Accelerate Science-Led Innovation for Competitive Advantage

COMPREHENSIVE MODELING AND SIMULATIONS FOR LIFE SCIENCES

3DEXPERIENCE®

訊聯基因數位股份有限公司

資深經理 陳冠文 (Gene)

DASSAULT SYSTEMES | The 3DEXPERIENCE® Company

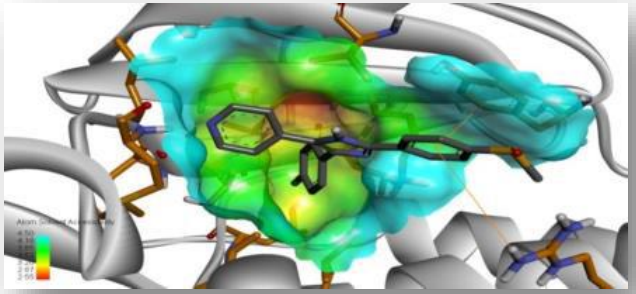
Copyright and Disclaimer

- Copyright © **2024** GGA corp. All rights reserved.
  - This presentation and/or any related documents contains statements regarding our plans or expectations for future features, enhancements or functionalities of current or future products (collectively "Enhancements"). Our plans or expectations are subject to change at any time at our discretion. Accordingly, GGA Corp. is making no representation, undertaking no commitment or legal obligation to create, develop or license any product or Enhancements.
  - The presentation, documents or any related statements are not intended to, nor shall, create any legal obligation upon GGA Corp., and shall not be relied upon in purchasing any product. Any such obligation shall only result from a written agreement executed by both parties.
  - In addition, information disclosed in this presentation and related documents, whether oral or written, is confidential or proprietary information of GGA Corp. It shall be used only for the purpose of furthering our business relationship, and shall not be disclosed to third parties.

Simulation

- Assume……that we have a molecule on the computer screen

- It could have been
  - Built from scratch
  - Prepared from a PDB file
  - Read in from a prepared file



- But…How do you know you have a reasonable structure?

Simulation

- Molecular Mechanism (MM), Molecular Dynamic (MD) simulation
  - Assignment of the force field
    - CHARMm, charmm19/22/27/36, CHARMM Polar H, MMFF, CFF
  - Additional solvent model
    - implicit/explicit water molecules, Implicit Membrane
  - Additional constraints
    - Fix, Dihedral, Harmonic, Distance
- Energy calculation
  - Time point energy calculation
  - Structural optimization
  - Structural optimization using QM (DFT)
  - MD（CHARMm, NAMD）
- Analyze trajectory
  - Plot of various properties – distance, angle, interaction energy, PCA, RDF, Radius of Gyration

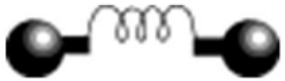
Molecular Mechanism (MM)

- Based on the classical molecular mechanics
  - Energy calculations
  - Energy minimizations
  - Molecular dynamics

- Widely used in
  - Structure-based ligand design
  - Structure generation from NMR experiments
  - Protein engineering
  - Substrate recognition studies

Force Field

"The molecule is considered to be a collection of atoms held together by simple elastic or harmonic forces…"

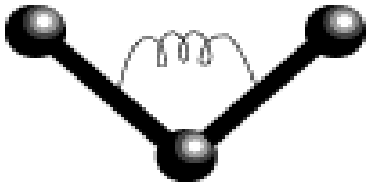


$$E = k_b (r - r_0)^2$$

Adopted from computer software applications in chemistry, Peter C. Jurs, 1986

Force Field

"The molecule is considered to be a collection of atoms held together by simple elastic or harmonic forces. The forces are defined in terms of potential energy described by the internal coordinates of the molecule…"



$E = k_b(r - r_0)^2$

$E = k_\theta(\theta - \theta_0)^2$

Adopted from Computer Software Applications in Chemistry, Peter C. Jurs, 1986

Force Field Energy Expression

- General form

$$E = E_{bond} + E_{angle} + E_{torsion} + E_{impr} + E_{nonbond}$$

$$(E_{vdw} + E_{elecs})$$

$$+ E_{other}$$

- Force fields are generally differ in:
  - Function expression of each term
  - Parameter calculation method
  - Function parameters
  - Cross terms

$$E_{\text{pot}} = \sum k_b(r - r_0)^2 + \sum K_\theta(\theta - \theta_0)^2 + \sum |k_\phi| - k_\phi \cos(n\phi) + \sum k_\chi(\chi - \chi_0)^2$$

$$+ \sum \frac{q_i q_j}{\varepsilon 4\pi \varepsilon_0 r_{ij}} + \sum \left(\frac{A_{ij}}{r_{ij}^{12}} - \frac{B_{ij}}{r_{ij}^{6}}\right) sw(r_{ij}^2, r_{\text{on}}^2, r_{\text{off}}^2) + E_{\text{constraint}} + E_{\text{user}}$$

CHARMm force field

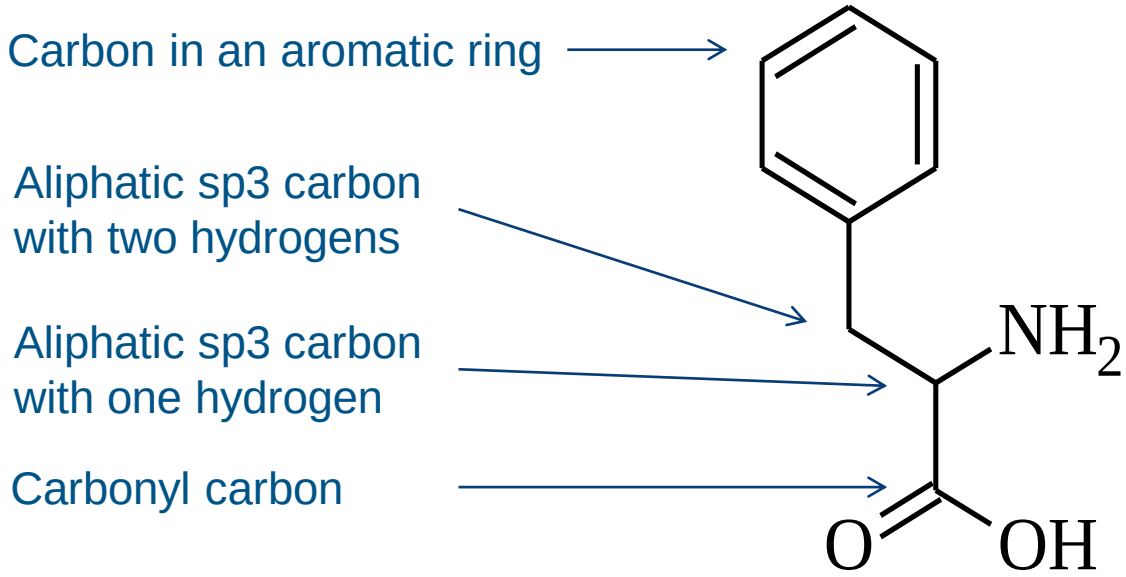
- 2 official version

2, 3 months late release

- Harvard academic version (Professor Martin Karplus, the 2013 Nobel laureate in Chemistry)
- DS integration (Graphical interface)
  - Integration with other programs that use the CHARMm engine (CDOCKER, MCSS, ZDOCK)
  - GBSA / IM solvation model

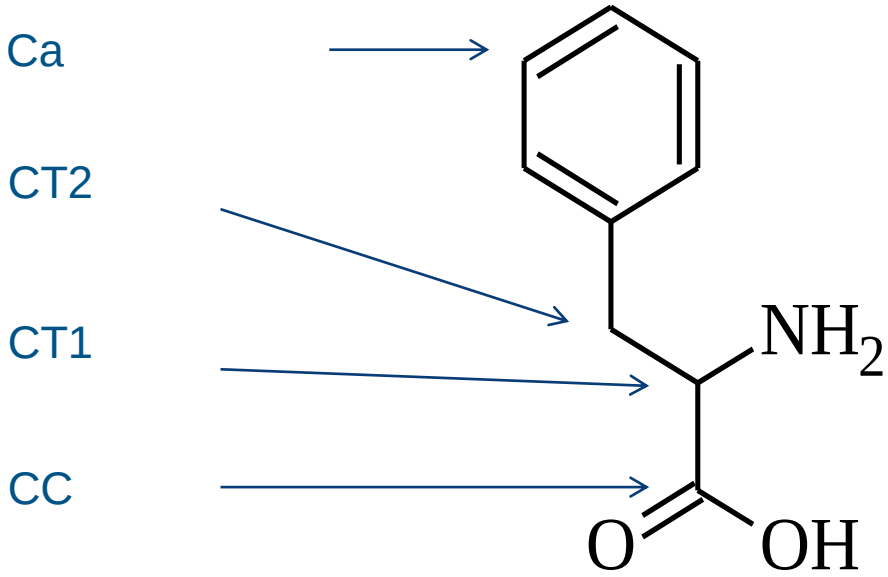
Force Field Parameters

- Atoms have different environments

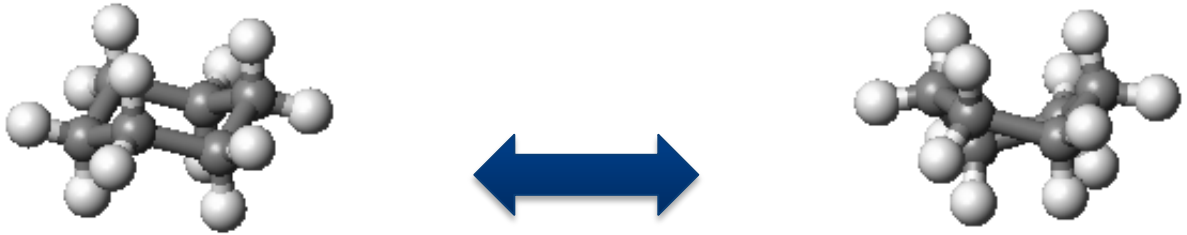


Force Field Parameters

- ...which yield different parameters.



