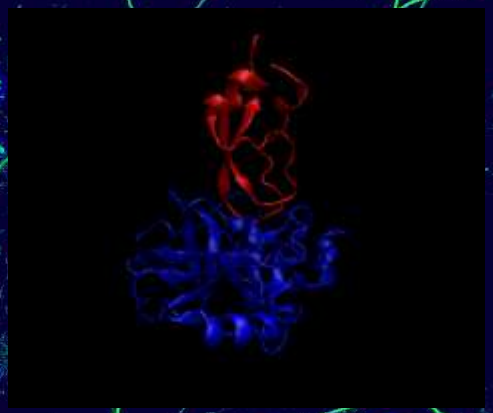
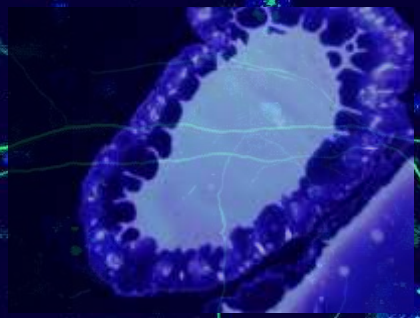
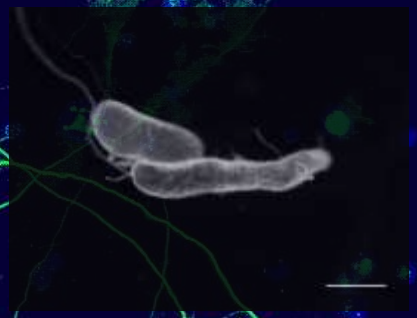
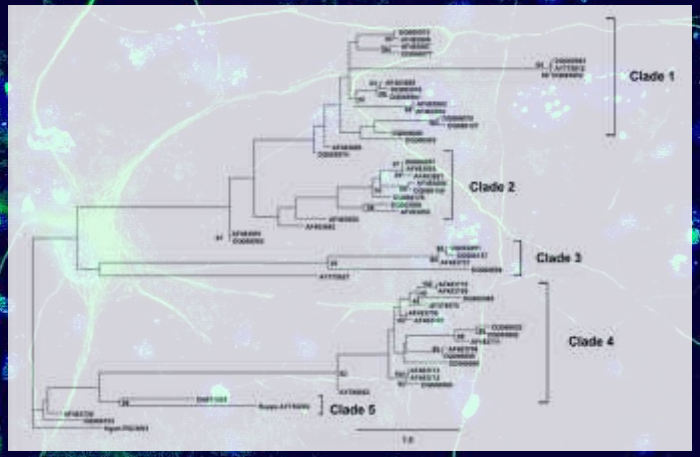
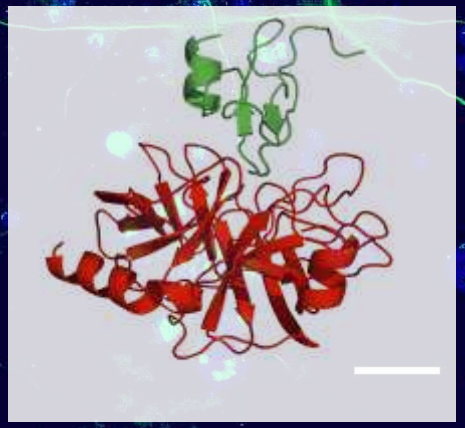






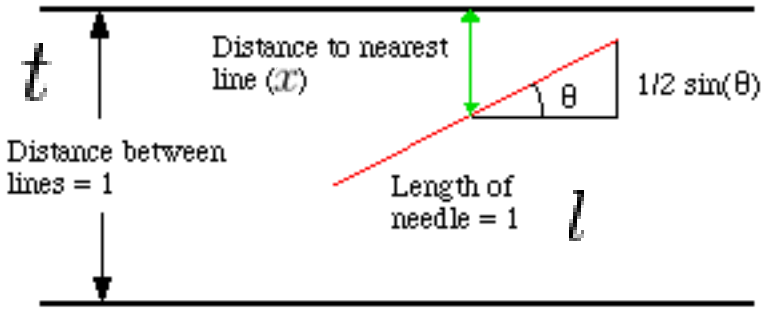

Clade 1
Clade 2
Clade 3
Clade 4
Clade 5

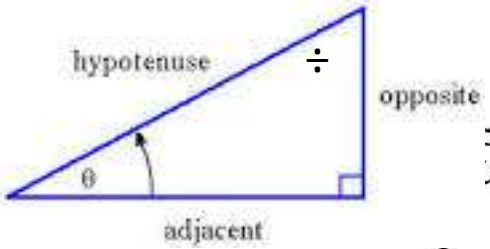


Molecular Dynamics
&
PELE

What is the Probability the Needle Crosses a Line?

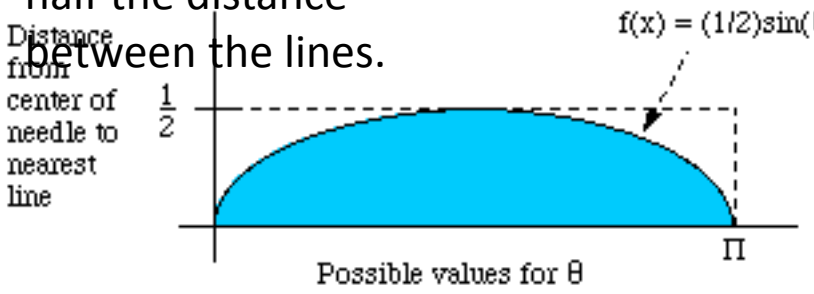
- The 1st Monte Carlo Experiment





The distance from the center to the closest line can never be more that half the distance between the lines.

