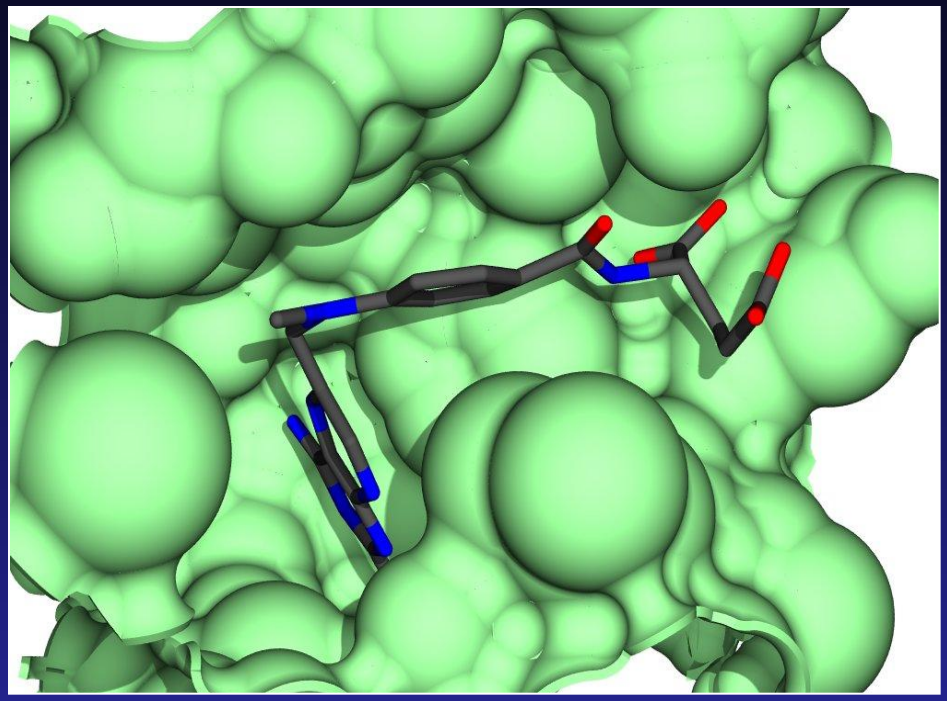
Receptor based Drug Design



Overview

- Introduction
- Preparation of Protein
- Preparation of Ligand
- Identification of Binding site
- Selecting Docking Program (Ligand Fit)
- Set up of Docking Parameters
- Analysis of Docking Results
- Demonstration in DS

Drug Discovery & Development

Mechanism of Disease

↓

Target (Enzyme, Receptor, Ion channel) Identification & Validation

↓

Lead compound (inhibitor, agonist/antagonist) identification

↓

Pre-clinical Testing in Animal Model

↓

Clinical Trials (Phase I, II, III)

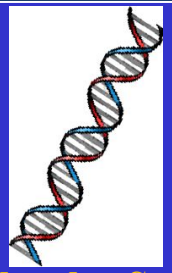
↓

Approval, Registration & Patent

Drug Discovery & Development

Target selection & validation

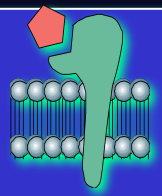
Studies of Disease Mechanisms



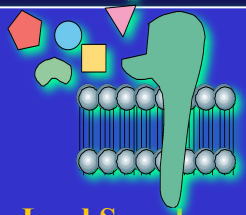
Molecular Studies

Animal Model Studies
- transgenic KO/KI mice
- agonists/antagonists
- antibodies
- antisense
- RNAi

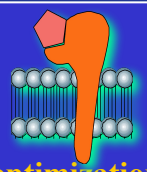
Discovery



Target
-receptor; -ion channel; -transporter; -enzyme; - signalling molecule



Lead Search
-Develop assays (use of automation)
-Chemical diversity
-Highly iterative process



Lead optimization
-selectivity
-efficacy in animal models
-tolerability: AEs mechanism-based or structure-based?
-pharmacokinetics
-highly iterative process

Development

Drug Candidate
safety testing

Human Studies
Phases I,II, III

Drug Approval & Registration

Target Identification & Validation

- Understand the molecular mechanism of the disease
- Identify a therapeutic target in that pathway (e.g gene, key enzyme, receptor, ion-channel, nuclear receptor)
- Demonstrate that target is relevant to disease mechanism using genetics, animal models, lead compounds, antibodies, RNAi, etc.

Discovery

- **Identification/Design of lead compound (*in vitro*):**
  - Screen chemical library (collection of compounds) from natural source or chemical synthesis against target

    (Chemical Combinatorial Synthesis & Robotics)
  - Design potential lead based on based on prior SAR information about target and/or known effectors.

