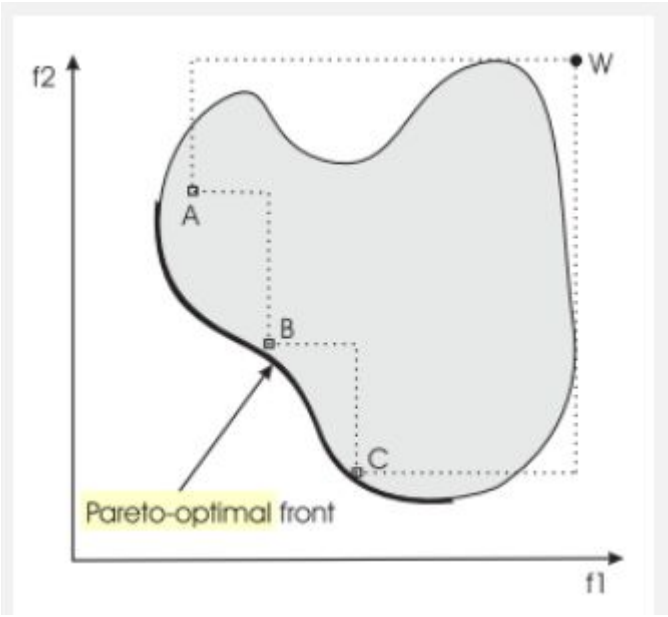
- The algorithm allows a good balance of the calculation time.
- It is possible to choose the anchor (rigid docking).

Simulated annealing.

$$P(\Delta E) = e^{\left(-\frac{\Delta E}{k_B T}\right)}$$

- It is based on the work of Metropolis.
- It is a Monte Carlo method, which seeks to leave local minima.
- It can be refined to become a Markov chain.

Metaheuristic



- An optimization flow is proposed that can be mono- or multi-objective.
- These methods have been very successful in various scientific applications.

Some implementations in docking.

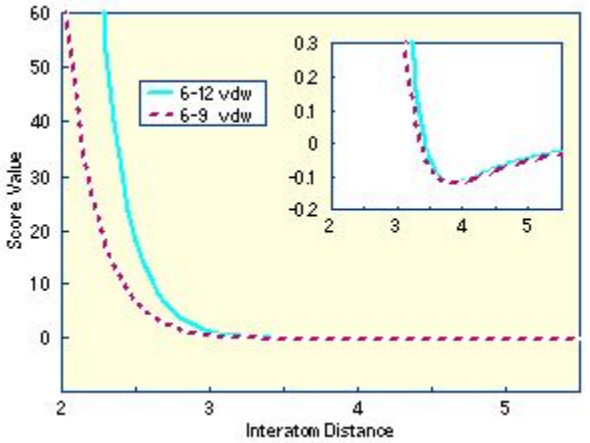
- Taboo search
- Genetic algorithms
- Particle swarm optimization
- Lamarckian Genetic algorithm
- Ant colony optimization
- Local search

Introduction to Structure-based drug discovery and molecular docking.

Docking.

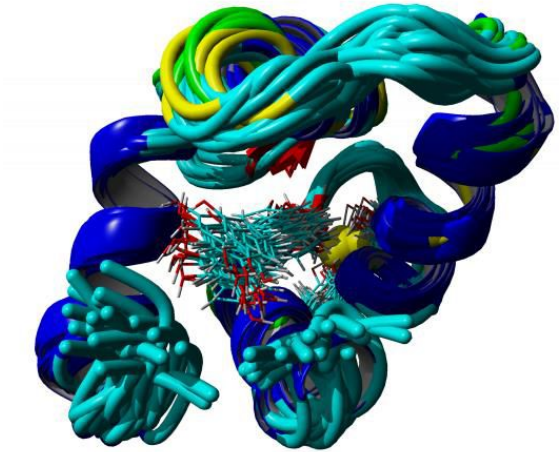
Receptor flexibility?

- **Library of rotamers**
- **Docking ensamble**
- **Soft receptor** ⇨ $E(\varepsilon) = \frac{A}{r^{12}} - \frac{B}{r^{6}}$ ⇨



Brás NF, Cerqueira NMFSA, Sousa SF, et al (2014) Protein ligand docking in drug discovery. Protein Model .