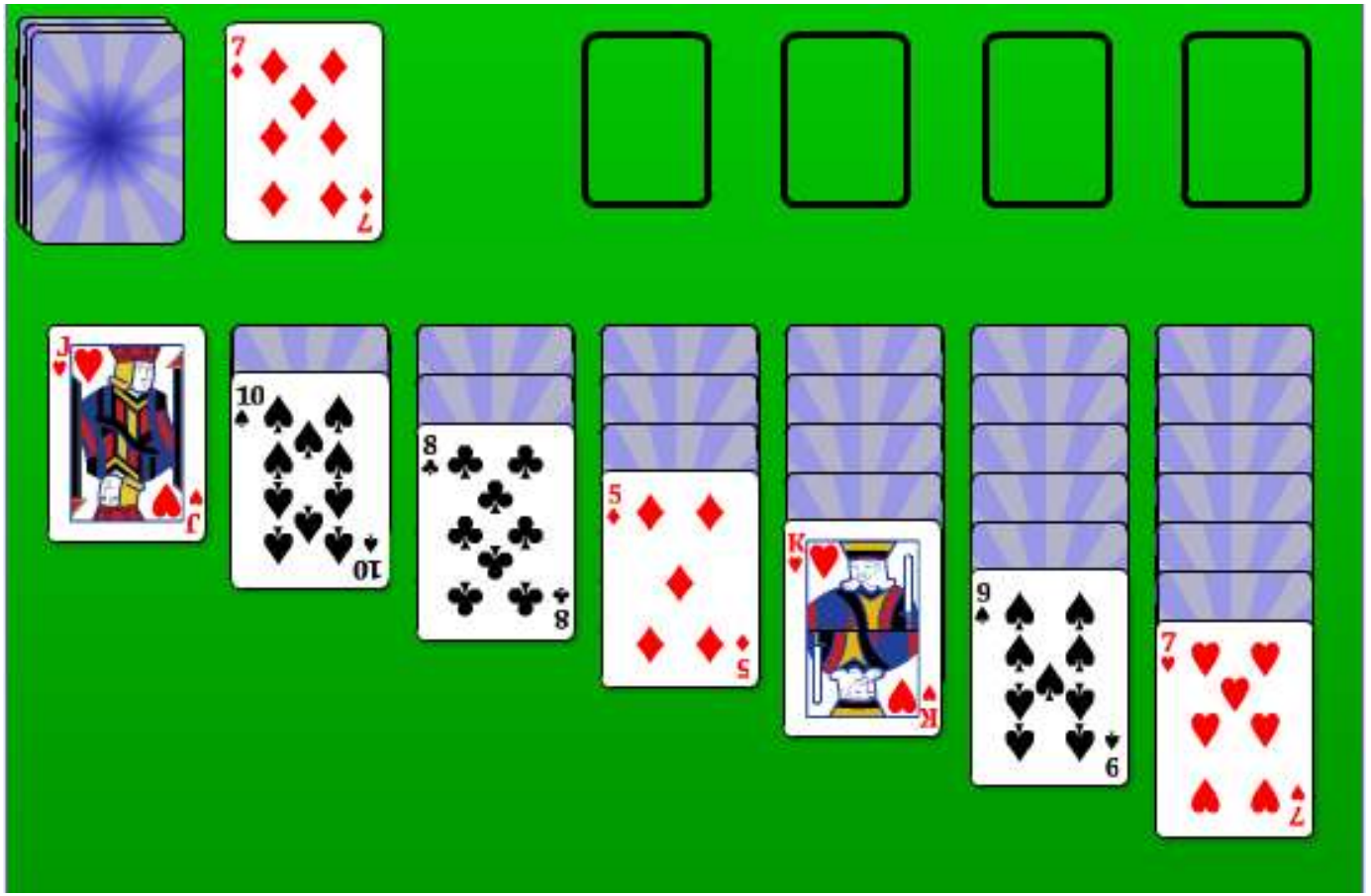
$$x \leq \frac{l}{2} \sin\theta.$$



2(n)/th ≈ π

Georges-Louis Leclerc, Comte de Buffon **(1777)**

Combinatorial Analysis
&
Theory of Probabilities



Metropolis Monte Carlo

1. Initial atom coordinates.
2. Random atom i moves by random displacement: ΔX_i, ΔY_i, and ΔZ_i
3. Calculate the change of potential energy ΔV corresponding to this displacement.
4. If $\Delta V < 0$ accept the new coordinates and go to step 2.
5. Otherwise, if $\Delta V \geq 0$, select a random number R in the range [0,1] and:
   A. If $e^{-\Delta V/kT} < R$ accept the new coordinates and go to step 2,
   B. If $e^{-\Delta V/kT} \geq R$ keep the original coordinates and go to step 2.

Phase Transition for a Hard Sphere System

B. J. Alder and T. E. Wainwright
University of California Radiation Laboratory, Livermore, California
(Received August 12, 1957)
J. Chem. Phys. 27, 1208

wall collis

$t = 0$

$t =$

$t = 4.04$

$t =$

03

.37

BPTI

Nature Vol. 267 16 June 1977 585

articles

Dynamics of folded proteins

J. Andrew McCammon, Bruce R. Gelin & Martin Karplus

Department of Chemistry, Harvard University, Cambridge, Massachusetts 02138

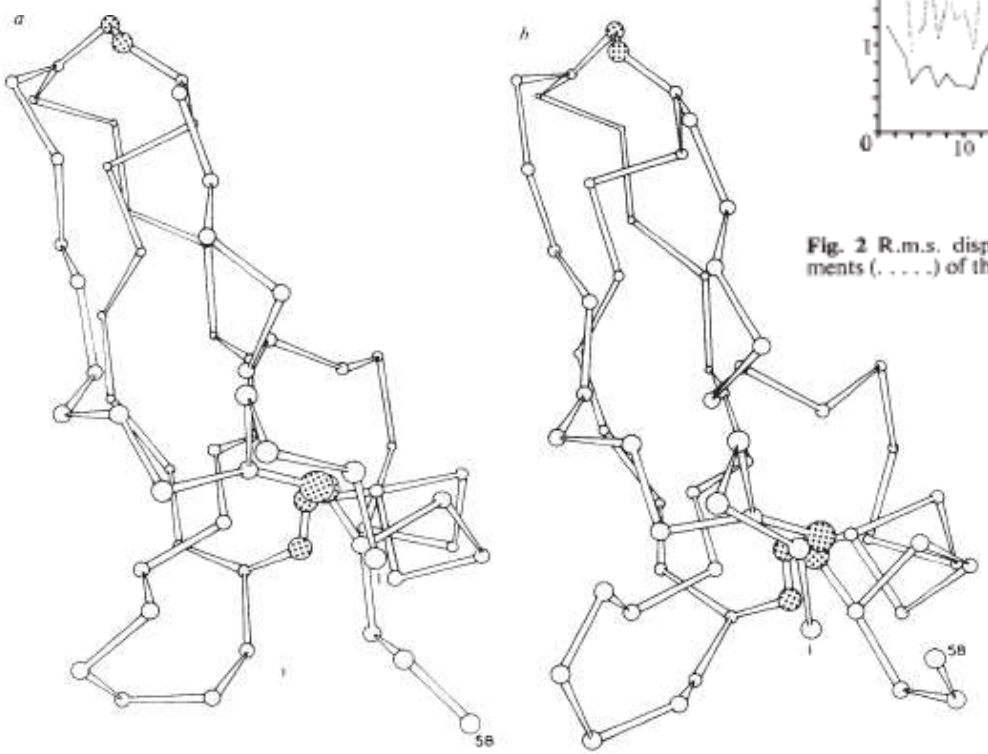


Fig. 1 The peptide backbone (α carbons) and disulphide bonds of PTI. *a*, X-ray structure[23]. *b*, Time evolved structure after 3.2 ps of dynamical simulation.

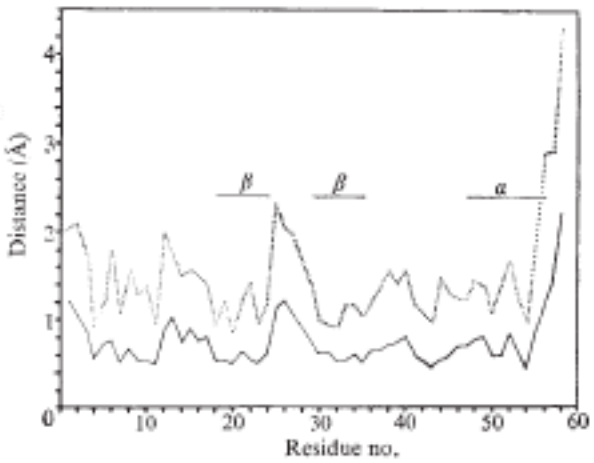


Fig. 2 R.m.s. displacements (———) and maximum displacements (. . . . .) of the α carbons, relative to the dynamical average structure.