    (Structure-based Drug Design or "Rational Drug Design")

- **Preclinical Testing (*in vivo*):**

Perform the efficacy & safety (Pharmacodynamics & Pharmacokinetics) of identified lead in animal model

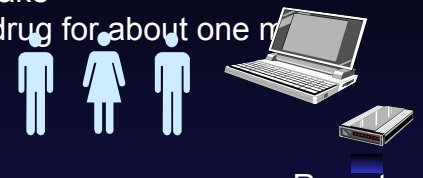
- **Lead Optimization for better PD/PK profile**

Development

Clinical Trials

Phase I

20 - 100 healthy volunteers take
drug for about one m



Remote data entry

Product Profile → Marketing
SOI

Information Learned
1. Absorption and metabolism
2. Effects on organs and tissue
3. Side effects as dosage is increased

Phase II

Several hundred health-impaired patients

Treatment Group

Control Group

Information Learned
1. Effectiveness in treating disease
2. Short-term side effects in health -impaired patients
3. Dose range

Phase III

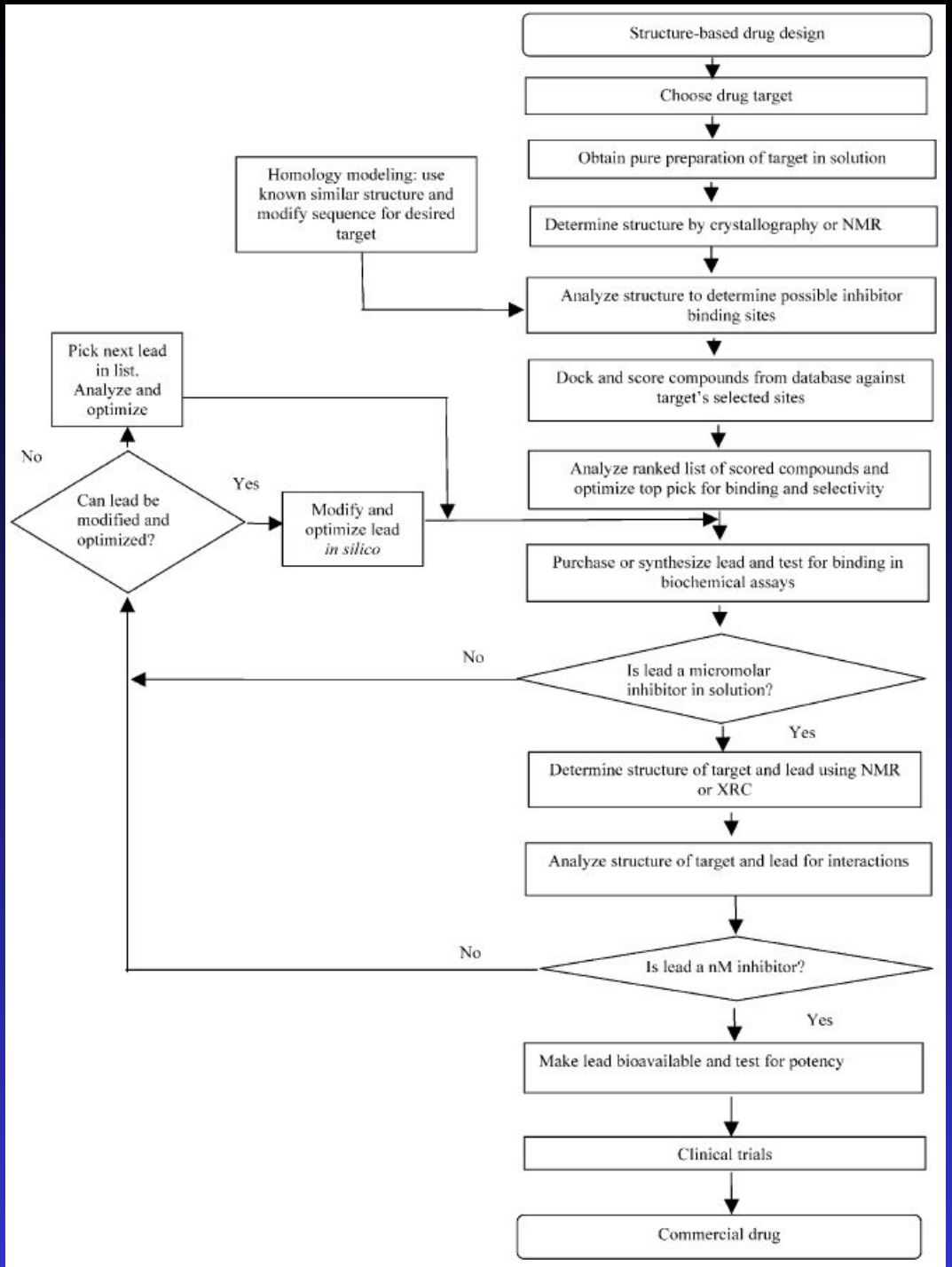
Hundreds or thousands of health-impaired patients

