UPORABA MOLEKULSKIH SIMULACIJ PRI NAČRTOVANJU ZDRAVILNIH UČINKOVIN

Andrej Perdih

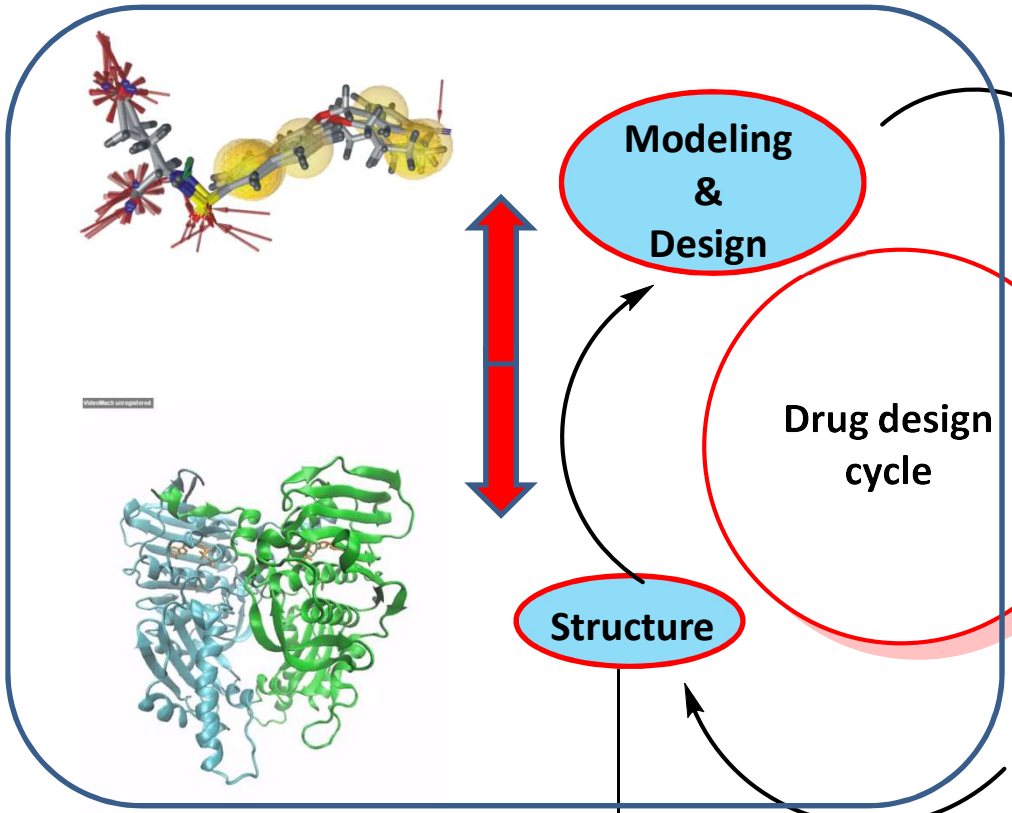
Delavnica: ***Računske in statistične metode v Kemiji***

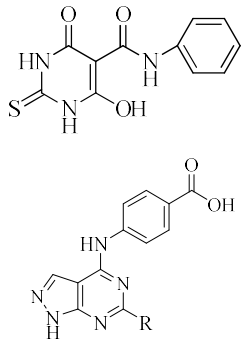
Ljubljana, 19. – 21. Oktober 2021

KEMIJSKI INŠTITUT

THE DRUG DESIGN CYCLE:

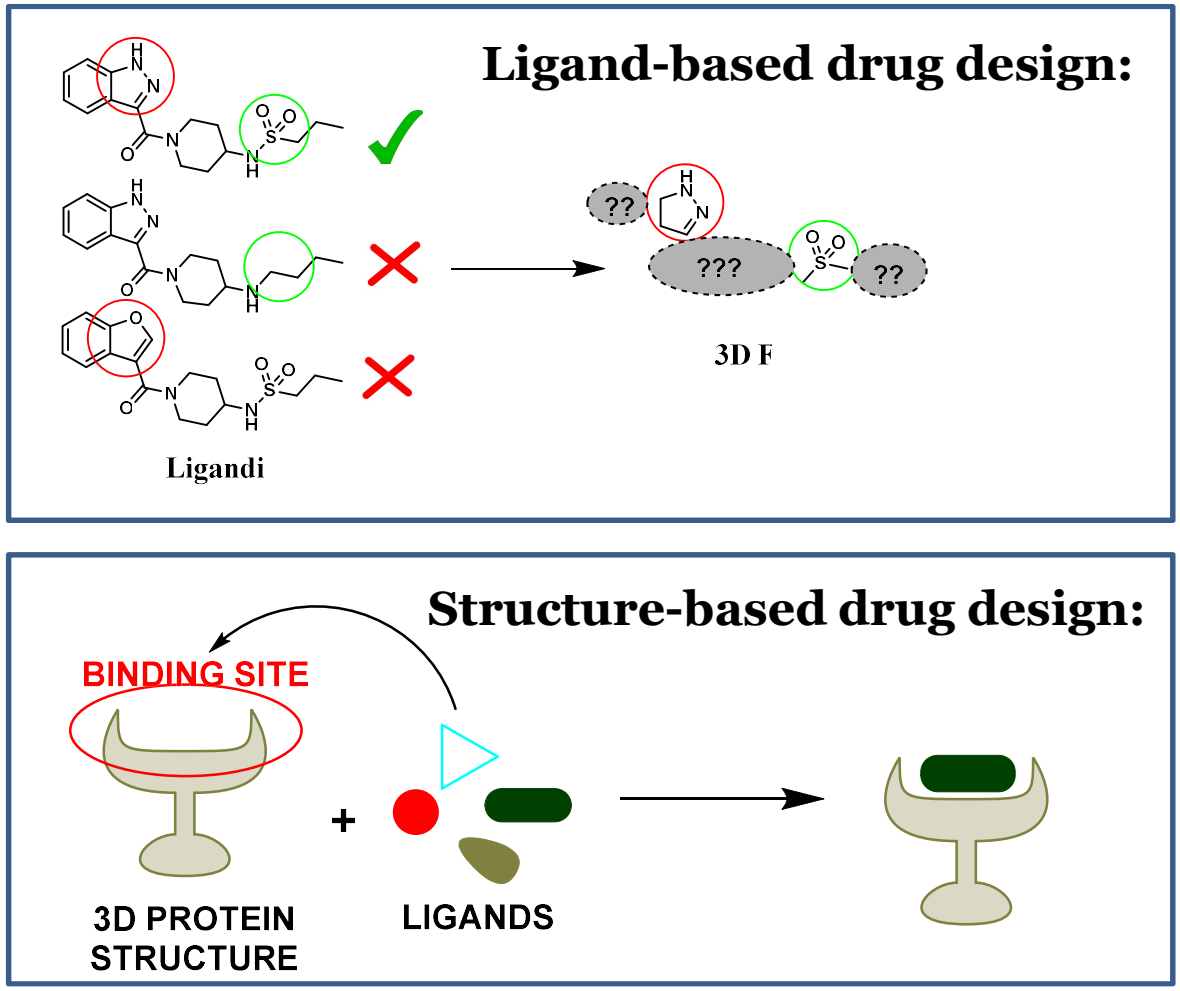
"THE BIG PICTURE"

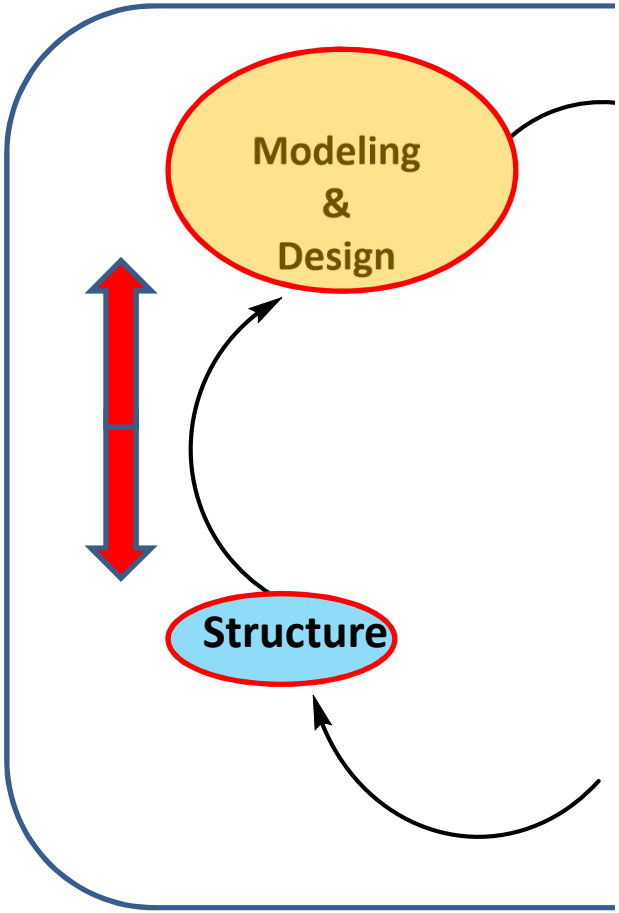




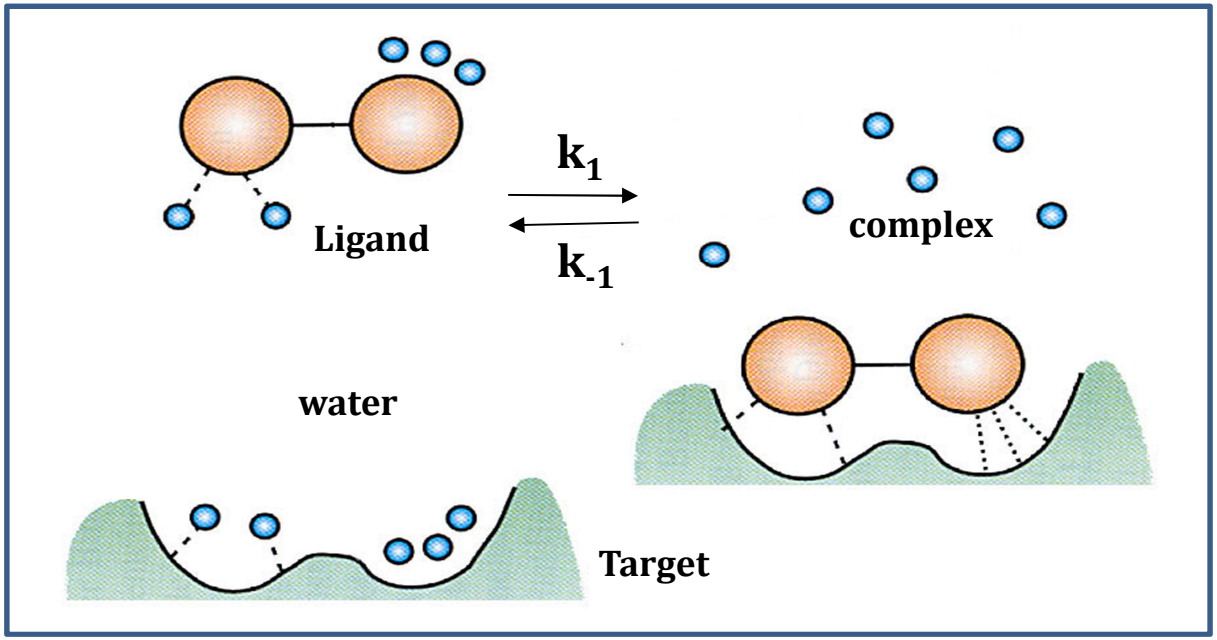
Warren, G.L.; et. al. *J. Med. Chem.*, **2006**, *49*, 5912-5931

COMPUTER-AIDED DRUG DESIGN:
METHODOLOGIES





MOLECULAR RECOGNITION:
THE BINDING EVENT



The free energy is the energy left for once you paid the tax to entropy:

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

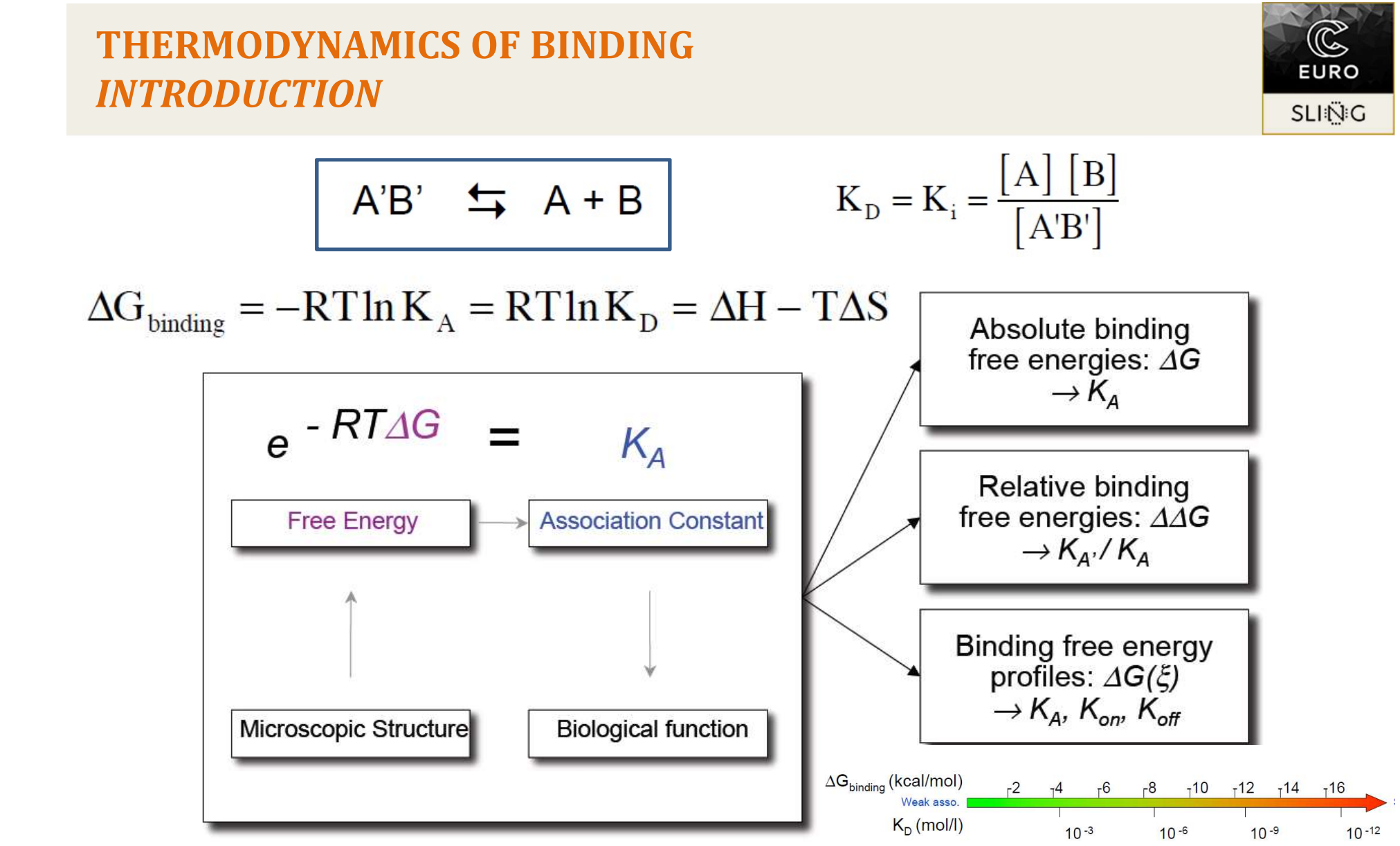
Enthalpic

- Hydrogen bonds
- Polar interactions
- Van der Waals interactions

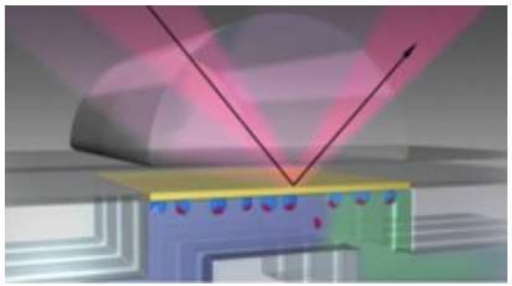
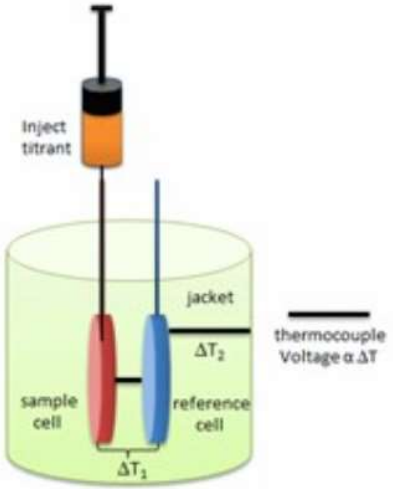
Entropic

- Loss of degrees of freedom
- Gain of vibrational modes
- Loss of solvent/protein structure

THERMODYNAMICS OF BINDING
INTRODUCTION



THERMODYNAMICS OF BINDING
EXPERIMENTAL DETERMINATION

SPR – Surface Plasmon Resonance	ITC – Isothermal Titration Calorimetry	FP – Fluorescence Polarization
		
		

MODELS OF MOLECULAR RECOGNITION

1890	lock-and-key	Emil Fischer

Lock and key

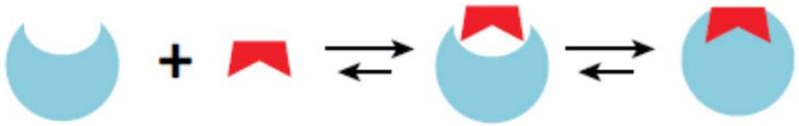
545. **Emil Fischer: Einfluss der Configuration auf die Wirkung der Enzyme.**

[Aus dem I. Berliner Universitäts-Laboratorium.]

(Vorgetragen in der Sitzung vom Verfasser.)

1958	induced-fit	Daniel Koshland

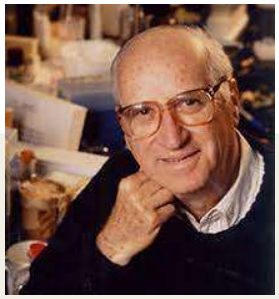
Induced fit



*APPLICATION OF A THEORY OF ENZYME SPECIFICITY TO PROTEIN SYNTHESIS**

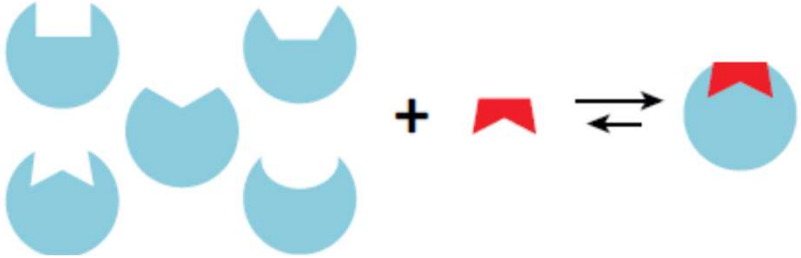
By D. E. Koshland, Jr.†

BIOLOGY DEPARTMENT, BROOKHAVEN NATIONAL LABORATORY, UPTON, NEW YORK



2003	conformation ensemble	Buyong Ma et al.

Conformational selection



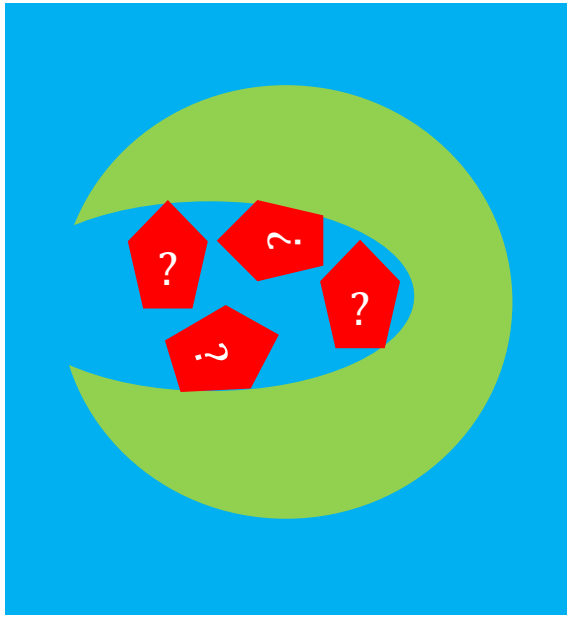
Protein Engineering vol.12 no.9 pp.713–720, 1999

Folding funnels and binding mechanisms

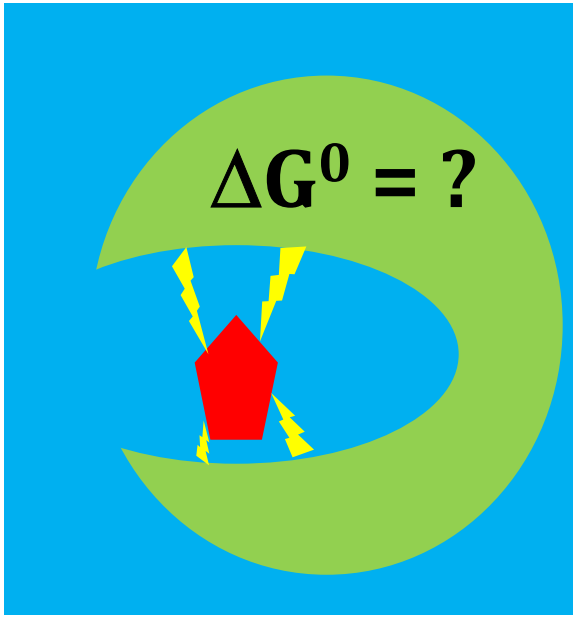
Buyong Ma[1], Sandeep Kumar[2], Chung-Jung Tsai[2] and Ruth Nussinov[2,3,4]

[1]Laboratory of Experimental and Computational Biology and [2]Intramural Research Support Program—SAIC, Laboratory of Experimental and Computational Biology, NCI–FCRDC, Frederick, MD 21702, USA and [3]Sackler Institute of Molecular Medicine, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv 69978, Israel

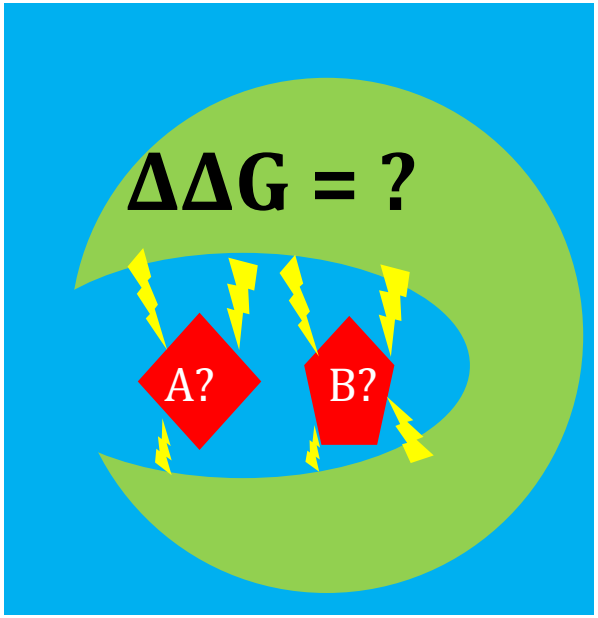
COMPUTATIONAL CHEMISTRY & LIGAND BINDING



Sampling/docking problem

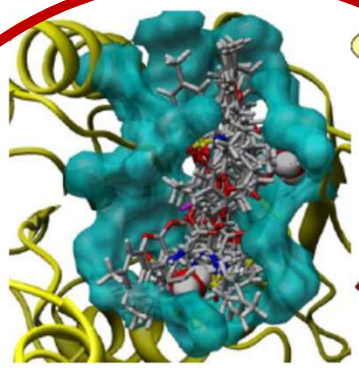


Scoring problem



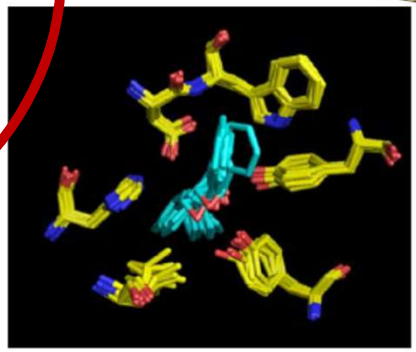
Relative binding affinity problem

COMPUTATIONAL STUDIES OF BINDING THERMODYNAMICS 1: MOLECULAR DOCKING



select initial structures

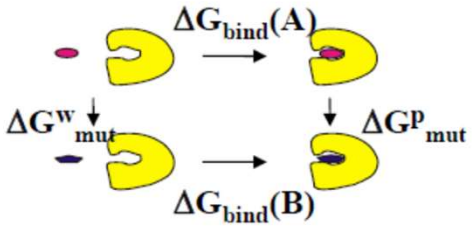
Automated docking + empirical scoring



structure refinement improved ranking

Simplified free energy calculations
Force Fields + MD/MC sampling
(LIE, MM-PBSA etc.)

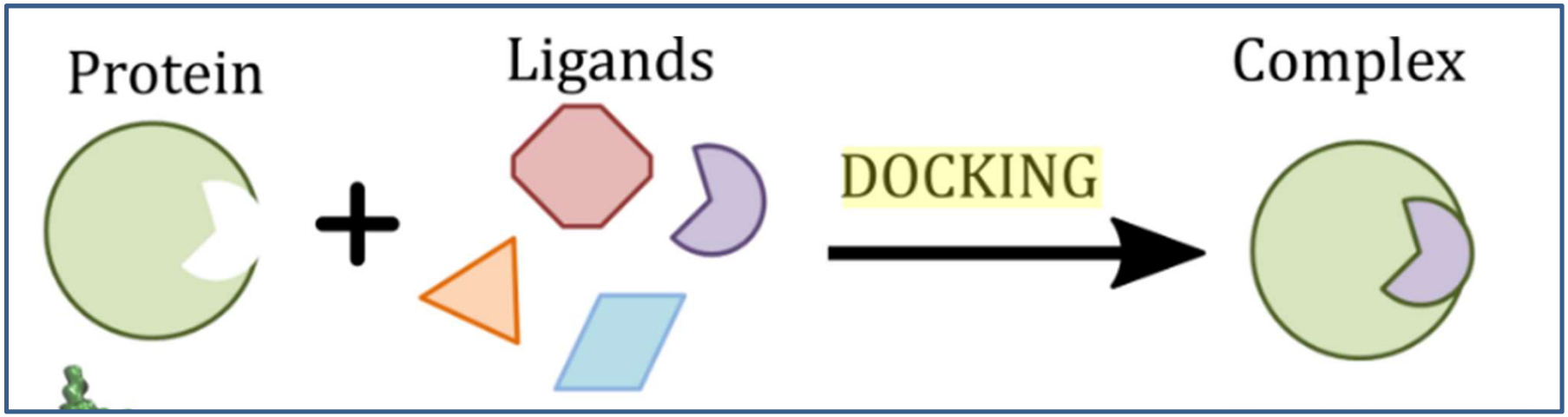
effects of small modifications



FEP/TI simulations

METHOD 1:
MOLECULAR DOCKING

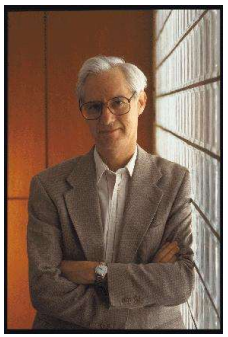
Molecular docking is a method of **structure-based drug design** that calculates the preferered conformation of a selected molecule (usually small ligand molecule, but also macromolecule) in a selected active/binding site of a biological macromolecule (target), assuming that they form a stable complex.



