

# Fyzika biopolymerů

## Základy simulací

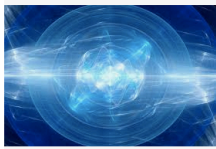
Robert Vácha

Kamenice 5, A4 2.13  
[robert.vacha@mail.muni.cz](mailto:robert.vacha@mail.muni.cz)



## Force Field

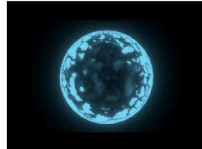
google



Force Field Analysis  
[free-management-ebooks.com](http://free-management-ebooks.com)



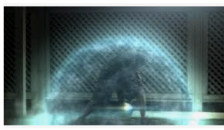
Force Field Development — L-P. Wang ...  
[lpwchem.org](http://lpwchem.org)



Blender - Force Field Texture Tutorial ...  
[youtube.com](http://youtube.com)



Into reality with force field patent ...  
[trustedreviews.com](http://trustedreviews.com)



An Accidentally Created Force Field ...  
[disclose.tv](http://disclose.tv)



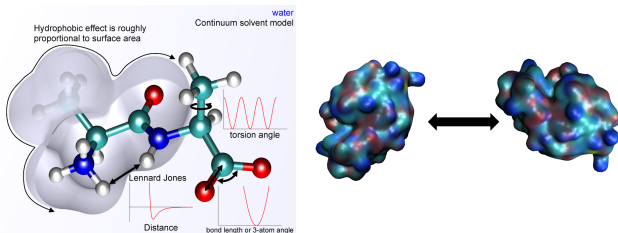
Magical Girl Lytical Nanoha Wiki ...  
[nanoha.fandom.com](http://nanoha.fandom.com)

2

## Force Field

$V(\vec{r})$  - force field

- interakční potenciál ... přibližný
- soubor rovnic a parametrů popisujících interakce mezi částicemi systému
- odvozen fitováním experimentálních a *ab initio* či jinak vypočtených dat
- označován za empirický potenciál (ale není 100% protože obsahuje *ab initio* data)
- často dobře funguje jen na oblasti blízké fitovaným podmínkám (nejčastěji 300K, 1atm, ve vodném roztoku, pH neutral, ...)
- obvykle se takto označuje atomový model, lze chápat obecněji i na zhrubené modely

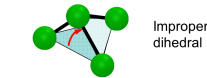
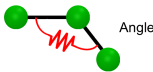
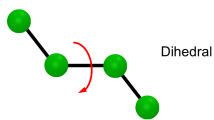
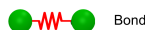


3

## Force Field all-atom

$$V_{tot} = V_{bonded} + V_{nonbonded}$$

$V_{bonded}$



$V_{nonbonded}$

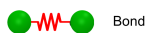


$$V_{tot} = \sum_{bonds} K_r(r - r_0)^2 + \sum_{angles} K_\theta(\theta - \theta_0)^2 +$$

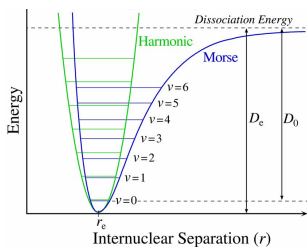
$$\sum_{dihed} \frac{V_n}{k} [1 + \cos(n\varphi - \gamma)] + \sum_{vdw} \left( \frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} \right) + \sum_{elstat} \frac{q_i q_j}{\epsilon R_{ij}}$$

4

## Force Field all-atom

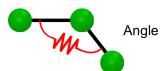


$$V_{morse}(r_{ij}) = D_{ij} [1 - \exp(-\beta_{ij}(r_{ij} - b_{ij}))]^2$$



$$V_b(r_{ij}) = \frac{1}{2} k_{ij}^b (r_{ij} - b_{ij})^2$$

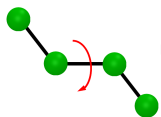
- v praxi často drženo (constrained)  
na pevné rovnovážné hodnotě pro  
urychlení simulací



$$V_a(\theta_{ijk}) = \frac{1}{2} k_{ijk}^\theta (\theta_{ijk} - \theta_{ijk}^0)^2$$