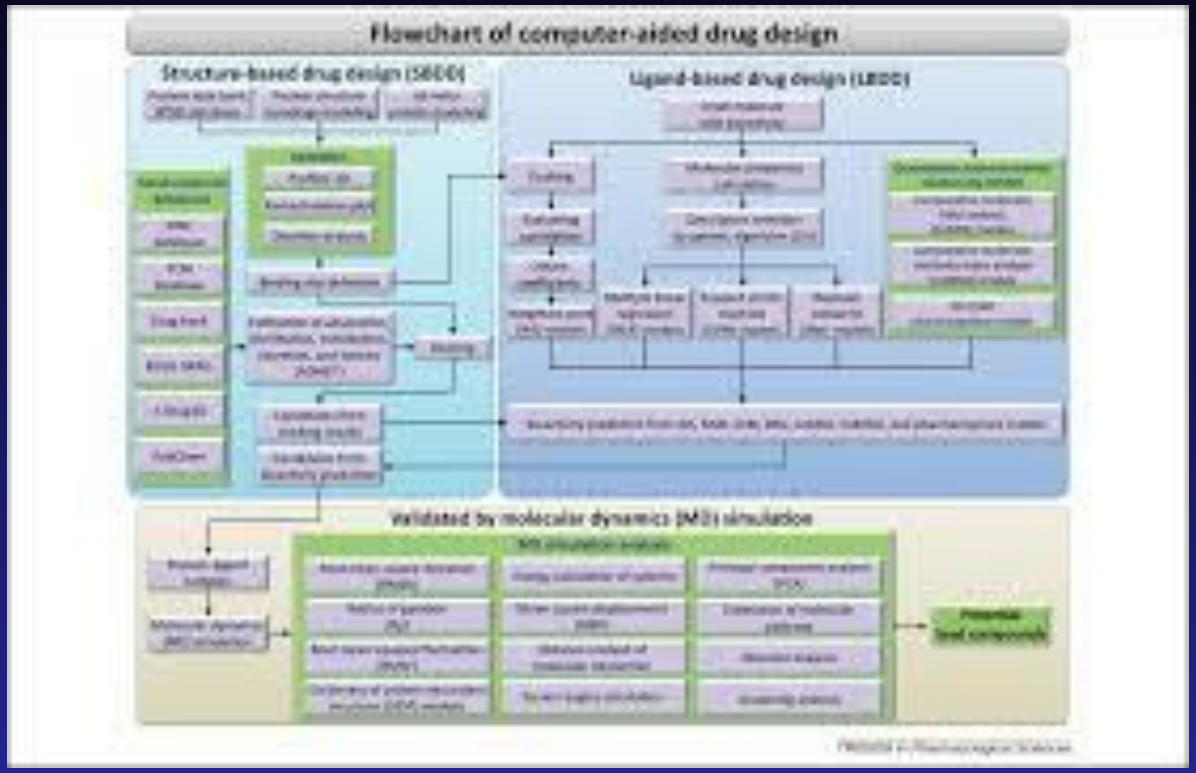
Information Learned
1. Benefit/risk relationship of drug
2. Less common and longer term side effects
3. Labeling information

Compassionate Use

Rational drug design

Based on prior knowledge about SAR




Flowchart of computer-aided drug design
Structure-based drug design (SBDD)
Ligand-based drug design (LBDD)
Validated by molecular dynamics (MD) simulation

Types of Computational Rational Drug Design

1. **Receptor-based Drug Design**
   (3-D structure of biological target)

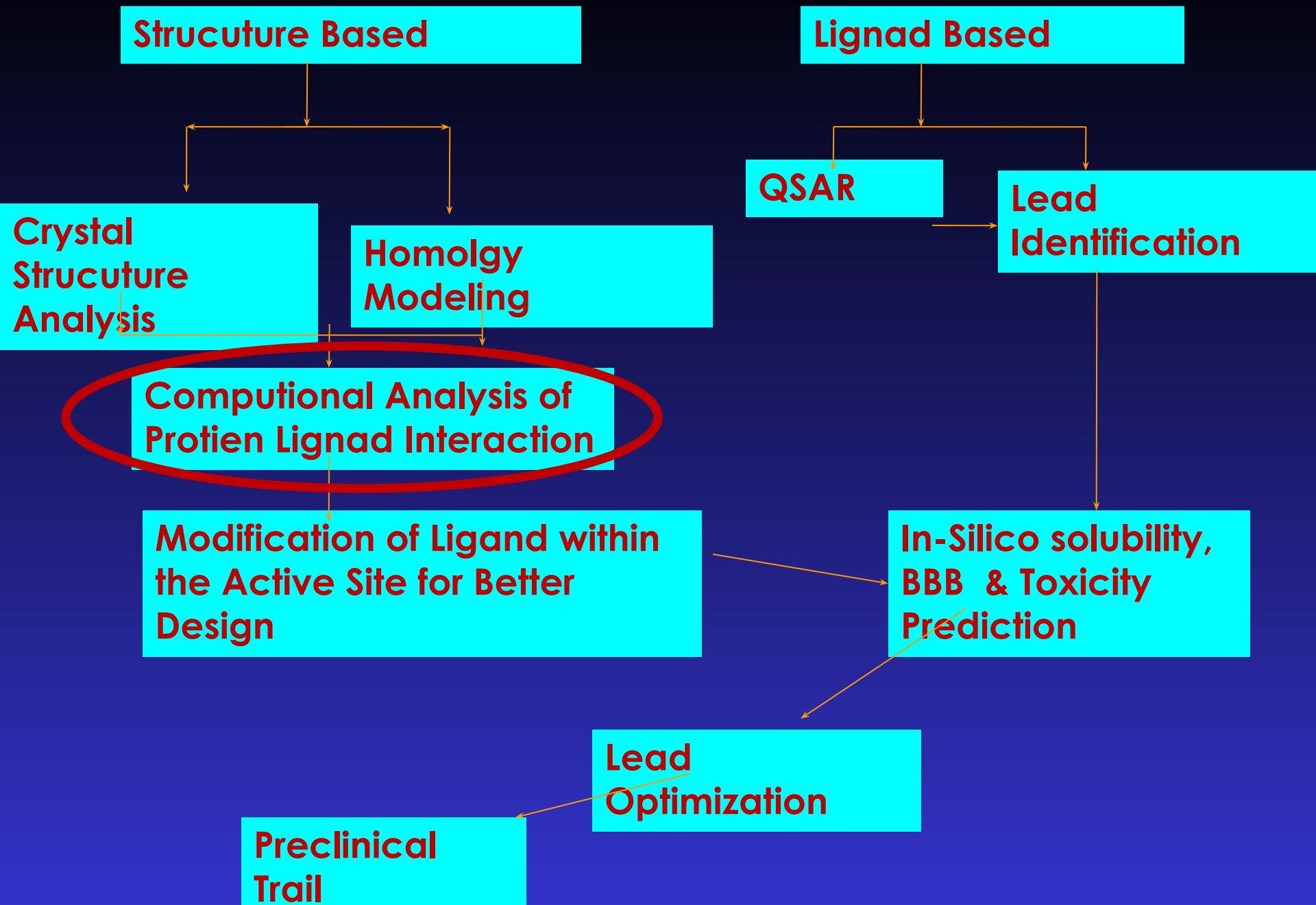
1. **Pharmacophore Modelling & QSAR**
   (Structure (s) of known active small molecules)