The main components of every docking software package are:

SEARCH ALGORITHM - serves to generate new conformations
SCORING FUNCTION - assessing the strength of binding interactions

Irwin Kuntz

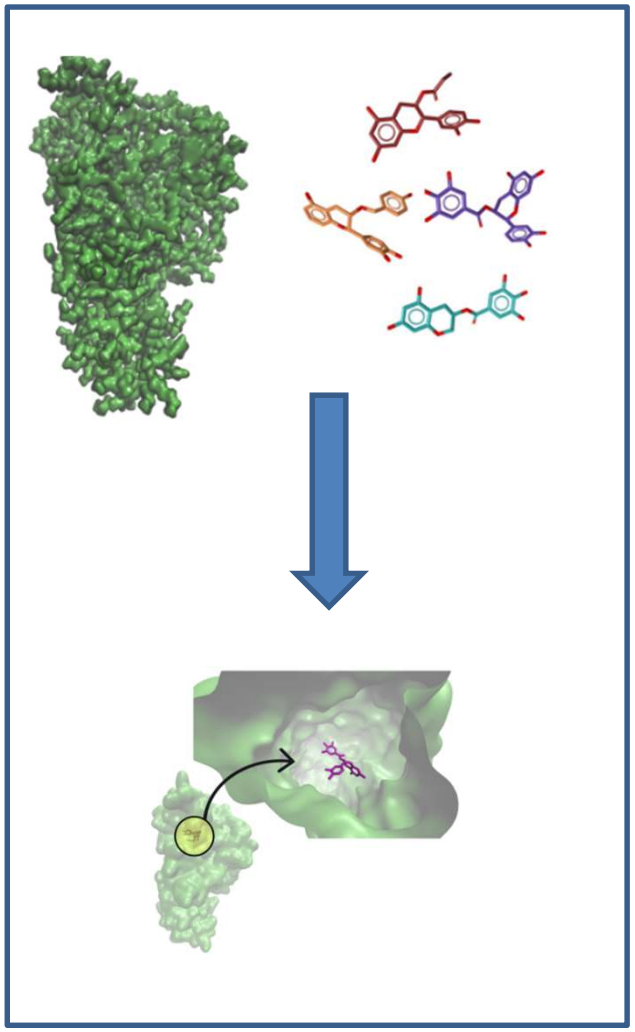


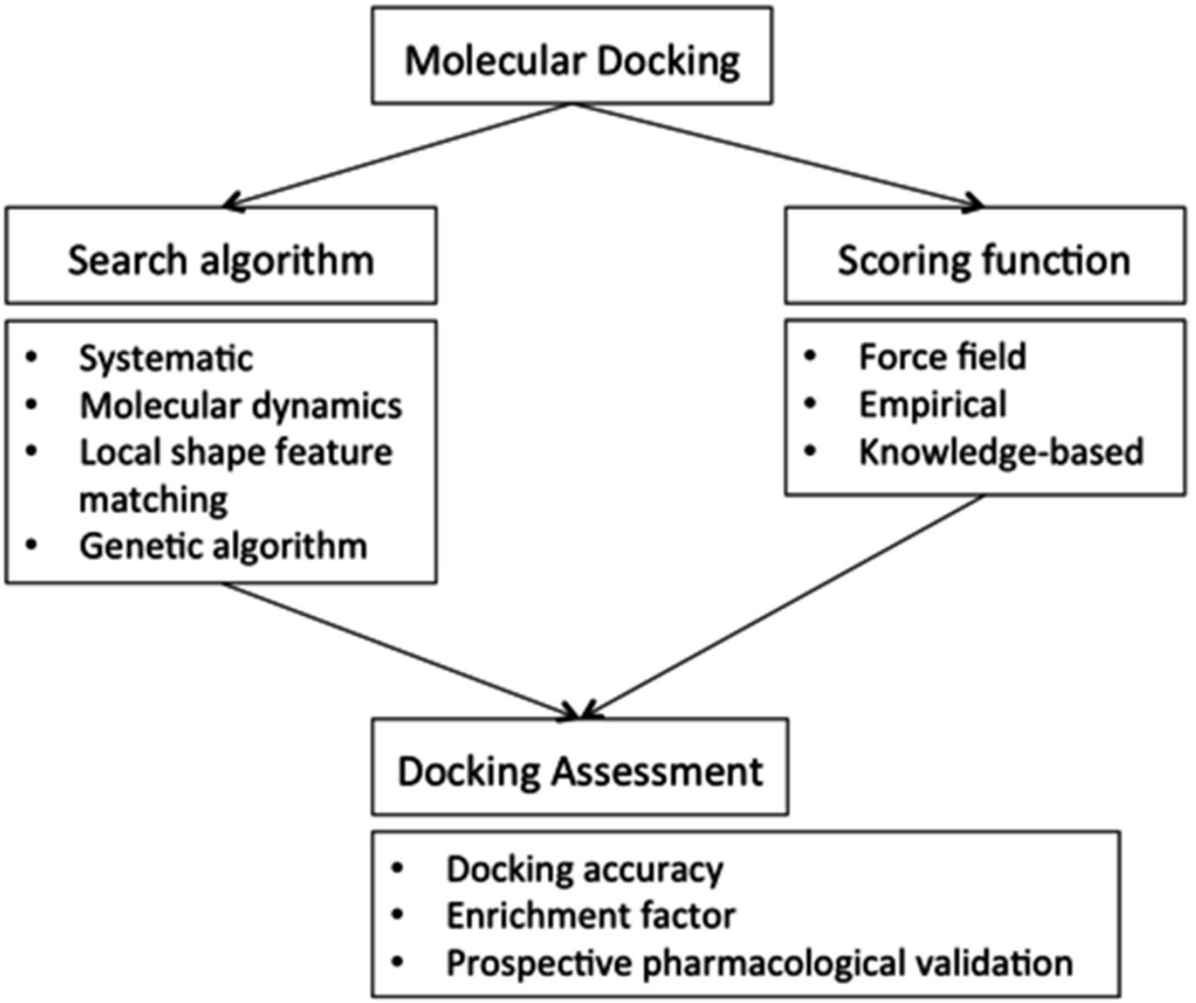
Automated Docking with Grid-Based Energy Evaluation

Elaine C. Meng, Brian K. Shoichet, and Irwin D. Kuntz*
Department of Pharmaceutical Chemistry, School of Pharmacy, University of California, San Francisco, California 94143-0446

Received 24 September 1991; accepted 4 December 1991

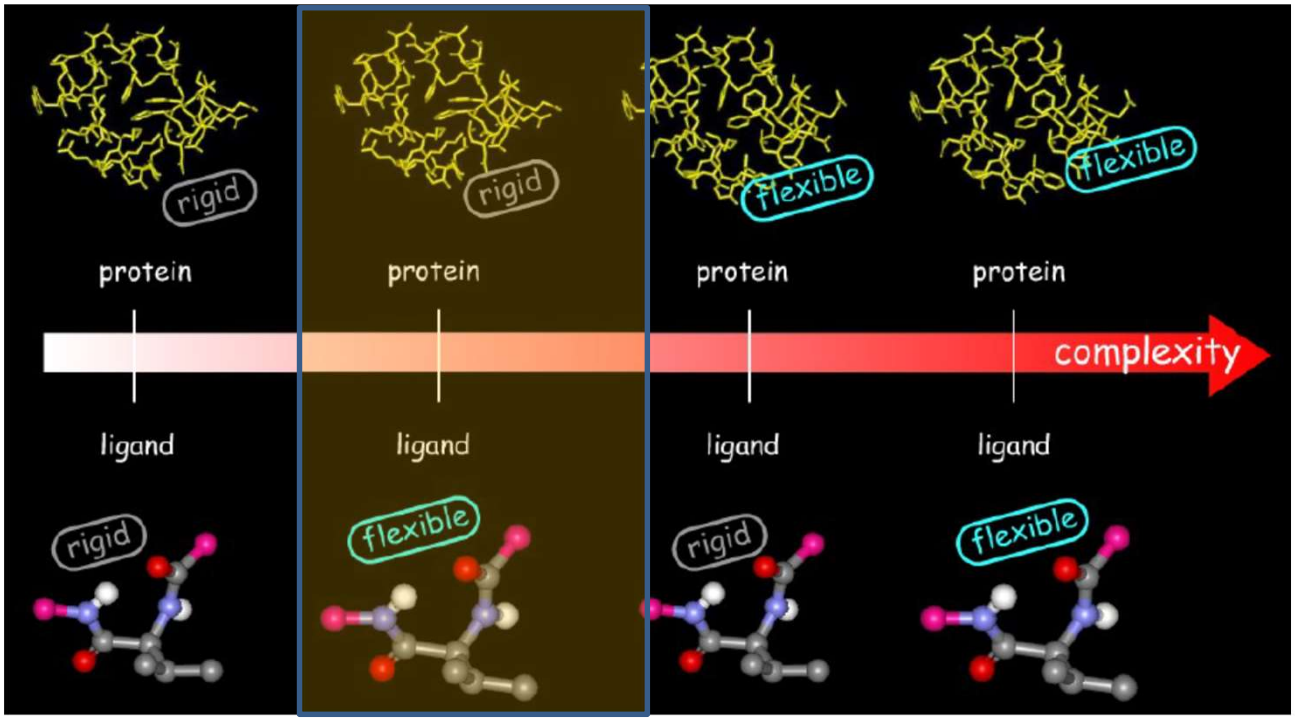
MOLECULAR DOCKING:
WORKFLOW





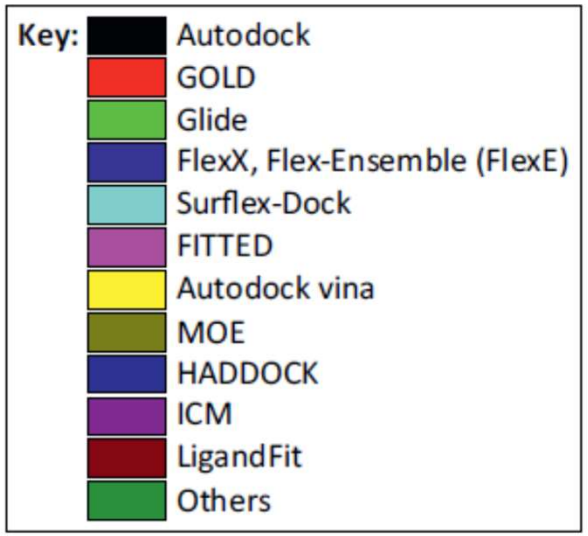
MOLECULAR DOCKING:
APPROCHES

1. Rigid body docking ignores the flexibility of the molecules and treats them like rigid objects
2. Rigid receptor – flexible ligand docking: only the ligand is treated as flexible, receptor is rigid
3. Flexible receptor – flexible ligand docking: both protein and ligand are treated as flexible.

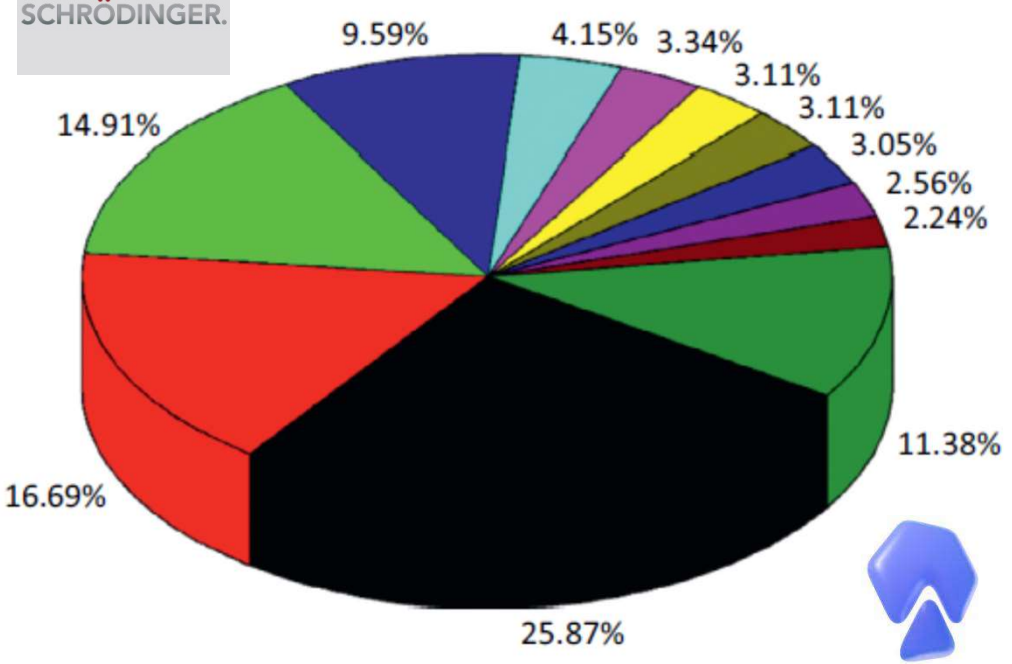


MOLECULAR DOCKING:
SOFTWARE

Software	Searching Algorithm	Native Scoring Function [1]
AutoDock [16]	Stochastic	Force-Field based
DOCK [17]	Systematic	Force-Field based
FlexX [18]	Systematic	Empirical
Glide [19]	Systematic	Empirical
GOLD [20]	Stochastic	Force-Field based
ICM [21]	Stochastic	Force-Field based
MOE [22]	Stochastic	Force-Field based



SCHRÖDINGER.