Docking Ensemble

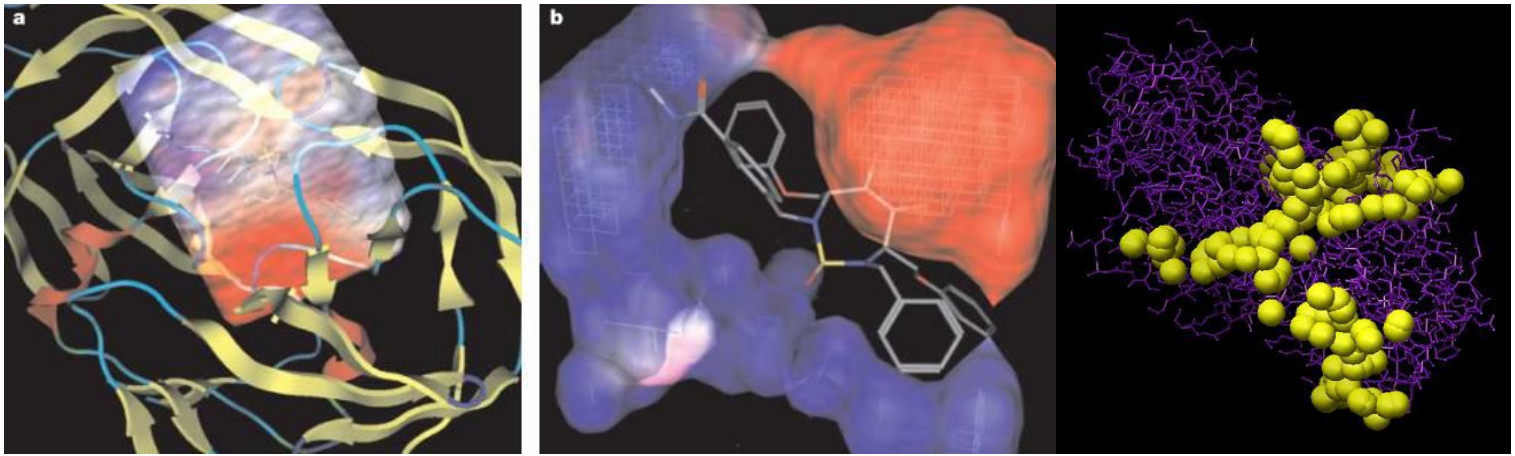


- NMR
- Differents PDB
- Monte Carlo
- Molecular Dynamic

Introduction to Structure-based drug discovery and molecular docking.

Docking.

Defining the search space



Curr. Comput Aided Drug Des. 7:146 57.

Tools to identify binding sites

- SiteHound
- COACH
- FINDSITE
- PocketFinder/PocketPicker
- SiteFinder
- SiteMap