CADD

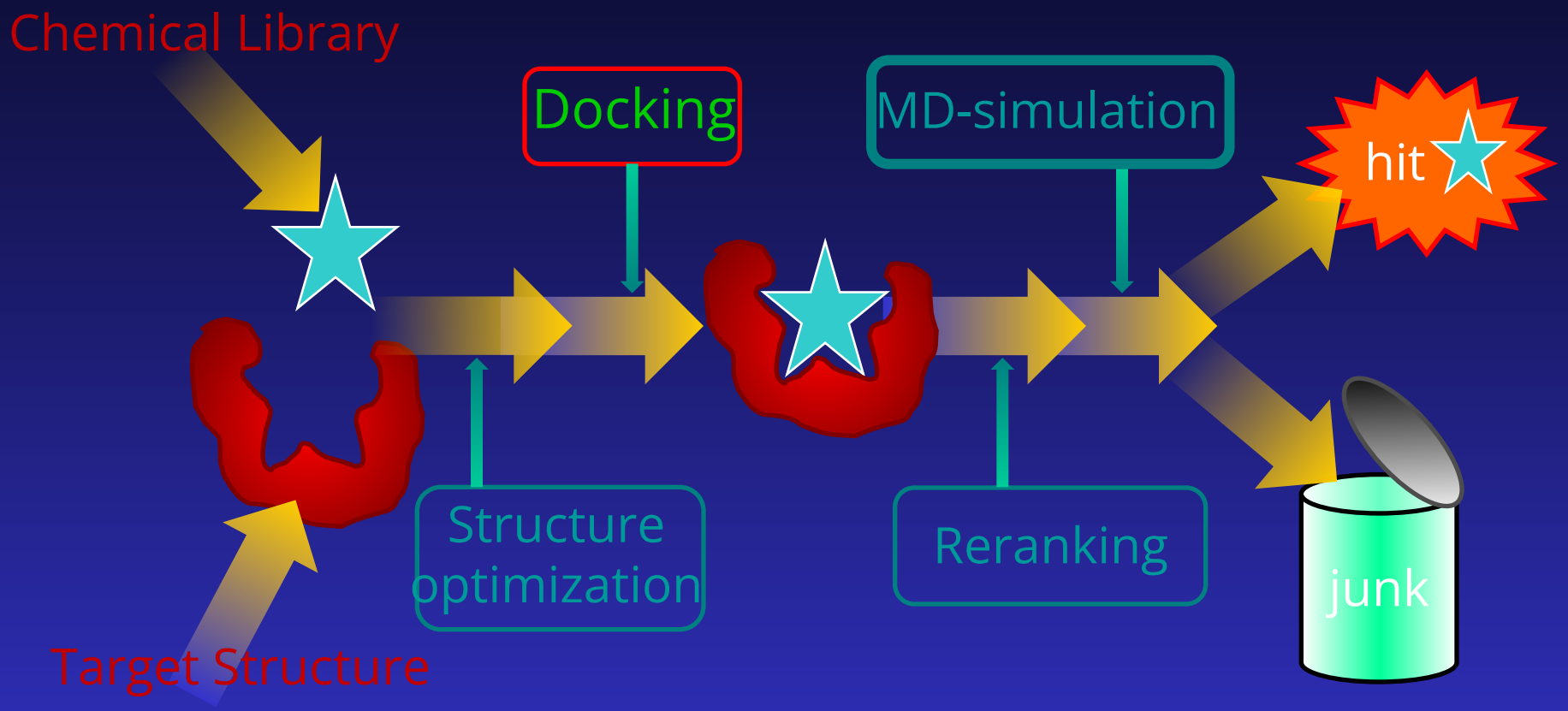
In silico Design based on prior knowledge about SAR

