# Monte Carlo Methods in Molecular Modeling

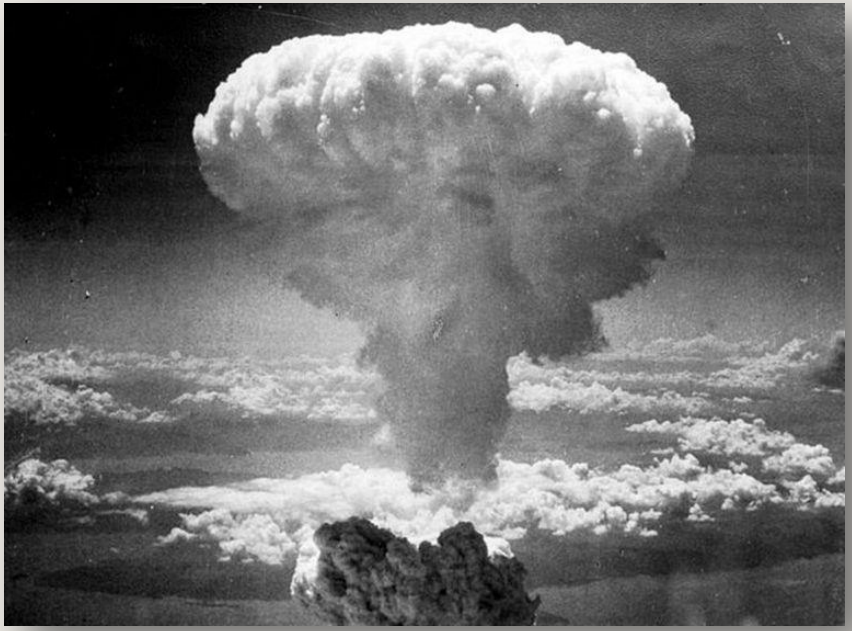
Alicja Nowakowska, Molecular Modeling Seminar

# TABLE OF CONTENTS

# 1. The Power of Monte Carlo

# A little bit of history

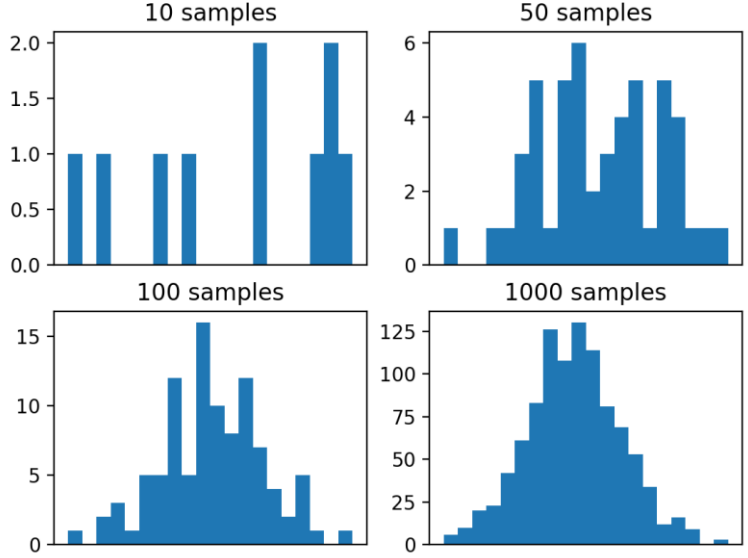
**P1**: Complex card game – how often can I win?



**P2**: Manhattan Project, 1946, tracking neutron chains

The Beginning of the Monte Carlo Method, N.Metropolis, Los Alamos Science, Special Issues 1987