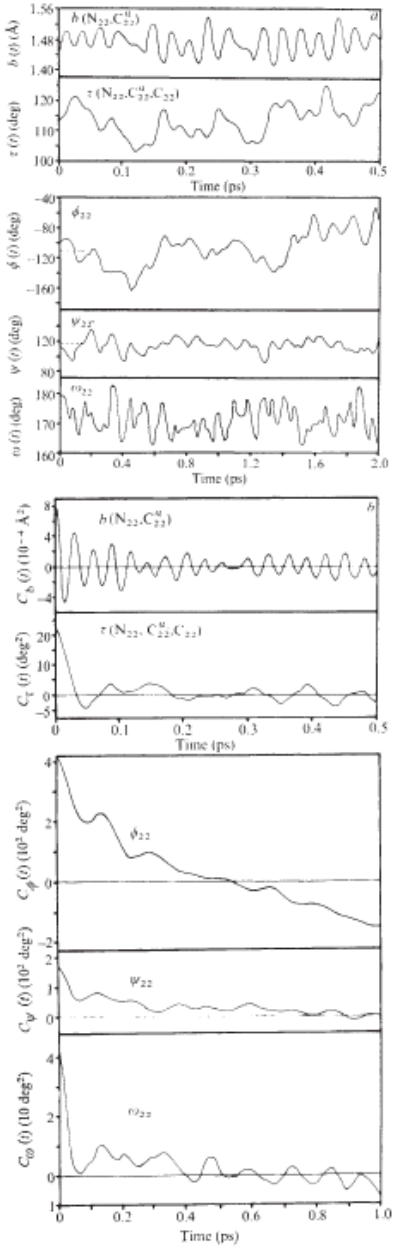
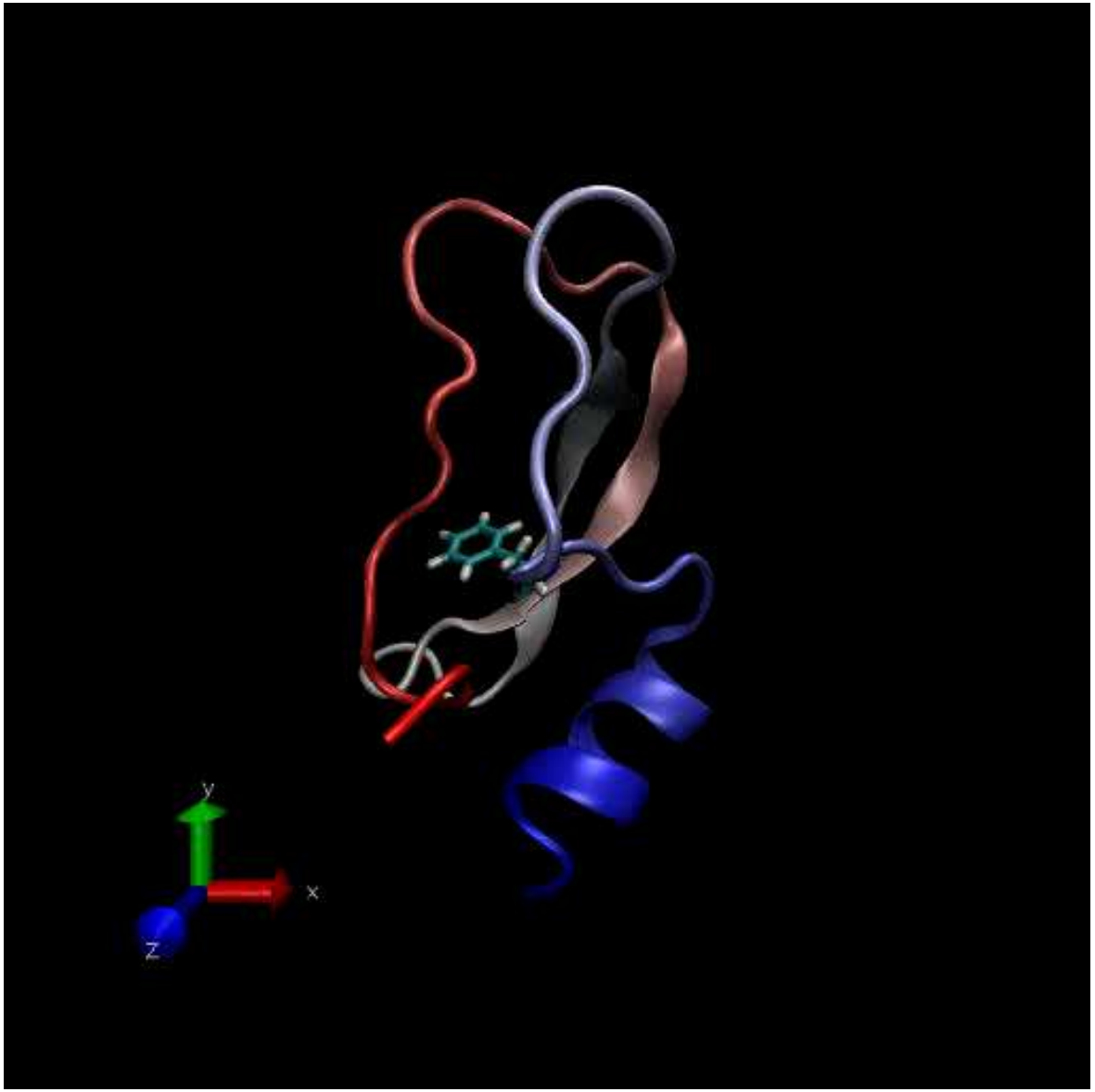


Fig. 3 *a*, Time development of selected Phe 22 internal coordinates; the initial time is 4.9 ps after the beginning of the dynamical simulation; average values are indicated by broken line segments. *b*, Time correlation functions for fluctuations of selected Phe 22 internal coordinates.

NMR of BPTI



Molecular dynamics simulations must pass through all possible states corresponding to the particular thermodynamic constraints

Molecular Dynamics

Statistical Mechanics

&

Classical Mechanics

Molecular Dynamics

- A computational method that calculates the time dependent behavior of a molecular system

- Provides detailed information on the fluctuations and conformational changes of ligands, proteins and nucleic acids

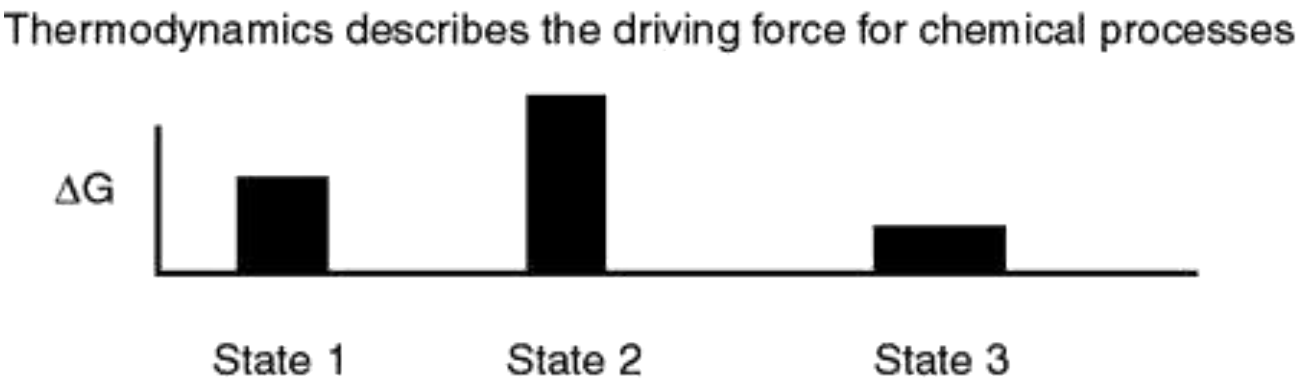
Definitions

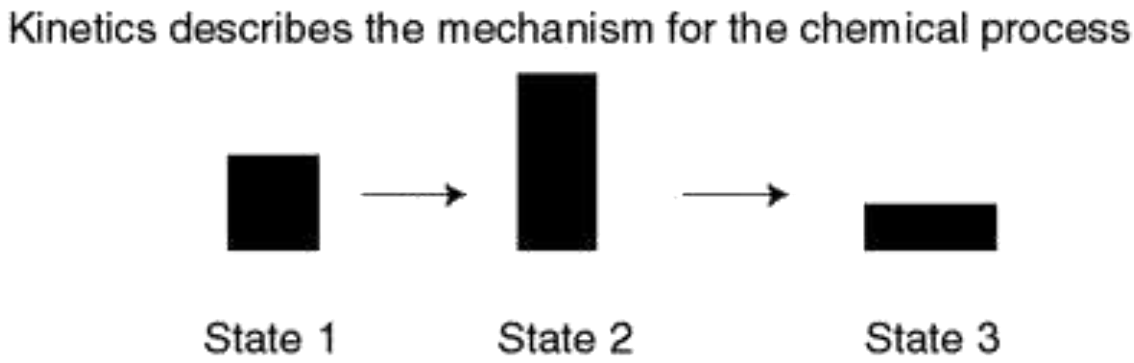
- *Thermodynamic state* is usually defined by the temperature (**T**), the pressure (**P**), and the number of particles (**N**).

