Modelling of ligand-protein binding I. Computation of thermodynamic quantities

1

Outline

- Molecular dynamics
- Sampling in molecular dynamics
- Calculation of thermodynamic quantities along the pathway
  - Techniques for calculating free energy:
    - Thermodynamic integration
    - Free energy perturbation
    - Potential of mean force
    - Non-equilibrium work
  - Enthalpy and entropy
  - Examples

2

Molekular dynamics

3

Introduction

- Molecular dynamics – link between microscopic and macroscopic quantities
  - structure
  - dynamics
  - thermodynamics

FnIII_10_preSMD.mpg

4

History

- Alder, B. J. and Wainwright, T. E.
  J. Chem. Phys. **27**, 1208 (1957)

- McCammon, J. A., Gelin, B. R., and Karplus, M.
  Nature (Lond.) **267**, 585 (1977)

5

Basic terms

- Thermodynamic/Macroscopic state
  - The system is characterized by few macroscopic parameters; e.g.: T, P, N
- Microscopic state
  - The system is characterized by the positions and momenta of atoms (phase space).
- Ensembles
  - Microscopic states corresponding to a macroscopic state
- Molecular dynamics simulations
  - Generation of microscopic states of an ensemble as a function of time

6

Thermodynamic ensembles

- Microcanonical – NVE (isolated system)
- Canonical – NVT (thermal equilibrium)
- Izotherm-izobar – NPT
- Grand canonical – μVT (equilibrium with a reservoir of particles)

7

7

Ergodic hypothesis

- Measurable quantities: ensemble average $\langle A \rangle_{ensemble}$
  - e.g. (non-covalent) binding of two molecules in solution
- Molecular dynamics: time average $\langle A \rangle_{time}$

$$\langle A \rangle_{time} = \langle A \rangle_{ensemble}$$

- „long enough" MD – appropriate sampling

8

8

Force field

- Molecular mechanics
  - basic entities: atoms
  - „classical"
  - simple, fast computations
  - includes parameters
    - Can be applied within the validity of the parameter space
    - Chemical reactions are typically outside the validity
- Quantum mechanics
  - basic entities: nuclei and electrons
  - accurate
  - time intensive computations

9

9

Molecular mechanics

Atoms are pointlike objects with mass and interactions

$$E = E_{str} + E_{bend} + E_{tors} + E_{vdw} + E_{el} + E_{cross}$$

intramolecular | intermolecular

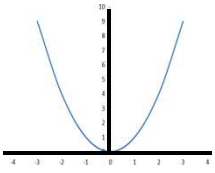
10

10

Bond stretching energy

$$E_{str} = k(r - r_0)^2$$

$$F_{str} = -2k(r - r_0)$$



good approximation in the vicinity of r_0
k and r_0 are atom dependent parameters

11

11

Bond stretching energy - parameters

$$E_{str} = k(r - r_0)^2$$

general class	atom type[1]	description
hydrogen types		
	H	amide or imino hydrogen
	HC	explicit hydrogen attached to carbon
	HO	hydrogen on hydroxyl oxygen
	HS	hydrogen attached to sulfur
	HW	hydrogen in water
	H2	amino hydrogen in NH_2
	HS	hydrogen of lysine or arginine (positively charged)
all-atom carbon types[2]		
	C	sp^2 carbonyl carbon and aromatic carbon with hydroxyl substituent in tyrosine
	CA	sp^2 aromatic carbon in 6-membered ring with 1 substituent
	CB	sp^2 aromatic carbon at junction between 5- and 6-membered rings
	CC	sp^2 aromatic carbon in 5-membered ring with 1 substituent and next to a nitrogen
	CK	sp^2 aromatic carbon in 5-membered ring between 2 nitrogens and bonded to 1 hydrogen (in purine)
	CM	sp^2 same as CJ but one substituent
	CN	sp^2 aromatic junction carbon in between 5- and 6-membered rings
	CQ	sp^2 carbon in 6-membered ring of purine between 2 NC nitrogens and bonded to 1 hydrogen
	CR	sp^2 aromatic carbon in 5-membered ring between 2 nitrogens and bonded to 1 H (in his)
	CT	sp^3 carbon with 4 explicit substituents
	CV	sp^2 aromatic carbon in 5-membered ring bonded to 1 N and bonded to an explicit hydrogen
	CW	sp^2 aromatic carbon in 5-membered ring bonded to 1 N-H and bonded to an explicit hydrogen
	C*	sp^2 aromatic carbon in 5-membered ring with 1 substituent