Blind Docking

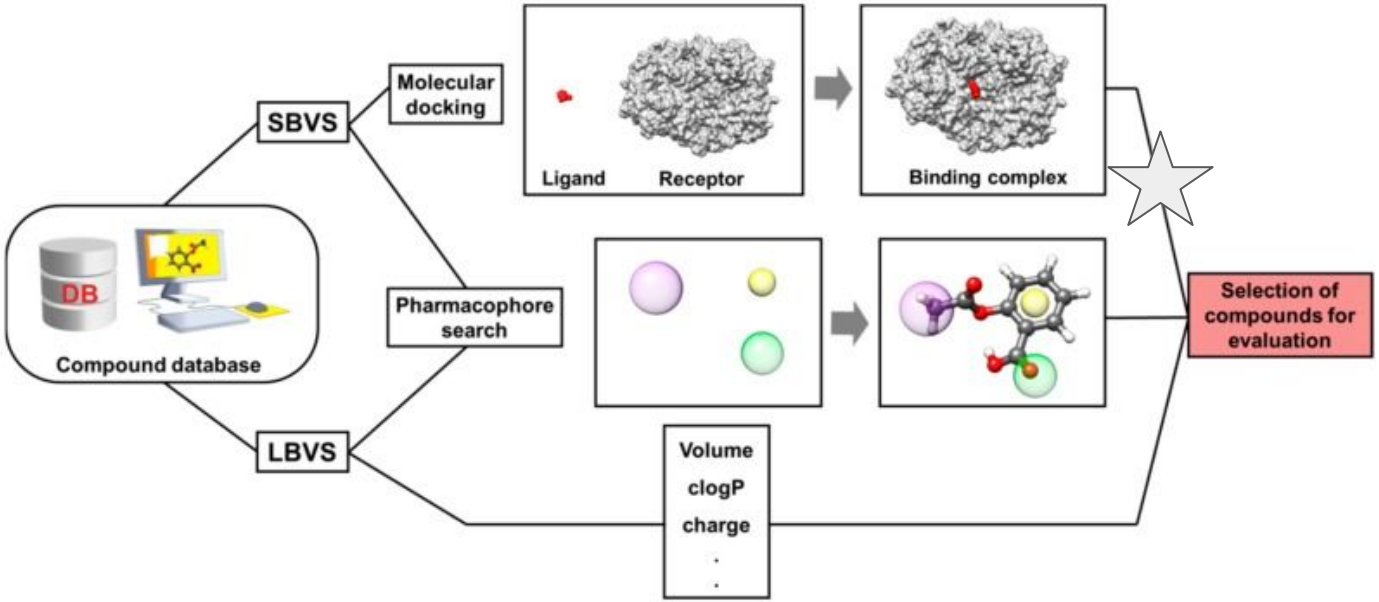
- ★ The search space is defined in the whole protein.
- ★ It is useful when the active site is unknown.
- ★ It can be used as a benchmark.
- ❑ Not all softwares can do it.
- ❑ It is more efficient to probe individual cavities.

About the grid size

- It is necessary to ensure that the ligand has enough space.
- It impacts accuracy and results.
- There are some tools and scripts to choose the right size.

Introduction to Structure-based drug discovery and molecular docking.

Docking.



Molecules .20(7): 13384–13421.

About Scoring Functions

- Forcefields based
- Empirical
- Knowledge based
- Consensus scoring

Introduction to Structure-based drug discovery and molecular docking.

Docking.

Forcefield based

$$E = W_{vdw} \sum_{i,j} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} \right) + W_{Hbond} \sum_{i,j} E(t) \left(\frac{C_{ij}}{r_{ij}^{12}} - \frac{D_{ij}}{r_{ij}^{10}} \right) + W_{elec} \sum_{i,j} \frac{q_i q_j}{\in(r_{ij}) r_{ij}}$$

- ★ They are robust.
- ★ They have a clear physical meaning.
- ★ They are reasonable calculation.
- ❑ They do not consider entropic terms.
- ❑ They are not "realistic."
- ❑ They require modifications / updates.

Curr. Comput . Aided Drug Des 7:146 57.

Introduction to Structure-based drug discovery and molecular docking.

Docking.

Empirical

$$E_{PLANTS} = f_{PLP} + f_{clash} + f_{tors} + c_{site} - 20$$

- ★ They are simpler (computationally).
- ★ They are more refined.
- ★ They can be almost universal.
- ❑ They are not transferable.
- ❑ They usually make incomplete descriptions.
- ❑ Its performance is affected by the size of the ligand.

Protein Model. pp 249-286 doi: 10.1007/978 3 319 09976 7_11

Introduction to Structure-based drug discovery and molecular docking.

Docking.

Knowledge based

$$F_{SMoG2001} = \sum_r \sum_p \sum_l F(r, \sigma_p, \sigma_l)\Delta(r, p, l) + F^{rot\,1} + N^{rot\,1} + F^{rot\,2} + N^{rot\,2}$$

- ★ Reproduce a conformation.
- ★ They have a good balance.
- ★ They are more robust and transferable.
- ★ They allow to describe unusual interactions.

- ❑ They depend on the reference state.
- ❑ They depend on the information available.

Phys Chem Chem Phys 12:12899.

Introduction to Structure-based drug discovery and molecular docking.

Docking.

