Organic chemistry laboratory – Pharmacy Faculty
University of Medicine and Pharmacy at Ho Chi Minh city

Monitor recruitment on June 2nd, 2023

Overview of Molecular Docking

Presenter: Msc. Phan Tieu Long

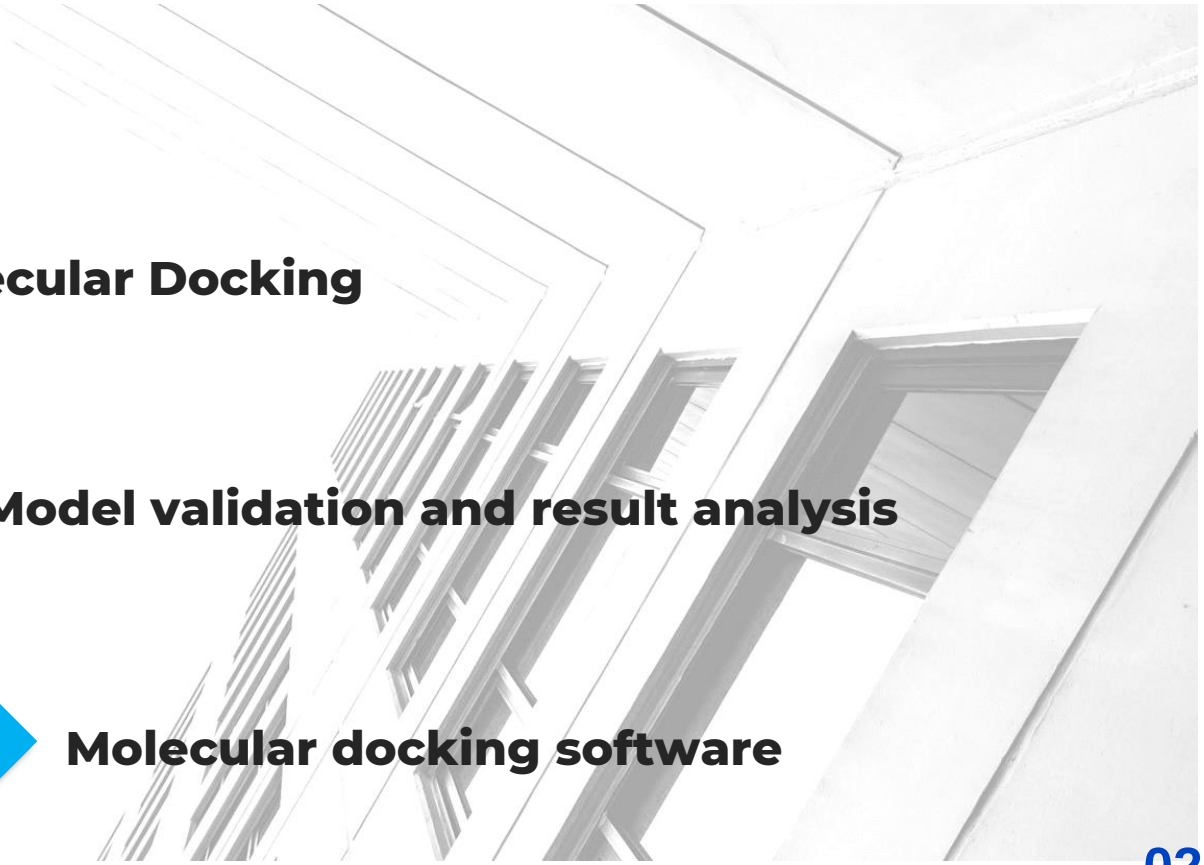
Agenda

01 Introduction

02 Molecular Docking

03 Model validation and result analysis

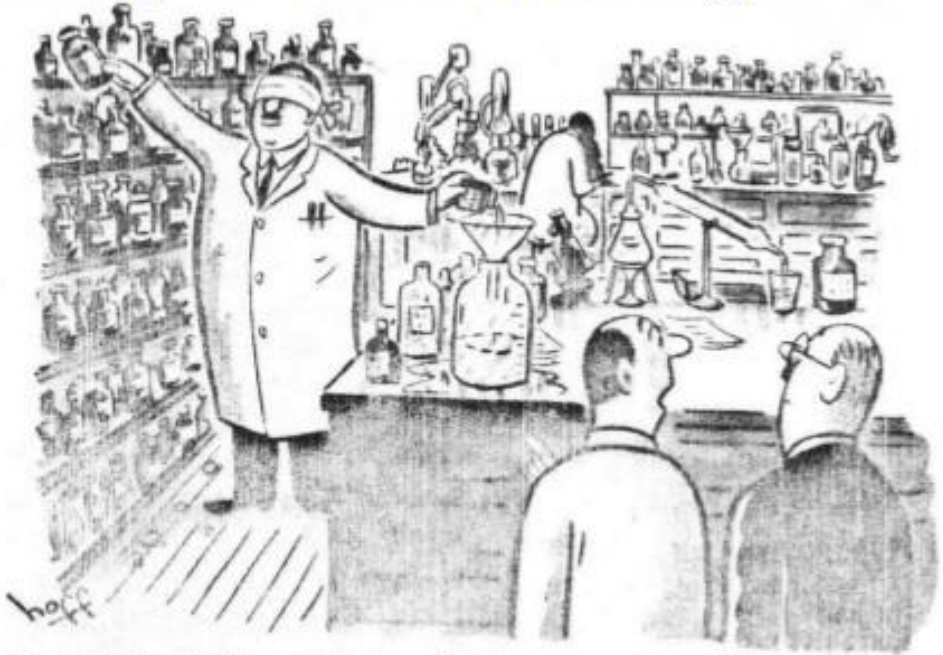
04 Molecular docking software



Introduction

Introduction

One way to "discover" drugs



'That's Dr Arnold Moore. He's conducting an experiment to test the theory that most great scientific discoveries were hit on by accident.'

Drawing by Hoff; © 1957
The New Yorker Magazine, Inc.

Introduction

Drug discovery

A good drug (e.g., kills virus)

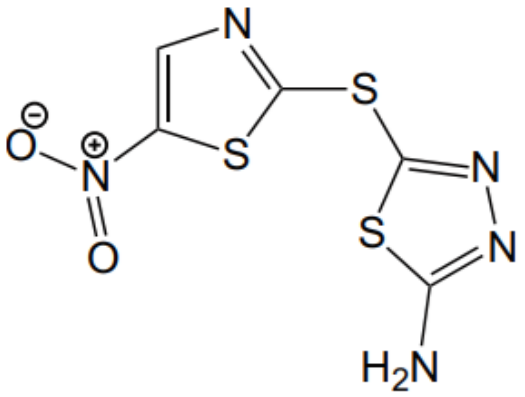
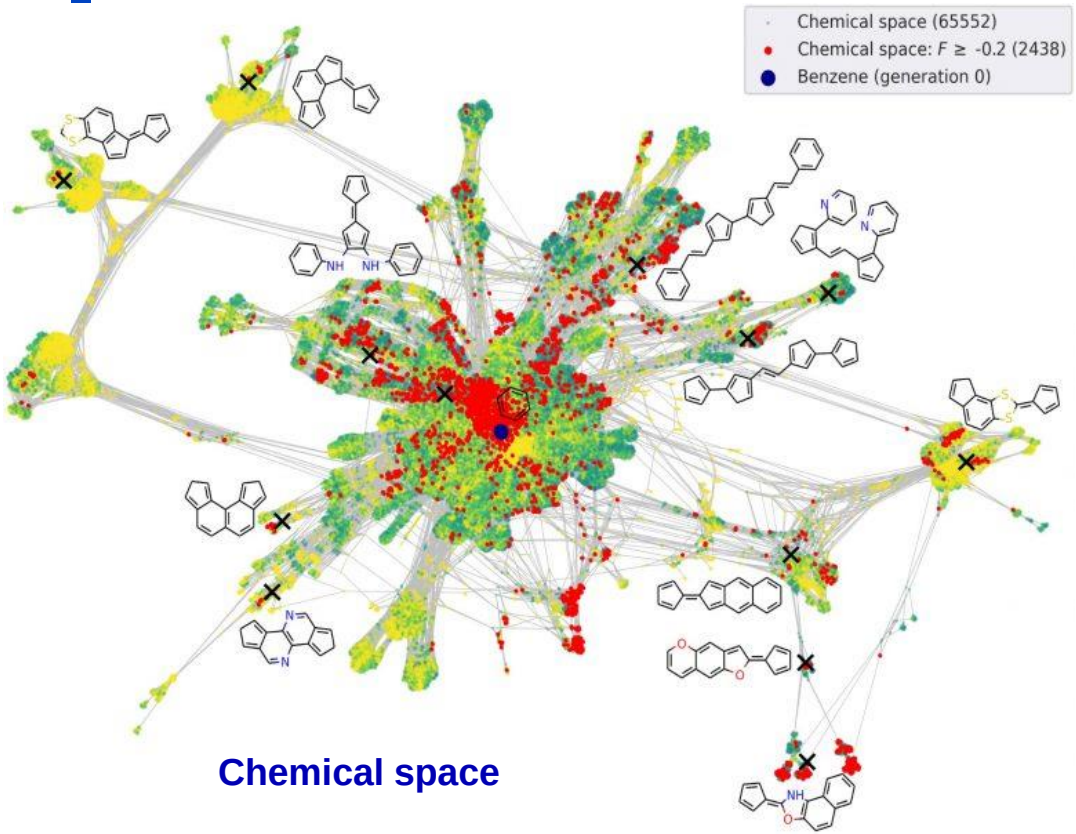