# Game outcomes distribution – How probable it is to win X $




**Background:** complex model defining the rules for events

**Problem:** how probable is each event? What events can we observe?

**Solution:** perform model simulation N times – outcomes = samples

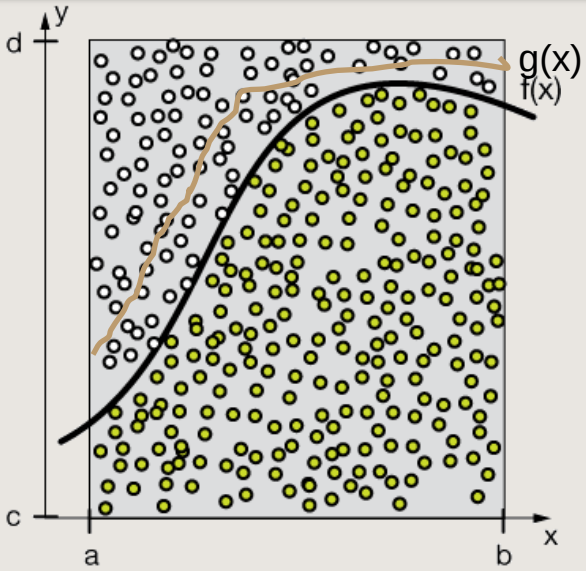
**Checklist:**

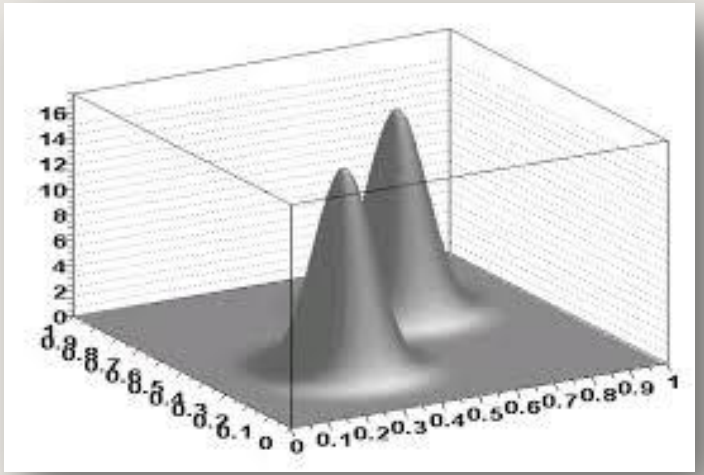
- Error of results ~ $\frac{1}{\sqrt{N}}$
- Random number generator

https://machinelearningmastery.com/monte-carlo-sampling-for-probability/

# Monte Carlo methods

**Definition**: any numerical technique that uses (pseudo)random numbers for solving the problem
**First Example:** Integral calculation, D dimensions: 1 sum instead of D sum



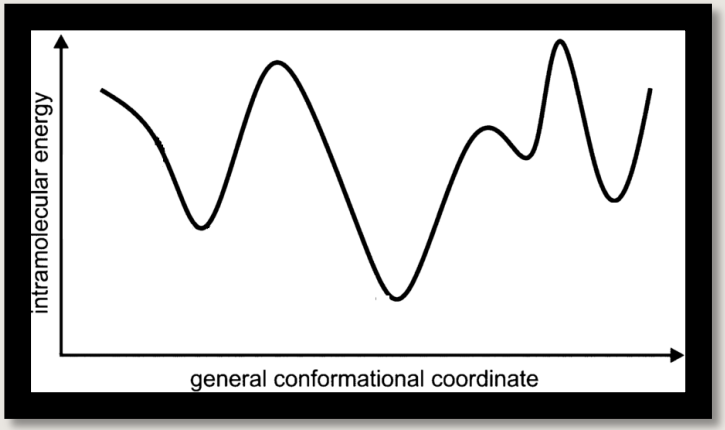



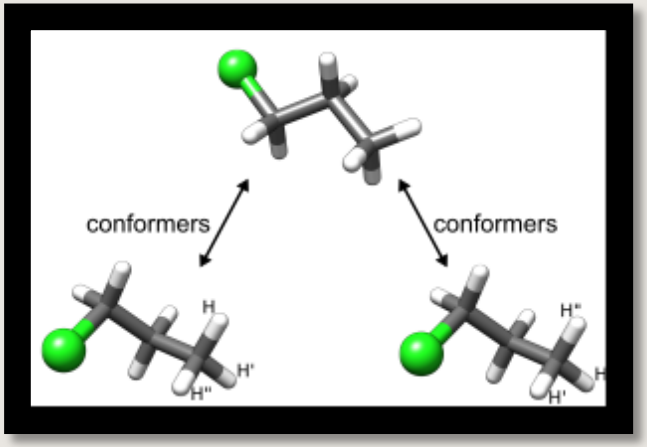

**Improve:** sample from a similar distribution and accept with some ratio – **Importance sampling**

[1] https://xaktly.com/NumericalIntegration.html [2] https://indico.cern.ch/event/568875/contributions/2397925/attachments/1459058/2253175/mcgbr-May12-2017.pdf