What options exist?

Cite this: *Phys. Chem. Chem. Phys.*, 2016, **18**, 12964

Comprehensive evaluation of ten docking programs on a diverse set of protein–ligand complexes: the prediction accuracy of sampling power and scoring power†

Zhe Wang,[a] Huiyong Sun,[a] Xiaojun Yao,[b] Dan Li,[a] Lei Xu,[c] Youyong Li,[d] Sheng Tian[d] and Tingjun Hou*[ae]

Novel Consensus Docking Strategy to Improve Ligand Pose Prediction

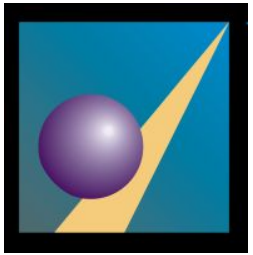
Xiaodong Ren,[†] Yu-Sheng Shi,[*,‡] Yan Zhang,[*,¶] Bin Liu,[¶] Li-Hong Zhang,[¶] Yu-Bo Peng,[¶] and Rui Zeng[§]

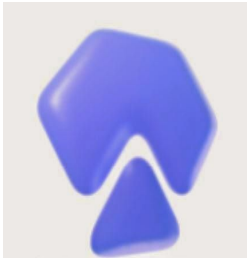
- AutoDock 4
- VINA
- rDock
- DOCK
- MOE
- LeDock
- Glide
- PLATS
- ...

Introduction to Structure-based drug discovery and molecular docking.

Docking.

Protein and Ligand preparation





Mol. mod. Prot. (443) 365-382.

About the types of atoms

- It is the most careful item, a minor error can be decisive.
- Determine compatibility with the program.
- They are especially careful when the input is **.mol2** or **.sdf**

- ❏ AMBER
- ❏ CHARMM
- ❏ MMFF94
- ❏ SYBYL
- ❏ AUTODOCK

About the charges

- They have an impact on scoring.
- They can improve the description of the system.

- ❏ Kollman
- ❏ Gasteiger
- ❏ AM1
- ❏ MMFF94
- ❏ PM6

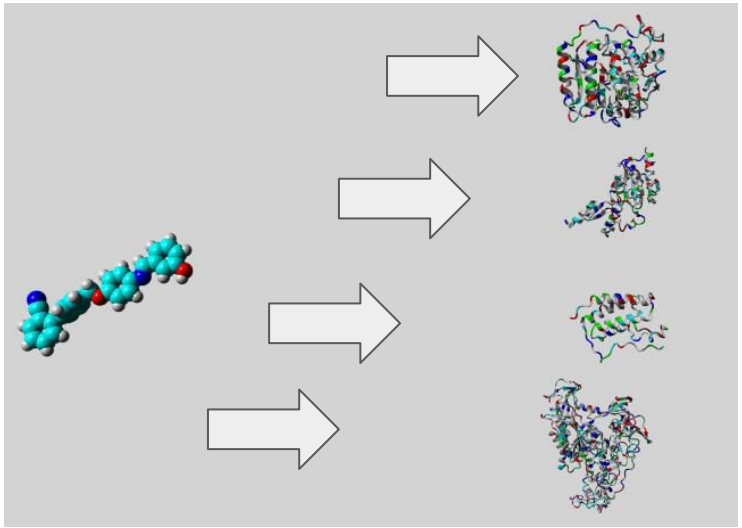
Special cases

- ➔ Covalent docking.
- ➔ Cross docking.
- ➔ Solvated Docking.
- ➔ Docking in metalloproteins.
- ➔ Docking with peptides.

Introduction to Structure-based drug discovery and molecular docking.

Docking.

Target-Fishing



Methods. 2015, 71:98 103.

CABANA

Capacity building for bioinformatics in Latin America

CABANA Workshop:

Chemoinformatics in Drug Discovery

Introduction to Structure-based drug discovery and molecular docking.

Edgar, López-López

edgar.lopez.593@hotmail.com

CINVESTAV

Unidad Irapuato

DIFACQUIM

UNAM

DIFACQUIM

COMPUTER-AIDED DRUG DESIGN AT UNAM