	Receptor-based Design	Ligand-based Design
	Target Structure Present	**Target Structure Absent**
Ligand Structure Present	Molecular Docking (Virtual Screening)	Pharmacophore-QSAR Modeling (Virtual Screening)
Ligand Structure Absent	*De Novo* Design	Library Design

CADD (Approaches)



Computer-aided Structure-based Rational Drug Design



Molecular Docking

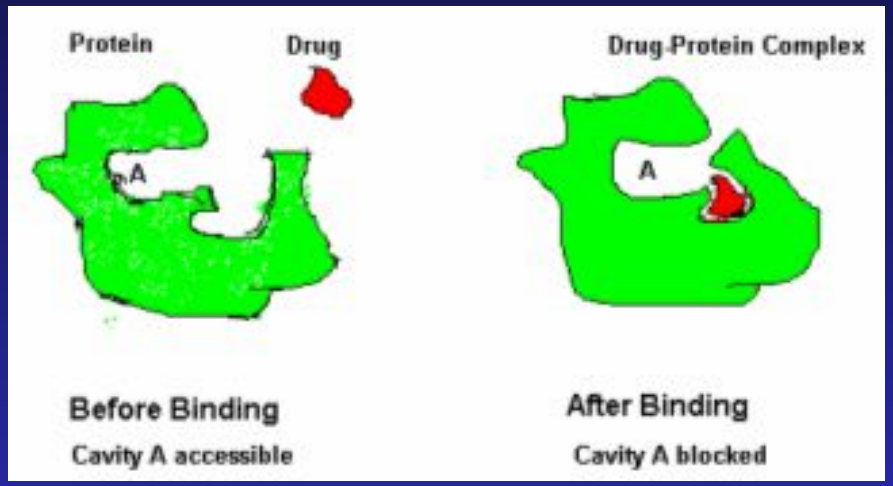
- **It is a computer program to predict the most probable mode of association of two molecules, receptor and ligand.**

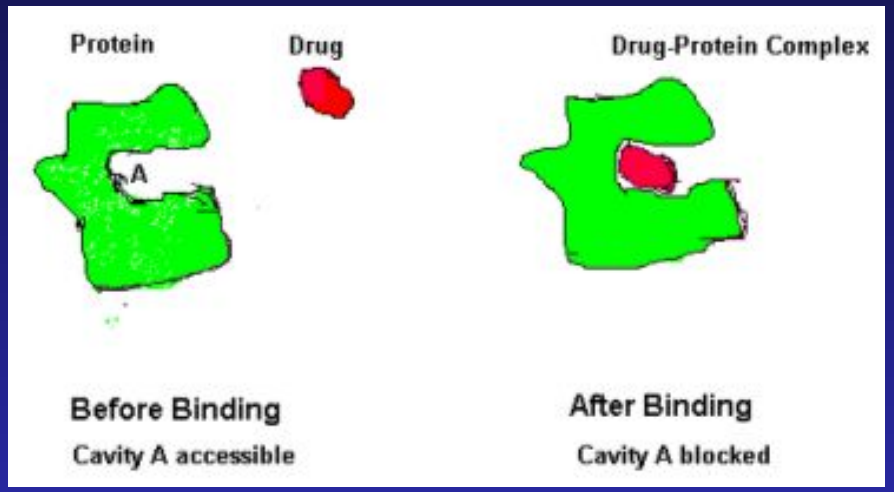
- Docking problem optimizes:
  - Binding between two molecules such that their orientation maximizes the interaction
  - Total energy of interaction such that for the best binding

Protein–Ligand Binding

Definition:

Computationally predict the structures of protein-ligand complexes from their conformations and orientations





Hence, the conformation and orientation that maximizes the interaction reveals the most accurate structure of the complex.