# 2. Conformational landscape

# **Problem:** equilibrium states and their populations




**According to Boltzmann distribution:**

$$< E > = \sum_i E_i e^{-\frac{E_i}{kT}} / \sum_i e^{-\frac{E_i}{kT}}$$

**Simplified problem:** Average energy

**First idea:** get N configurations by randomly generating 3N Cartesian coordinates

**Effect:** many of unphysical meaningless configurations with low weight

**Other solution needed!**

1sr image: Which conformations make stable crystal structures? Mapping crystalline molecular geometries to the conformational energy landscape, Hugh P. G. Thompsona and Graeme M. Day, : Chem. Sci., 2014, 5, 3173 2nd image: https://xtb-docs.readthedocs.io/en/latest/crest.html Theory from: Equilibrium Sampling in Biomolecular Simulation, Daniel M. Zuckerman, Annu Rev Biophys. 2011 ; 40: 41–62.Molecular Modelling: Principles and Applications 2nd edition, Andrew R. Leach. 2001, ISBN 0582382106

# Metropolis Algorithm

NVT Ensemble

**Ratio:**

$$\frac{p_i}{p_j} = e^{-\frac{\Delta E}{kT}}$$

Proportional to Boltzmann dsitribution

**Checklist:**

- Burn-in period
- Risk of getting trapped in a local min
- Initial configuration
- Time of convergence

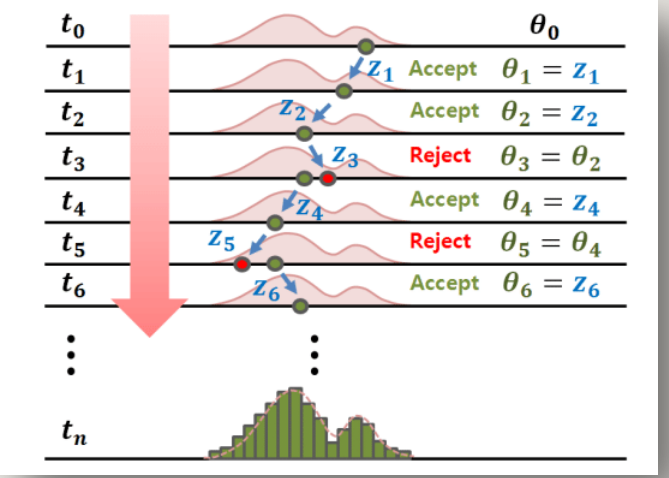
01 Set the initial configuration and force field