Figure source: PhRMA.org

Introduction

Challenges



Chemical space

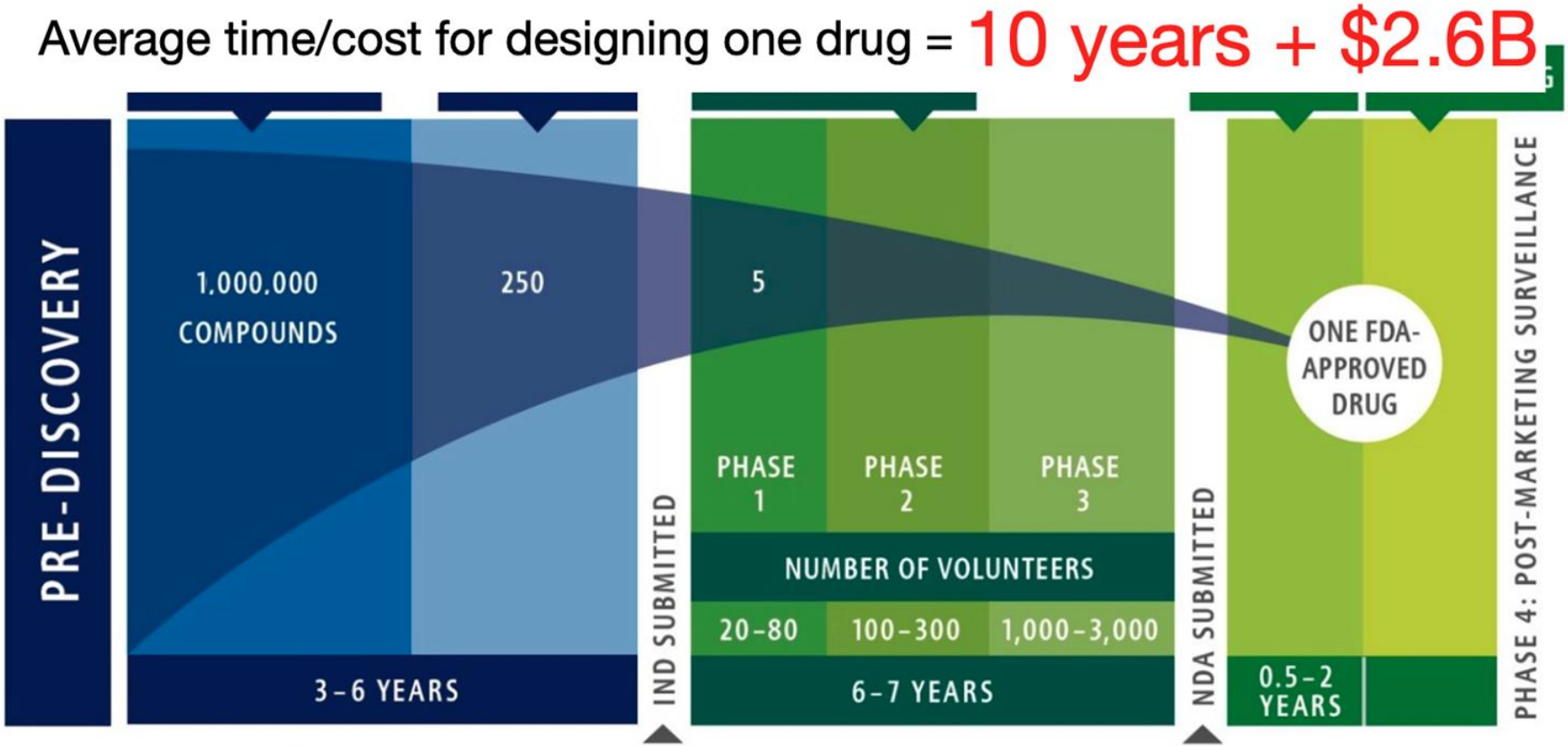
Experimental facilities in industry can only test 10^5 compounds/day

Figure source: © Kunkel/FHI

Kirkpatrick, et al., Nature. 2004

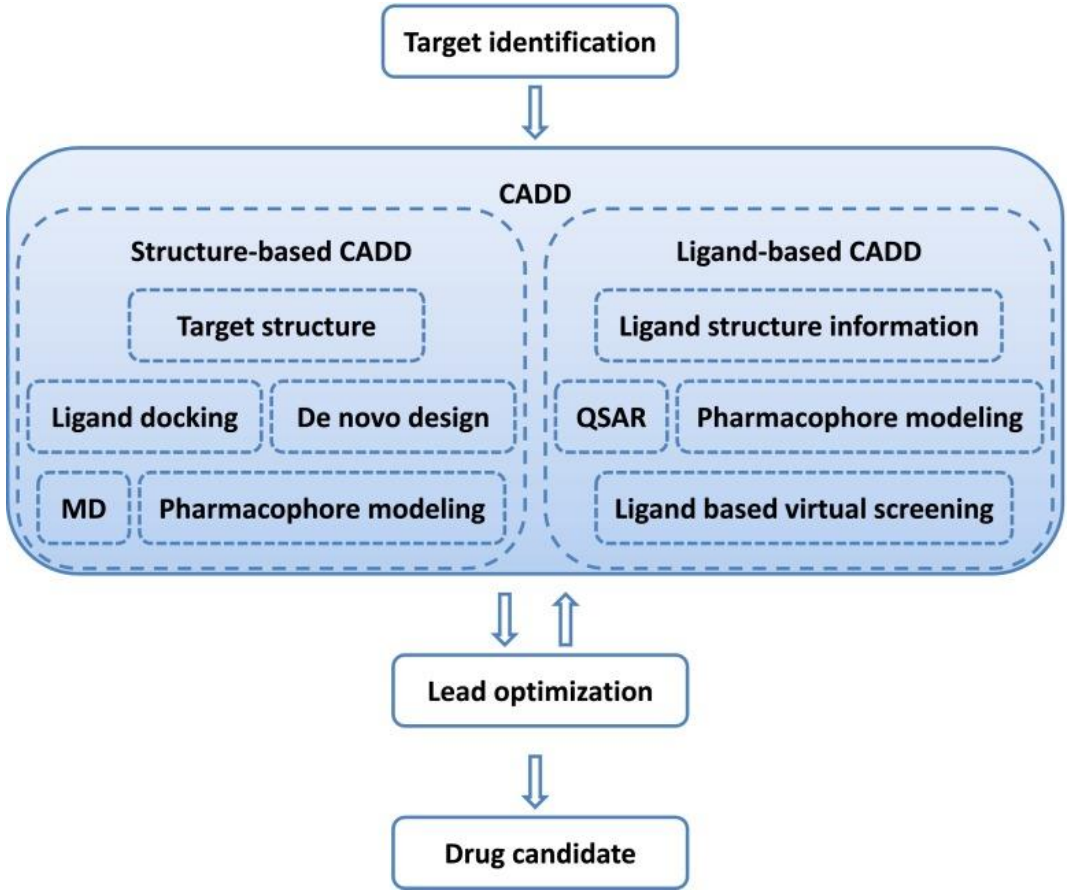
Introduction

Drug discovery and development timeline



Introduction

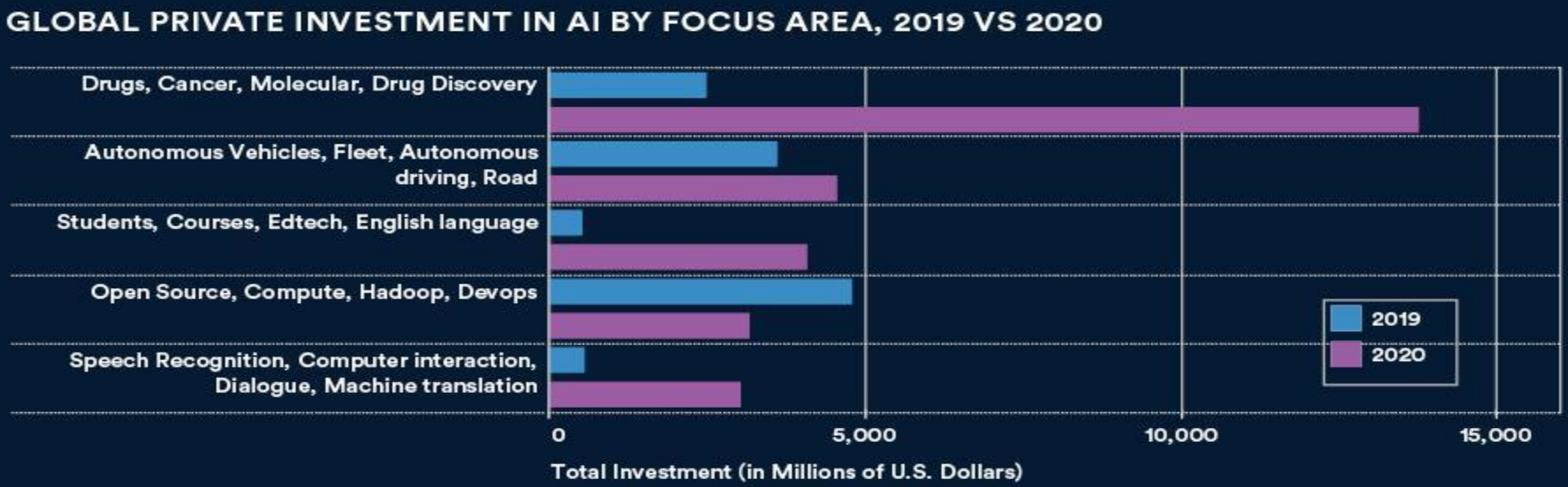
In silico methods



Sliwoski G, *et al*. Computational methods in drug discovery (2013)

3 AI Industry Shifts Focus to Drug Discovery

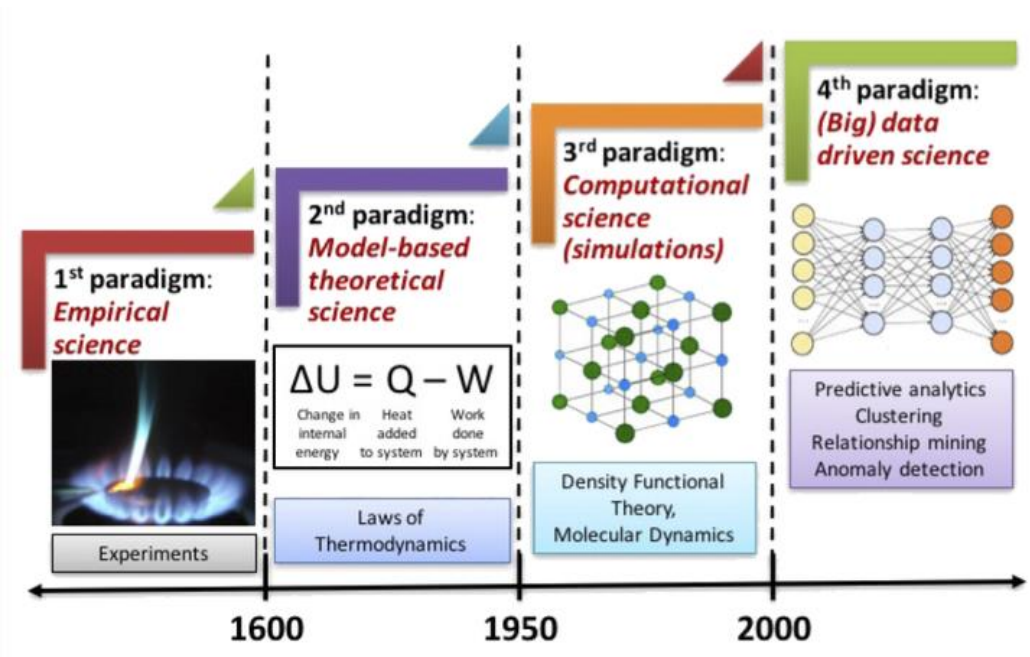
Global investment spikes in reaction to pandemic.