- The *mechanical* or *microscopic state* is defined by the atomic positions, (**r**), and momenta, (**p**); these can also be considered as coordinates in a multidimensional space called **phase space**.

Molecular Dynamics

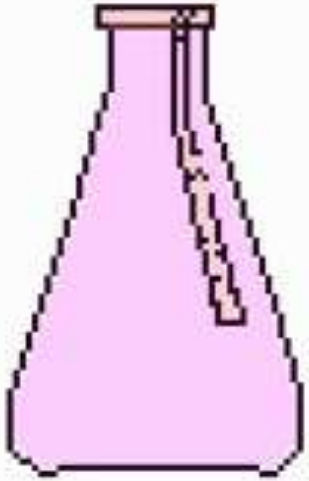
- An ensemble is a collection of all possible systems which have different microscopic states but have an identical macroscopic or thermodynamic state.





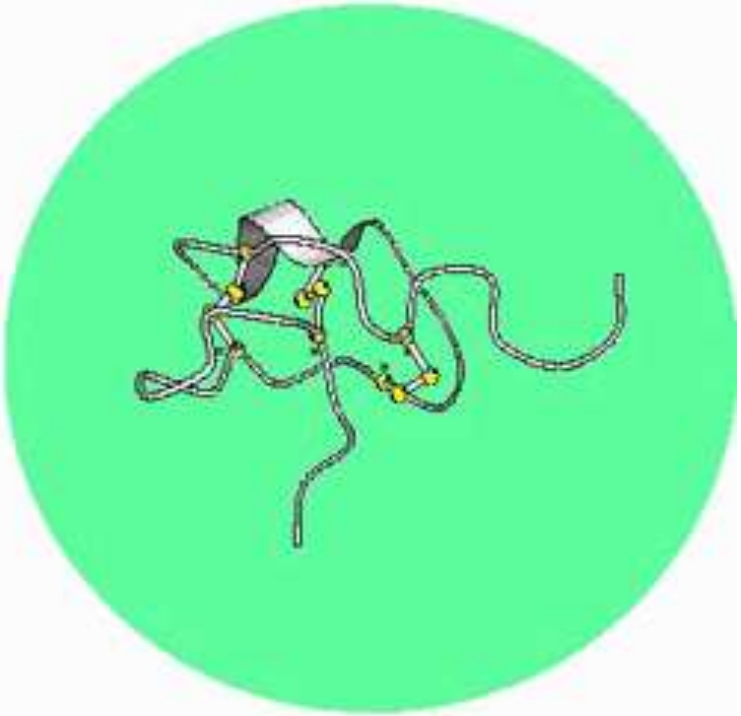
Molecular Dynamics

Experiment



Macroscopic

Molecular Simulation



Microscopic

Molecular Dynamics

- Statistical mechanics

  Ensemble average = Time average

$$\langle A \rangle_{ensemble} = \iint dp^N dr^N A\left(p^N, r^N\right) \rho\left(p^N, r^N\right)$$

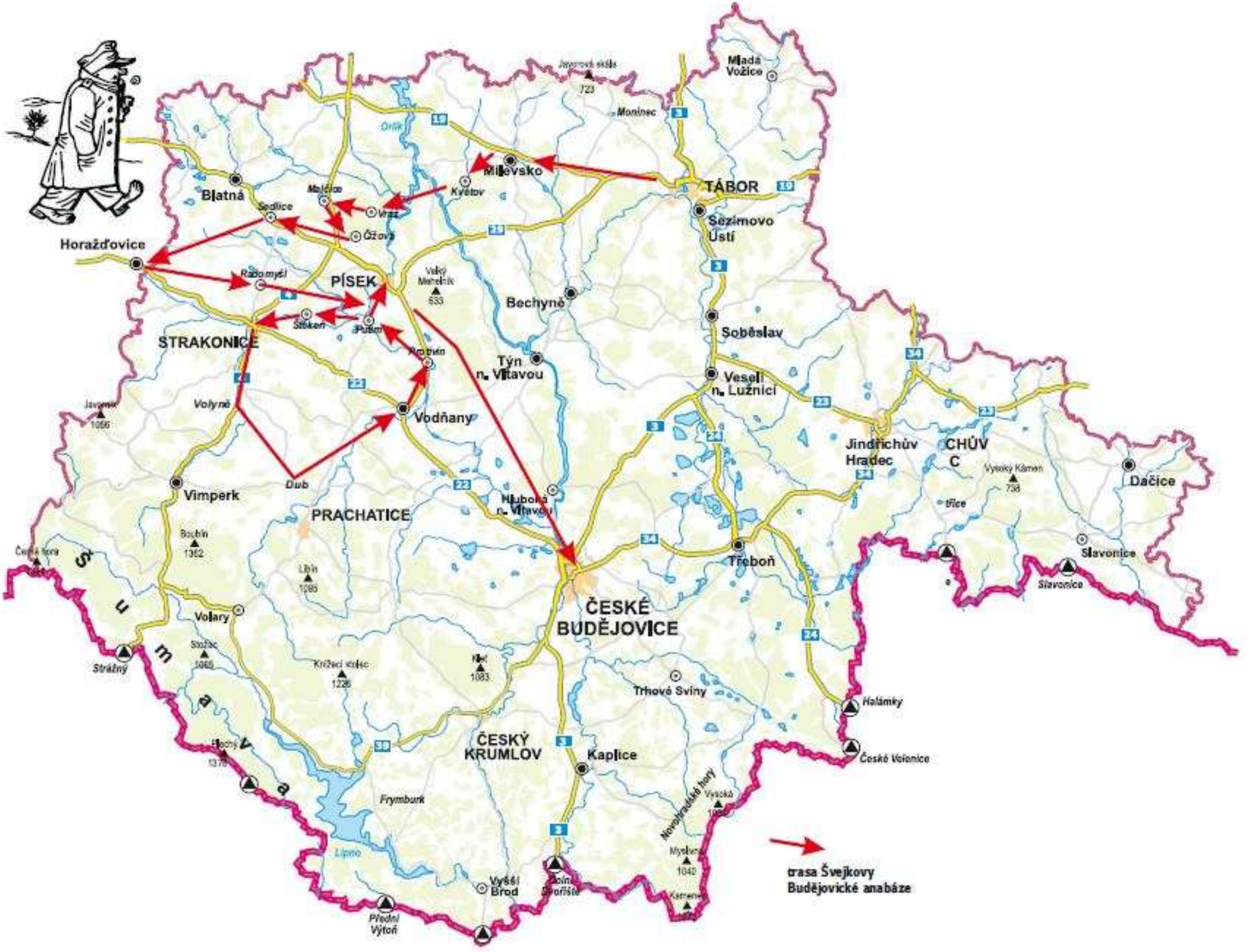
$$\rho\left(p^N, r^N\right) = \frac{1}{Q} \exp\left[-H\left(p^N, r^N\right) / k_B T\right]$$

$$\langle A \rangle_{time} = \lim_{\tau \to \infty} \frac{1}{\tau} \int_{t=0}^{\tau} A\left(p^N(t), r^N(t)\right) dt \approx \frac{1}{M} \sum_{t=1}^{M} A\left(p^N, r^N\right) = \frac{R}{N_A}.$$