Molecular Mechanics

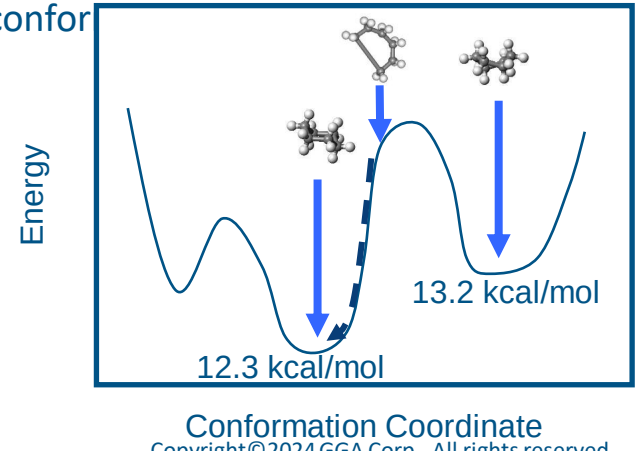


12.3 kcal/mol **13.2 kcal/mol**

It is the difference in energy that matters!

Energy Minimization

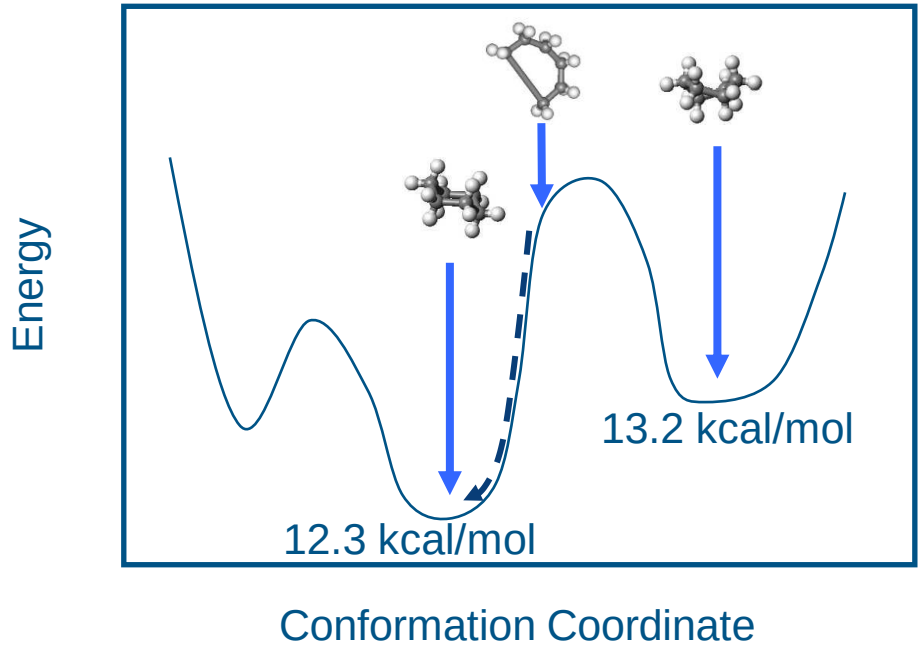
- If lower energy conformations are more favorable, we may want to find them.
- Could perform an energy minimization
  - A mathematical technique to find a local minimum for the energy expression shown earlier
  - Adjust Cartesian coordinates to attain the more favored confor
- When to perform an energy minimization?
  - Clean a structure that you build
  - Prepare a structure from a PDB file
  - Relax a modified structure
  - Relax a docked protein-ligand complex



Molecular Dynamics

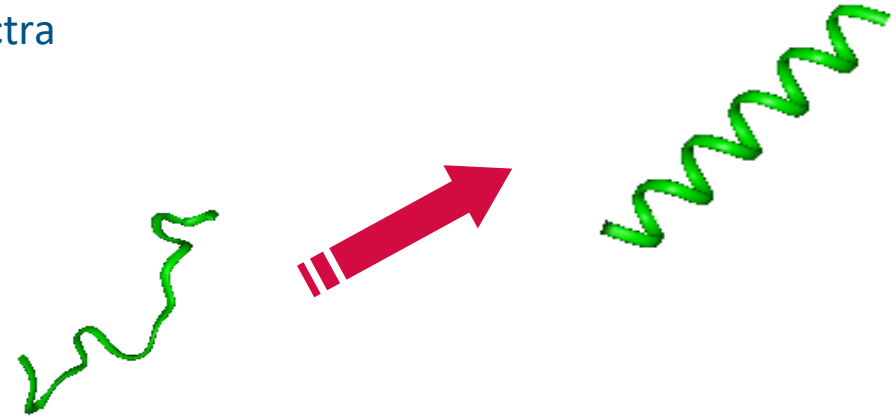
- Global vs Local minima

How to explore other minima or other areas of the potential energy surface?

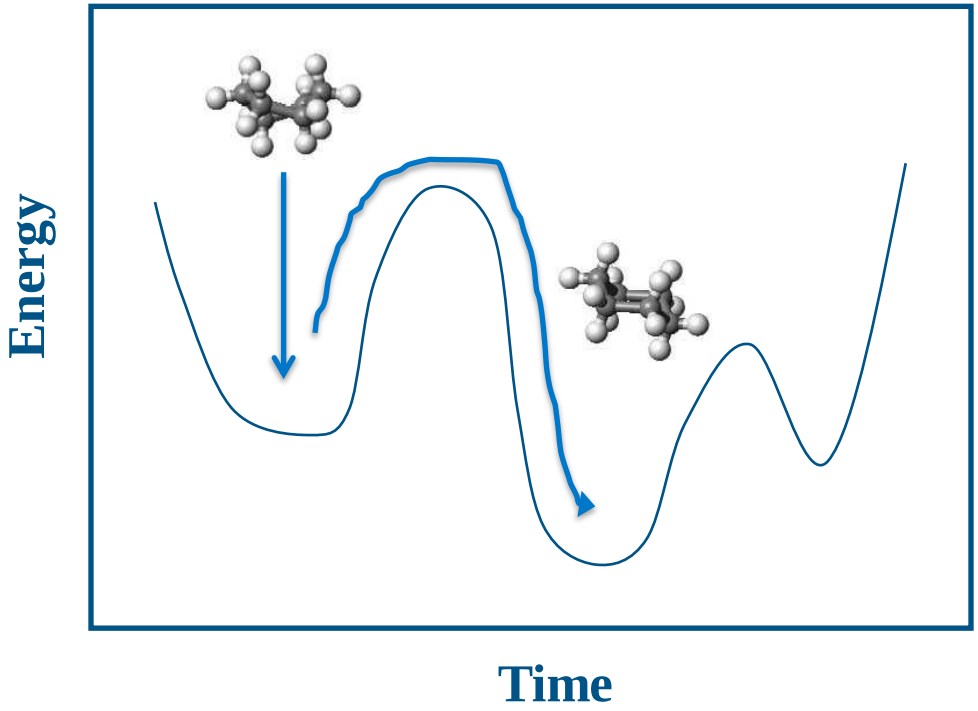


Use of Molecular Dynamics

- Conformational changes of molecule
  - Simple vibrations that produce spectra
  - Breathing of DNA
  - Hinge bending
  - Allosteric transitions
  - Protein folding
  - Refinement of docking results
  - Induced fit
- Free energy changes
  - Calculation to binding energies
- Determination of thermodynamic properties
  - Enthalpy and entropy changes



Example for MD simulation



Temperature and Velocities

- From kinetic theory, temperature is proportional to the kinetic energy of the particles.

$$KE = \frac{3}{2}kT$$

- From physics, kinetic energy is proportional to velocity.