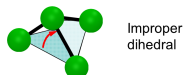
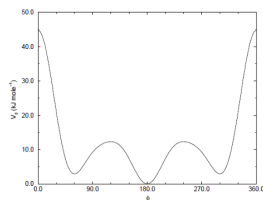
5

## Force Field all-atom



$$V_d(\phi_{ijkl}) = \sum_{n=0}^5 k_{\phi_n} (1 + \cos(n\phi_n - \phi_{sn}))$$

$$V_{rb}(\phi_{ijkl}) = \sum_{n=0}^5 C_n (\cos(\psi))^n$$



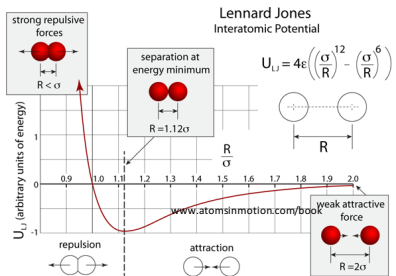
$$V_{id}(\xi_{ijkl}) = \frac{1}{2} k_\xi (\xi_{ijkl} - \xi_0)^2$$

6

## Force Field all-atom



$$V_{LJ}(r_{ij}) = \frac{C_{ij}^{(12)}}{r_{ij}^{12}} - \frac{C_{ij}^{(6)}}{r_{ij}^6}$$



$$U_{LJ} = 4\epsilon \left( \left( \frac{\sigma}{R} \right)^{12} - \left( \frac{\sigma}{R} \right)^6 \right)$$



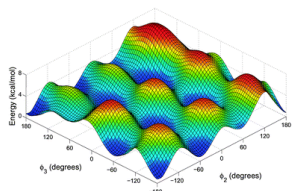
Coulombic

$$V_{ij} = \frac{q_i q_j}{4\pi\epsilon_0 r_{ij}}$$

7

## Potential energy surface

Mnoharozměrný (počet stupňů volnosti) povrch potenciální energie po kterém se částice pohybují ... může být popsán pomocí kvantové teorie, forcefieldu, modelu...



Pro řešení pohybu v kvantové teorii separujeme pohyb elektronů a jader

**Born-Oppenheimerova aproximace:** elektrony se okamžitě přizpůsobují pohybu jader (jádra jsou mnohem těžší než elektrony), pohyb jader tedy lze řešit klasicky a pro každou konfiguraci jader vyřešit Schrodingerovu rovnici = vlnovou funkci rozložení elektronů

8

## Molekulová dynamika (MD)

- řeší klasické Newtonovy rovnice

$$\vec{F}_i = m_i \vec{a}_i$$

$$-\frac{dV}{d\vec{r}_i} = m_i \frac{d^2 \vec{r}_i}{dt^2} \quad V(\vec{r}) - \text{force field}$$

- výsledkem je trajektorie = set po sobě jdoucích konfigurací systému s definovaným časovým odstupem

- vlastnosti systému se počítají analýzou trajektorie

- musí být zadány počáteční podmínky - polohy a rychlosti všech částic

- používá forcefield, který musí mít definované síly (první derivace potenciálu)

- deterministická

- časově reversibilní

- zachování momentů hybnosti

- numerické řešení rovnic pomocí integrátoru

$$\vec{r}_i(t=0), \vec{v}_i(t=0)$$



$$\vec{r}_i(t), \vec{v}_i(t)$$

9

## Molekulová dynamika (MD)

- Historie

1957 - Alder and Wainwright - basics of MD

1964 - Rahman - MD with Lennard-Jones potential, NVE, liquid argon

1967 - Verlet - Verlet integration algorithm and Verlet neighbor list

1974 - Stillinger and Rahman - MD of water

1981 - Andersen and Parrinello-Rahman - NPT

1986 - Nose and Hoover - NVT with Nose-Hoover thermostat

1985 - Car and Parrinello - ab initio MD

10

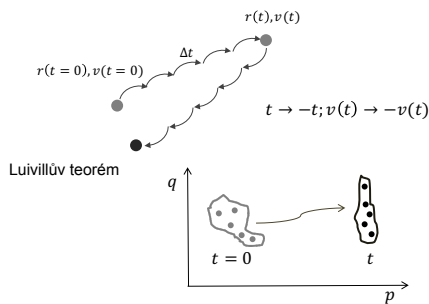
## Integrátory = numerické řešení pohybových rovnic