Th
pc
momenta (p) and the positions

Partition function the thermodynamic state variables (temperature and volume)

The average simulation time (t) is the number of time steps (M) in the simulation and the instantaneous value of A ($A(p^N,r^N)$)



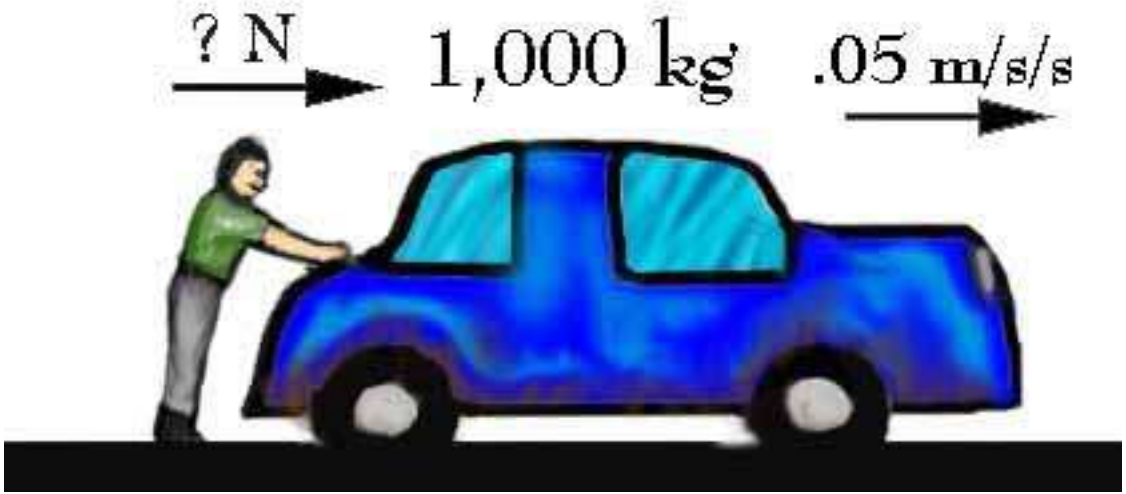
TÁBOR
Milevsko
Květov
Vráz
Malčice
Čížová
Sedlice
Blatná
Horažďovice
Radomyšl
PÍSEK
Putim
Štěkeň
STRAKONICE
Protivín
Vodňany
Volyně
Dub
Vimperk
PRACHATICE
Bechyně
Týn n. Vltavou
Hluboká n. Vltavou
ČESKÉ BUDĚJOVICE
Mladá Vožice
Monínec
Orlík
Sezimovo Ústí
Soběslav
Veselí n. Lužnicí
Jindřichův Hradec
Dačice
Slavonice
Třeboň
Trhové Sviny
Kaplice
ČESKÝ KRUMLOV
Volary
Frymburk
Lipno
Vyšší Brod
Strážný
Halámky
České Velenice
Přední Výtoň
Šumava
trasa Švejkovy Budějovické anabáze

Molecular Dynamics

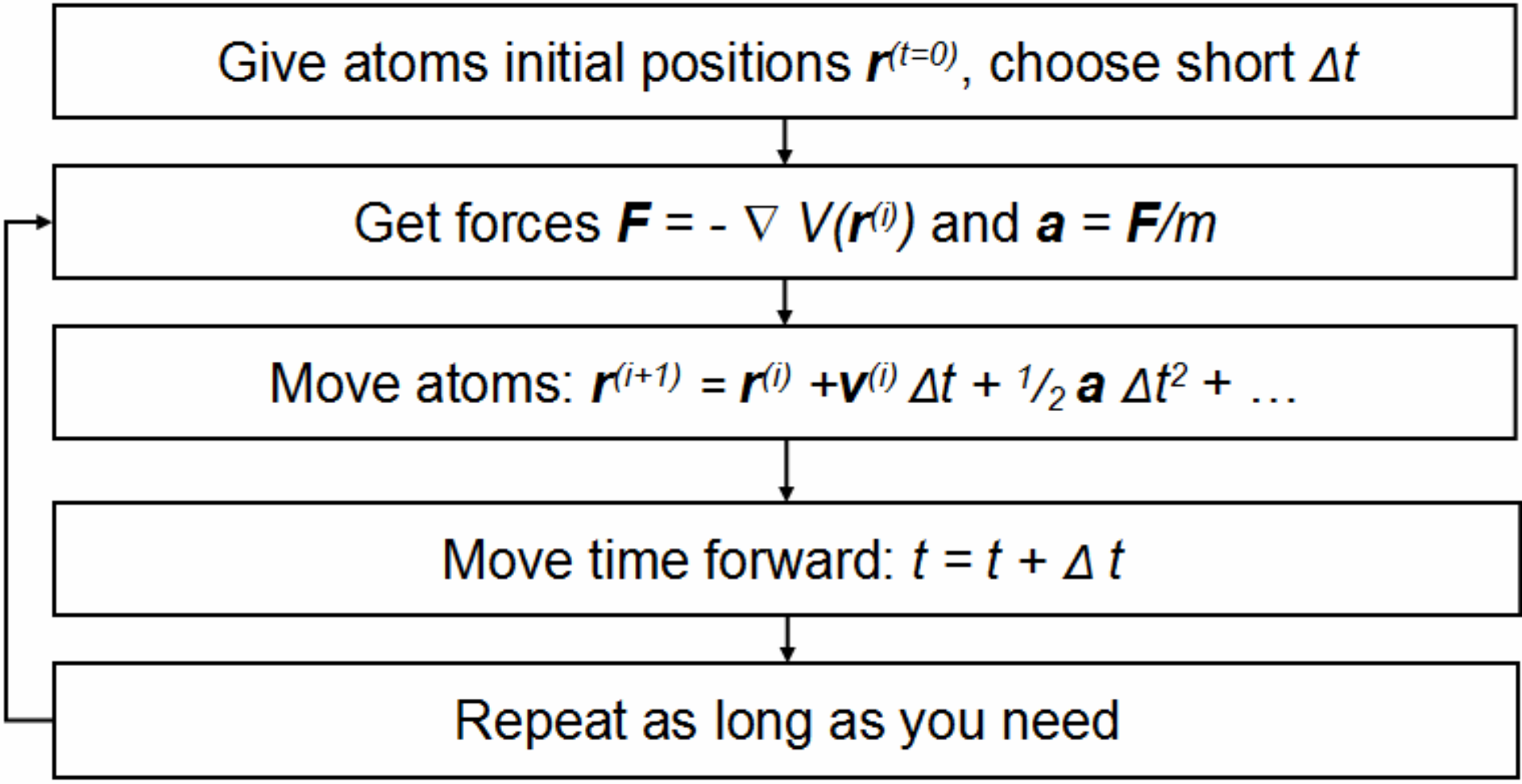
- **Ergodic hypothesis =** over a long period of time, the time spent by a system in some region of the phase space of microstates with the same energy is proportional to the volume of this region