Bond Stretching Potential Parameters

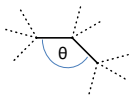
Bond	used for	K_r/ kcal mol^{-1} Å^{-2}	R_{eq}/ Å
CT-CT	*BMF_EMF*	310.0	1.526
CT-H1	*BMF_EMF*	340.0	1.090
CT-HC	*BMF_EMF*	340.0	1.090
CT-NA	*BMF_EMF*	337.0	1.475
CR-H5	*BMF_EMF*	367.0	1.080
CR-NA	*BMF_EMF*	477.0	1.343
CW-H4	*BMF_EMF*	367.0	1.080
CW-NA	*BMF_EMF*	427.0	1.381
CW-CW	*BMF_EMF*	549.0	1.350
AL-Cl	*TCA*	116.1	2.170
P-F	PF_6^-	260.3	1.646
NN-ON	NO_3^-	300.0	1.260

12

12

Bending energy

$$E_{bend} = k(\theta - \theta_0)^2$$

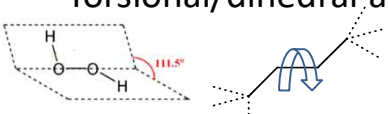


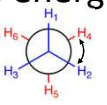
k and θ_0 are atom dependent parameters

13

13

Torsional/dihedral angle energy



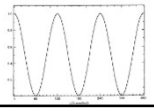


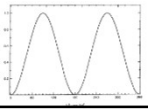
$$E_{tors} = \frac{V_n}{2}[1 + cos(n\varphi - \varphi_0)]$$

V_n – barrier height
n - periodicity

Dihedral angle (atom types)				V_n, kcal/mol	n	ϕ_0	comments
C	CT1	NH1	C	0.2000	1	180.00	! backbone phi
NH1	C	CT1	NH1	0.6000	1	0.00	! backbone ksi
CT1	C	NH1	CT1	1.6000	1	0.00	! backbone omega
CA	CA	CA	CA	3.1000	2	180.00	! Phe side chain
H	OH1	CT2	CT1	0.4200	3	0.00	! Ser side chain

~15000 parameters

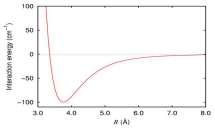




14

14

van der Waals energy



short range:	repulsive;	exp(-r) or r^{-12}	Pauli repulsion
middle range:	attractive;	r^{-6}	dispersion
long range:	disappears		

$$E_{vdw} = 4\epsilon\left[\left(\frac{\sigma}{r}\right)^{12}\right] - \left[\left(\frac{\sigma}{r}\right)^{6}\right]$$

$$\sigma_{ij} = \frac{1}{2}(\sigma_i + \sigma_j)$$

$$\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}$$

$$c * exp\left(\frac{-r}{\sigma}\right)$$

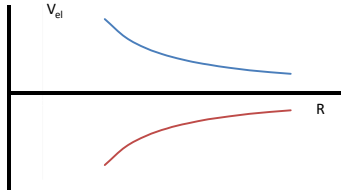
15

15

Electrostatic energy

$$V_{el} = \frac{q_i q_j}{\epsilon r_{ij}}$$

Coulomb



16

16

MM parameters

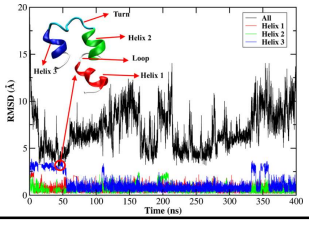
- Derivation
  - Quantum mechanical calculations
  - Experimental data
  - Extension based on analogy
- Validation by comparing computed and experimental data
  - Macromolecular structure
  - NMR data
  - Structure and energy of van der Waals complexes
- Error compensation; mutual interdependence of parameters

17

17

Quality of MM force field

- Protein structure
- DNA, RNA structure
- Conformation of organic molecules
- Ligand-protein interactions
- Structure and interactions of lipids and membranes
- …



RMSDs of backbone atoms from the native structure as a function of MD simulation time

Duan J Mol Model 2014,20,2195

18

18

Selected MM force fields

- Charmm (Chemistry at HARvard Macromolecular Mechanics)
- AMBER (Assisted Model Building with Energy Refinement)
- OPLS (Optimized Potentials for Liquid Simulations)
- GROMOS (GROningen Molecular Simulation)
- MMFF (Merck Molecular Force Field)

19

19

MD algorithm

Newtonian mechanics

$$r^N(r_1, r_2 \dots r_N) \qquad p^N(p_1, p_2 \dots p_N)$$

$$U(\underline{r}) \qquad K(\underline{p}) = \sum_i \frac{|p_i|^2}{2m_i}$$

$$H = K + U \qquad \dot{r}_i = \frac{p_i}{m_i} \qquad \dot{p}_i = f_i$$

20

20

Verlet algorithm

Calculation of p_i and r_i at δt time steps

$$p_i(t), r_i(t) \rightarrow p_i(t+\delta t), r_i(t+\delta t) \qquad \rightarrow f_i(t)$$

$$p_i\left(t+\frac{1}{2}\delta t\right) = p_i(t) + \frac{1}{2}\delta t \cdot f_i(t)$$

$$r_i(t+\delta t) = r_i(t) + \frac{\delta t \cdot p_i\left(t+\frac{1}{2}\delta t\right)}{m_i} \qquad \rightarrow f_i(t+\delta t)$$

$$p_i(t+\delta t) = p_i\left(t+\frac{1}{2}\delta t\right) + \frac{1}{2}\delta t \cdot f_i(t+\delta t)$$