Three Components of Docking

3-D co-ordinates:
Structure

Representation of receptor binding site and ligand

Docking Program:
Sampling function

Sampling of configuration space of the ligand-receptor complex

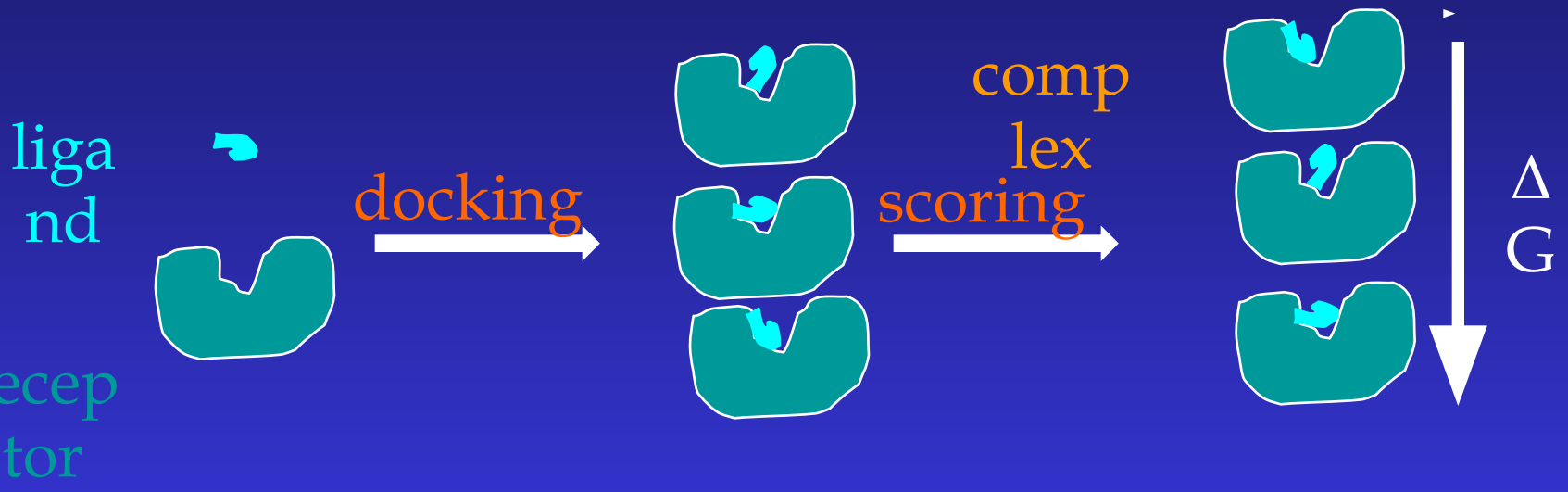
Docking Program:
Scoring function

Evaluation of ligand-receptor interactions

Molecular Docking: two components

1. **Search strategy for Conformation Sampling:** To position the ligand into the binding site of the receptor and look for various possible conformation as well as orientations of the ligand in the manners appropriate for optimal interactions with a receptor and, hence generates various docked poses.

2. **Fitness or Scoring function:** To evaluate and rank the docked poses of ligand-receptor base on scoring functions that depends on the interactions. (This can also estimate the binding affinity of the ligand.)



Docking scenario

Rigid Receptor, Rigid Ligand

Lock and Key

Rigid Receptor, Flexible Ligand

Currently used Docking programs

Flexible Receptor, Flexible Ligand

Induced Fit, under Development stage

Types of docking

(Based on conformation of ligand)

1) **Rigid ligand docking**

(only translation and rotation of ligand in the binding site, no change in conformation of ligand)

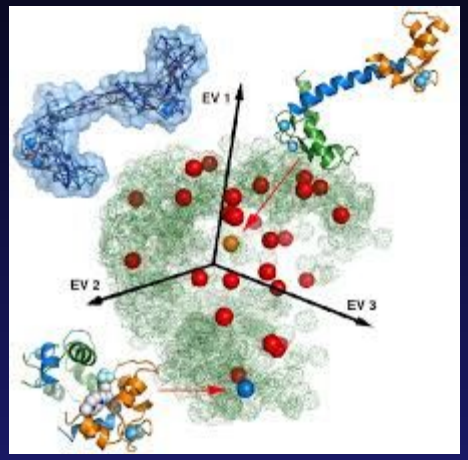
1) **Flexible ligand docking**

(translation and rotation of ligand in the binding site as well as changes in conformation of ligand)

Issue with flexible docking

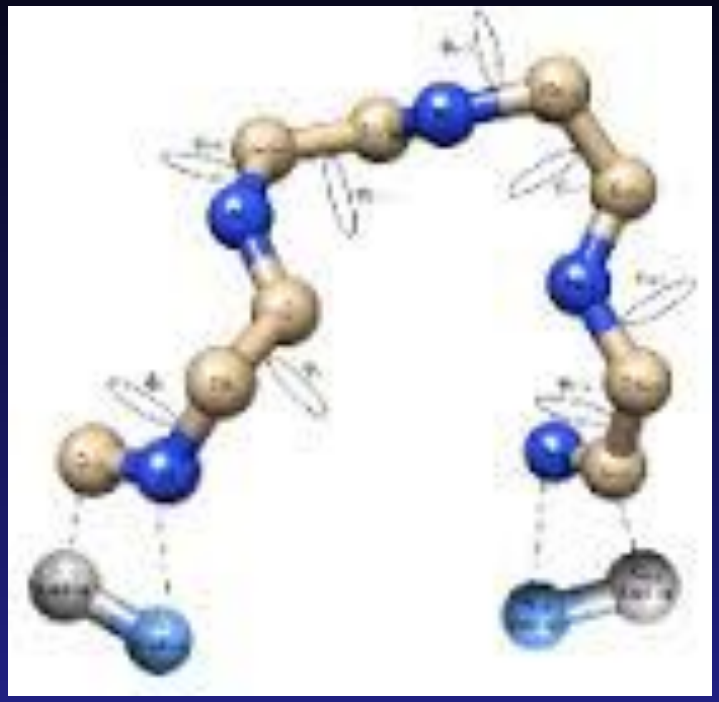
- 3 translational Degrees of freedom
- 3 rotational degrees
- Dihedral rotation i.e rotatinal degree of freedom
- 6 + n Dimensional space ligand
- *n* number of rotatable
- Systematic and random generation of conformers