02 Propose a move

03 Accept the move with

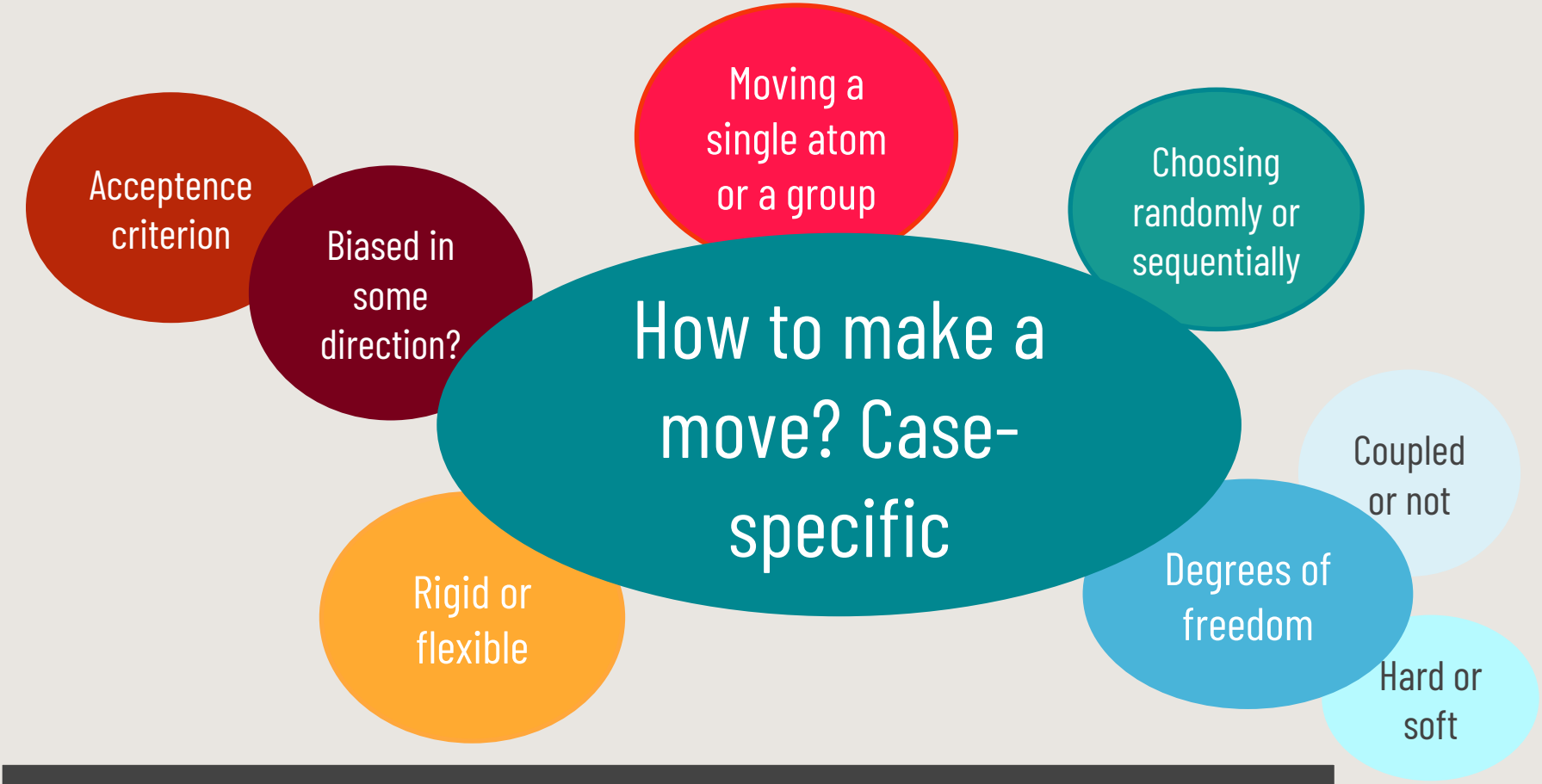
$$p=\min\left(1, e^{-\frac{\Delta E}{kT}}\right)$$