Molecular Dynamics

- Classical mechanics
  - Newton's 2nd Law of Thermodynamics
  - *F=ma* (the acceleration of an object is dependent on the net force acting upon the object and the mass of the object)



Molecular Dynamics



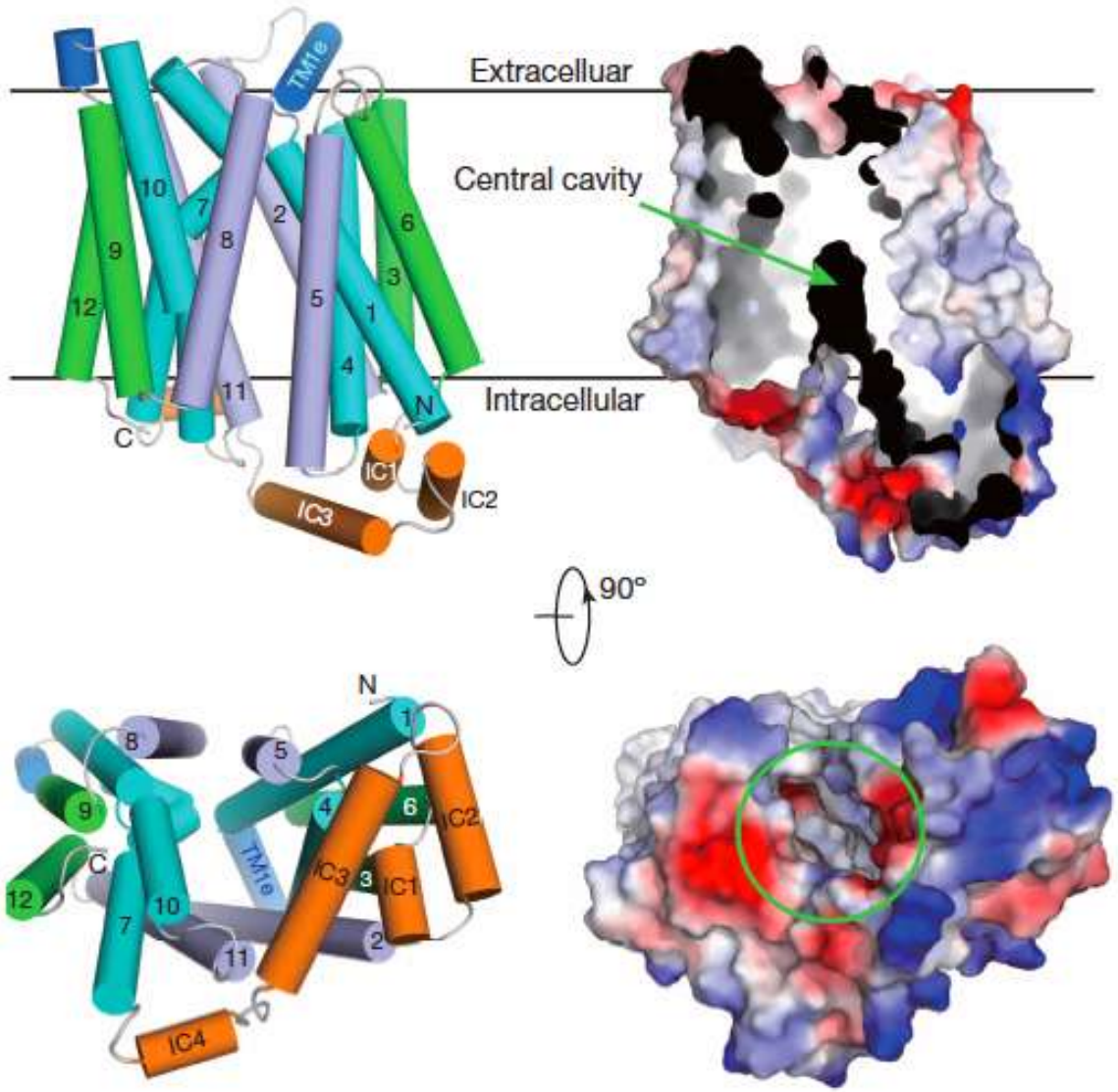
Application

Glucose Transporter System

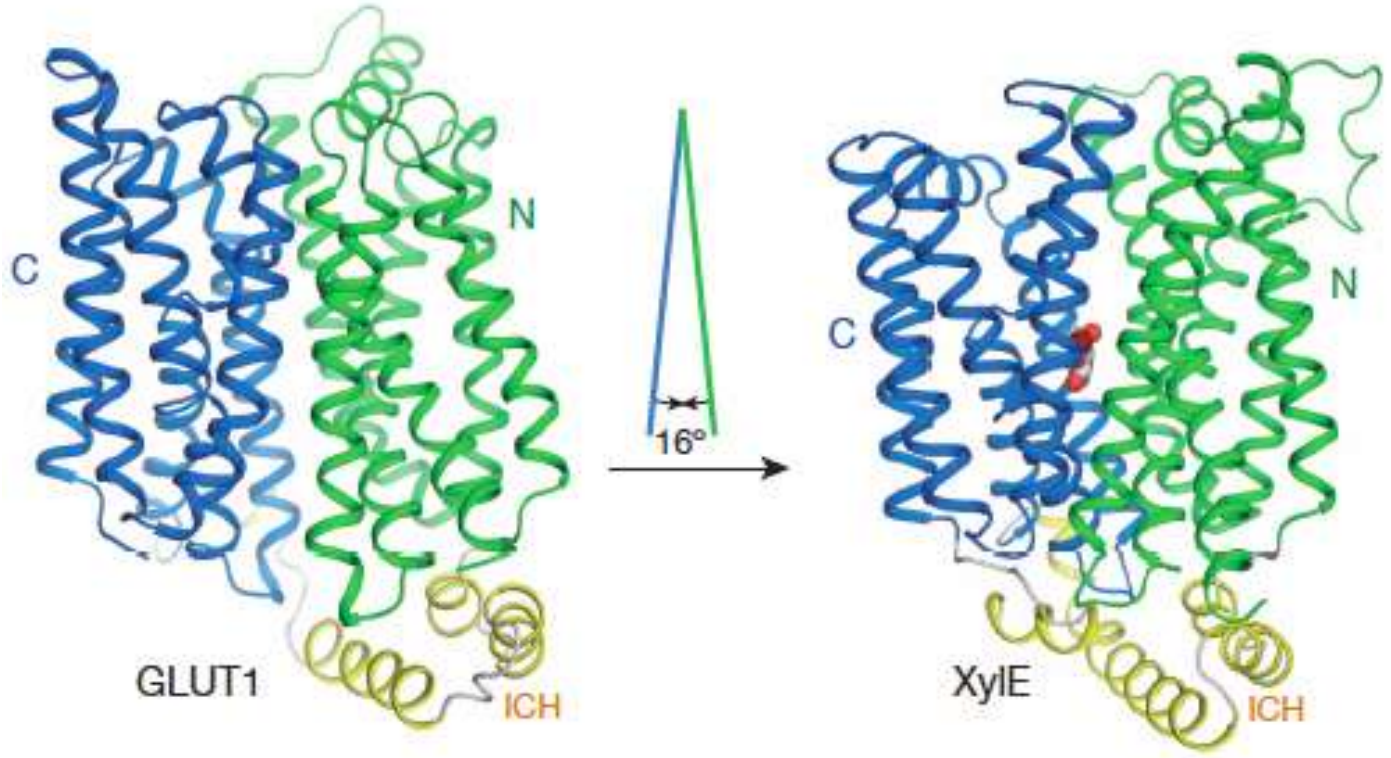
Crystal structure of the human glucose transporter GLUT1