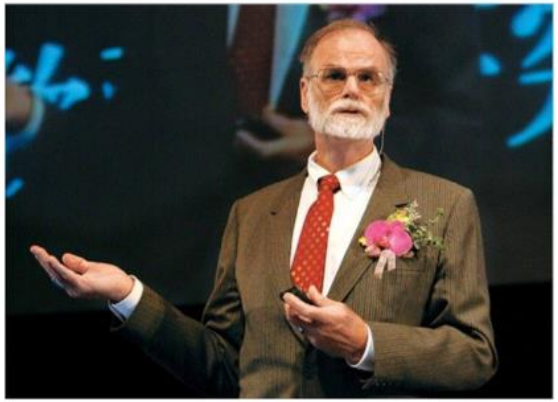


Data Source: CAPIQ, Crunchbase, and NetBase Quid
Adapted from the AI Index 2021 Annual Report

Introduction

The fourth paradigm of science

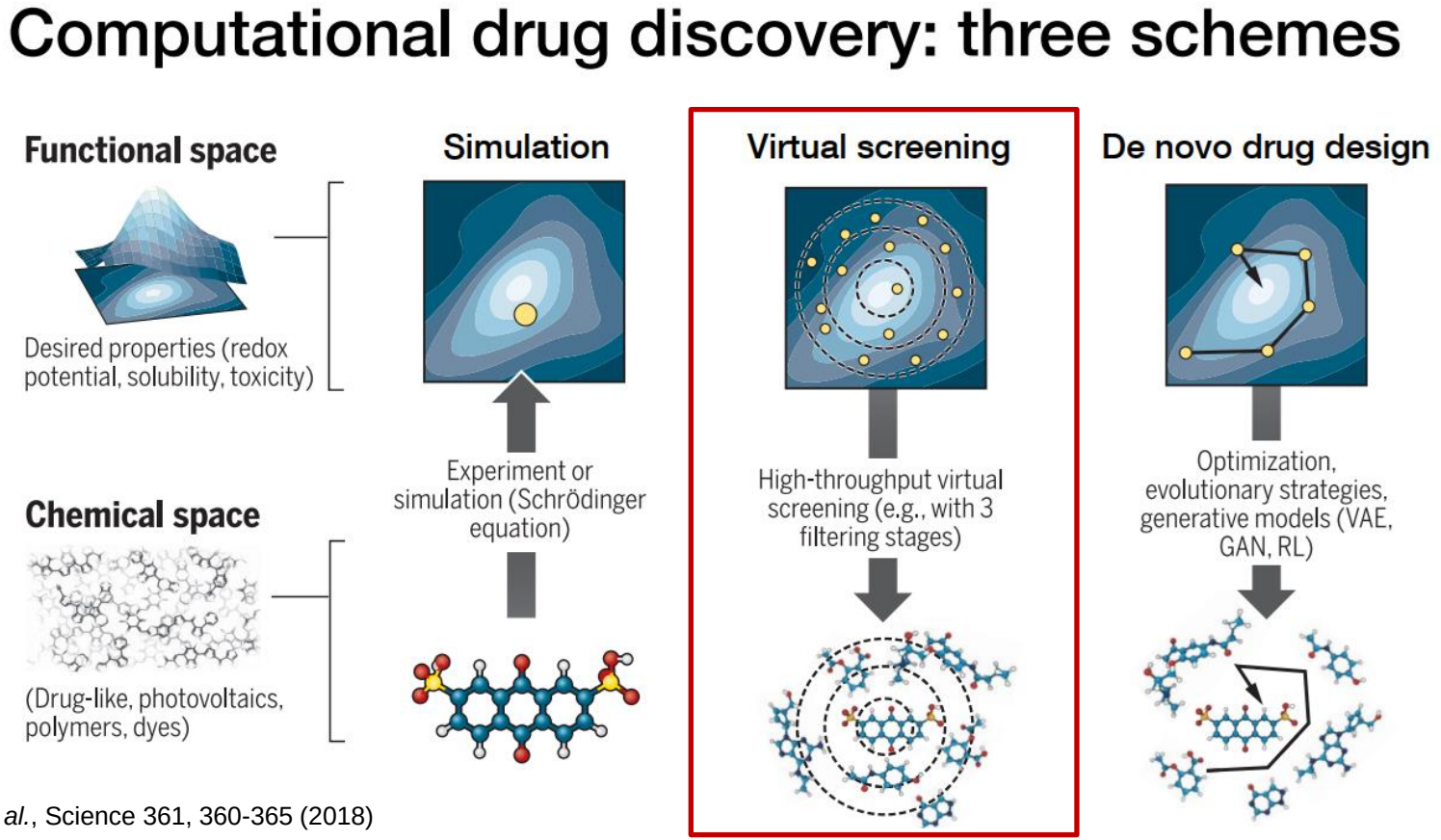




Jim Gray, Turing Award 1998 (1944-2007)
Honoured as father of *The 4th Paradigm*

Introduction

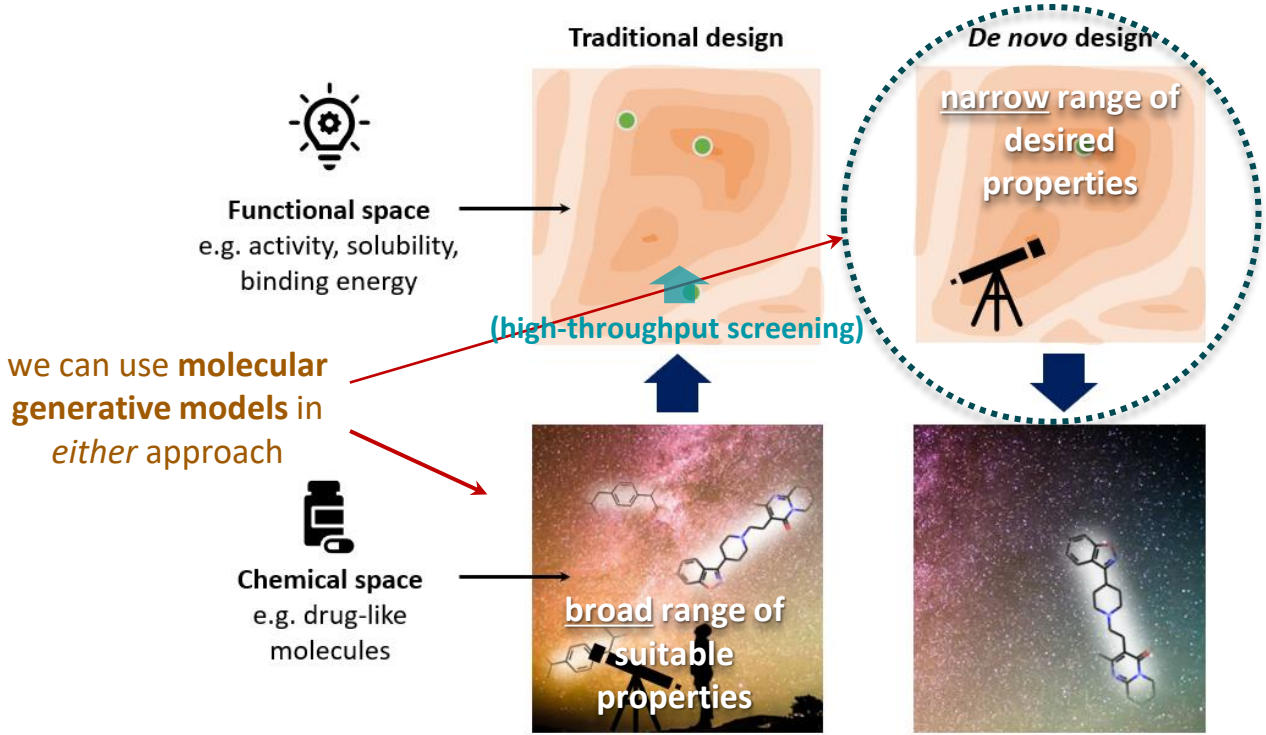
In silico methods



Sanchez-Lengeling *et al.*, Science 361, 360-365 (2018)

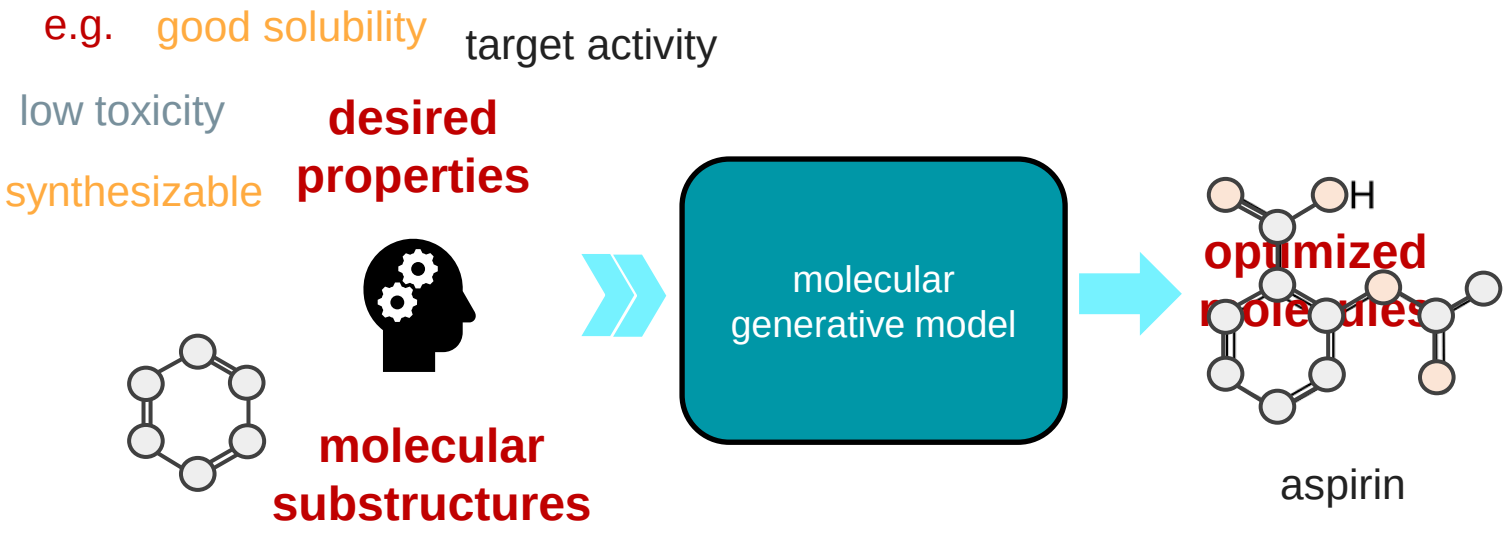
Introduction

In silico methods



Introduction

De novo design



Save chemists time by narrowing down the chemical space of viable molecules *faster and more efficiently*.

Can be used in conjunction with existing (traditional) methods.