$$KE = \frac{3}{2}mv^2$$

- Therefore, temperature is proportional to the velocity of the atoms.

$$T = \frac{mv^2}{k}$$

Temperature and Dynamics

At a higher temperature…

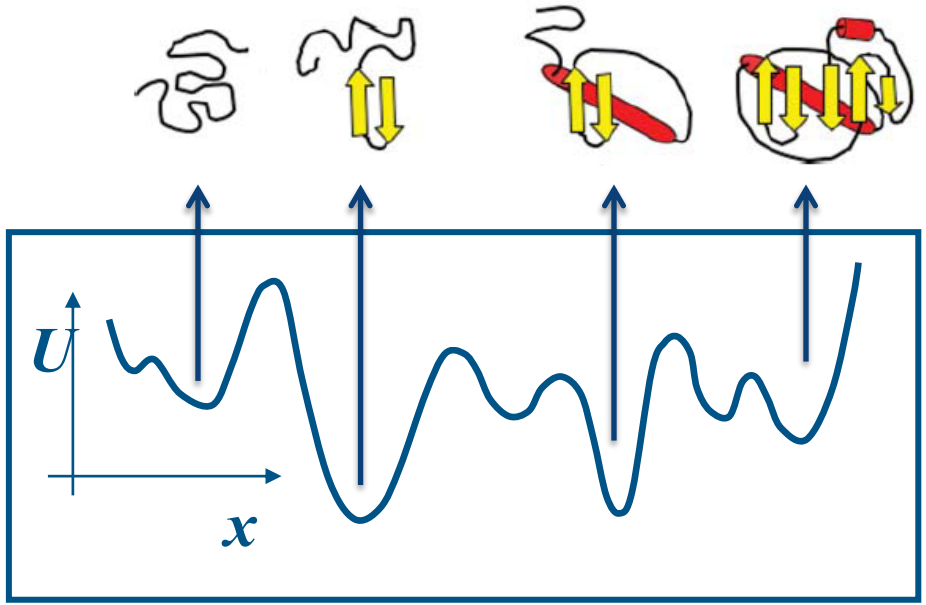
Atoms move more…

Cover more conformational space!

- High temperature dynamics can be used to overcome barriers to explore new local minima
- Effective for reasonable length peptides and other oligomers (12 residues or less)

- Must be aware of
  - Possible loss of secondary structure
  - Trans-cis isomerization of peptide bonds

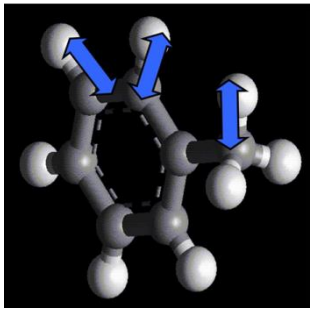
Energy landscape:
searching the possible conformations

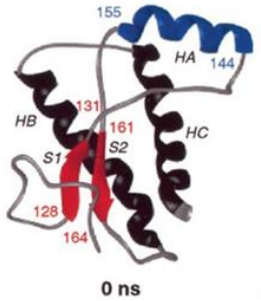


Local minimum

Time scale of Biological events

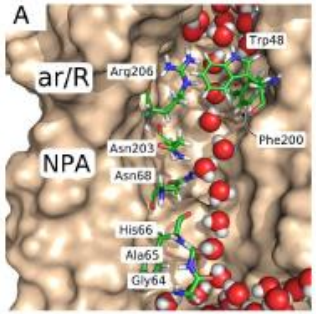
Motion	Time scale (Sec)
Bond stretching	10^{-14} to 10^{-13}
Elastic vibrations	10^{-12} to 10^{-11}
Rotations of surface side chains	10^{-11} to 10^{-10}
Hinge bending	10^{-11} to 10^{-7}
Rotation of buried side chains	10^{-4} to 1 sec
Allosteric transitions	10^{-5} to 1 sec
Local denaturations	10^{-5} to 10 sec

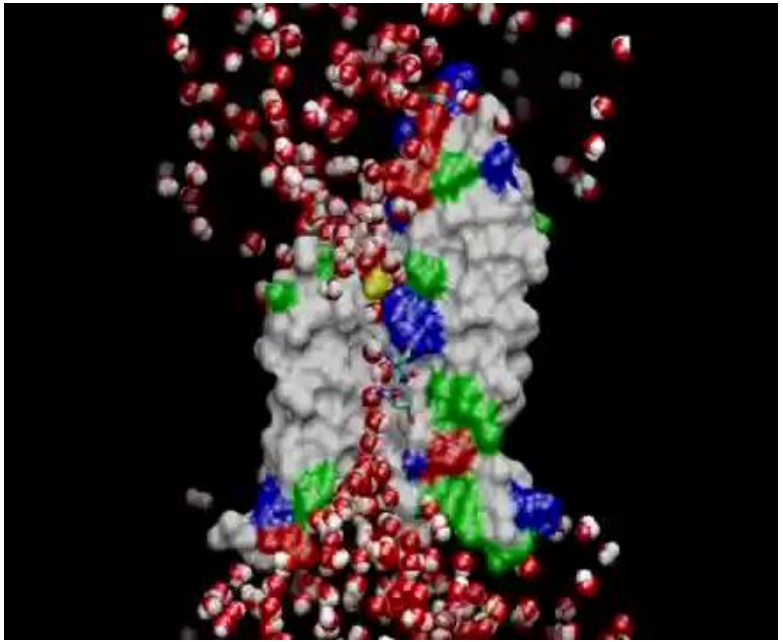




Usage of MD simulation

- Example: Aquaporin
- Aquaporins form a family of pore proteins that facilitate the efficient and selective flux of small solutes across biological membranes

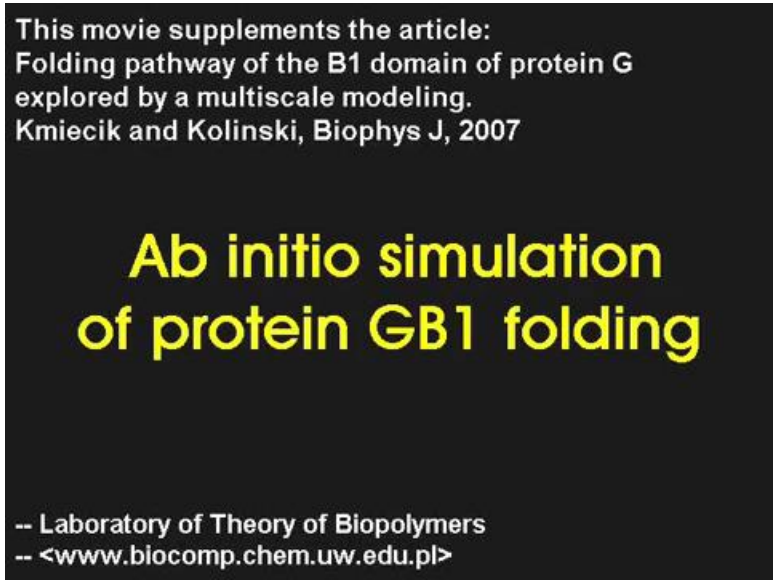




Mechanism of selectivity in aquaporins and aquaglyceroporins. Hub JS, de Groot BL. Proc Natl Acad Sci U S A. 2008

Usage of MD simulation

- Example: GB1 folding
- IgG-binding protein G

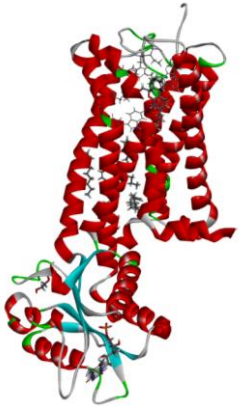


Folding pathway of the b1 domain of protein G explored by multiscale modeling. *Kmiecik S, Kolinski A. Biophys J. 2008*

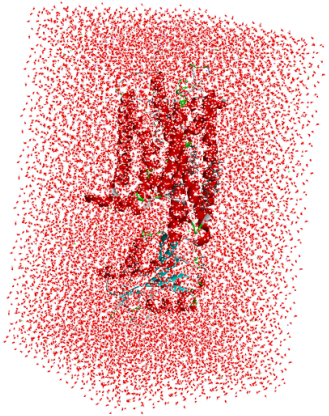
Molecular Dynamics in DS

- Before Run MD
  - Solvation
  - Explicit Membrane-Based Solvation
- MD in DS
  - Standard Dynamic Cascade (SDC)
  - NAMD
    - Free Energy Perturbation
  - Steered Molecular Dynamics (SMD)
  - Multi-Site Lambda Dynamics (MSLD)