CCDC

AutoDock 4

CALCULATIONS OF BINDING THEROMDYNAMICS AND MOLECULAR SIMULATIONS

$$\Delta G = G_A - G_B = -k_B T \ln\left(\frac{Z_A}{Z_B}\right)$$

Free energy simulations techniques aim at computing ratios of partition functions using various techniques.

$$Z = \sum_i e^{-\beta E i}$$

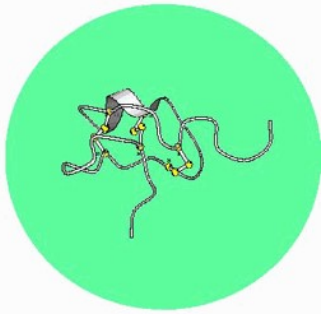
Sampling of important microstates of the system (MD, MC, GA, …)

Computation of energy of each microstate (force fields, QM, CP)

STATISTICAL MECHANICS:
PART 1

Thermodynamic properties are averages of microscopic quantities over the accessible microscopic states of the system

Molecular Simulation



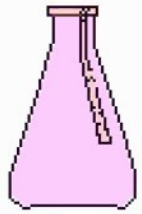
Microscopic

The partition function: normalisation factor of the probability

$$Q(N,V,T) = \frac{1}{h^{3N}N!}\int \dots \int \exp\left(\frac{-E_i(r_N,p_N)}{k_BT}\right) dr_N dp_N$$

The partition function contains all the information necessary to compute average quantities

Experiment



Macroscopic

We use formalisms of **statistical mechanics** to link the world of molecules with the macroscopic world of measurable quantities (e.g thermodynamic quantities).

STATISTICAL MECHANICS:
PART 2

$$E_i(r_N, p_N) = K_i(p_N) + U_i(r_N)$$

Kinetic energy

$$K_i(p_N) = \sum_j \frac{p_j^2}{2m_j}$$

Potential energy:
molecular interactions
Internal degrees of freedom

$$Q(N,V,T) = \left[\frac{1}{h^{3N}N!}\int \exp\left(\frac{-K_i(p_N)}{k_BT}\right)dp_N\right] \times \left[\int \exp\left(\frac{-U_i(r_N)}{k_BT}\right)dr_N\right]$$

"Ideal" partition function Q_{id}

Configurational partition function: Q_c

Easy to compute

Analytical expression:

$$Q_{id} = \prod_j \left(\frac{2\pi m_j k_B T}{h^2}\right)^{3/2} = \prod_j \frac{1}{\Lambda_j^3}$$

Comes from intermolecular forces and internal degrees of freedom

Analytical expression available only for few models: hard spheres, van der Waals fluids

Need to be evaluated numerically

EVALUATION OF CONFIGURATIONAL PARTITION FUNCTION Qc

$$\int \exp\left(\frac{-U_i(r_N)}{k_B T}\right) dr_N$$

Direct evaluation:
Generation of all the possible configurations of the system

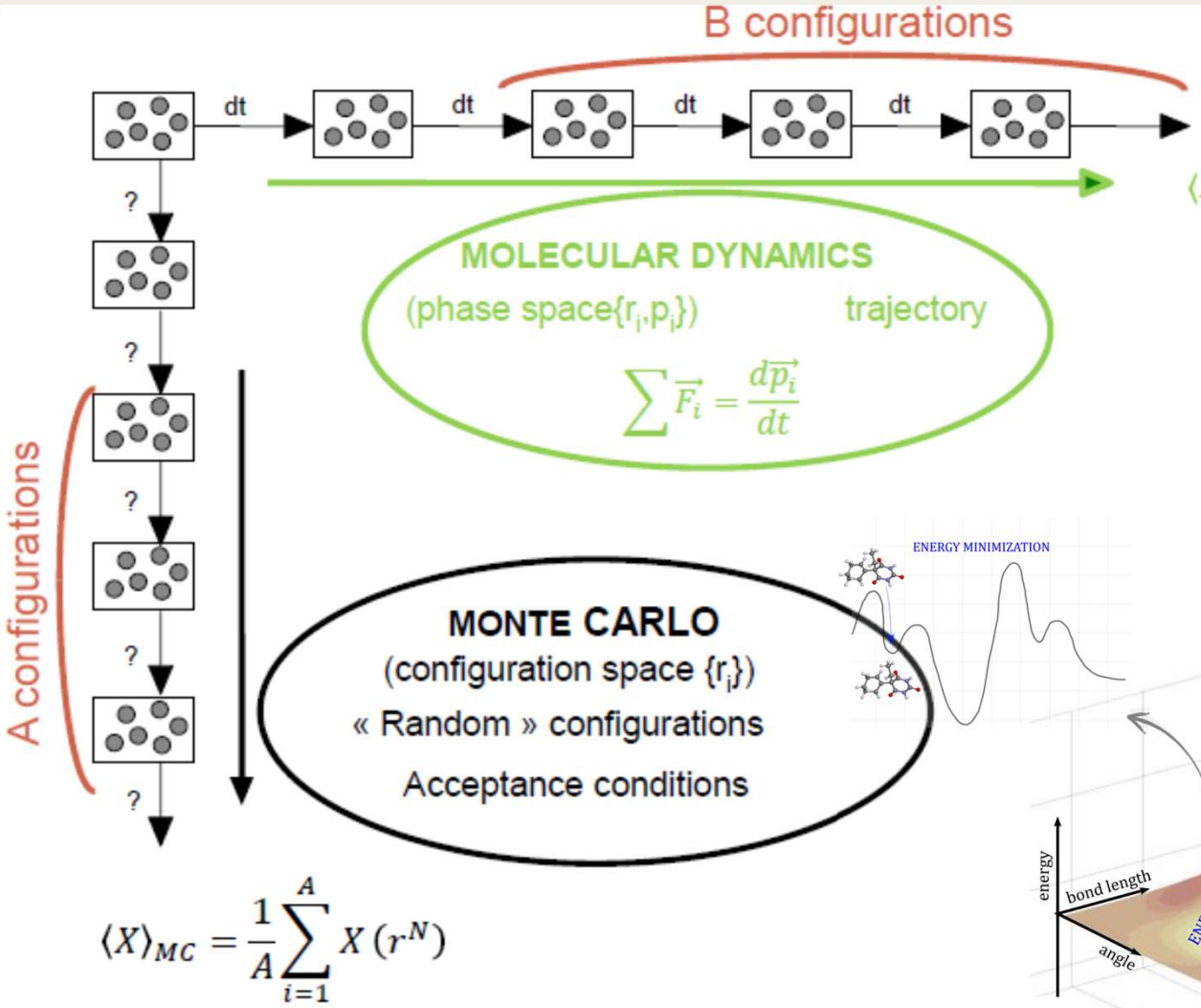
→ Impossible because of high dimensionality

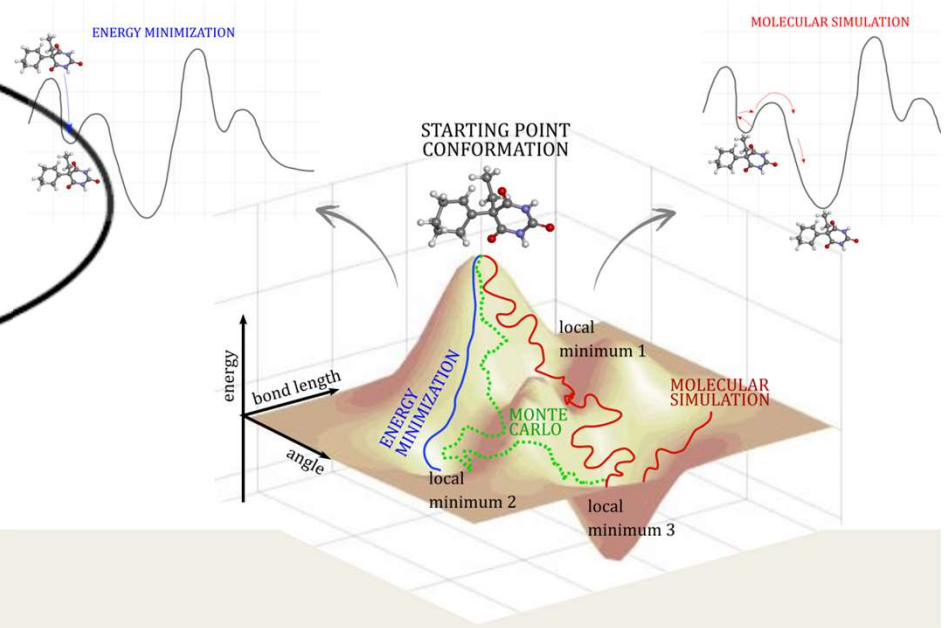
Over all the possible configurations, only few have a non-negligible contribution to Q_c

→ Generation of **representative** configurations

Molecular simulation

MOLECULAR SIMULATION:
METHODS





ERGODIC HYPOTHESIS

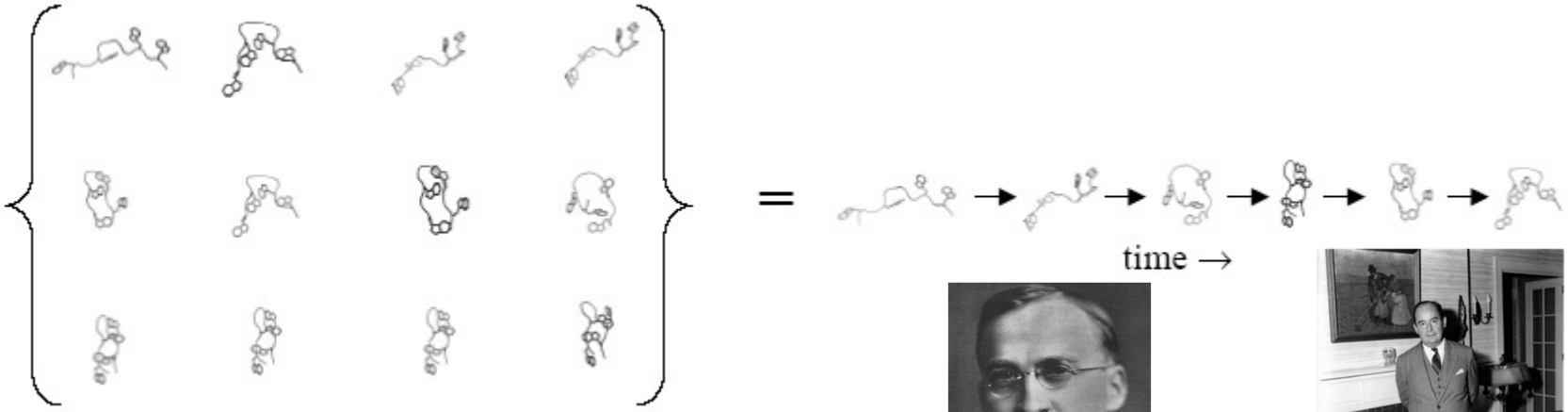
Ergodic hypothesis assumes that the **average** of a process parameter over **time** and the **average** over the **statistical ensemble** are the same.

- The statistical average of the quantity

$$\overline{X(r_N, p_N)} = \int \dots \int X(r_N, p_N)\, P(r_N, p_N)\, dr_N dp_N$$

The time average of the quantity

$$\langle X(t) \rangle = \frac{1}{\tau} \int_{t_0}^{t_0+\tau} X(t) dt$$



$$\overline{a(r_N, p_N)} = \langle a(t) \rangle$$

George D. Birkhoff

John von Neumann

PROOF OF THE ERGODIC THEOREM

By George D. Birkhoff

Department of Mathematics, Harvard University

PROOF OF THE QUASI-ERGODIC HYPOTHESIS

By J. v. Neumann

Department of Mathematics, Princeton University

MOLECULAR DYNAMICS: *REVISION*

We start with in initial coodinates of the molecular system **r**, and we select a **time interval Δt** (1 fs)

We calculate forces on each atom and consequently **acceleration**:

$$F(X) = -\nabla E(X) = -\frac{\partial E}{\partial X} \longrightarrow F_i = m_i a_i$$