Pros and cons

Simulation	
Pros	**Cons**
- Replace/provoke/explain experiments - The most accurate method among 3 aforementioned schemes - Provides initial hypotheses about the binding mode of the compounds	- Time-consuming and usually slower than other schemes - Require decent computational resources - The performance is highly dependent on the biological system under study

Molecular docking

Molecular docking

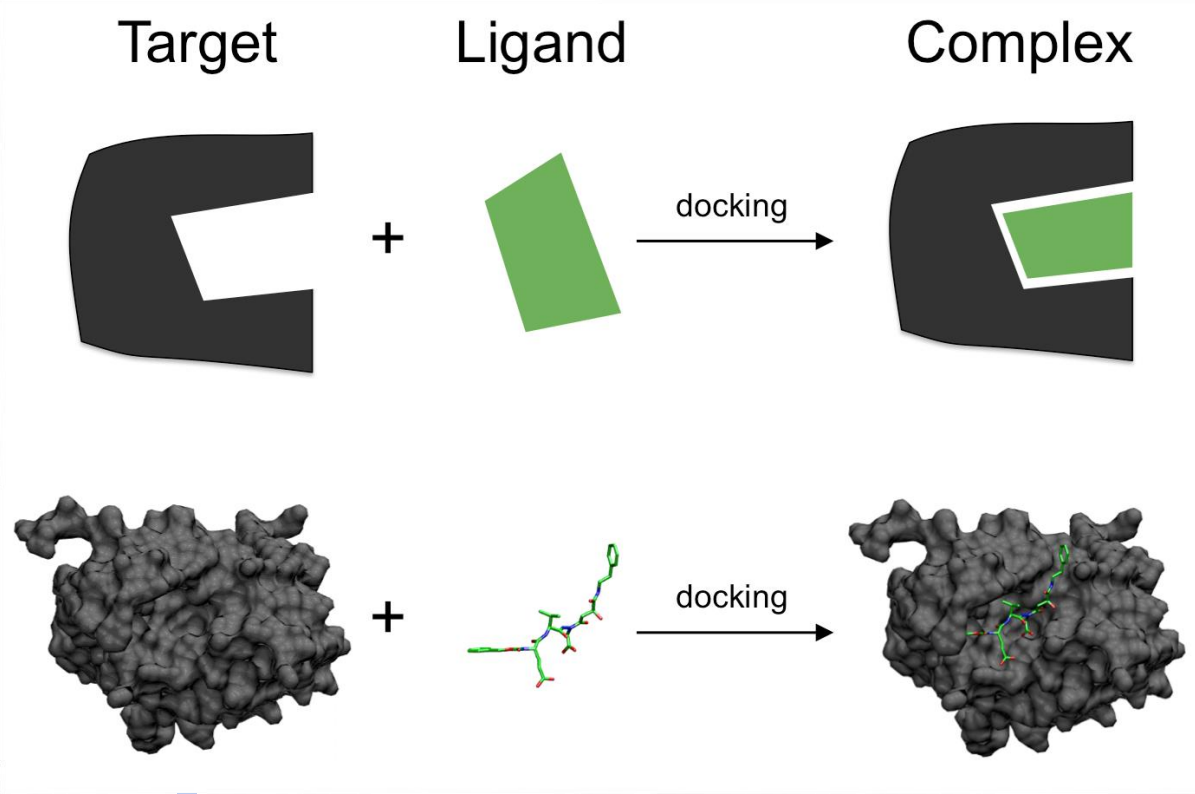
- **One of the most frequently used methods** in structure-based drug design

- **Predict the binding-conformation** of small molecule ligands to the appropriate target binding site to form a stable complex. *Docking does not predict bioactivity!*

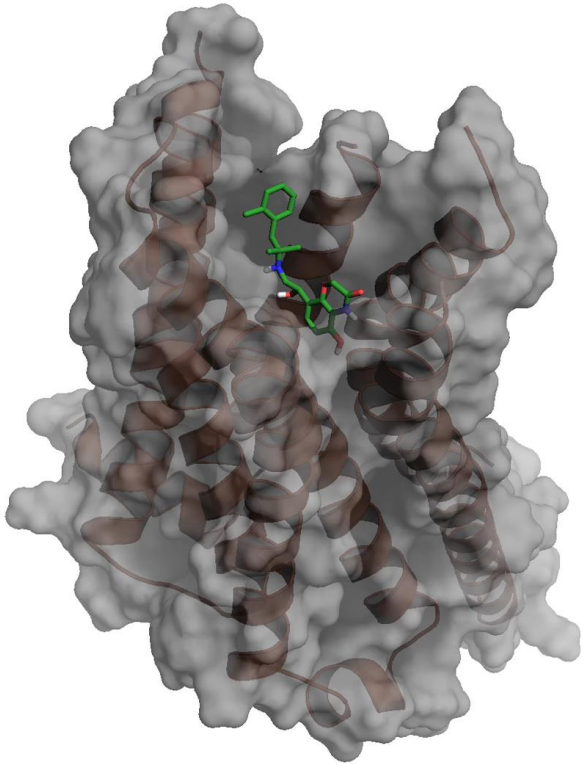
- Based on "**Lock-and-key theory "** *(rigid docking)* and **"Induced fit theory"** *(flexible docking).*

- Docking can be achieved through two interrelated steps: 1. sampling conformations of the ligand in the active site of the protein; 2. ranking these conformations via a **scoring function**.

Molecular docking



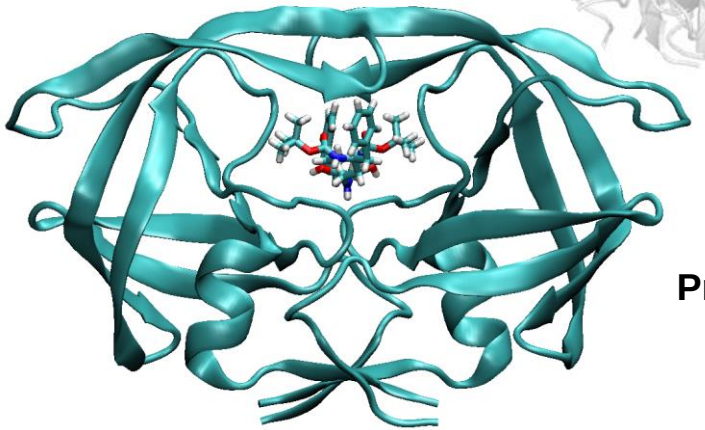
Schematic illustration of docking a small molecule ligand (green) to a protein target (black) producing a stable complex



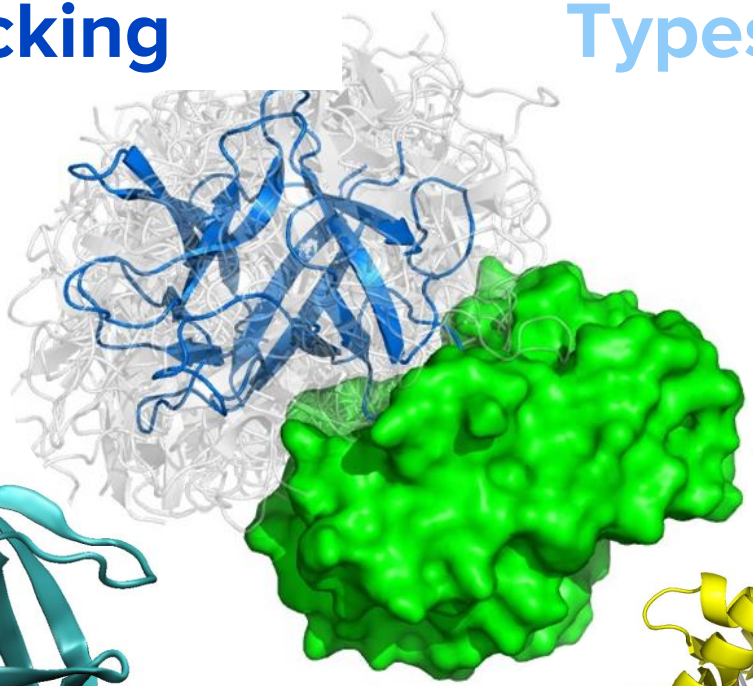
Docking of a small molecule (green) into the crystal structure

Molecular docking

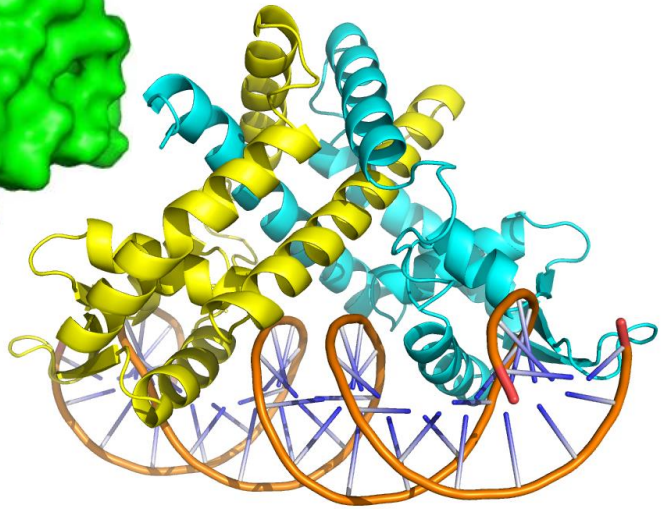
Types of complexes



Protein – Small molecule complex



Protein – Protein complex



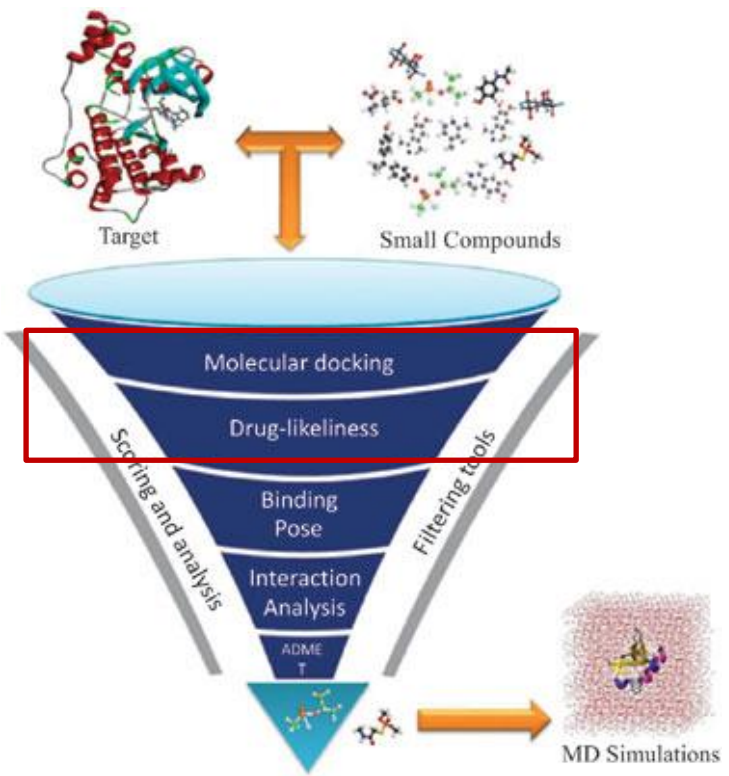
Protein – Nucleotide complex

Molecular docking

Applications

Molecular docking 4 main applications:

1. Reproduce the binding mode of X-ray complex
2. Predicting the binding mode of a known active ligands
3. Predicting the binding affinities of related compounds from a known active series
4. Identifying new ligands using Virtual screening



Molecular docking

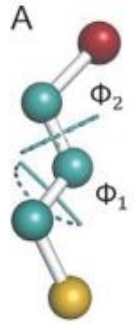
Algorithms	Characteristic
Matching algorithms	Geometry-based, suitable to VS and database enrichment for its high speed
Incremental construction	Fragment-based and docking incrementally
MCSS	fragment-based methods for the de novo design
LUDI	fragment-based methods for the de novo design
Monte Carlo	Stochastic search
Genetic algorithms	**Stochastic search**
Molecular dynamics	For further refinement after docking

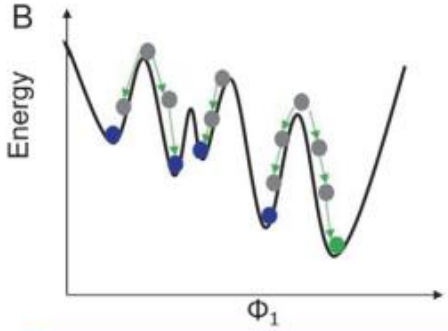
Current algorithms and their characteristics used in molecular docking

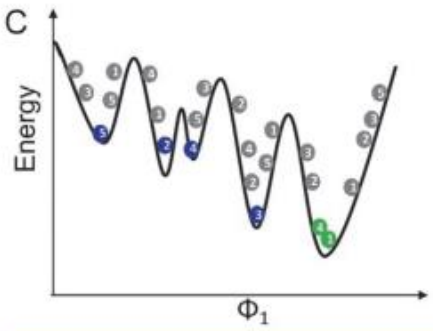
Molecular Docking: A powerful approach for structure-based drug discovery (2011)

Molecular docking

Conformational search







Systematic methods

- Glide
- FlexX
- Dock
- Superflex-Dock
- FRED

…

Stochastic methods

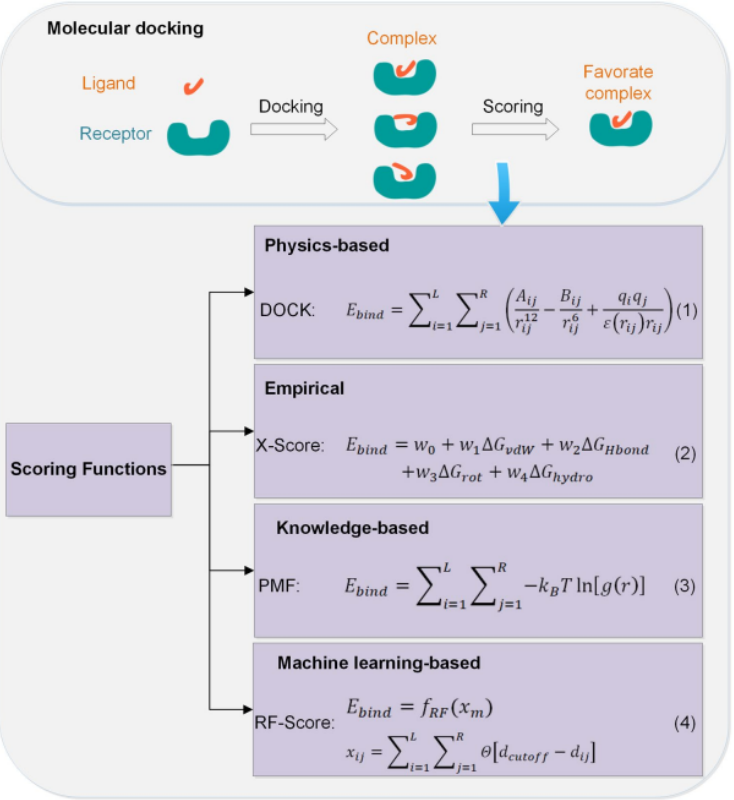
- GOLD
- DockVision
- PLANTS
- AutoDock
- PSI-DOCK

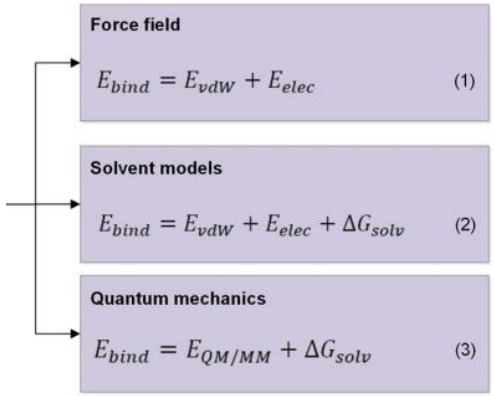
…

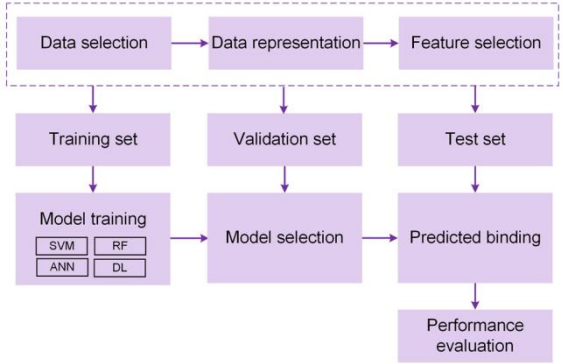
Small molecule conformational search methods