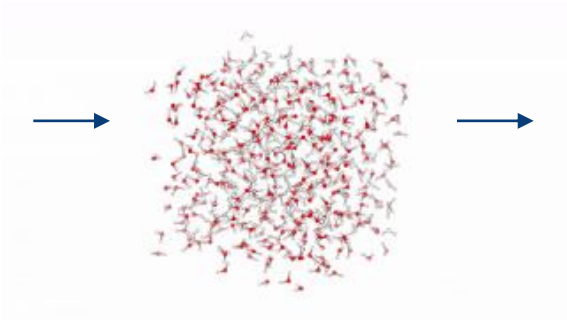
MD steps



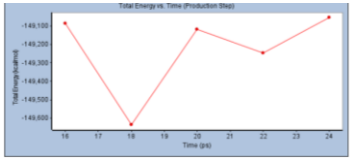
Build Model



Solvation



Run MD



Analysis

Build Model

- Model for MD simulation can be obtain by…
  - Protein Data Bank
  - Alphafold
  - Theoretical Model
    - Homology Modeling
    - Docking
    - Others…

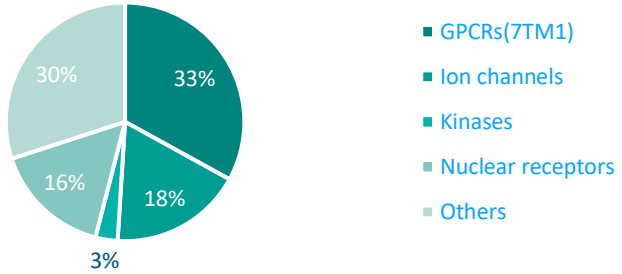
Solvation

- Implicit solvent model
  - Distance –Dependent Dielectrics
  - Generalized Born
  - GBMV, GBSW, GBIM
- Explicit solvent model
  - Cell shape
  - Periodic Boundary Distance
  - Counterion

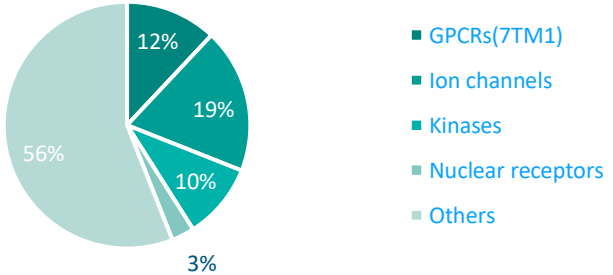
Explicit Membrane-Based MD Simulations

- Membrane proteins account for approximately 1/3 of the human proteome and account for ~60% of pharmaceutical targets+
- These membrane proteins include a large number of signaling receptors, transporters, ion channels and enzymes that are vital to cellular regulation, metabolism and homeostasis, and are high-priority, pharmaceutically relevant research
- CHARMm has the best and most consistent set of lipid parameters, including sterols, unusual bacterial, mitochondrial, and endosomal lipids, and a consistent set of protein parameters, to model complex membranes and protein membrane interactions+

Proportion of human protein drug targets in major families



Proportion of small-molecule drugs that target major families

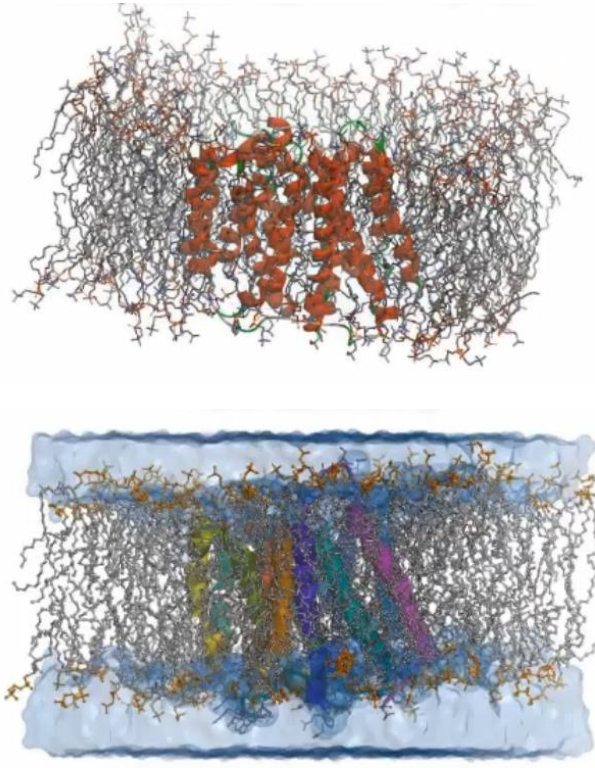


Nat Rev Drug Discov. 2006, *5*, 993-6
Chem Rev. 2019, *119*, 6184-6226

訊聯基因數位

Explicit Membrane-Based MD Simulations

- Membrane proteins account for approximately 1/3 of the human proteome and account for ~60% of pharmaceutical targets+
- These membrane proteins include a large number of signaling receptors, transporters, ion channels and enzymes that are vital to cellular regulation, metabolism and homeostasis, and are high-priority, pharmaceutically relevant research
- CHARMm has the best and most consistent set of lipid parameters, including sterols, unusual bacterial, mitochondrial, and endosomal lipids, and a consistent set of protein parameters, to model complex membranes and protein membrane interactions+

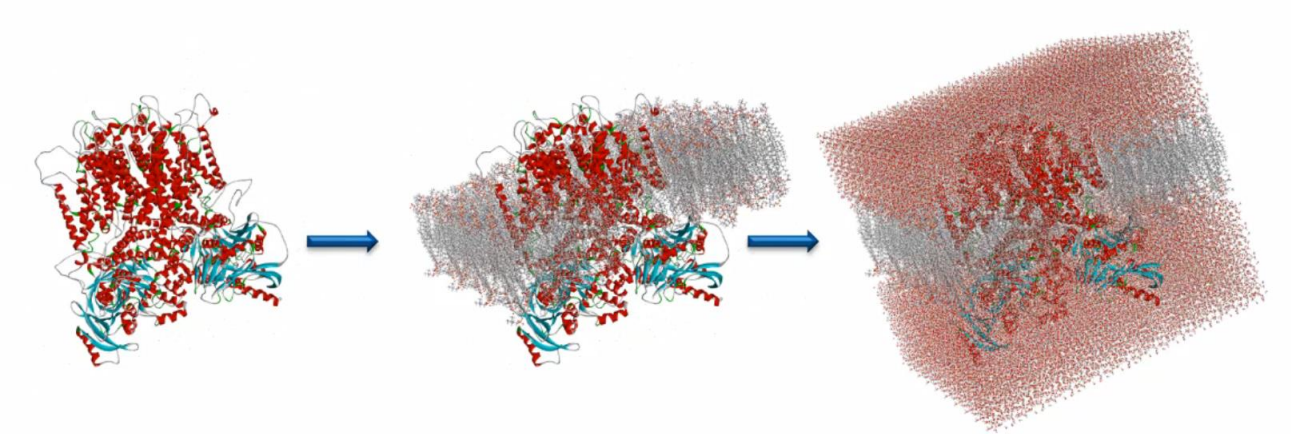


Nat Rev Drug Discov. 2006, *5*, 993-6
Chem Rev. 2019, *119*, 6184-6226

訊聯基因數位

Explicit Membrane-Based MD Simulations

- New protocol adds an explicit bilayer of lipids, water and counterions to a transmembrane protein and allows the option to equilibrate the system, prior to further simulation
- Supports a number of prepared equilibrated homologenous membranes, as well as complex custom membranes (e.g. http://www,charm-gui.org)
  - POPC, POPE, DPPC, DMPC, and DLPC

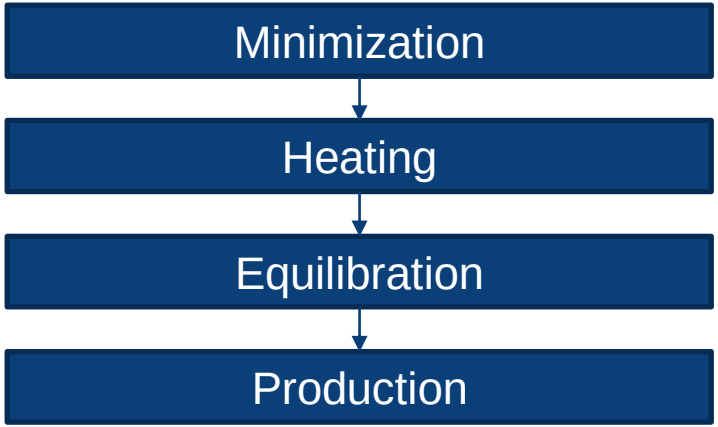


Molecular Dynamics in DS

- Before Run MD
  - Solvation
  - Explicit Membrane-Based Solvation
- MD in DS
  - Standard Dynamic Cascade (SDC)
  - NAMD
    - Free Energy Perturbation
  - Steered Molecular Dynamics (SMD)
  - Multi-Site Lambda Dynamics (MSLD)

Standard Dynamic Cascade (SDC)