Typical δt for simulation of biochemical systems: 1-4 fs

1 μs MD ~ 10^9 evaluations

21

21

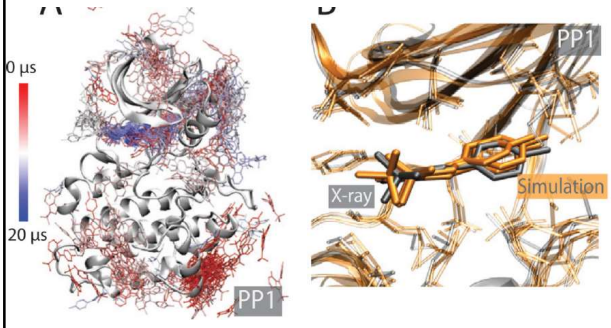
Selected applications

22

22

Application 1

process of ligand binding



PP1 molecule finds the binding site of Src kinase in a 15μs simulation

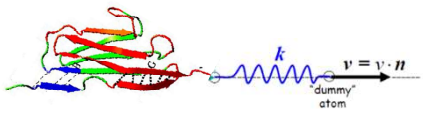
JACS 2011 133 9181

23

Application 2

Steered MD

- Constant velocity pulling



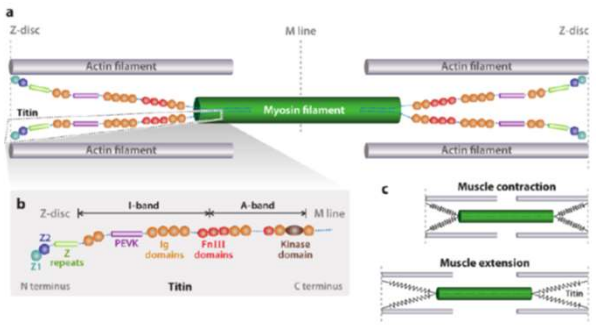
$$U = \frac{1}{2}k[vt - (\vec{r} - \vec{r}_0) \cdot \vec{n}]^2$$

- Constant force pulling

24

Application 2

Titin structure and function



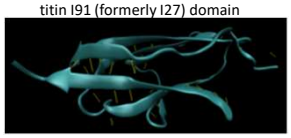
25

25

Application 2

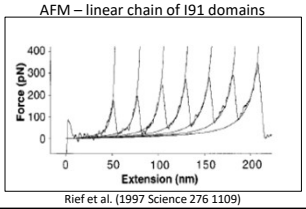
Unfolding and force

titin I91 (formerly I27) domain

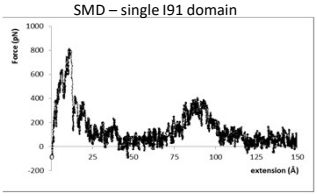


Constant velocity pulling

AFM – linear chain of I91 domains



SMD – single I91 domain

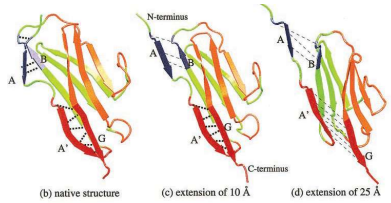


Rief et al. (1997 Science 276 1109)

26

Application 2

Unfolding and structure

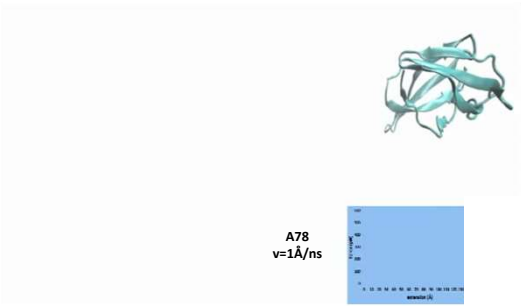


Lu, H., and Schulten, K. (2000).. Biophys. J. 79, 51–65

27

Application 2

Titin Fn domain unfolding mechanism



28

28

MD scope and limitation

- Scope
  - Structural study; structure refinement
  - Dynamics
    - conformations, ligand-protein binding, steered processes,…
  - Thermodynamics see later
    - Free energy changes
      - solvation, ligand-protein binding,…
- Limitations
  - Sampling see later
  - Accuracy of force field
  - Chemical reactions cannot be routinely studied

29