Molecular docking

Scoring function

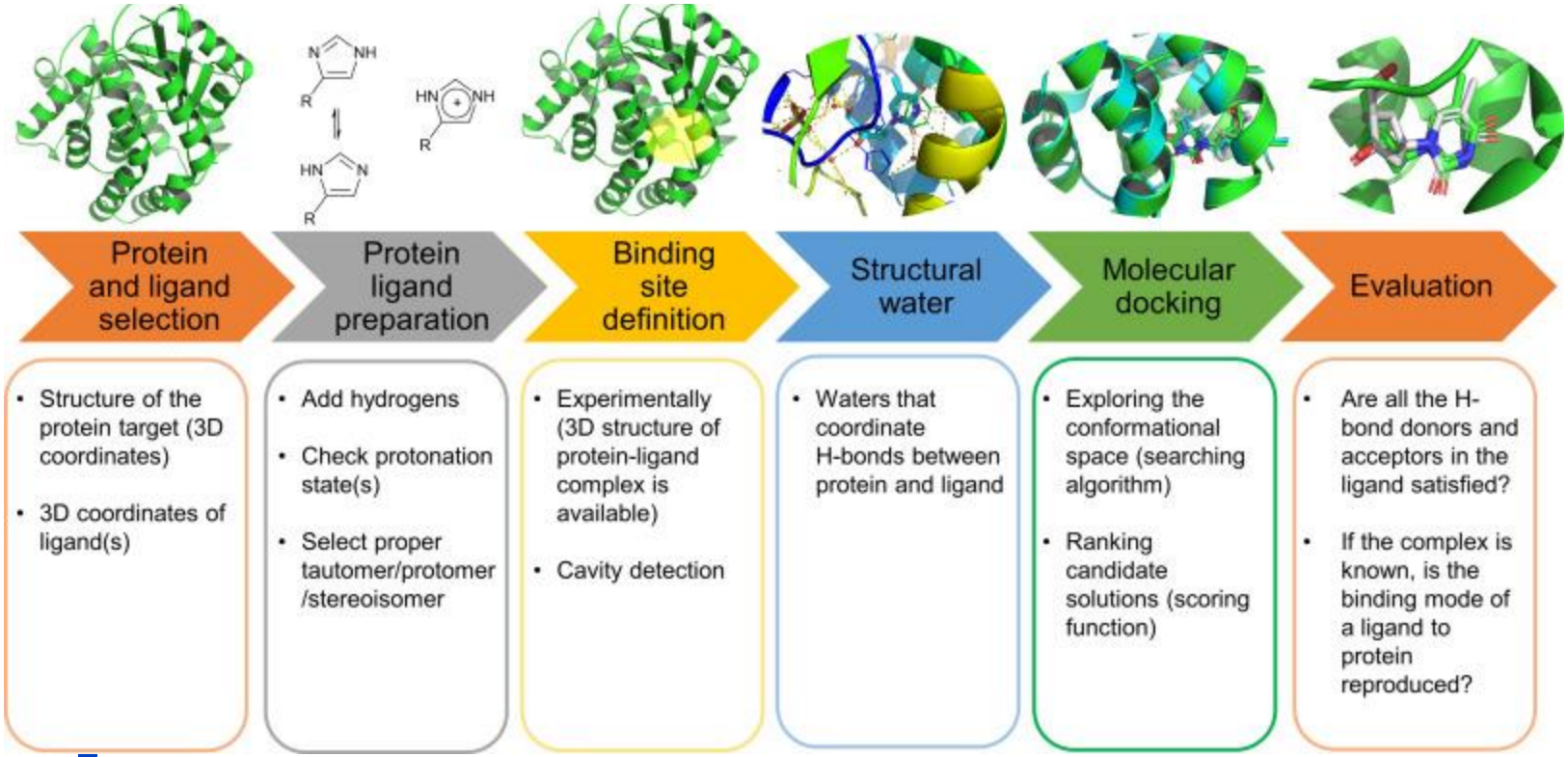






Molecular docking

Docking procedure

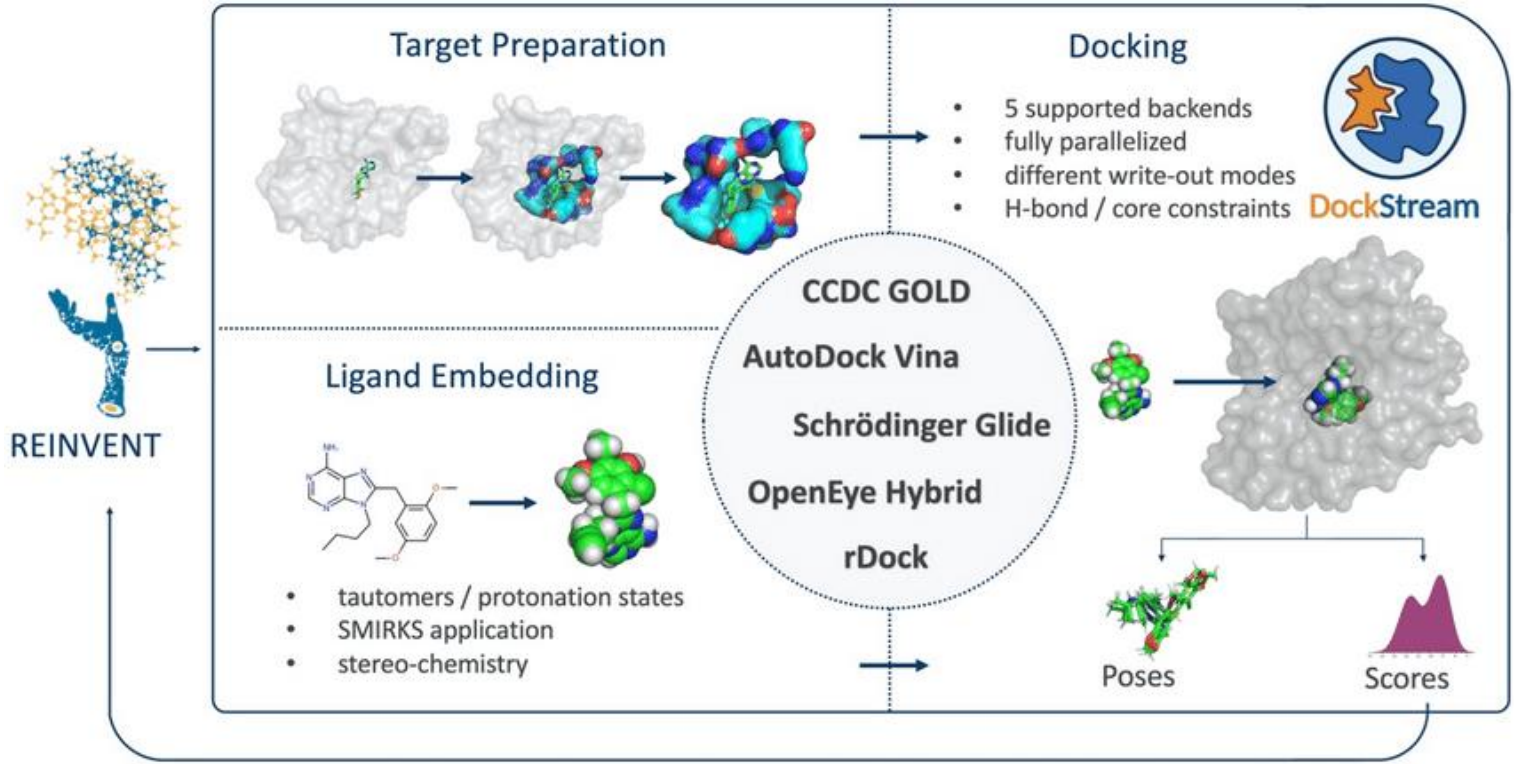


Molecular docking: step by step

DockStream: a docking wrapper to enhance de novo molecular design. Journal of Cheminformatics (2021)

Molecular docking

New approach



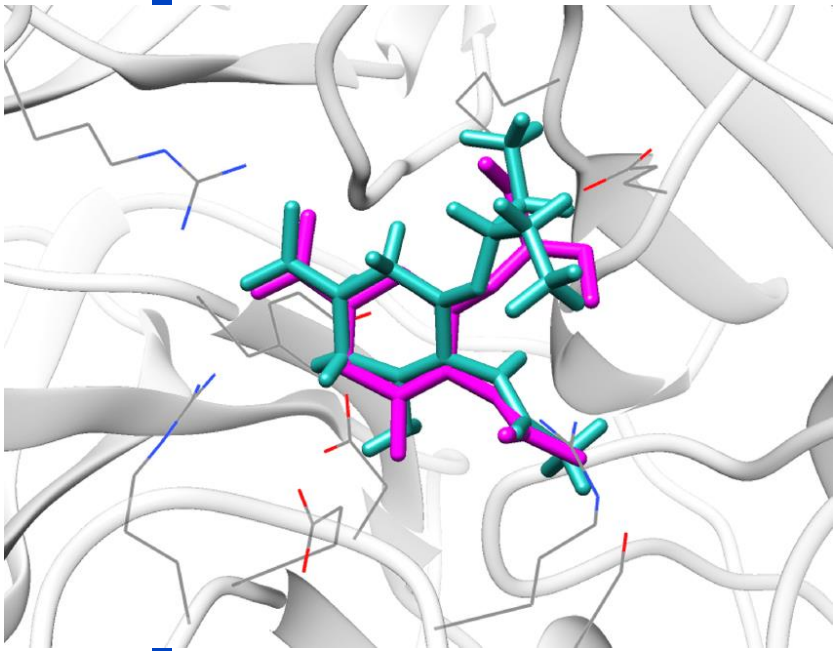
Overview of molecular docking procedure

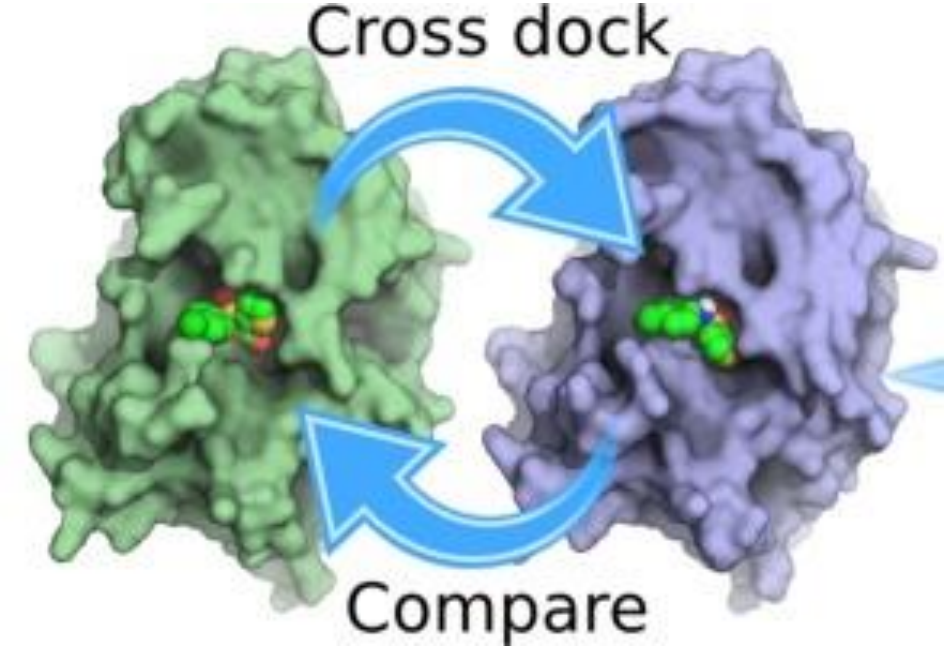
DockStream: a docking wrapper to enhance de novo molecular design. Journal of Cheminformatics (2021)