- General MD in DS
- Based on CHARMm Forcefield



To obtain reasonable starting model

Adds energy to make atom moves

Initializes the atom velocities using *Target Temperature* and performs equilibration molecular dynamics

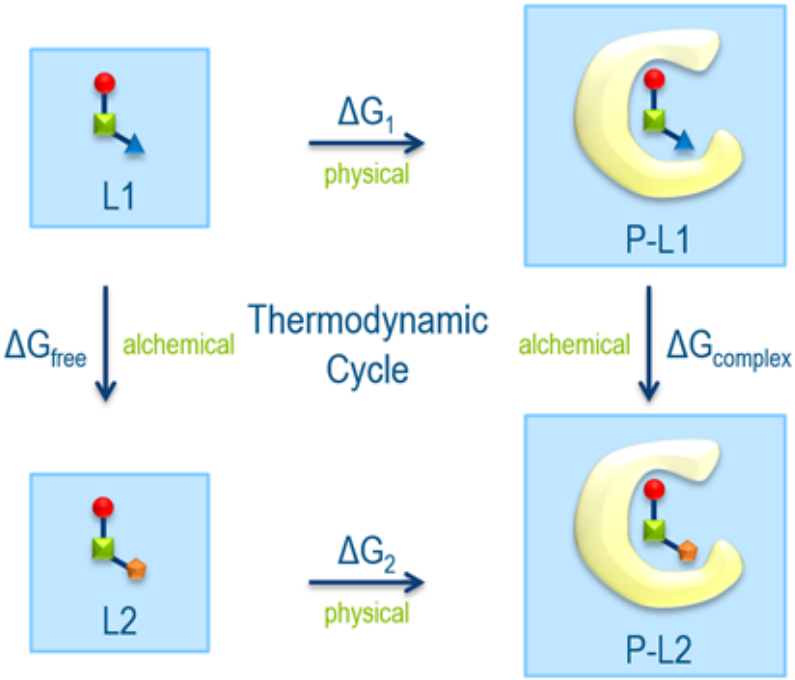
Dynamics provides insights into the motions of a molecule with time

Molecular Dynamics in DS

- Before Run MD
  - Solvation
  - Explicit Membrane-Based Solvation
- MD in DS
  - Standard Dynamic Cascade (SDC)
  - NAMD
    - Free Energy Perturbation
  - Steered Molecular Dynamics (SMD)
  - Multi-Site Lambda Dynamics (MSLD)

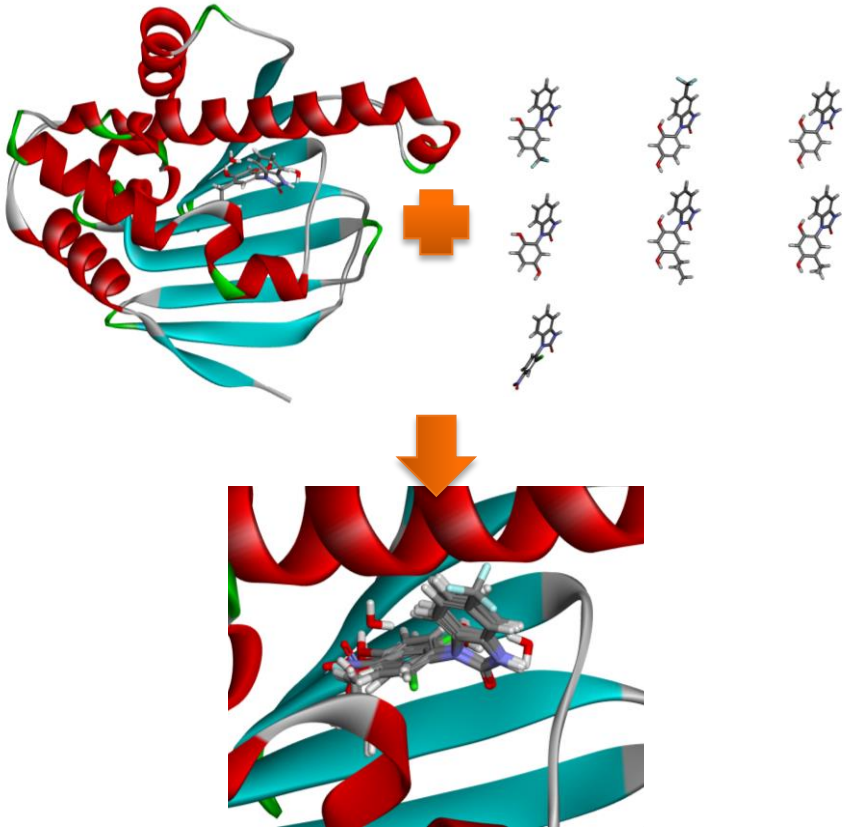
Free Energy Perturbation

- The ability to understand and quantify the strength of interactions between a ligand and protein is essential in drug discovery projects to discern how changes to a ligand might alter affinity.
- Alchemical free energy perturbation (FEP) is based on a thermodynamic cycle.



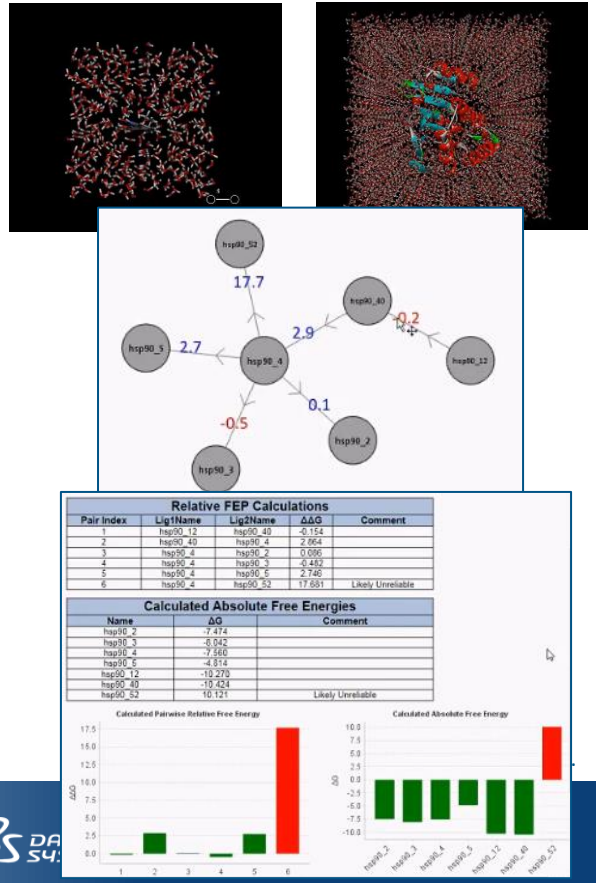
Generate Analog Conformation

- Receptor
  - Receptor should be prepared well before run FEP
- Docked Lead
  - Used as reference to align analogues to the binding site
- Analog
  - Superimpose is recommended before run FEP



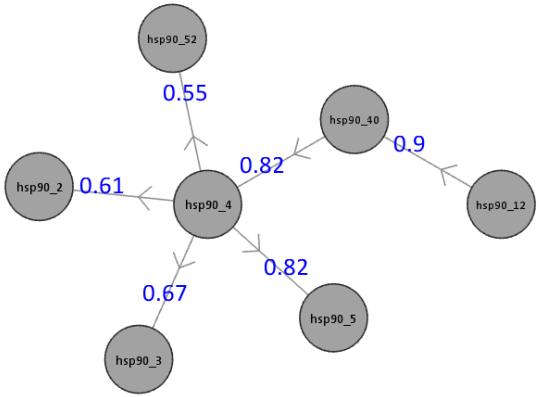
Predict Relative Free Energy of Binding

- Apply FEP to a congeneric analog series in four steps using NAMD
  - Use either docking results, or *in-situ* analog series as input
  - Optional: Type analogs with charmm36/GCenFF using MATCH engine
- Step 1: Generate a ligand pairs network map
  - Builds a pairs network using maximal common substructures
- Step 2: Set up relative FEP calculations
  - Generates explicit solvent models for each ligand pair and protein-ligand pair complex in the network map
- Step 3: Run free energy perturbation calculations
  - Perform FEP simulations on each solvated ligand pair model
- Step 4: Collate the simulation results
  - Summarises both relative free energy calculations and absolute free energies (if a free energy value is available for the reference lead)