Taylorův rozvoj

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \mathbf{v}_i(t) + \frac{\Delta t^2}{2m_i} \mathbf{f}_i(t) + \mathcal{O}(\Delta t^3)$$

$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i(t) + \frac{\Delta t}{m_i} \mathbf{f}_i(t) + \mathcal{O}(\Delta t^2)$$

není časově reverzibilní a nezachovává fázový objem



11

## Integrátory

řešením reverzibility je symetrizace

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \mathbf{v}_i(t) + \frac{\Delta t^2}{2m_i} \mathbf{f}_i(t) + \frac{\Delta t^3}{3!} \ddot{\mathbf{r}}_i(t) + \mathcal{O}(\Delta t^4)$$

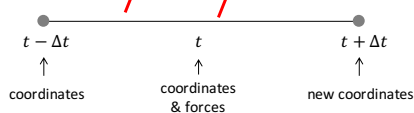
$$\mathbf{r}_i(t - \Delta t) = \mathbf{r}_i(t) - \Delta t \mathbf{v}_i(t) + \frac{\Delta t^2}{2m_i} \mathbf{f}_i(t) - \frac{\Delta t^3}{3!} \ddot{\mathbf{r}}_i(t) + \mathcal{O}(\Delta t^4)$$

sečtením rovnic....

Verletův algoritmus

$$\mathbf{r}_i(t + \Delta t) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t - \Delta t) + \frac{\Delta t^2}{m_i} \mathbf{f}_i(t) + \mathcal{O}(\Delta t^4)$$

$$\mathbf{v}_i(t) = \frac{\mathbf{r}_i(t + \Delta t) - \mathbf{r}_i(t - \Delta t)}{2\Delta t} + \mathcal{O}(\Delta t^3)$$



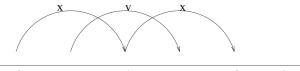
12

## Integrátory

Leap-frog (Skákající žáby) algoritmus

algebraicky stejné řešení jako Verlet

- časově reverzibilní
- ale nezachovává fázový objem


$$\mathbf{v}_i(t + \frac{\Delta t}{2}) = \mathbf{v}_i(t - \frac{\Delta t}{2}) + \frac{\Delta t}{m_i} \mathbf{f}_i(t)$$
$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \mathbf{v}_i(t + \frac{\Delta t}{2})$$

Rychlostní-Verletův algoritmus (Velocity-Verlet)

$$\mathbf{r}_i(t + \Delta t) = \mathbf{r}_i(t) + \Delta t \mathbf{v}_i(t) + \frac{\Delta t^2}{m_i} \mathbf{f}_i(t) + \mathcal{O}(\Delta t^3),$$

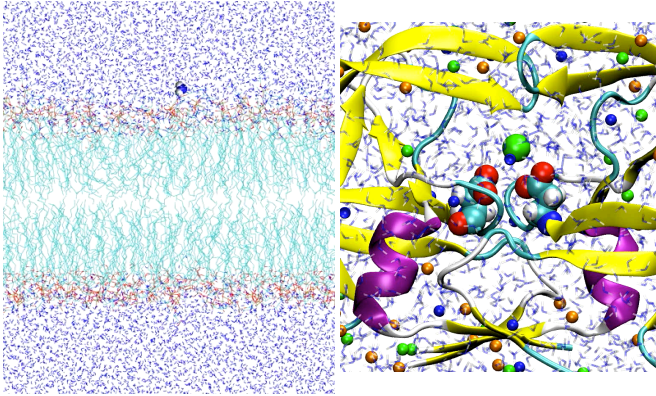
$$\mathbf{v}_i(t + \Delta t) = \mathbf{v}_i(t) + \frac{\Delta t}{2m_i} (\mathbf{f}_i(t) + \mathbf{f}_i(t + \Delta t)) + \mathcal{O}(\Delta t^3)$$

- je časově reverzibilní
  - zachovává fázový objem => ideální řešení
  - velmi stabilní
- ale zkontrolujte doporučení  
používaného programu