Positional Sampling of Ligand



- Translation (DoF=3)
- Rotation (DoF=3)

Conformational Sampling Strategy



DoF=number of rotational tortion

- Systematic
- Stochastic

Systematic Conformation Search

- Exhaustive
- Deterministic
- Dependent on granularity of sampling
- Feasible only for low-dimensional problems
- DOF, number of rotatable bond in ligand

Stochastic Conformation Search

- Random
- Outcome varies
- Repeat to improve chances of success
- May miss real low energy conformation
- Faster as compared to systematic search
- Feasible for higher-dimensional problems

Conformation Search Algorithm

- Monte Carlo methods (MC)
- Molecular Dynamics (MD)
- Simulated Annealing (SA)
- Genetic Algorithms (GA)
- Lamarckian Genetic Algorithm (LGA)

- Docking Program:
  - AutoDock (MC,SA,GA, LGA)
  - GOLD (GA)
  - GLIDE (Systematic)
  - LigandFit (MC)
  - CDocker (MD)

Scoring Function

Free Energy of Binding

DG = H + TD S
DG = -RTlnKd

$$\Delta G_{binding} = \Delta G_{vdW} + \Delta G_{elec} + \Delta G_{hbond} + \Delta G_{desolv} + \Delta g_{tors}$$

- ΔG_{vdW} 12-6 Lennard-Jones potential Δg_{hbond} 12-10
- Δg_{elec} Coulombic with Solmajer-dielectric
- Potential with Goodford Directionality
- Δg_{desolv} Stouten Pairwise Atomic Solvation Parameters
- Δg_{tors} Number of rotatable bonds

Scoring Function

- Empirical Scoring
- Force Field Scoring
- Knowledge-based Scoring

- Consensus Scoring

Force Field-Based Scoring Function

- Describe only enthalpic contributions (ΔH).
- No estimate of ΔG.
- Time consuming
- Use non-bonded interactions

$$E_{non\_bonded} = \sum_{i}^{lig} \sum_{j}^{prot} \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + 332 \frac{q_i q_j}{D r_{ij}} \right)$$

- Augment force field terms with solvation and entropy terms

Empirical Scoring Functions

- Data: A set of protein-ligand complexes with known 3D structures and binding affinities (ΔG)
- Scoring parameters fit to reproduce
- Use regression to fit coefficients to a set of physically motivated terms in order to reproduce the experimental binding affinity of a training set of known protein-ligand complexes.

$$\Delta G = \Delta G_0 + \Delta G_{rot} N_{rot} + \Delta G_{hb} \sum_{neutral\_Hbonds} f(\Delta R, \Delta\alpha)$$

$$+ \Delta G_{io} \sum_{ionic\_\text{int}} f(\Delta R, \Delta\alpha) + \Delta G_{aro} \sum_{aro\_\text{int}} f(\Delta R, \Delta\alpha)$$

$$+ \Delta G_{lipo} \mid A_{lipo} \mid$$

- ΔG_0: Lose of translational/rotational entropy.
- Δg_{rot}: Lost of conformational DOF (ligand entropy).
- N_{rot}: Number of rotatable bonds immobilized during complex formation.
- ΔG_{hb}: ΔG_{io}: Hydrogen bonds (neutral, charged).
- ΔG_{aro}: Interaction between aromatic groups.
- ΔG_{lipo}: Accounts for lipophilic interactions.
- A_{lipo}: Receptor-ligand lipophilic contact surface area.

Empirical Scoring Functions: Training Data Set

Protein–ligand complex	PDB entry	Exp.
Adenosinedeaminase – deazaadenosine	1ADD	6.74
Carbonic anhydrase – methazolamide	1BZM	6.03
Carboxypeptidase – benzylsuccinate	1CBX	6.30
Carboxypeptidase – sulfodiimide	1CPS	6.66
Cytidine deaminase – 4 dehydrozebularine	1CTT	4.52
Elastase – TFA-Lys-Pro-*p*-isopropylanilide	1ELA	6.35
Elastase – TFA-Lys-Phe-*p*-isopropylanilide	1ELC	7.15
FKFB – FK506	1FKF	9.70
HIV protease – VX478	1HPV	9.22
HIV protease – XK263	1HVR	9.51
Lysozyme (L99A mutant) – benzofuran	1L82	3.95
Lysozyme (L99A mutant) – benzene	1L83	3.75
Lysozyme (L99A mutant) – phenylbutane	1L86	4.85
Lysozyme (L99A mutant) – *p*-xylene	1L87	3.37