04 Repeat 2 and 3



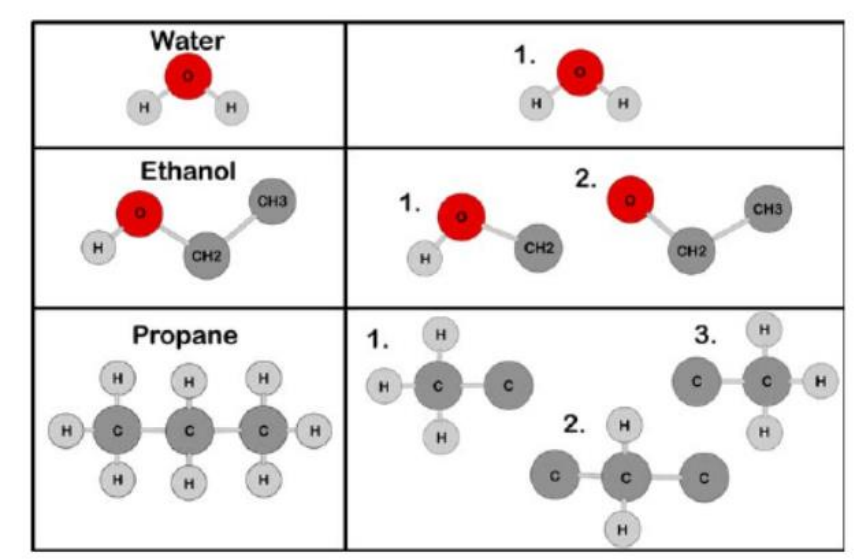

DICE: A Monte Carlo Code for Molecular Simulation Incldуing the Configurational Bias Monte Carlo Method, H. M. Cezar, S. Canuto, K. Coutinho, J.Chem. Inf. Model. 2020, 60, 3472-3488, Adaptive Markov chain Monte Carlo algorithms for Bayesian inference: recent advances and comparative study, Seung-Seop Jin, Heekun Ju, Hyung-Jo Jung

DICE: A Monte Carlo Code for Molecular Simulation Inclduing the Configurational Bias Monte Carlo Method, H. M. Cezar, S. Canuto, K. Coutinho, J.Chem. Inf. Model. 2020, 60, 3472-3488




Balance between displacement and acceptence rate ~50%

# Simulation




**New configuration generation:**

**First part: Standard Metropolis**

- Molecule translation
- Molecule rotation
- Bond angle perturbation
- Dihedral angle perturbation

**Second part (not obligatory): Configuration Biased**

Split the molecule according to the rotable parts bond, choose a new conformation for the part, join again

**Reject or Accept**

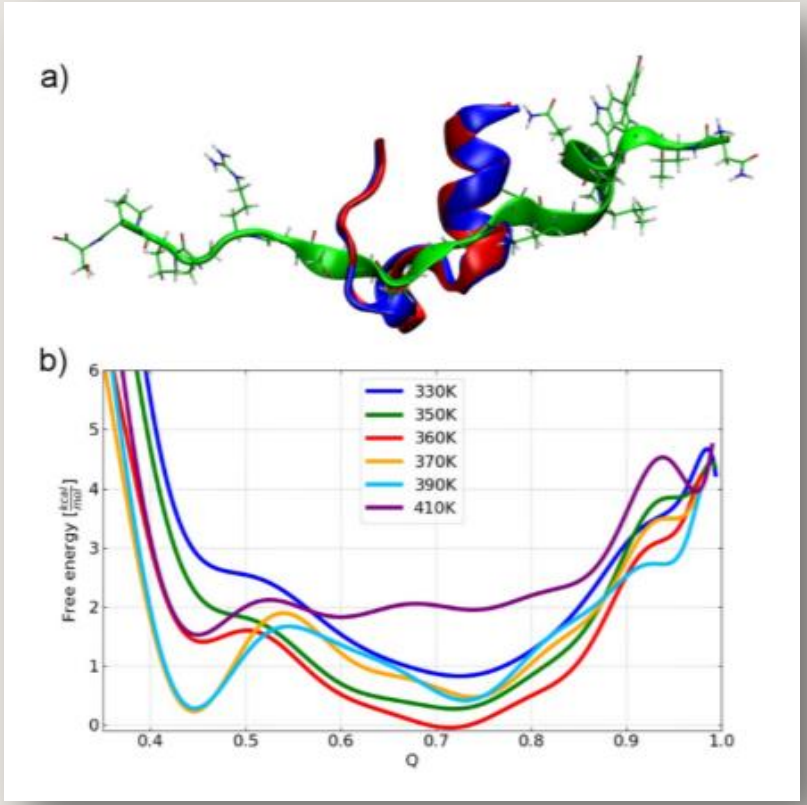
Cassandra: An Open Source Monte Carlo Package for Molecular Simulation, Jindal K. Shah, Eliseo Marin-Rimoldi, Ryan Gotchy Mullen, Brian P. Keene, Sandip Khan, Andrew S. Paluch, Neeraj Rai, Lucienne L. Romanielo, Thomas W. Rosch, Brian Yoo,Edward J. Maginn, Journal of Computational Chemistry 2017, 38, 1727–1739

# Other things to consider

- Force Field
- Ensemble
- Solvent – implicit treatment
- Random number generator