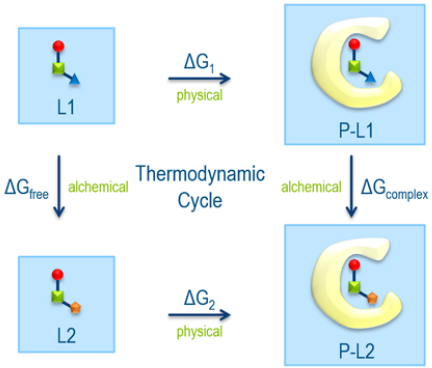
Generate a ligand pairs network map

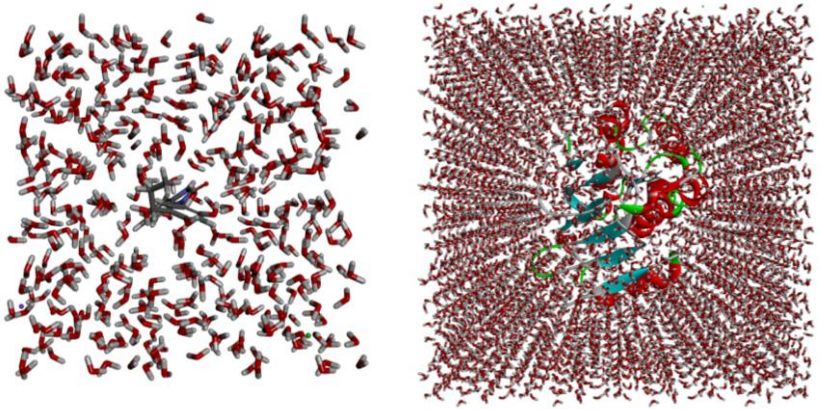
- Generates a set of optimal pairs based on chemical and topological similarity in order to maximize the total similarity of a network spanning the entire set.
- Similarity matrix using a Maximal Common Substructure (MCS) method.
- The similarity between any two ligands L1 and L2 is computed as exp[-Beta*(nL1+nL2-2*nMCS)]



Set up relative FEP calculations

- Generates explicit solvent models for each ligand pair and protein-ligand pair complex in the network map
- The calculation is set up to transform the first ligand in each pair into the second. For each pair, a merged dual topology molecule is generated based on the identified common and different set of atoms in the topologies and nonbonded parameters.

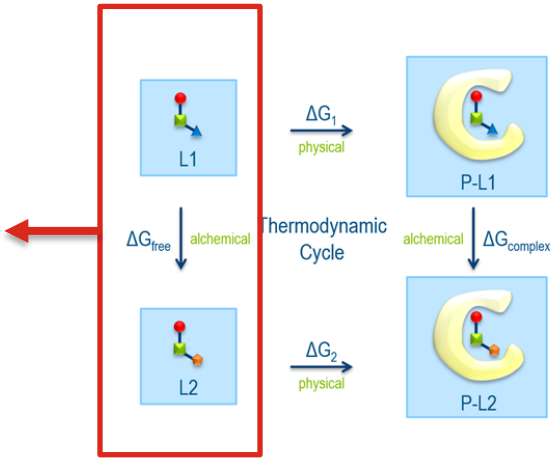




Run free energy perturbation calculations

- Runs alchemical Free Energy Perturbation (FEP) dynamics simulations using **NAMD**
- Solvated Ligand and Solvated Complexes are run separately.

FEP Simulation Results							
Name	Lambda Range	Calc ΔG	StDev	SimulationType	FEP Report	Simulation Trajectory	Simulation Trajectory Without Solvent
hsp90_12--hsp90_40	0-1	1.02	0.04	Ligand	FEP Report	Trajectory	Trajectory
hsp90_40--hsp90_4	0-1	0.90	0.05	Ligand	FEP Report	Trajectory	Trajectory
hsp90_4--hsp90_5	0-1	-10.01	0.09	Ligand	FEP Report	Trajectory	Trajectory
hsp90_4--hsp90_3	0-1	4.04	0.02	Ligand	FEP Report	Trajectory	Trajectory
hsp90_4--hsp90_2	0-1	2.46	0.12	Ligand	FEP Report	Trajectory	Trajectory
hsp90_4--hsp90_52	0-1	20.30	0.63	Ligand	FEP Report	Trajectory	Trajectory



Collate the simulation results

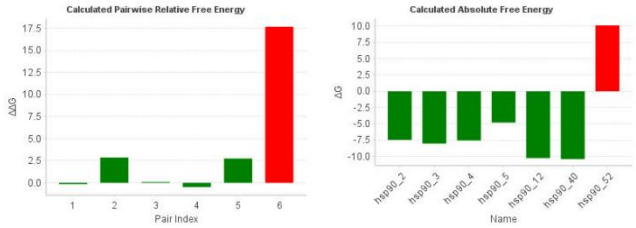
- The *Collate FEP Simulation Results* protocol also generates a network plot (pictured) that graphically presents the ΔΔG values of the ligand pairs. The edges are labeled and color-coded with the relative free energy of the ligand pair, and the arrow head indicates the direction of the alchemical change.

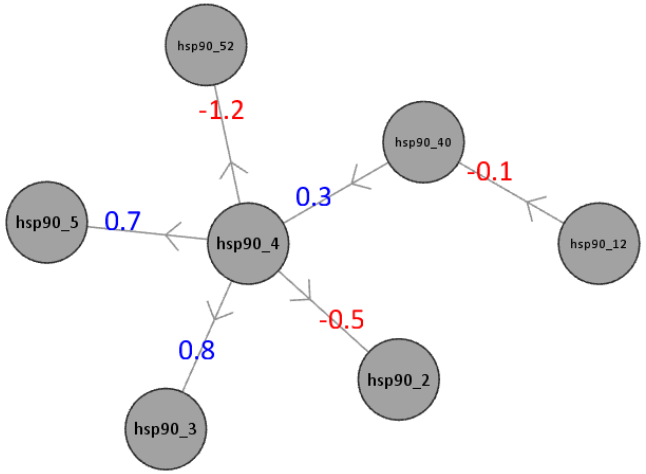
Relative FEP Calculations

Pair Index	Lig1Name	Lig2Name	ΔΔG	StDev	Calc ΔG (Ligand)	Calc ΔG (Complex)	Comment
1	hsp90_12	hsp90_40	-0.154	0.101	7.16	7.01	
2	hsp90_40	hsp90_4	2.864	0.054	-9.39	-6.53	
3	hsp90_4	hsp90_2	0.086	0.175	20.03	20.11	
4	hsp90_4	hsp90_3	-0.482	0.083	13.85	13.37	
5	hsp90_4	hsp90_5	2.746	0.084	-0.64	2.11	
6	hsp90_4	hsp90_52	17.681	0.503	1.54	19.22	Likely Unreliable

Calculated Absolute Free Energies

Name	ΔG	Comment
hsp90_2	-7.474	
hsp90_3	-8.042	
hsp90_4	-7.560	
hsp90_5	-4.814	
hsp90_12	-10.270	
hsp90_40	-10.424	
hsp90_52	10.121	Likely Unreliable



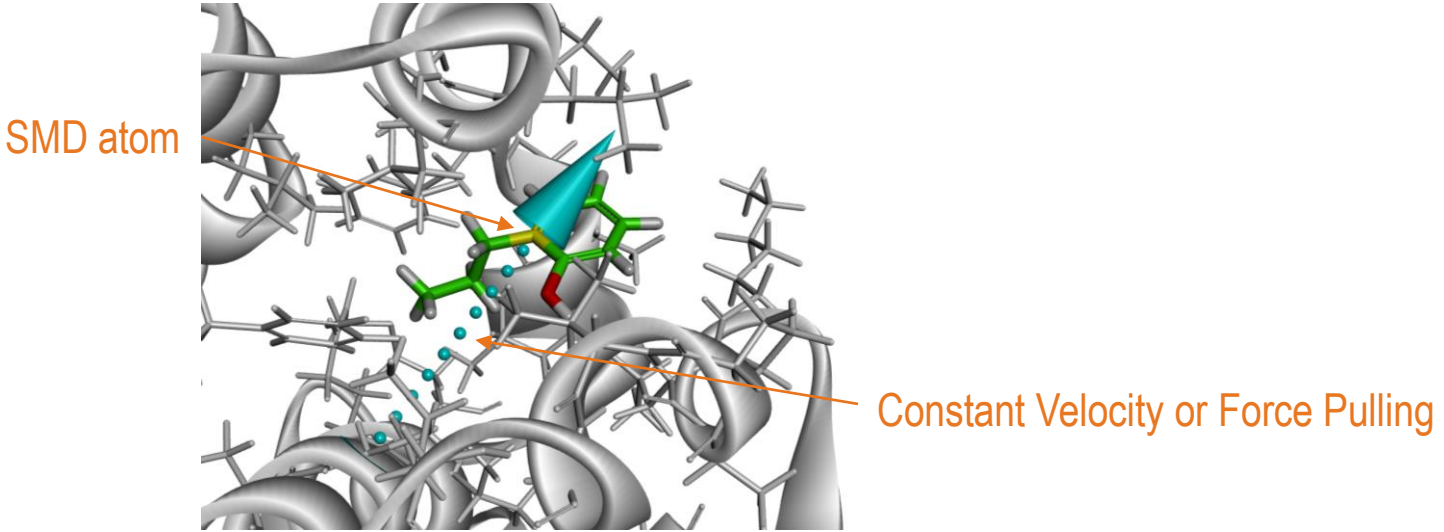


Molecular Dynamics in DS

- Before Run MD
  - Solvation
  - Explicit Membrane-Based Solvation
- MD in DS
  - Standard Dynamic Cascade (SDC)
  - NAMD
    - Free Energy Perturbation
  - Steered Molecular Dynamics (SMD)
  - Multi-Site Lambda Dynamics (MSLD)

Steered Molecular Dynamics

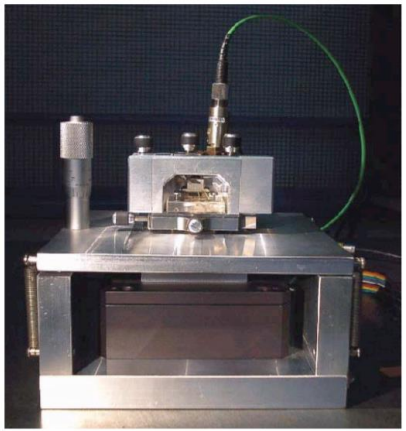
- The basic idea behind any SMD simulation is to apply an external force to one or more atoms, which we refer to as SMD atoms.