Validation - Analysis

Molecular docking

Validation



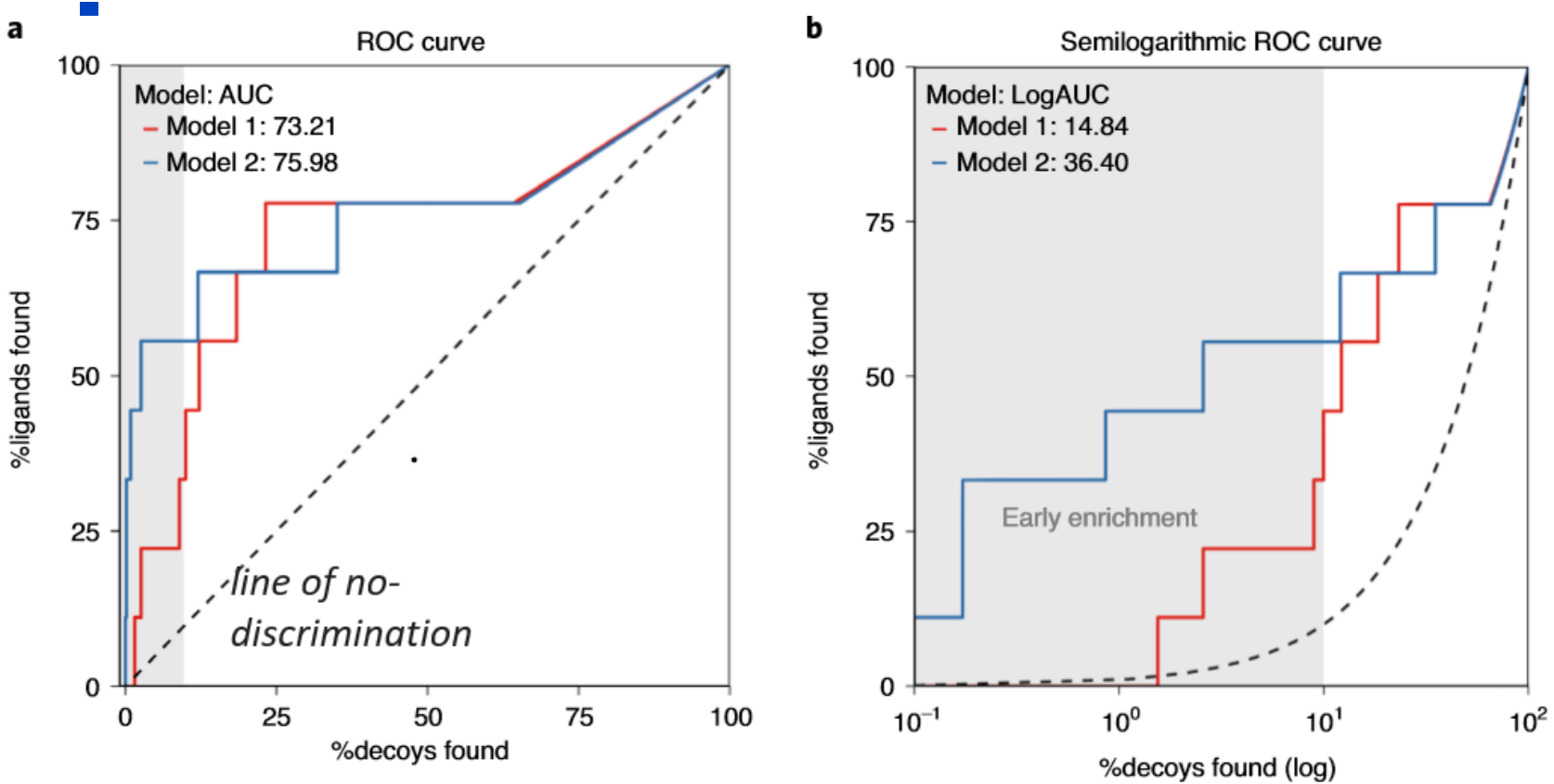


Redocking for *docking power*

Crossdocking for *screening power*

Molecular docking

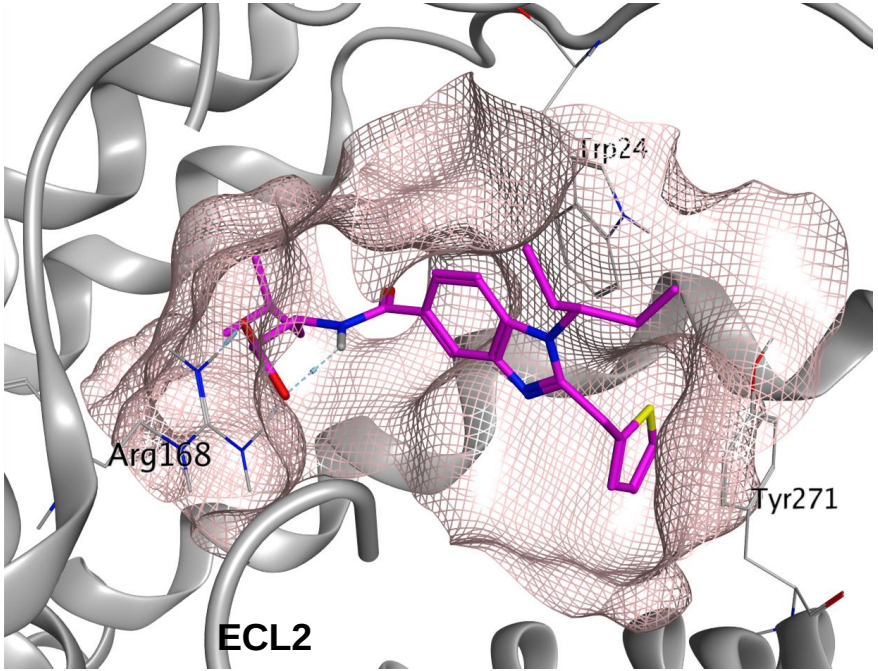
Validation

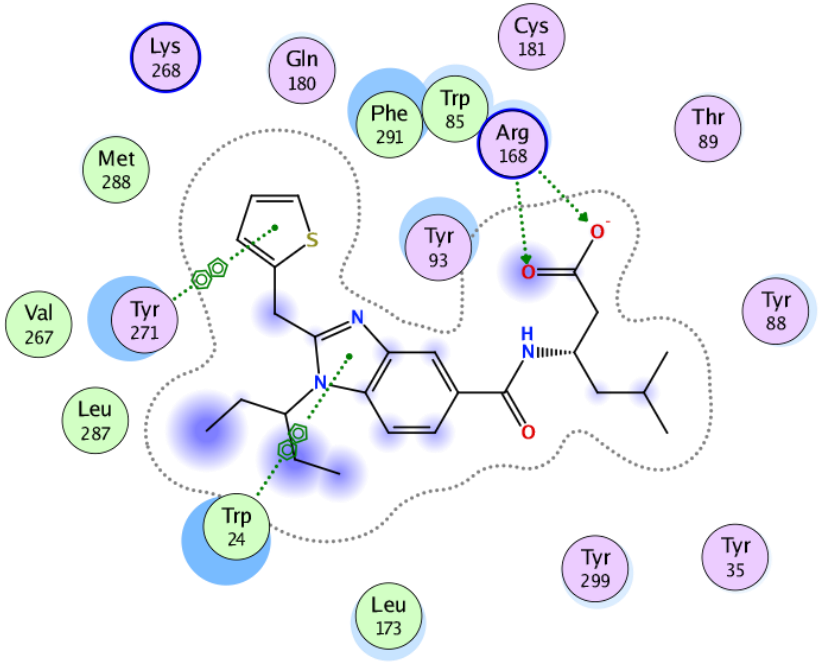


ROC AUC curve for performance validation

Molecular docking

Analysis





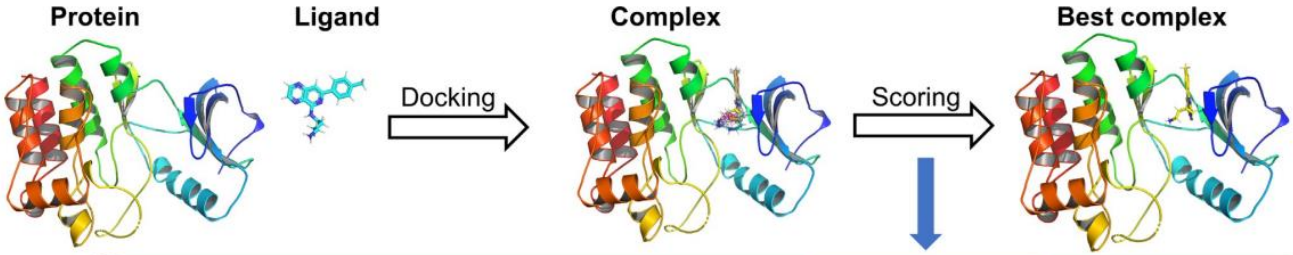
Binding energy: -10.10 kcal/mol

Molecular docking

Title	Docking score	Glide Gscore	Glide Emodel
33	-10.849	-10.849	-126.527
36	-10.670	-10.670	-126.302
31	-10.598	-10.598	-119.240
42	-10.512	-10.512	-116.183
32	-10.468	-10.468	-118.528
61	-10.446	-10.446	-109.465
9	-10.429	-10.429	-122.974
39	-10.041	-10.041	-113.930
62	-9.517	-9.517	-106.098
37	-9.470	-9.477	-96.283
19	-9.407	-9.407	-99.511
14	-9.385	-9.385	-98.541
15	-9.383	-9.383	-116.632
35	-9.352	-9.352	-93.823
18	-9.268	-9.268	-97.559
34	-9.258	-9.272	-90.379
41	-9.091	-9.097	-92.741

Docking score table

Molecular docking



Scoring functions

Classical Scoring functions

$$E_{bind} = \sum_{m=1}^{M} w_m x_m$$

Machine Learning-based

$$E_{bind} = f_{ML}(x_m)$$

ML= RF, SVM,DL *etc*

Physics-based

$$E_{bind} = \sum_i \sum_j \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + \frac{q_i q_j}{\varepsilon(r_{ij}) r_{ij}} \right)$$

Empirical

$$E_{bind} = w_0 + w_1 \Delta G_{vdW} + w_2 \Delta G_{Hbond} + w_3 \Delta G_{rotor} + w_4 \Delta G_{hydroph}$$

Knowledge-based

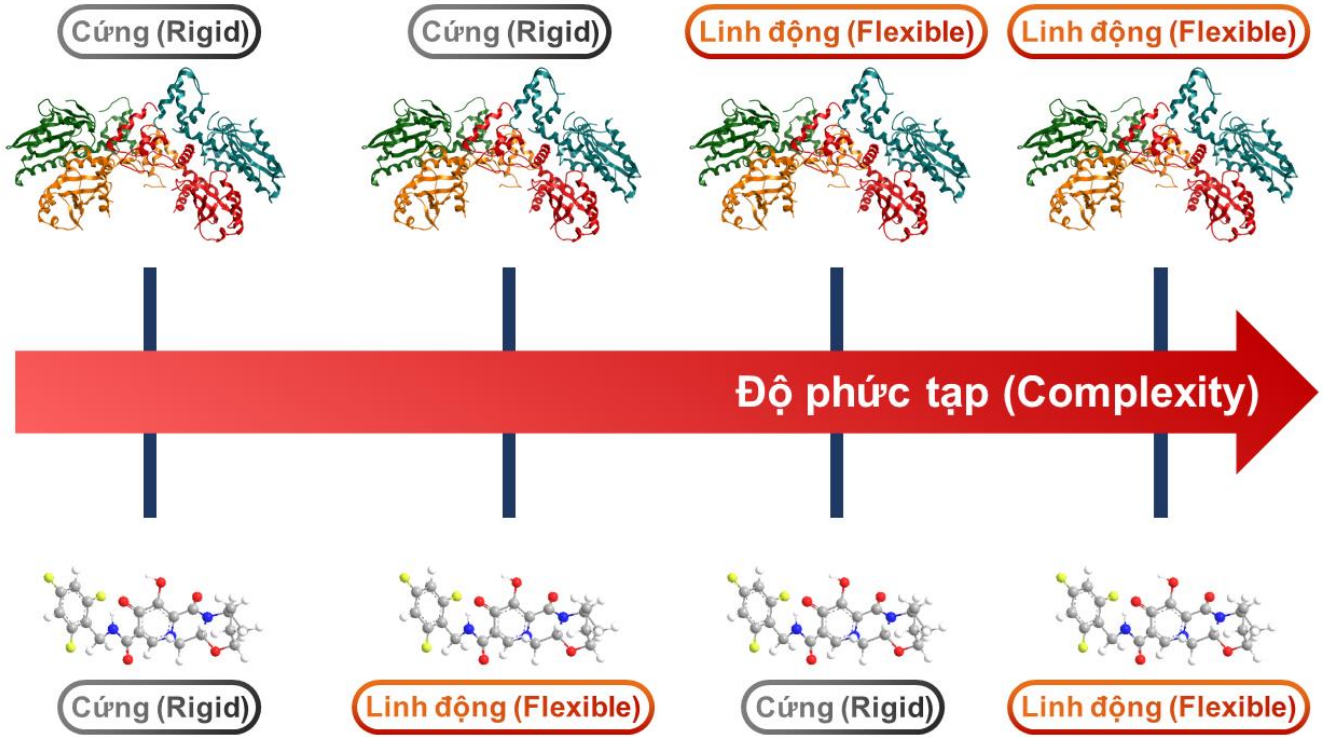
$$E_{bind} = \sum_{i=1}^{Ligand} \sum_{j=1}^{Protein} w_{ij}(r)$$