Potřeba termostatu, barostatu, atd. !!! specifické přednášky

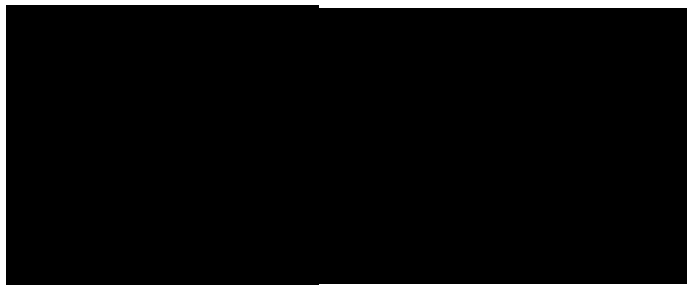
13

## Příklady Molekulové dynamiky



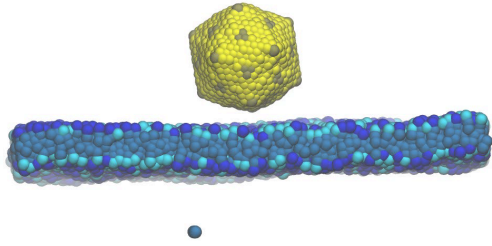
14

## Příklady Molekulové dynamiky



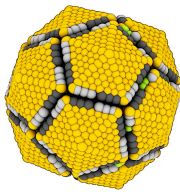
15

## Příklady Molekulové dynamiky



16

## Příklady Molekulové dynamiky



17

## Monte Carlo (MC)

- obecná metoda na vzorkování pomocí náhodných čísel
- molekulární x nemolekulární (výpočet integrálů, diferenciálních rovnic atd.) používá se ve finančnictví, počítačové grafice nebo umělé inteligenci

### Historie

- 1930s & 1940s – Fermi, Ulam, von Neumann: Los Alamos, the Manhattan project = procházení neutronů skrz materiál
- 1950s – Teller: vývoj vodíkové bomby
- 1953 – Metropolis (Metropolis-Hastings algorithm) "Equation of State Calculations by Fast Computing Machines"
- 1966 – Whittington et al. "Effect of density on configurational properties of long-chain molecules using a Monte Carlo method"
- 1987 – Li & Scheraga: Monte Carlo - minimalizace pro hledání mnoha minim v proteinovém sbalování (foldingu)
- 1992 – Leach: Dockování ligandu do proteinů s flexibilitou postranních řetězců
- 2011 – Chodera: Pohyb nerovnovážného kandidáta
- 2011 – Martinez Veracochea & Frenkel : Hybridizační pohyb vytvoření vazby mezi ligandy a receptory

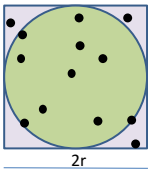
18

## Příklad

Vypočítejte pi pomocí nahodného vzorkování bodu ve čtverci a jemu vepsanému kruhu

19

## Řešení



Pravděpodobnost, že náhodný bod je v kruhu

$$P = \frac{A_{kruh}}{A_{ctverec}}$$

$$P = \frac{\pi R^2}{(2R)^2} = \pi/4$$

Obdobně lze spočítat objem velmi komplikovaných objektů

20

## Kanonický soubor

- vzorkování konfigurací pomocí nahodných změn
- pravděpodobnost konfigurací je dán Boltzmannovým vztahem

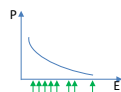
$$P(E) \approx e^{-\frac{E}{kT}}$$

- střední hodnota veličiny

$$\langle A \rangle = \frac{\sum A_i \exp\left(-\frac{E_i}{kT}\right)}{\sum \exp\left(-\frac{E_i}{kT}\right)}$$

- problém se vzorkováním - většinu času se vzorkují stavy, které přispívají minimálně, důležité stavy jsou málo vzorkovány

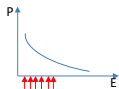
$$\langle A \rangle = \frac{\sum A_i \exp\left(-\frac{E_i}{kT}\right)}{\sum \exp\left(-\frac{E_i}{kT}\right)}$$