N steps

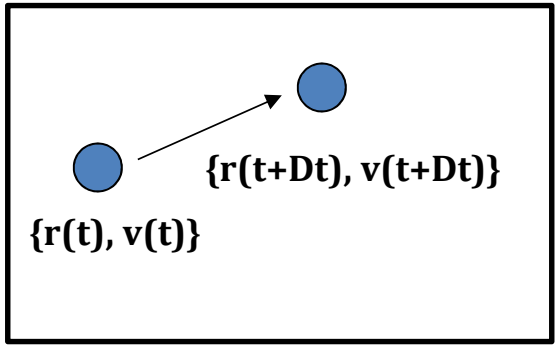
We move atoms in the direction of the acceleration using various **integration algorithm**

$$r^{i+1} = r^i + v^i \Delta t + 1/2 a \Delta t^2 + \ldots$$

We move time forward by **Δt**

Simulation time = N x Δt

A. Rahman

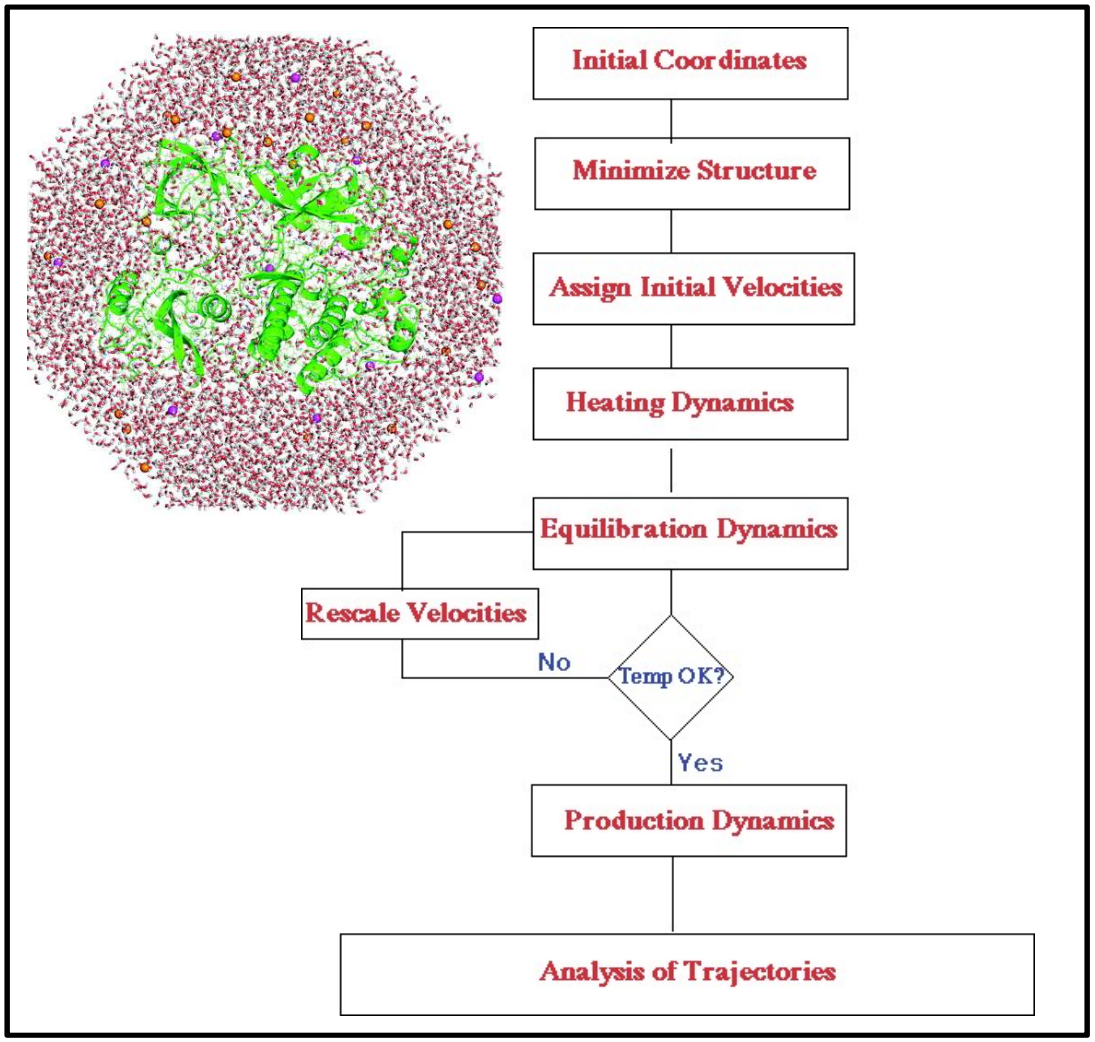


PHYSICAL REVIEW VOLUME 136, NUMBER 2A 19 OCTOBER 1964

Correlations in the Motion of Atoms in Liquid Argon*

A. RAHMAN
Argonne National Laboratory, Argonne, Illinois
(Received 6 May 1964)

MOLECULAR DYNAMICS SIMULATION:
USUAL SETUP



System preparation

Equlibration MD stage

Production MD stage

Trajectory analysis

MOLECULAR DYNAMICS:
SOFTWARE

Academic Packages

1) CHARMM
http://charmm.org/

2) AMBER
http://ambermd.org/

3) GROMACS
http://www.gromacs.org/

4) NAMD
http://www.ks.uiuc.edu/Research/namd/

Commercial Packages

5) DESMOND
http://www.deshawresearch.com/resources.html

6) **IMPACT**
https://www.schrodinger.com/

MOLECULAR SIMULATIONS & DRUG DESIGN

1. **GEOMETRY ASPECT**

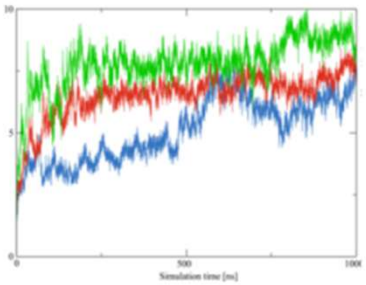
INCREASED UNDERSTANDING OF THE STRUCTURE-FUNCTION RELATIONSHIPS

2. **ENERGY ASPECT**

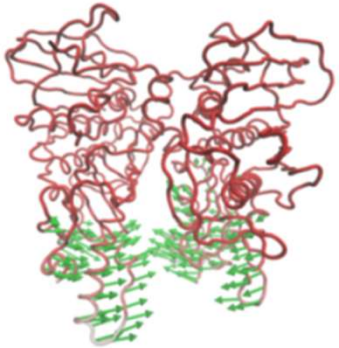
INCREASED UNDERSTANDING OF THE BINDING AND MOLECULAR RECOGNITION

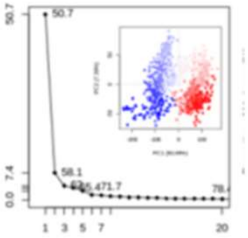
MOLECULAR SIMULATIONS
GEOMETRY ASPECT

RMSD parameter

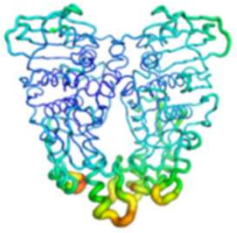


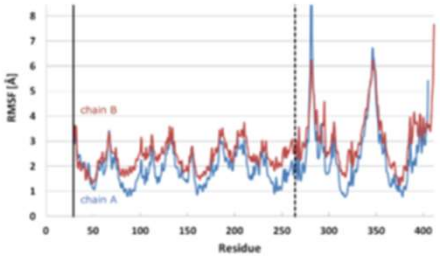
PRINCIPAL COMPONENT ANALYSIS



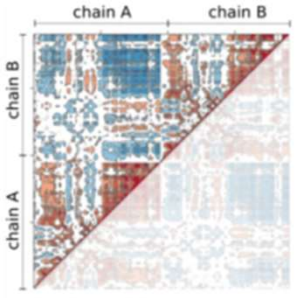


RMSF parameter





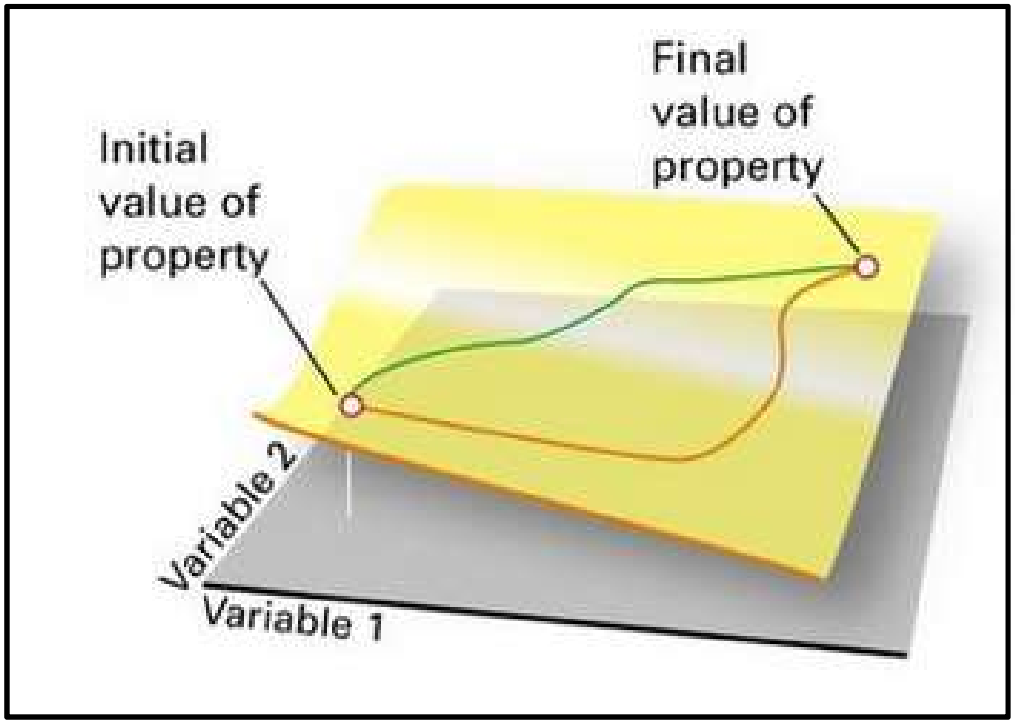
CROSS-CORRELATION ANALYSIS



MOLECULAR SIMULATIONS
ENERGETIC ASPECT

In simulations we are still limited with the representativeness of the generated molecular ensemble **quality of sampling**.

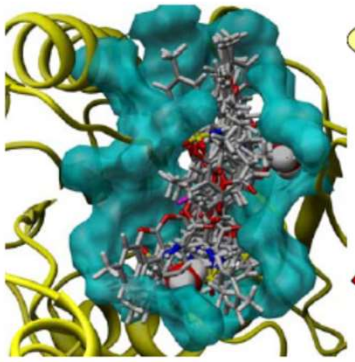
THERMODYNAMIC CYCLES enable that free energy is calculated from non-physical events.



$$\oint U = U_{i=f} - U_{i=f} = 0$$

Free energy is a state function not dependent of its pathway.

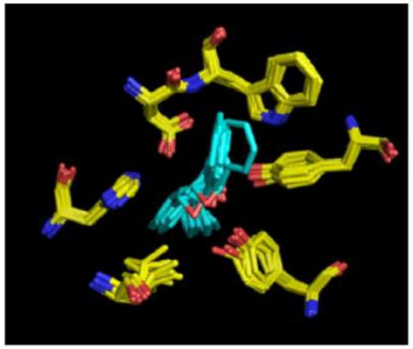
COMPUTATIONAL STUDIES OF BINDING THERMODYNAMICS 2: FREE ENERGY PERTRUBATION



select initial structures

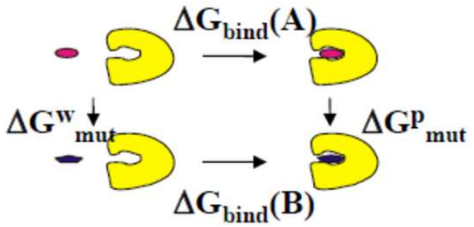
Automated docking + empirical scoring

structure refinement improved ranking



Simplified free energy calculations
Force Fields + MD/MC sampling
(LIE, MM-PBSA etc.)

effects of small modifications

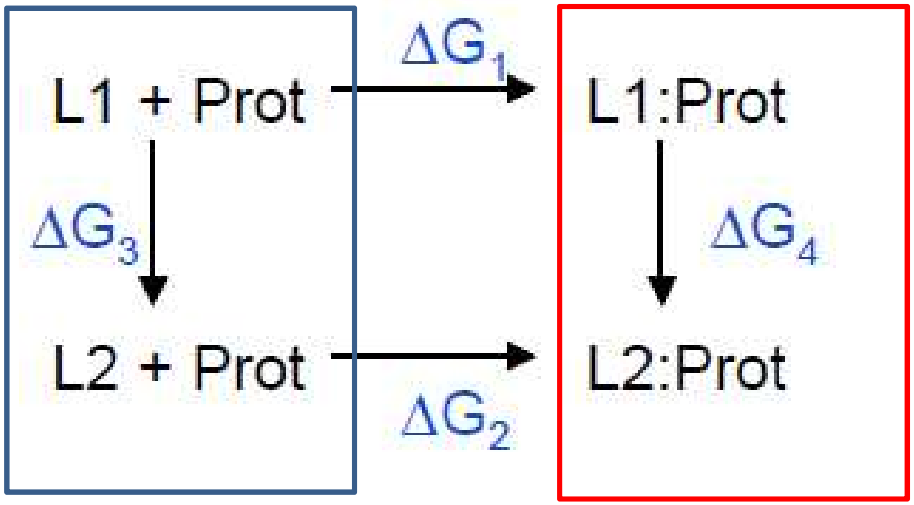


FEP/TI simulations

METHOD 2:
FREE ENERGY PERTRUBATION (FEP)

Robert W. Zwanzig

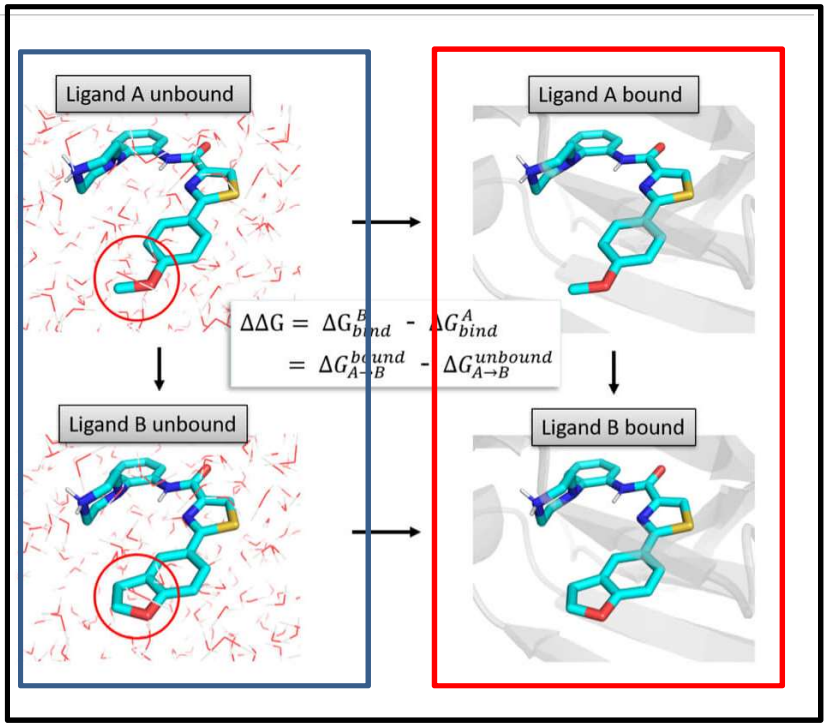
$$\Delta F(\mathbf{A} \rightarrow \mathbf{B}) = F_{\mathbf{B}} - F_{\mathbf{A}} = -k_{\mathrm{B}}T \ln \left\langle \exp\left(-\frac{E_{\mathbf{B}} - E_{\mathbf{A}}}{k_{\mathrm{B}}T}\right)\right\rangle_{\mathbf{A}}$$