$$E_{bind} = E_{bond} + E_{no-bond} + [E_{H-bond}] + E_{solv}$$

Scoring functions in docking

Molecular docking

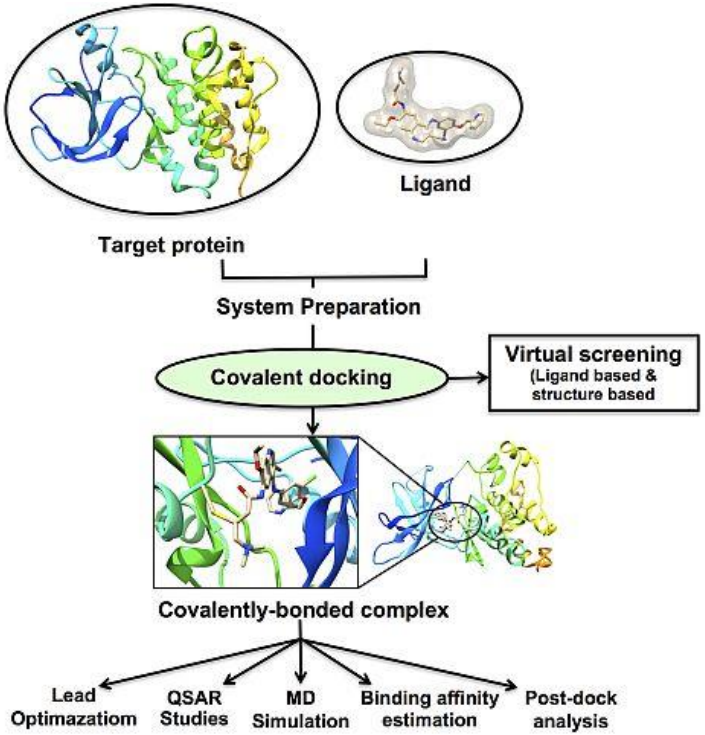
Molecular docking types



Types of molecular docking methods according to complexity

Molecular docking

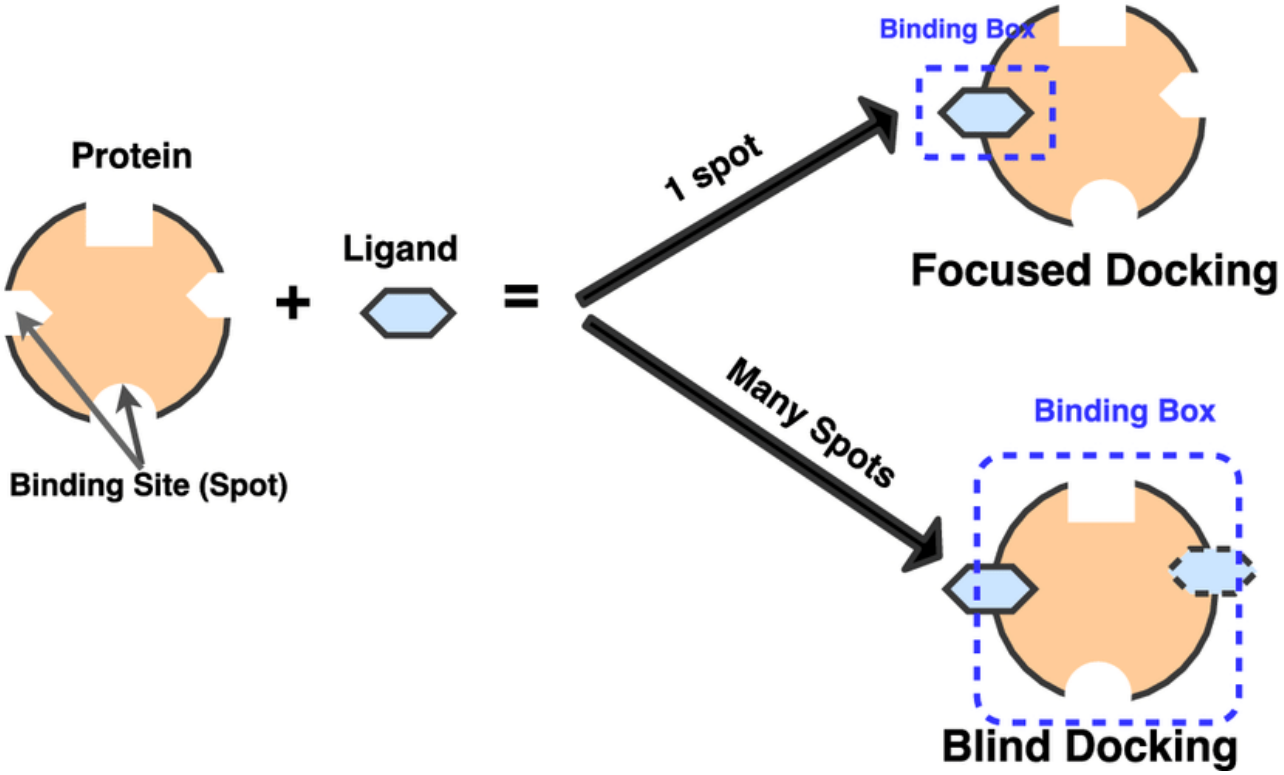
Other types of docking



Covalent docking

Molecular docking

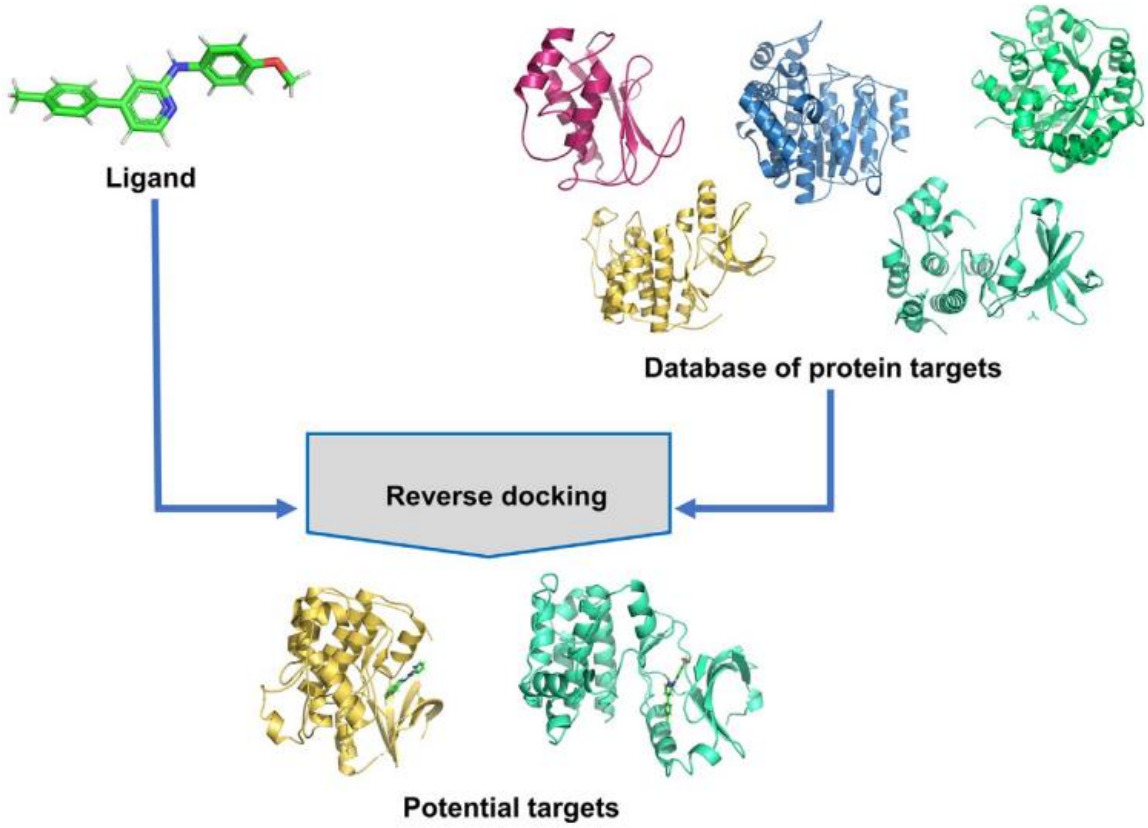
Other types of docking



Blind docking

Molecular docking

Other types of docking



Reverse docking

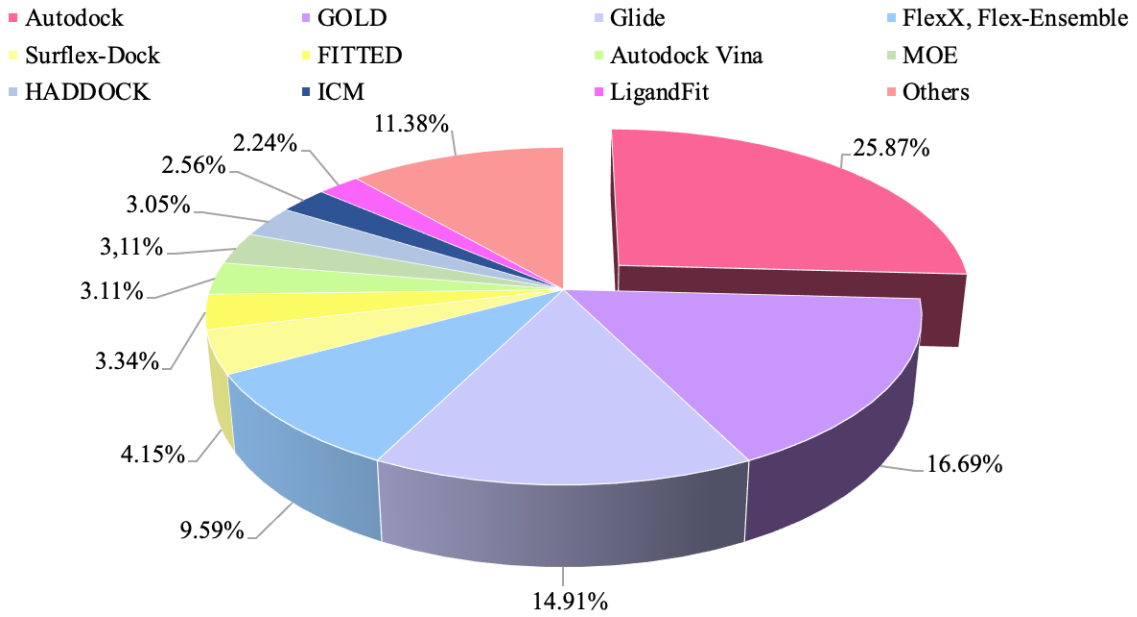
Molecular docking

Pros and cons

Molecular docking	
Pros	**Cons**
- Faster and simpler than other methods in simulation - Adaptable to different virtual screening protocols - Strong computational resources are not required - Suitable for homology approach	- Time-consuming (flexible docking), under-performance (rigid docking) - Often results in a high false positive rate - Require post-processing approaches (flexible minimization of the complexes, MD) for better results - Depend on experimental 3D structure

Molecular docking software

Software



Most used docking software from 1990 to 2013

Chen Y.-C. J. T. i. p. s. (2015), "Beware of docking!"

Software

Overview

1989 01

Autodock

Physical-based scoring functions

2010 02

Autodock vina 1.1.2

Empirical scoring functions; multicore processing

2013 03

Smina

Empirical scoring functions; could optimize SF

2017 04

Qvina2

Accelerate with BFGS optimization

2021 05

Autodock vina 1.2.3

Python scripting; combine with AD4

2021 06

Autodock-GPU

Batch-dock
GPU utilizing

2022 07

Vina-GPU

GPU utilizing

Thank you for your attention!

Does **anyone** have any questions?