21

## Metropolis-Hastings algorithm

- vzorkuje se oblast (kofigurace), které jsou v kanonickém souboru důležité



- nové konfigurace jsou přijímány s pravděpodobností

$$p(\text{old} \rightarrow \text{new}) = e^{-\frac{E_{\text{new}} - E_{\text{old}}}{kT}} \quad \text{if } E_{\text{new}} > E_{\text{old}}$$

$$= 1 \quad \text{if } E_{\text{new}} \leq E_{\text{old}}$$

Příklad algoritmu:

1. náhodně vyber částici
2. vyber náhodný směr
3. pohni částici
4. přijmi novou polohu na základe pravděpodobnosti výše

22

## Jak navrhnout nový krok

- Podmínka detailní rovnováhy (detailed balance condition)

$$P(\text{old})a(\text{old} \rightarrow \text{new}) = P(\text{new})a(\text{new} \rightarrow \text{old}) \quad \text{v rovnováze se zastoupení stavů nemění}$$

$P$  je pravděpodobnost stavu,  $a$  je pravděpodobnost přechodu skládající se z pravděpodobnosti výběru daného pohybu krát pravděpodobnost přijetí daného pohybu  $p$

- pokud je pravděpodobnost výběru daného pohybu vpřed a zpěz stejná existuje jednoduché pravidlo splující podmínku rovnováhy (mohou být i další)

$$p(\text{old} \rightarrow \text{new}) = \frac{P(\text{new})}{P(\text{old})} \quad \text{if } P(\text{new}) < P(\text{old})$$

$$= 1 \quad \text{if } P(\text{new}) \geq P(\text{old})$$

např. v NPT (isobaricko-isochorickém) souboru

$$P(N, V) = V^N e^{-\frac{pV}{kT}} e^{-\frac{E}{kT}}$$

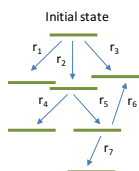
$$p(\text{old} \rightarrow \text{new}) = \min \left\{ 1, \exp \left[ -\frac{E_{\text{new}} - E_{\text{old}}}{kT} + \frac{p(V_{\text{new}} - V_{\text{old}}) - NkT \ln(V_{\text{new}}/V_{\text{old}})}{kT} \right] \right\}$$

23

## Kinetické Monte Carlo

- musíme znát rychlostní konstanty

$$\text{např. } \approx e^{-\frac{E_B}{kT}}$$



Algoritmus:

1. náhodně se vybere proces
2. proces se zrealizuje s pravděpodobností

$$P_i = \frac{r_i}{\sum r_k}$$

3. aktualizuj čas

- existuje algoritmus bez zamítání procesů
- lze simulovat nerovnovážné procesy

## Dynamické Monte Carlo

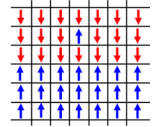
- musíme znát difuzní koeficienty (translační, rotační,...)
- parametry pohybů (maximální velikost pohybu, pravděpodobnost pohybu) se nastaví podle difuzních koeficientů
- konverguje k difuzivnímu (brownovskému) pohybu

24

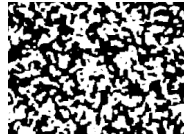


## Příklady Monte Carla

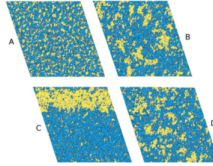
Ising model (originally for spins)



$$H(\sigma) = - \sum_{\langle i, j \rangle} J_{ij} \sigma_i \sigma_j - \mu \sum_j h_j \sigma_j$$