Thermodynamic cycle perturbation approach:

$$\Delta\Delta G_{bind} = \Delta G_2 - \Delta G_1 = \Delta G_4 - \Delta G_3$$

$\Delta G_4 - \Delta G_3$ is computationally accessible

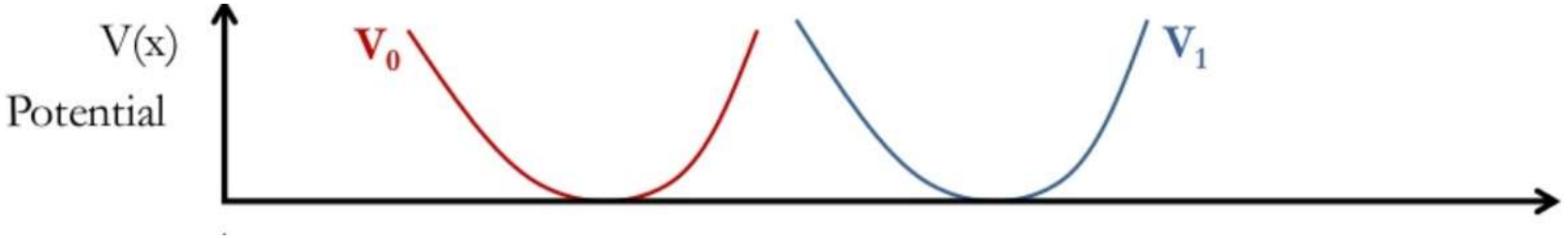


THE JOURNAL OF CHEMICAL PHYSICS VOLUME 22, NUMBER 8 AUGUST, 1954

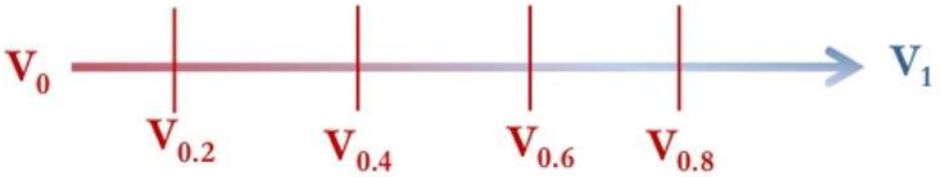
High-Temperature Equation of State by a Perturbation Method. I. Nonpolar Gases*

ROBERT W. ZWANZIG
Sterling Chemistry Laboratory,† Yale University, New Haven, Connecticut
(Received March 2, 1954)

METHOD 2:
FREE ENERGY PERTRUBATION (FEP)



Hence we can break the journey from transforming from potential V_0 to V_2 by splitting it into M sections:



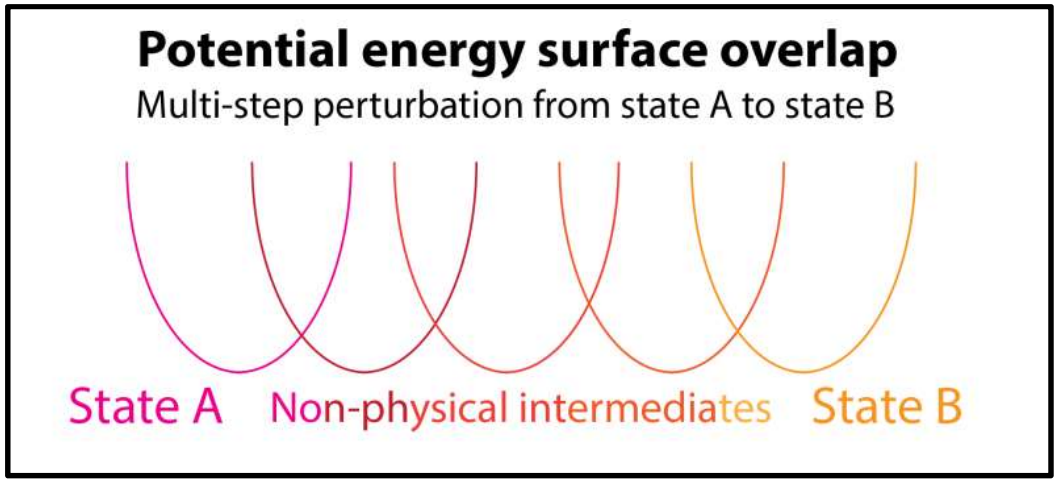
Here M=5

We can then calculate the free energy change for each section and add them together to generate the total free energy change on going from 1 to 2:

$$\Delta A_{0\to1} = \underbrace{\Delta A_{0\to0.2}} + \Delta A_{0.2\to0.4} + \underbrace{\Delta A_{0.4\to0.6}} + \Delta A_{0.6\to0.8} + \Delta A_{0.8\to1}$$

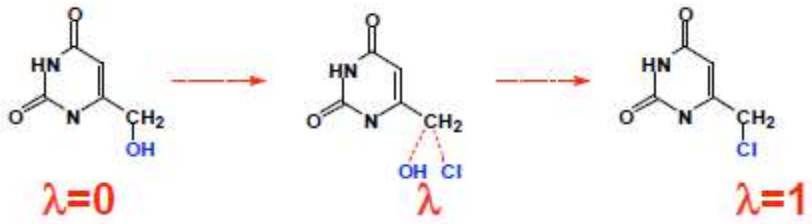
$$\Delta A_{0\to0.2} = -k_B T \ln \left\langle e^{-\beta(V_{0.2}(\mathbf{r})-V_0(\mathbf{r}))} \right\rangle_0 \qquad \Delta A_{0.4\to0.6} = -k_B T \ln \left\langle e^{-\beta(V_{0.6}(\mathbf{r})-V_{0.4}(\mathbf{r}))} \right\rangle_{0.4}$$

METHOD 2:
FREE ENERGY PERTRUBATION (FEP)



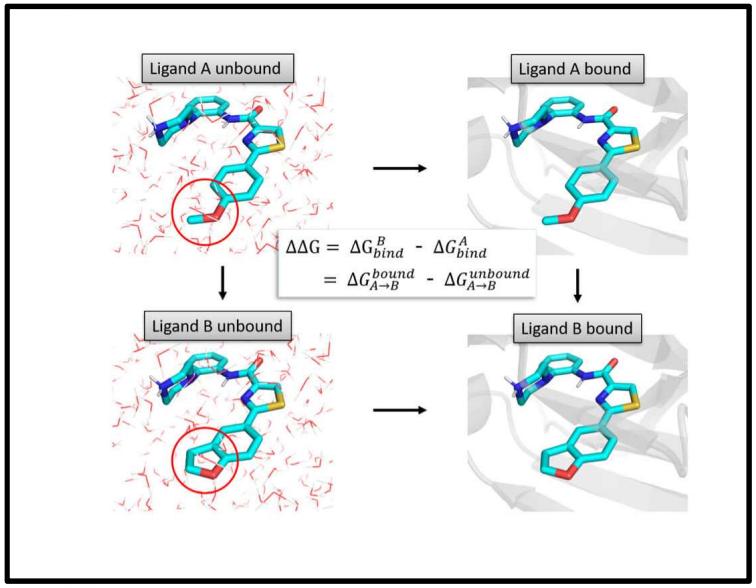
Coupling parameter λ

$$H_\lambda = H_0 + \lambda H_{L_1} + (1-\lambda) H_{L_2}$$



$$\Delta\Delta G_{bind} = -RT \sum_{i=0}^{n-1} \ln \left\langle \exp\left(-\left(H_{\lambda_i} - H_{\lambda_{i+1}}\right)/RT\right) \right\rangle_{\lambda_i}$$

METHOD 2:
FREE ENERGY PERTRUBATION (FEP)



$$\Delta\Delta G_{bind} = -RT \sum_{i=0}^{n-1} \ln \left\langle \exp\left(-\left(H_{\lambda_i} - H_{\lambda_{i+1}}\right)/RT\right) \right\rangle_{\lambda_i}$$

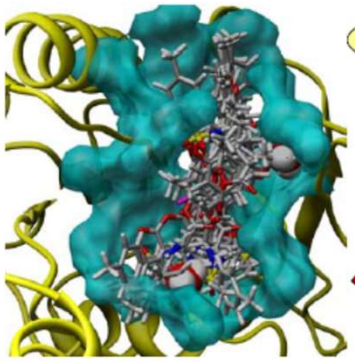
Advantages :

- Rigorous
- Estimates influence of small modifications
- No parameter to be fitted
- Partitioning of the free energy (TI)

Drawbacks :

- Restricted to small mutations of ligand or protein
- Most often: relative ΔG_{bind}
- Time consuming

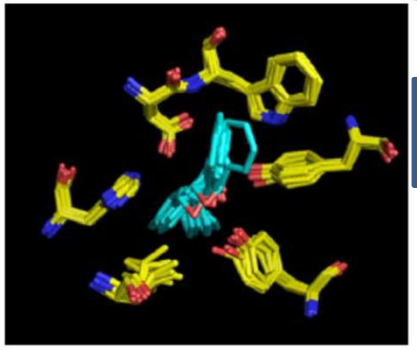
COMPUTATIONAL STUDIES OF BINDING THERMODYNAMICS 3: MM/PBSA METHOD



select initial structures

Automated docking + empirical scoring

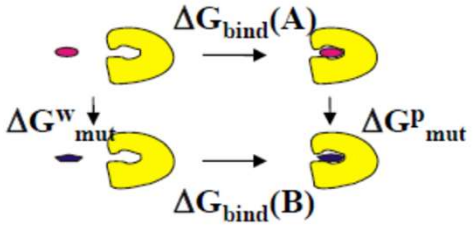
structure refinement
improved ranking



$$\Delta G = \alpha\left(\left\langle V_{L-P}^{vdW}\right\rangle - \left\langle V_{L-W}^{vdW}\right\rangle\right) + \beta\left(\left\langle V_{L-P}^{ES}\right\rangle - \left\langle V_{L-W}^{ES}\right\rangle\right)$$

Simplified free energy calculations
Force Fields + MD/MC sampling
(LIE, MM-PBSA etc.)

effects of small modifications



FEP/TI simulations

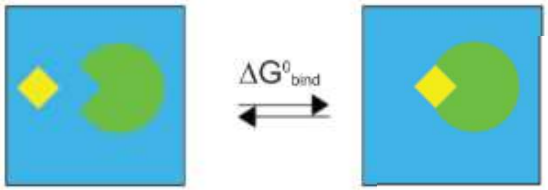
METHOD 3:
MM/PBSA METHOD

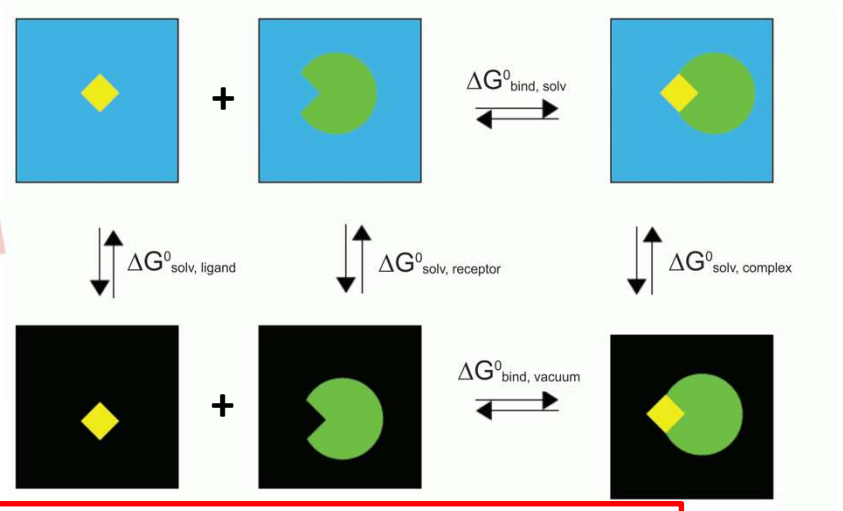
The acronym ***MM-PBSA*** stands for Molecular Mechanics- Poisson Bolzmann Surface Area

$$[A]_{aq} + [B]_{aq} \Leftrightarrow [A^*B^*]_{aq}.$$

Ideally we would like to calculate this free energy of binding directly.