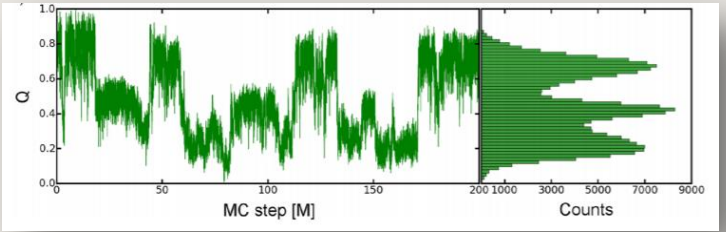
# Applications

# Protein Folding – Trp-cage protein




**Simulation features:**

- Implicit solvent
- Force field from MD
- 200 mln MC steps, 10% burn-in
- 2 times faster than MD (~180 h)
- Move: rather grouped, mainly rotations
- Metropolis criterion

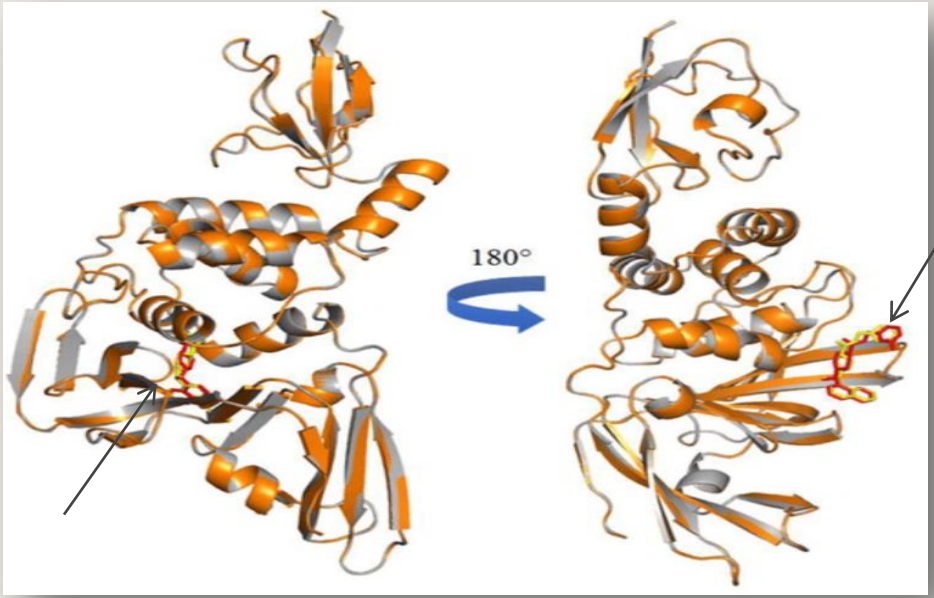



Other protein from this paper

Sampling of the conformational landscape of small proteins with Monte Carlo methods, Nana Heilmann, MoritzWolf, Mariana Kozlowska, Elaheh Sedghamiz, Julia Setzler, Martin Brieg & WolfgangWenzel, Scientifc Reports | (2020) 10:18211

# Simulated Annealing and Docking

**Simulated Annealing:** Global min search, go from high to low T in Metropolis Algorithm
**Application:** docking, computational drug screening




Chemical-informatics approach to COVID-19 drug discovery: Monte Carlo based QSAR, virtual screening and molecular docking study of some in-house molecules as papain-like protease (PLpro) inhibitors, Sk. Abdul Amin,Kalyan Ghosh,**Shovanlal Gayen** &**Tarun Jha,** https://doi.org/10.1080/07391102.2020.1780946, 2020, Journal of Biomolecular Structure and Dynamics

# Pros and Cons and Summary

+ more capable of crossing energy barrier
+ more efficient
+ numerically stable
+ no need to calculate a derivative of energy

Hybrid MC-MD

- no information about the dynamics
- no explicit solvent treatment
- little software
- need of designing MC-specific force fields
- moves case-specific

Methods for Monte Carlo simulations of biomacromolecules, Andreas Vitalis, Rohit V.Pappu, Annu. Rep. Comp.Chem. 201-, 1; 5:49-76

# Thank you!