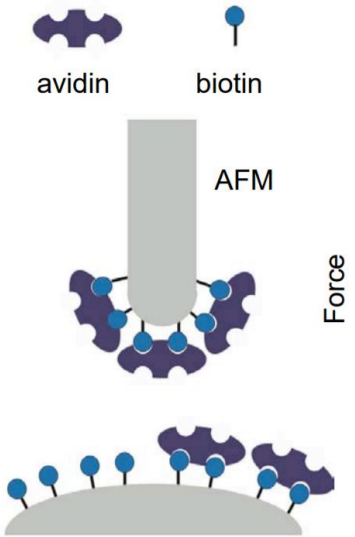


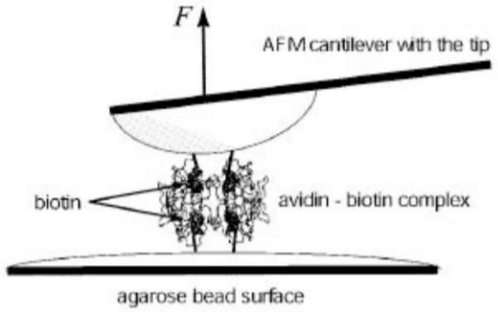
Why SMD

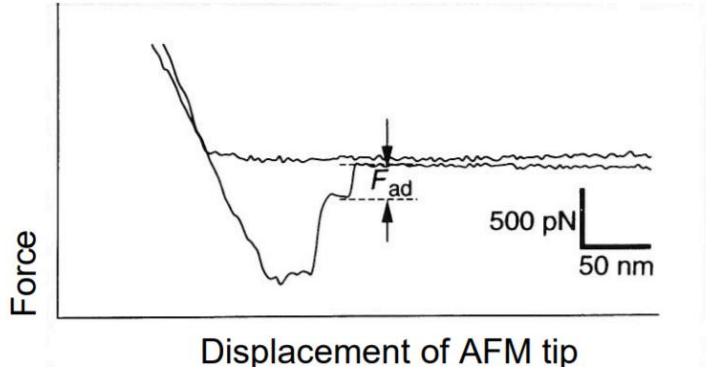
- Accelerates processes to simulation time scales (ns)
- Yields explanations of biopolymer mechanics
- Complements Atomic Force Microscopy
- Finds underlying unbinding potentials
- Generates and tests Hypotheses

Atomic Force Microscopy Experiments of Ligand Unbinding



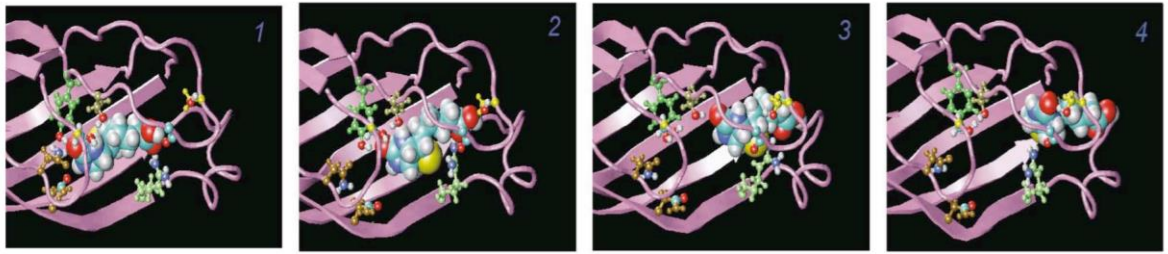


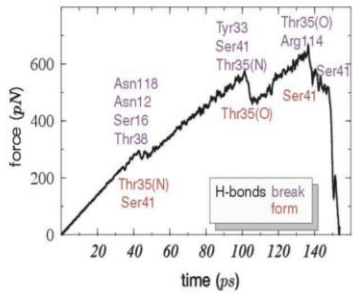


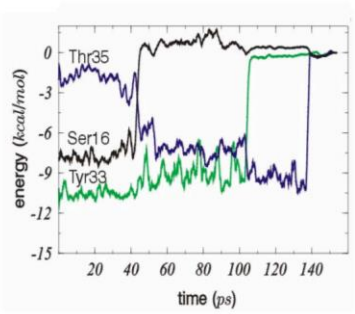


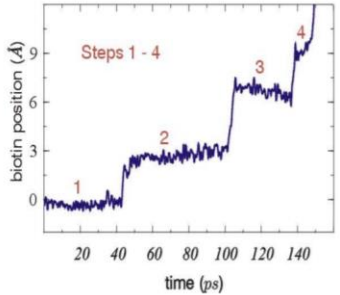
SMD of Biotin Unbinding: What We Learned

- biotin slips out in steps, guided by amino acid side groups, water molecules act as lubricant, MD overestimates extrusion force





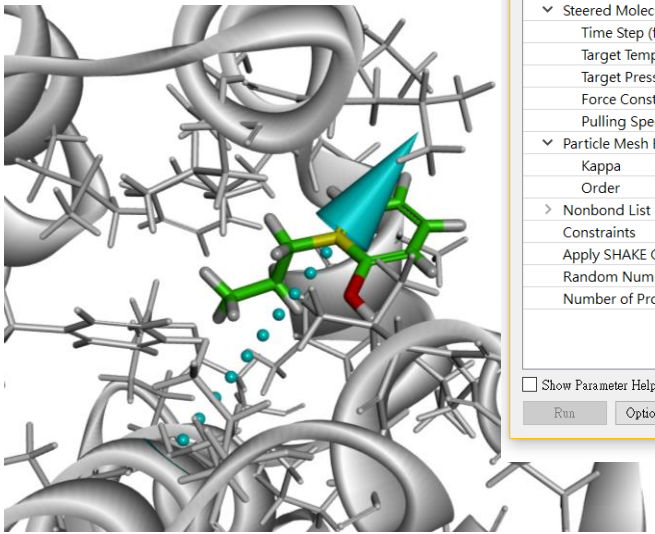




Israilev et al., Biophys. J., 72, 1568-1581 (1997)

General Steps in MD

- Prepare Entities
- Solvation
- Set SMD atom
- Set Constant Velocity or Force Pulling
- Set Simulation Time
- Run



Steered Molecular Dynamics

Parameter Name	Parameter Value
Input Typed Molecule	1HK0:1HK0
Simulation Time (ps)	1000
Save Results Interval (ps)	5
Pulling Atom Pair	
Advanced	
Steered Molecular Dyn...	
Time Step (fs)	2
Target Temperature	300
Target Pressure	1.0
Force Constant	500.0
Pulling Speed	0.02
Particle Mesh Ewald	
Kappa	0.34
Order	4
Nonbond List Radius	14.0
Constraints	
Apply SHAKE Constraint	True
Random Number Seed	
Number of Processors	1

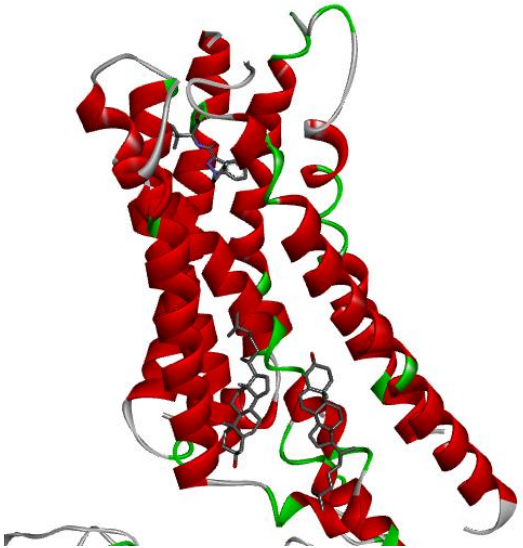
Show Parameter Help

Run | Options ▼ | Cancel | Help

Hands on

- Cannabinoid Receptor 1-G Protein Complex
- methyl N-{1-[(4-fluorophenyl)methyl]-1H-indazole-3-carbonyl}-3-methyl-L-valinate