Peter Kollman





Thus a more effective method is to divide up the calculation according to the following thermodynamic cycle:

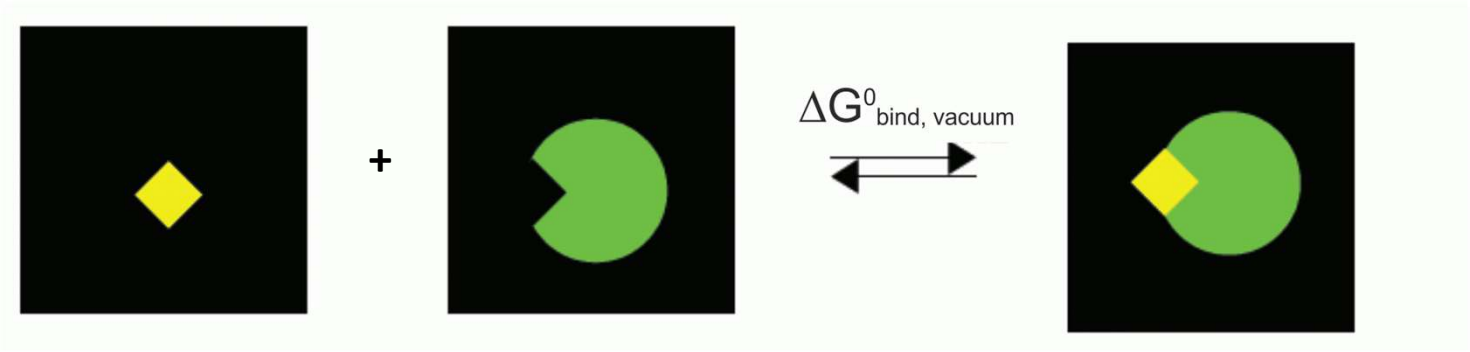
$$\Delta G^0_{bind,\,solv} = \Delta G^0_{bind,\,vacuum} + \Delta G^0_{solv,\,complex} - (\Delta G^0_{solv,\,ligand} + \Delta G^0_{solv,\,receptor})$$

Acc. Chem. Res. 2000,

Calculating Structures and Free Energies of Complex Molecules: Combining Molecular Mechanics and Continuum Models

PETER A. KOLLMAN *,† IRINA MASSOVA †,×

METHOD 3: MM/PBSA METHOD
FREE ENERGY OF LIGAND BINDING IN VACUUM



$$\Delta G^0_{vacuum} = \Delta E^0_{molecular\ mechanics} - T \cdot \Delta S^0_{normal\ mode\ analysis}$$

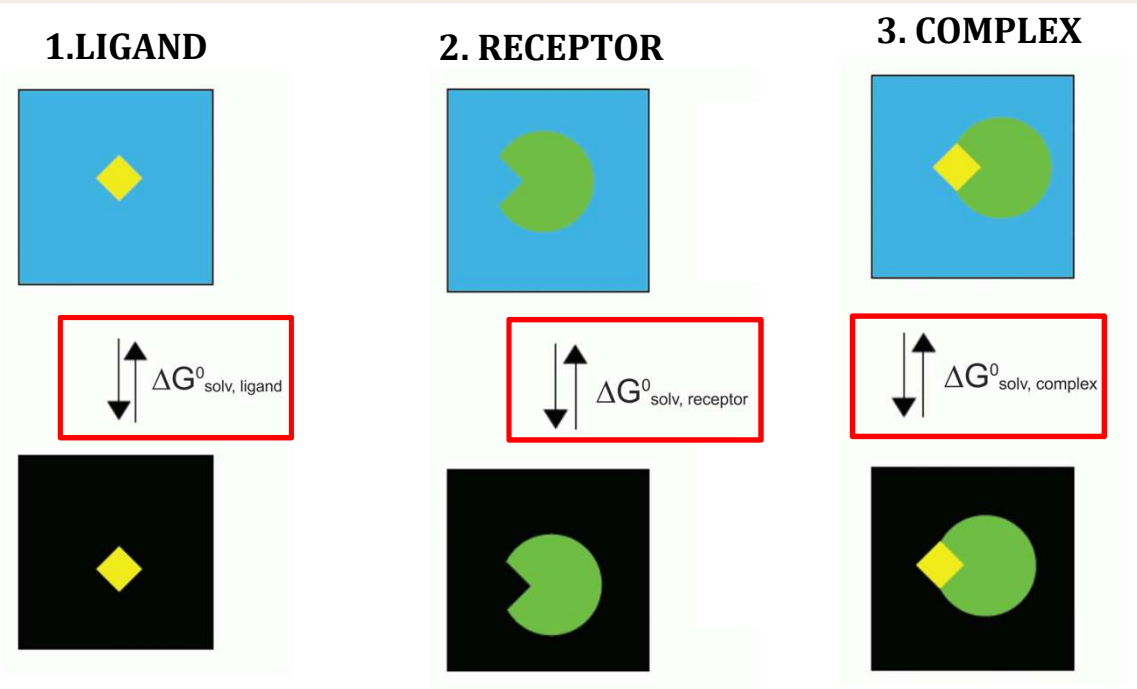
Interaction energy between the receptor and ligand is calculated using molecular mechanics (MM)

$$E^{total} = \sum_{bonds} K_r (r - r_{eq})^2 + \sum_{angles} K_\theta (\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)] + \sum_{i<j} \left[\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}} + \frac{q_i q_j}{\varepsilon r_{ij}} \right]$$

stretch terms | bend terms | torsional terms | non-bonded interactions

Entropy change upon binding is estimated by stimated by a normal-mode analysis of the vibrational frequencies.

METHOD 3: MM/PBSA METHOD
SOLVATION FREE ENERGIES CALCULATIONS



1. linearized Poisson-Boltzmann Eqaution - continuum solvent

$$\vec{\nabla}\cdot(\epsilon\vec{\nabla}U) = -4\pi\sum_i c_i^{\infty} z_i q \exp\left(\frac{-z_i qU}{kT}\right) - 4\pi\rho^{f}$$

$$\Delta G^{el} = \int_0^{\tau} qU(\tau')d\tau'$$

2. Solvent-Accessible Surface Area

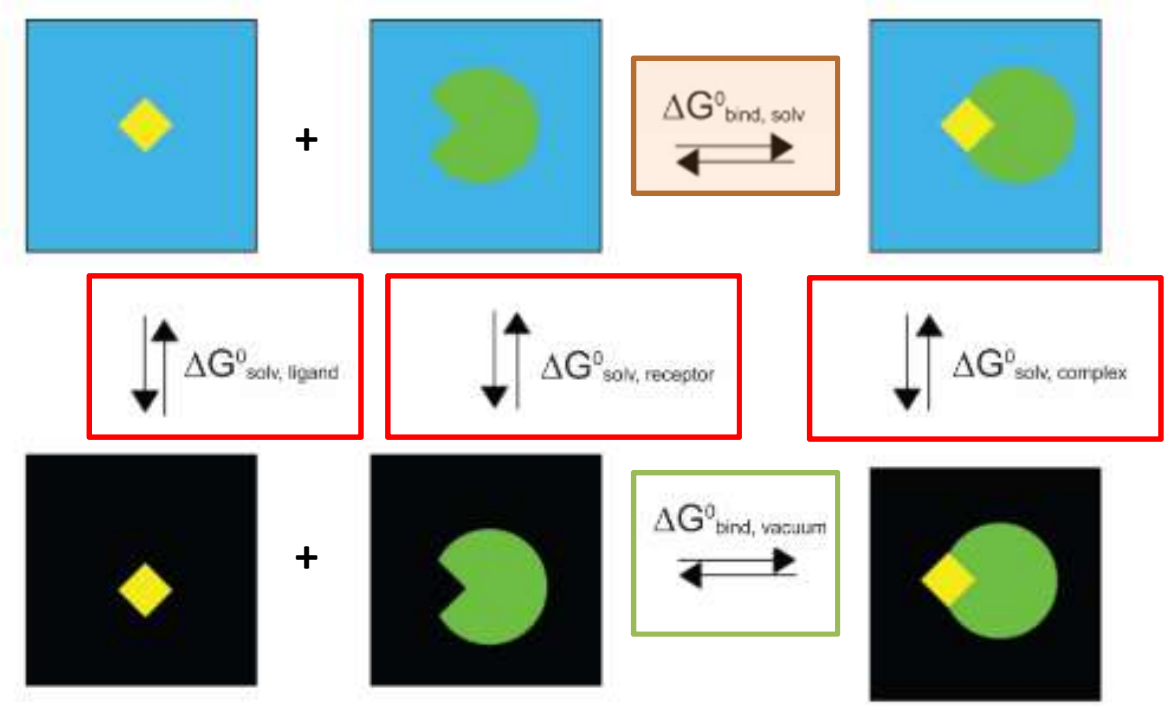
$$\Delta G^{SA}_{non-polar} = \gamma * SASA + b$$

$$\Delta G^{0}_{solv} = G^{0}_{electrostatic,\ \epsilon=80} - G^{0}_{electrostatic,\ \epsilon=1} + \Delta G^{0}_{hydrophobic}$$

A) Polar (electrostatic) contributions are calculated by **Linearized Poisson-Boltzmann** or **Generalized Born equation** for each of the three states: **LIGAND, RECEPTOR** and **COMPLEX**.

B) Hydrophobic contributions are usually estimated by **Solvent-Accessible Surface Area (SASA)**.

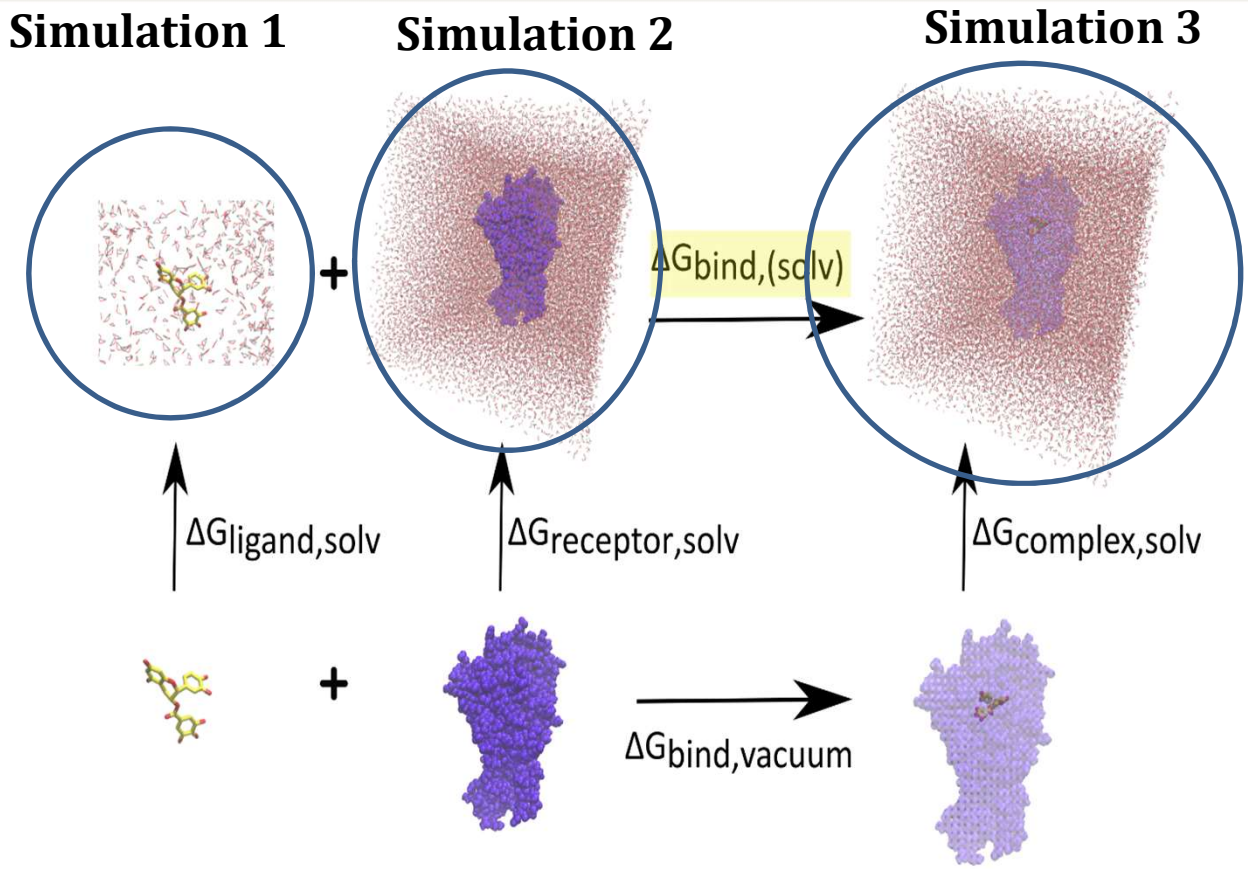
METHOD 3: MM/PBSA METHOD
FINAL OVERVIEW



$$\Delta G^0_{bind,\ solv} = \Delta G^0_{bind,\ vacuum} + \Delta G^0_{solv,\ complex} - (\Delta G^0_{solv,\ ligand} + \Delta G^0_{solv,\ receptor})$$

METHOD 3: MM/PBSA METHOD

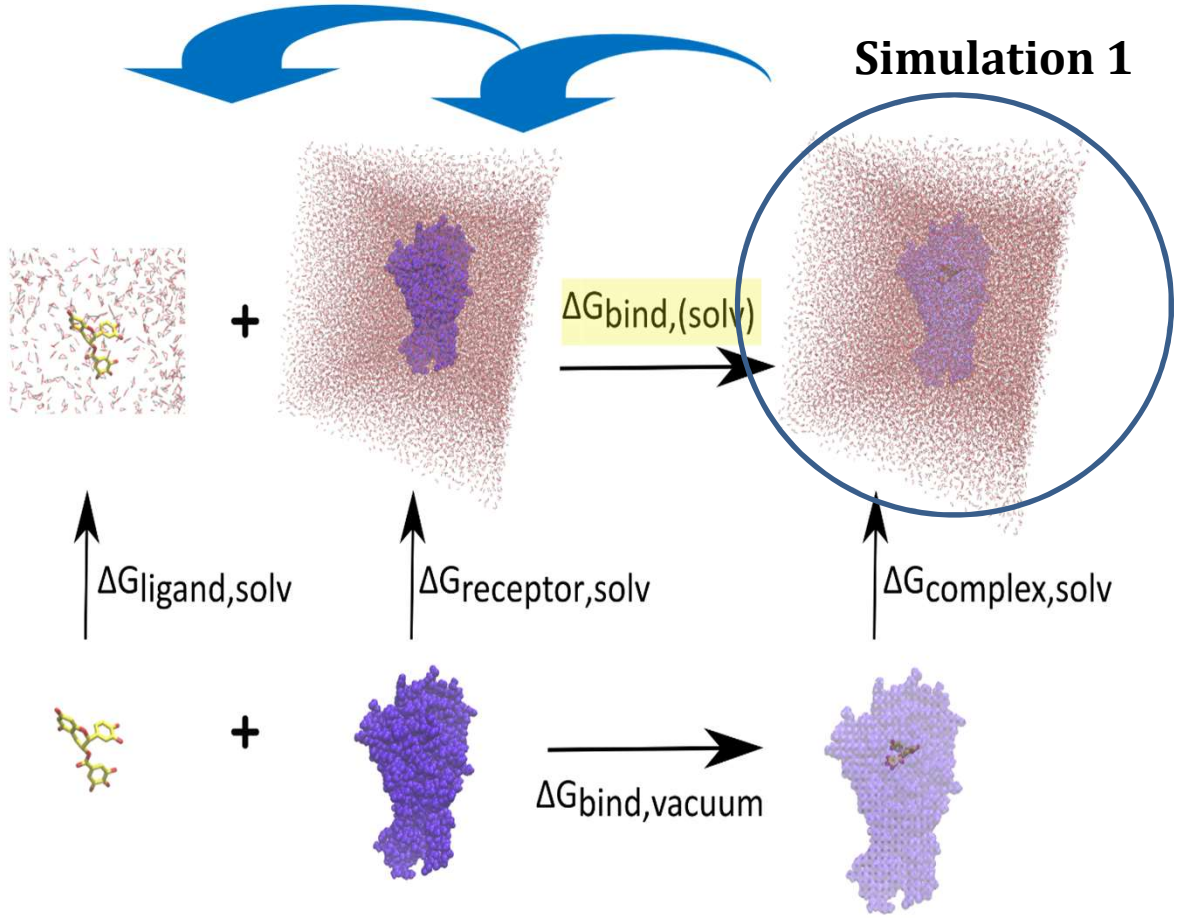
TECHNICAL IMPLEMENTATION: THREE MD SIMULATIONS