2014 | VOL 510 | NATURE | 121

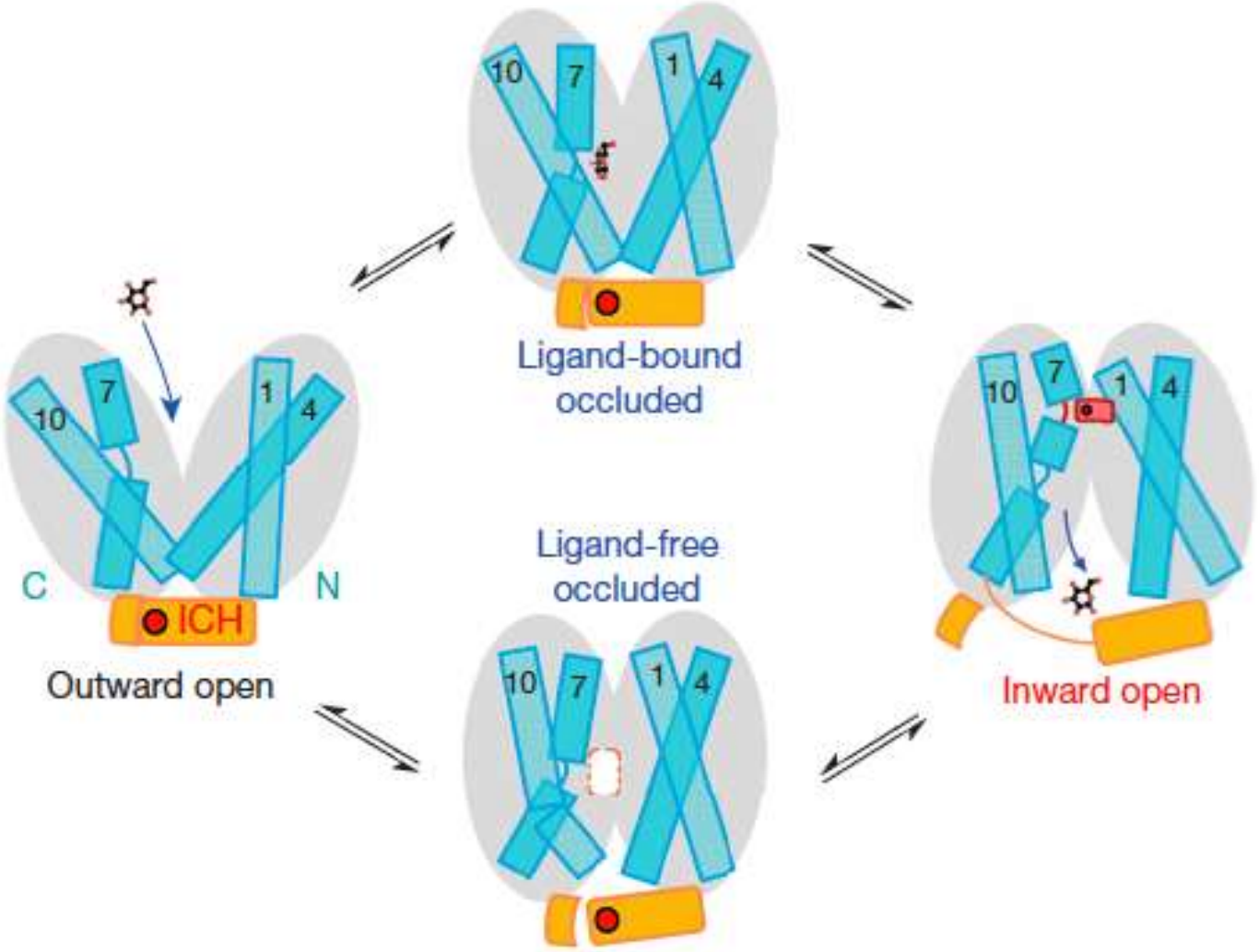
Dong Deng[1,2,3]*, Chao Xu[1,2,3]*, Pengcheng Sun[1,2]*, Jianping Wu[1,2,3]*, Chuangye Yan[1,2], Mingxu Hu[1,2,3] & Nieng Yan[1,2,3]