- Predict the unbounding pathway and binding free energy of the *in situ* ligand



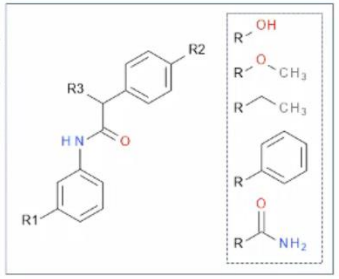
Molecular Dynamics in DS

- Before Run MD
  - Solvation
  - Explicit Membrane-Based Solvation
- MD in DS
  - Standard Dynamic Cascade (SDC)
  - NAMD
    - Free Energy Perturbation
  - Steered Molecular Dynamics (SMD)
  - Multi-Site Lambda Dynamics (MSLD)

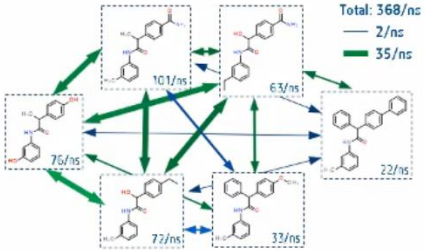
Free Energy: Multi-Site Lambda Dynamics+

- Explore large congeneric chemistry space in early stage Lead Optimization in a way that FEP can't
  - Based on Lambda Dynamics with more than two decades of development and refinement, unique to CHARMm
  - Multiple ligands in a single simulation, mimicking a competitive binding assay
  - Up t 20 times more efficient than FEP
  - User-friendly and easy to setup, run and analyze

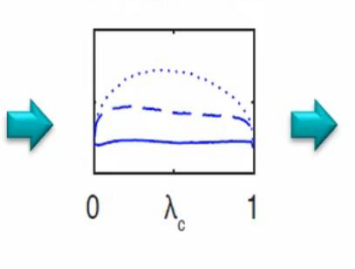
Simulate a competitive binding assay of a combinatorial library of ligands



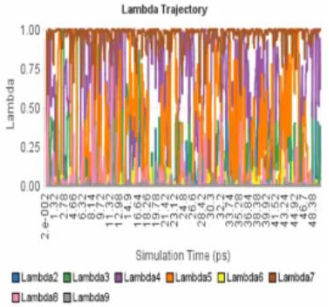
E.g., 5x5x5=125 analogs



Flattening free energy landscape enhances sampling and leads to efficient calculation of binding affinities



Obtain ligand affinities from occupancy ratios



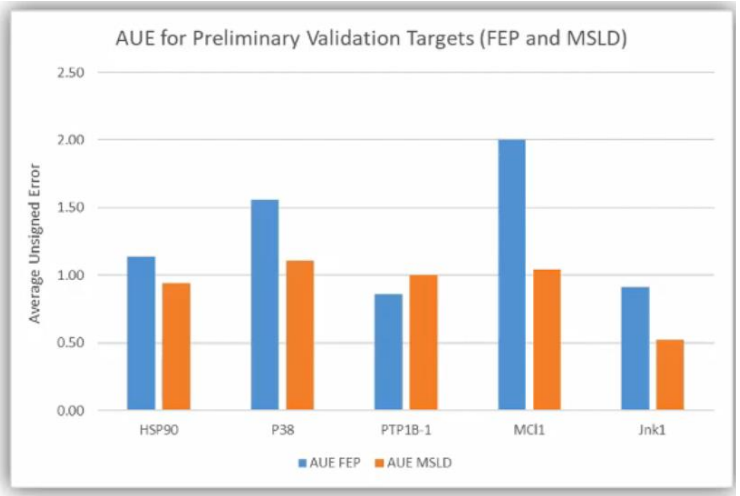
J Phys Chem B. 2017, 727, 3626-3635

訊聯基因數位

MSLD: Preliminary Validation

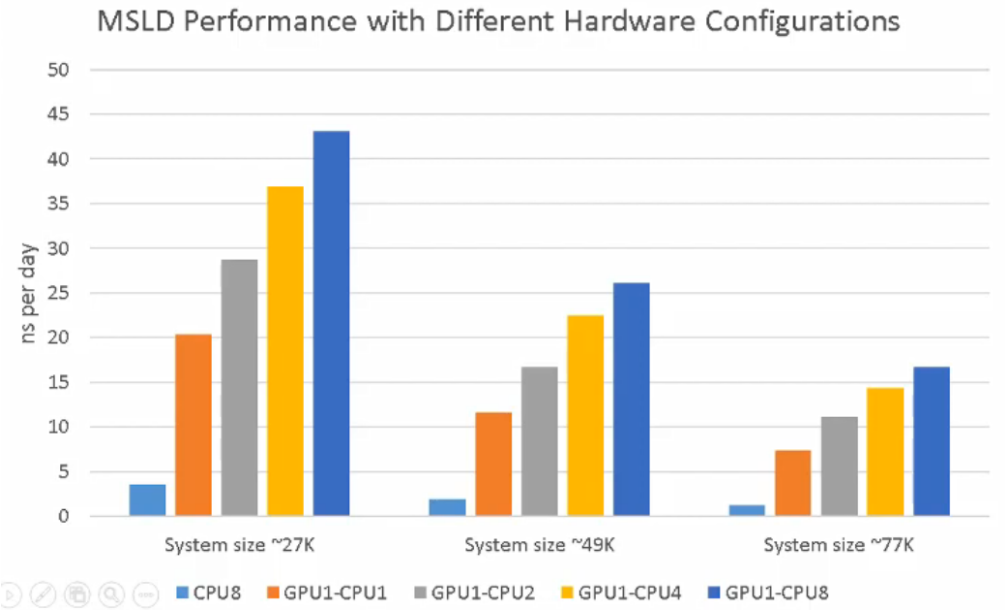
- MSLD-domain datasets with experimental data for validation are difficult to find
  - Permutation of multiple sites on a common core
  - Large dynamic range of relative binding free energy for realistic ranking challenge
- Preliminary validation
  - 5 protein targets, 48 ligands
  - Compare to existing FEP results (up to 32ns sampling per pair)
- MSLD validation parameters
  - Setup systems in cubic boxes and run on domdec-GPU (no restraints)
  - Bias Optimization:~100 iterations 100ps length each
  - 30ns of production, in triplicate (no replica exchange)
  - Compute simple average of $\triangle G_{solv}$ and $\triangle G_{prot}$ to calculate $<\triangle G>$ and σ
- MSLD average unsigned error of predicted affinities less than FEP
  - Approximately 1 kcal/mol

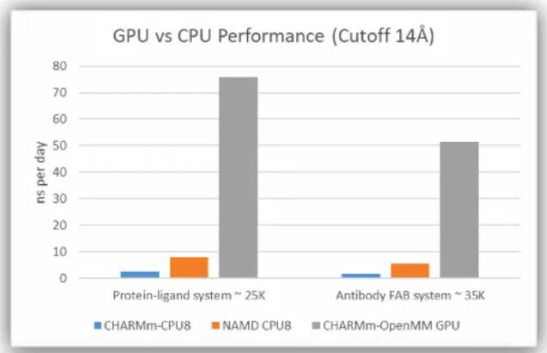
Set	Sites	nLigands	Range (kcal/mol)
HSP90	2	9 (6x5)	4.3
P38	1	6	1.5
PTP1B-1	1	14	2.8
MCl1	1	8	2.2
Jnk1	1	11	2.1

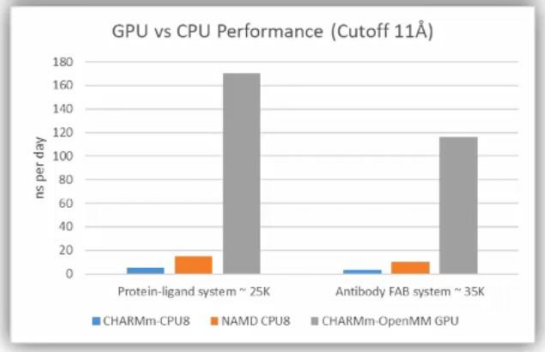


Multi-Site Lambda Dynamics - MSLD

- In DS2019 CHARMm-OpenMM GPU on Linux accelerated MD production
- In DS2020 MSLD is available through CHARMm-DOMDEC on Linux for both CPU and GPU







Q&A

訊聯基因數位

For more information please contact…

訊聯基因數位

台北市114內湖區新湖一路36巷28號

中華民國 台灣

Ph: (02) 2795 1777 x 3014 Fax: (02) 2793 8009

msc-support@gga.asia

www.gga.asia

訊聯基因數位