In general case, one carries out **three** independent MD simulations: for *ligand, receptor, and complex*

METHOD 3: MM/PBSA METHOD

TECHNICAL IMPLEMENTATION: ONE MD SIMULATION



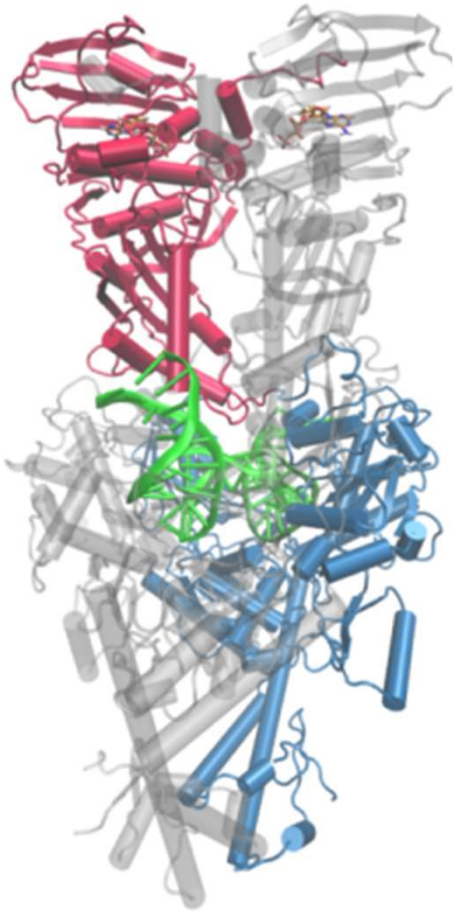
Single trajectory approach: one makes the approximation that no significant conformational changes occur upon binding so that the snapshots for all three species can be obtained from a **single** trajectory for a complex

CASE STUDY:

ENERGETICS OF MOLECULAR RECOGNITION BETWEEN ATP MOLECULE AND ITS BINDING SITE ON THE HUMAN DNA TOPOISOMERASE IIα

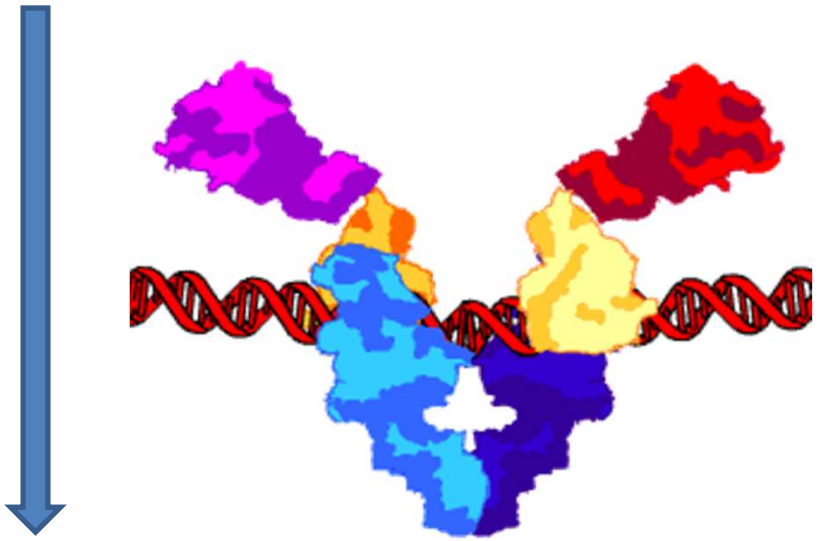
Barbara Herlah

TARGET: HUMAN DNA TOPOISOMERASE IIα
A DNA TOPOLOGY MODIFYING MOLECULAR MOTOR



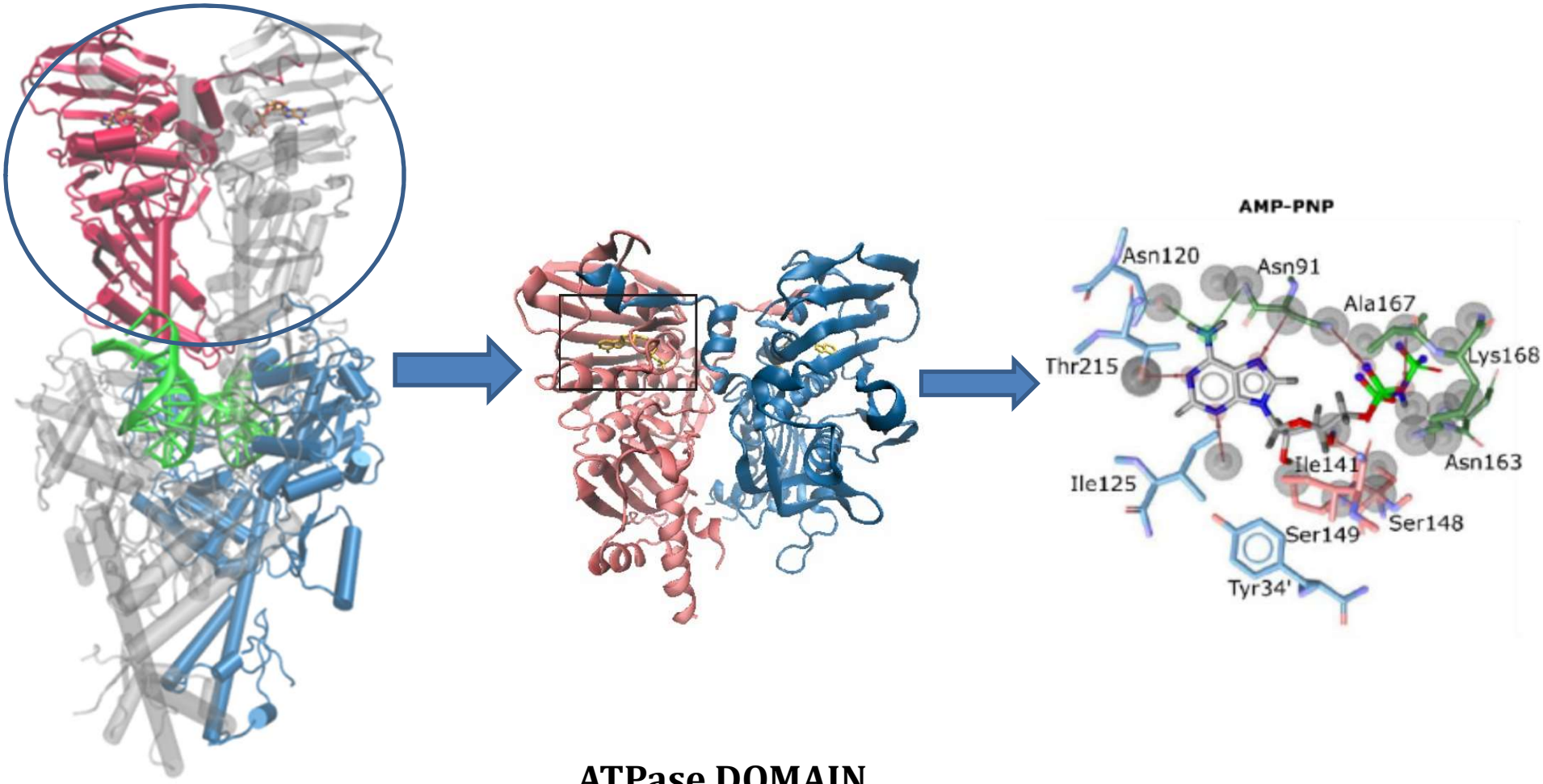
TARGET OF ANTICANCER DRUGS

COMPLEX DNA TOPOLOGY



SIMPLIFIED DNA TOPOLOGY

ATPase DOMAIN AND ATP BINDING SITE OF HUMAN DNA TOPOISOMERASE IIα



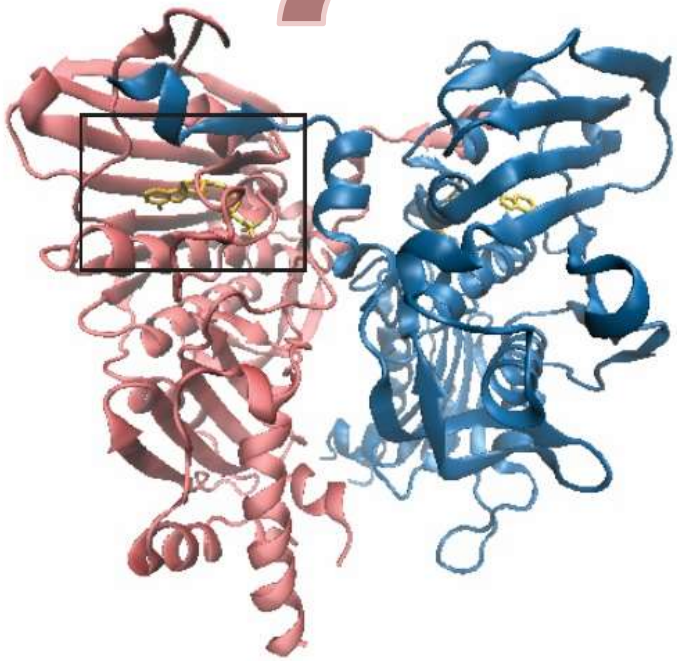
FULL ENZYME

ATPase DOMAIN (dimer)

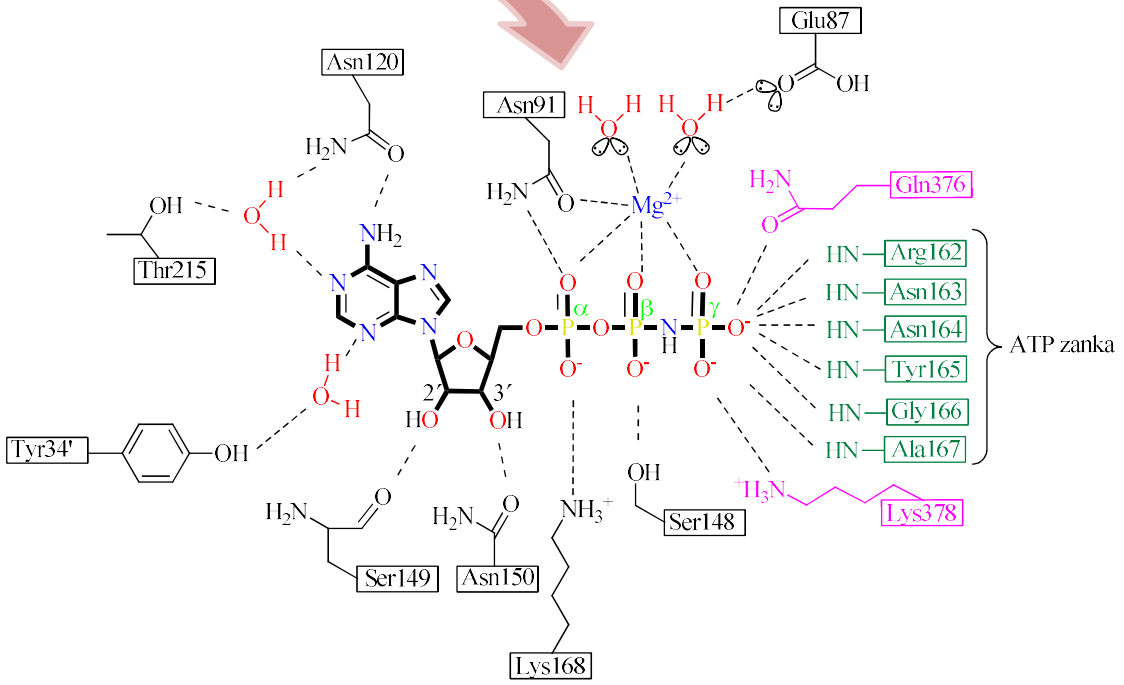
ATP BINDING SITE

RESEARCH QUESTION:
FOR OUR CASE STUDY

ATP BINDING SITE



ATPase DOMAIN OF HUMAN DNA TOPOISOMERASE IIα (PDB:1ZXM)



WHICH RESIDUES ARE (energetically) MOST IMPORTANT FOR ATP BINDING ?

LET´S START SIMULATING