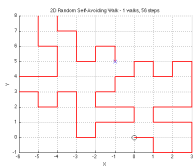
Ising model for mixed lipid membranes



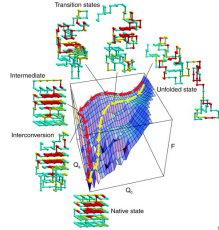
25

## Příklady Monte Carla

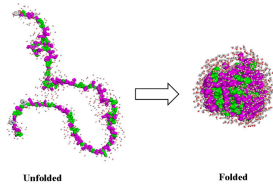
Polymer on a lattice



Polymer folding

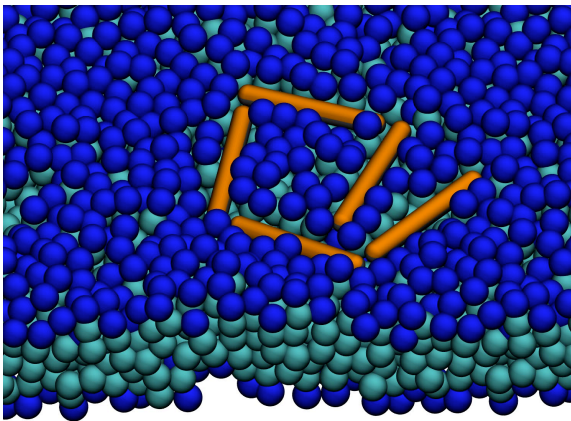


Protein folding



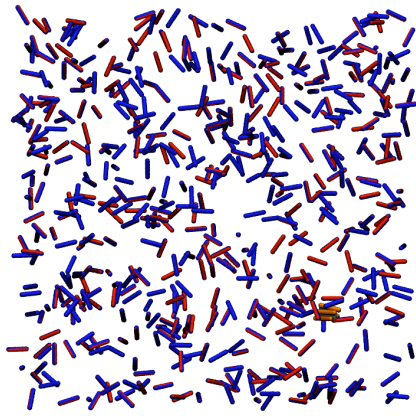
26

## Příklady Monte Carla



27

## Příklady Monte Carla



28

## MD x MC

MD

výhody:

- dynamika vývoje systému automatickou součástí
- informace o rychlostech momentu atd
- rovnovážná i nerovnovážná
- jednoduchý algoritmus
- spousta optimalizovaného softwaru

nevýhody:

- pouze malé změny v jednotlivých krocích = pomalost
- nutnost sil = pouze spojitě diferencovatelné potenciály
- potřeba dodatečných algoritmů a nastavení parametrů pro jiné než mikroaonické soubory

MC

výhody:

- obecnost a variabilita (systém, potenciály, termodynamický soubor, pohyby a změny systému)
- nevyžaduje síly
- možnost přesakovat bariéry, snadné úniky minim
- algoritmus více flexibilní
- snadno různé termodynamické soubory

nevýhody:

- systémy v rovnováze a bez dynamiky (nerovnováha a dynamika vyžadují speciální přístup)
- programy nejsou optimalizované, často vyžadují vlastní úpravu

29

## Ergodická hypotéza

$$\langle X \rangle = \lim_{\tau \rightarrow \infty} \frac{1}{\tau} \int_0^{\tau} X(t) dt = \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{i=1}^M X_i$$

= průměr získaný pro malý počet molekul přes dlouhý čas je ekvivalentní průměrování přes velký počet molekul a krátký čas (v limitě: časový průměr pro jednu molekulu je ekvivalentní průměru pro velký počet molekul)

- výsledky simulací a experimentů představují průměrné hodnoty (průměrování přes počet molekul, čas)

30

## Statistical ensembles

- thermodynamic statistical ensembles describe macroscopic conditions
- **NVE** – microcanonical
- **NVT** – canonical (other names: isothermal, Helmholtz canonical)
- $\mu$ **VT** – grand-canonical
- **NPE** – isobaric
- **NPT** – isobaric-isothermal (other name: Gibbs canonical)

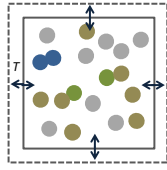
NVE



NVT



NPT



$$\Omega(N, V, E) = \frac{1}{N!} \sum_X \delta[\mathcal{H}(X) - E]$$

$$Z(N, V, T) = \frac{1}{N!} \sum_X e^{-\beta \mathcal{H}(X)}$$

$$Q(N, P, T) = \int dV e^{-\beta PV} \frac{1}{N!} \sum_X e^{-\beta \mathcal{H}(X)}$$

$$S(N, V, E) = k_B \ln \Omega(N, V, E)$$

$$F(N, V, T) = -k_B T \ln Z(N, V, T)$$

$$G(N, P, T) = -k_B T \ln Q(N, P, T)$$

---

---

---

---

---

---

---

---

---

---