Empirical Scoring Function: Ludi

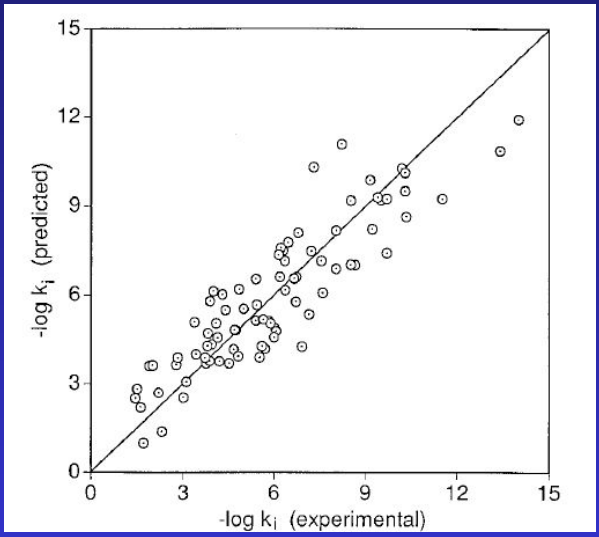
$$\Delta G = \Delta G_0 + \Delta G_{rot} N_{rot} + \Delta G_{hb} \sum_{neutral\_Hbonds} f(\Delta R, \Delta\alpha)$$

$$+ \Delta G_{io} \sum_{ionic\_int} f(\Delta R, \Delta\alpha) + \Delta G_{aro} \sum_{aro\_int} f(\Delta R, \Delta\alpha)$$

$$+ \Delta G_{lipo} \mid A_{lipo} \mid$$

#	1	2	3	4	5	6	7[a]	8
ΔG_0	+5.4	−1.4	−2.6	−2.9	−1.5	−1.8	−2.8	0.0[b]
ΔG_{hb}	−4.7	−3.1	−3.3	−3.2	−3.3	−3.7	−3.2	−3.4
ΔG_{ionic}	−8.3	−6.6	−7.0	−6.1	−6.1	−6.2	−5.7	−5.9
ΔG_{lipo}	−0.17	−0.15	−0.15	−0.14	−0.10	−0.09	−0.09	−0.10
ΔG_{rot}	+1.4	+1.0	+1.0	+0.9	+1.1	+1.1	+1.0	+1.1
ΔG_{aro}	–	–	–	−3.1	−2.5	−2.8	−2.6	−2.6
$\Delta G_{lipo\ water}$	–	–	–	–	−1.1	−1.2	−1.3	−1.4
ΔG_{esrep}	–	–	–	–	–	+0.6	+0.5	+0.4
α	0	0	0	0	0	0	0.5	0.5
β	1.0	1.0	1.0	1.0	1.0	1.0	1.2	1.2
TOL (Å)	0.20	0.20	0.20	0.20	0.20	0.20	0.25	0.25
s	9.5	8.8	8.6	8.1	7.5	7.4	7.3	7.4
r	0.835	0.841	0.837	0.859	0.882	0.887	0.890	0.890

Protein–ligand complex	PDB entry	$-\log K_i$	
		Pred.	Exp.
Acetylcholinesterase – tacrine	1ACJ	7.58	7.30
Carbonic anhydrase – dorzolamide	1CIL	8.30	9.43
Thrombin – benzamidine	1DWB	3.64	2.92
Thrombin – MQPA	1DWC	7.36	7.40
Rhinovirus coat protein – R61837	1R09	7.50	4.90[a]
Purine nucleoside phosphorylase – guanine	1ULB	4.25	5.30
PHBH – *p*-hydroxybenzoic acid	2PHH	8.27	4.68
Rhinovirus coat protein – Cmpd. IV	2R04	6.38	6.22[b]
Antibody – fluorescein	4FAB	8.60	10.53
Hemagglutinin – sialic acid	4HMG	3.86	2.55
HIV protease – A74704	9HVP	8.48	8.35



Knowledge based Scoring Functions

- Similar to force field based scoring function but statistically weighted.
- Free energies of molecular interactions derived from structural information on Protein-ligand complexes contained in PDB.

$$P(\sigma_p, \sigma_l) = P_{ref} \exp\left[-\beta F(\sigma_p, \sigma_l)\right]$$

Consensus Scoring

- No single scoring function performs best in all cases (although PMF seems to overall outperform other scoring functions).
- Studies comparing the performance of different scoring functions are sparse and conflicting.

Solutions

- Use two functions, one to identify the correct binding mode and one to score it.
- Use multiple scoring functions.
  - A molecule's consensus score is equal to its frequency of occurrence in the top rank percentile of each scoring function.

Type of Scoring Functions

FORCE FIELD BASED

- Dock-Score
- G-Score
- Gold Score
- Autodock free energy

EMPIRICAL

- Ludi
- X-Score
- Chem Score
- LigScore

KNOWLEDGE BASED

- PMF
- Drug Score
- SmoG

CONSENSUS Score

Binding Affinities

- **Ligand receptor binding affinity can be experimentally determined.**
- **Experimental errors lie in the range of 0.1-0.25 kcal/mol.**

$$PL \rightleftharpoons P + L$$

$$K_D = \frac{[P][L]}{[PL]}$$

$$\Delta G_{binding} = RT \ln K_D$$

$$\Delta G_{binding} = -RT \ln K_A$$

$\Delta G = -2.303 \times 8.3$ kJ/mol x 298K x log K_A

$\Delta G = -10.95$

$\Delta G = 8.8 \times 10^{-9}$ M