Structural Conformations



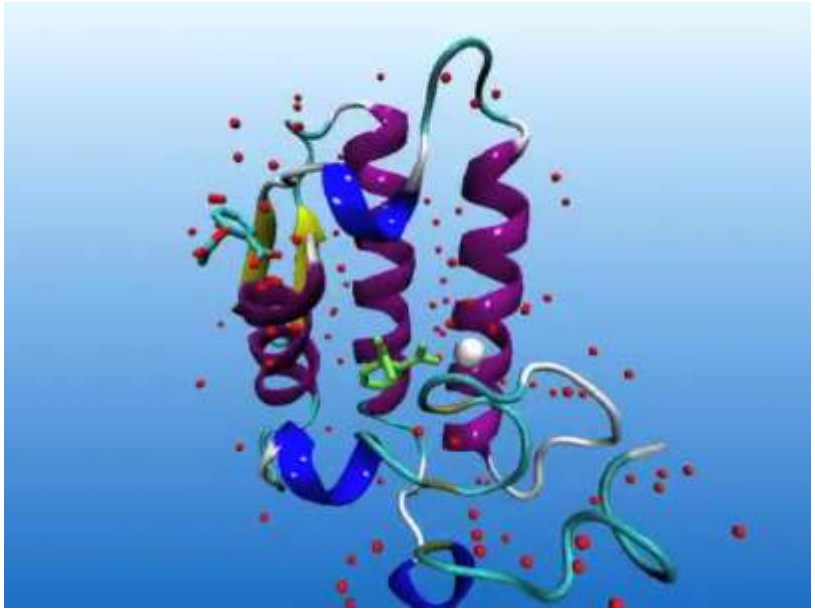
Structural Conformations



PELE

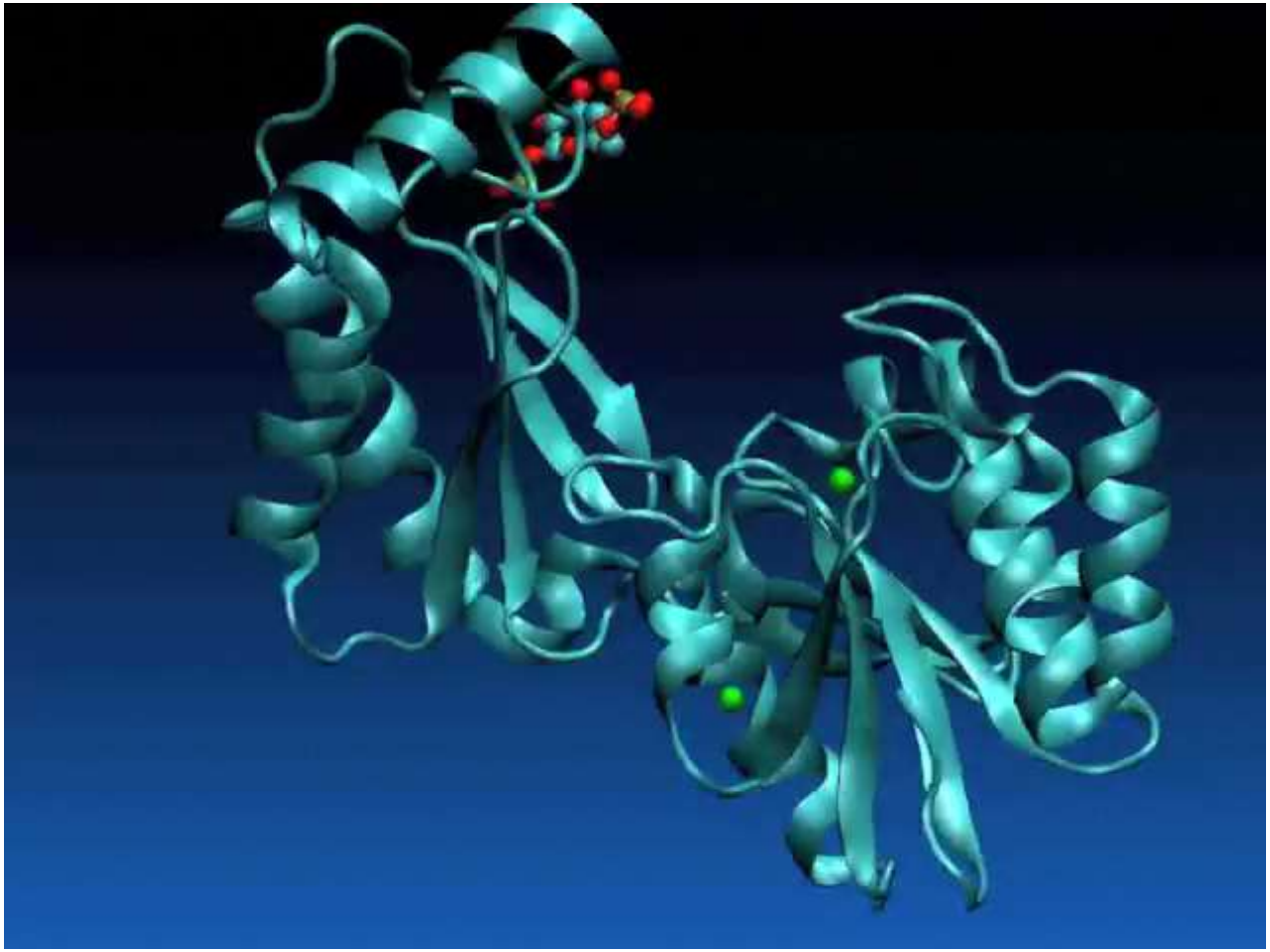
A Metropolis Monte-Carlo Software

PELE (PROTEIN ENERGY LANDSCAPE EXPLORATION)

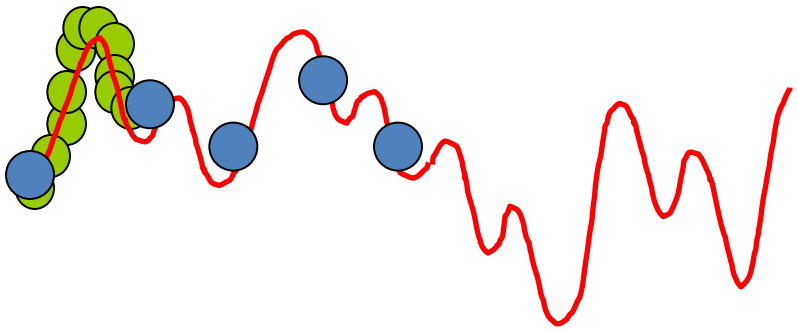


- Kenneth W. Borrelli, Andreas Vitalis, Raul Alcantara and Victor Guallar.Protein energy landscape exploration. A novel Monte Carlo technique. Implications on camphor ligand binding. *Journal of Chemical Theory and Computation,* **6**:1304-1311, (2005)
- Kenneth Borrelli, Benjamin Cossins, and Victor Guallar, Exploring hierarchical refinement techniques for induced fit docking with protein and ligand flexibility *J. Comp. Chem*., **31**: 1224-35 (2010)
- Benjamin Coussins, Ali Hosseini and Victor Guallar, Exploration of Protein Conformational Change with PELE and Meta-Dynamics, *Journal of Chemical Theory and Computation*, **8**:959-965 (2012)
- Fatima Lucas and Victor Guallar, An Atomistic View on Human Hemoglobin Carbon Monoxide Migration Processes, *Biophysical Journal*, **102**:887-896 (2012)
- Hosseini, A., Espona-Fiedler, M., Soto-Cerrato, V., Quesada, R., Pérez-Tomás, R. & Guallar, V.Molecular Interactions of Prodiginines with the BH3 Domain of Anti-Apoptotic Bcl-2 Family Members. *PloS one* **8**, e57562 (2013)
- Takahashi, R., Gil, V.A. and Guallar, V., Monte Carlo Free Ligand Diffusion with Markov Sate Model analysis and Absolute Binding Free Energy Calculations. *Journal of Chemical Theory and Computation*, **10:**282-288. (2014)

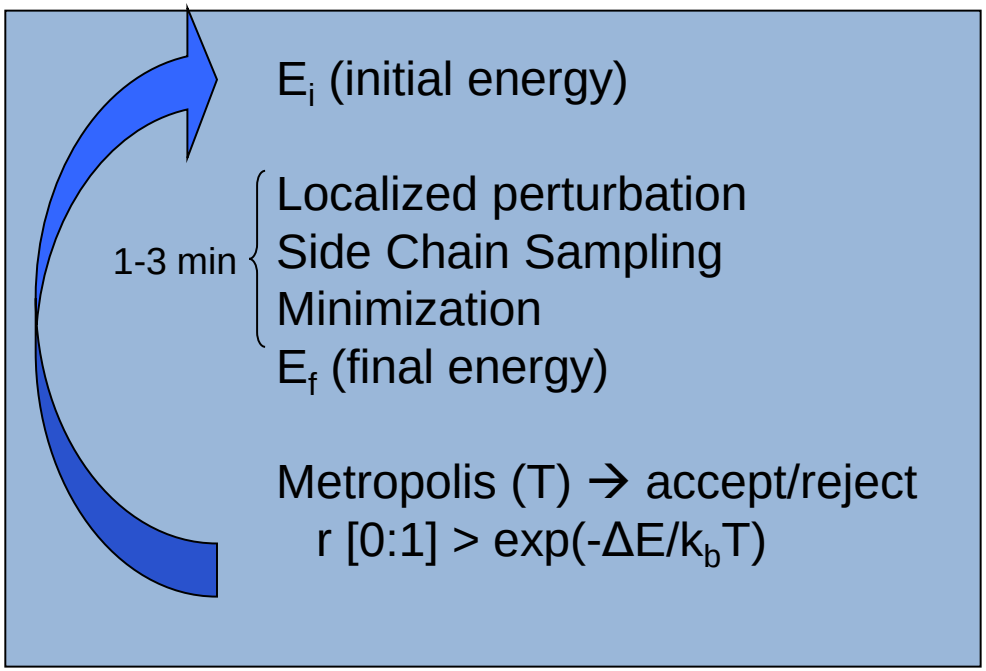
You collect what you input!



PELE: combining Monte Carlo with protein prediction algorithms

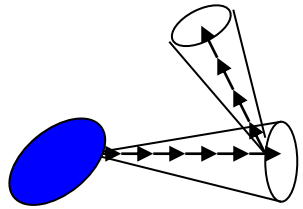


The algorithm (PELE):

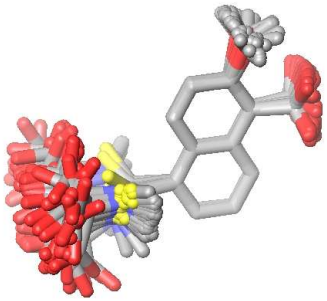


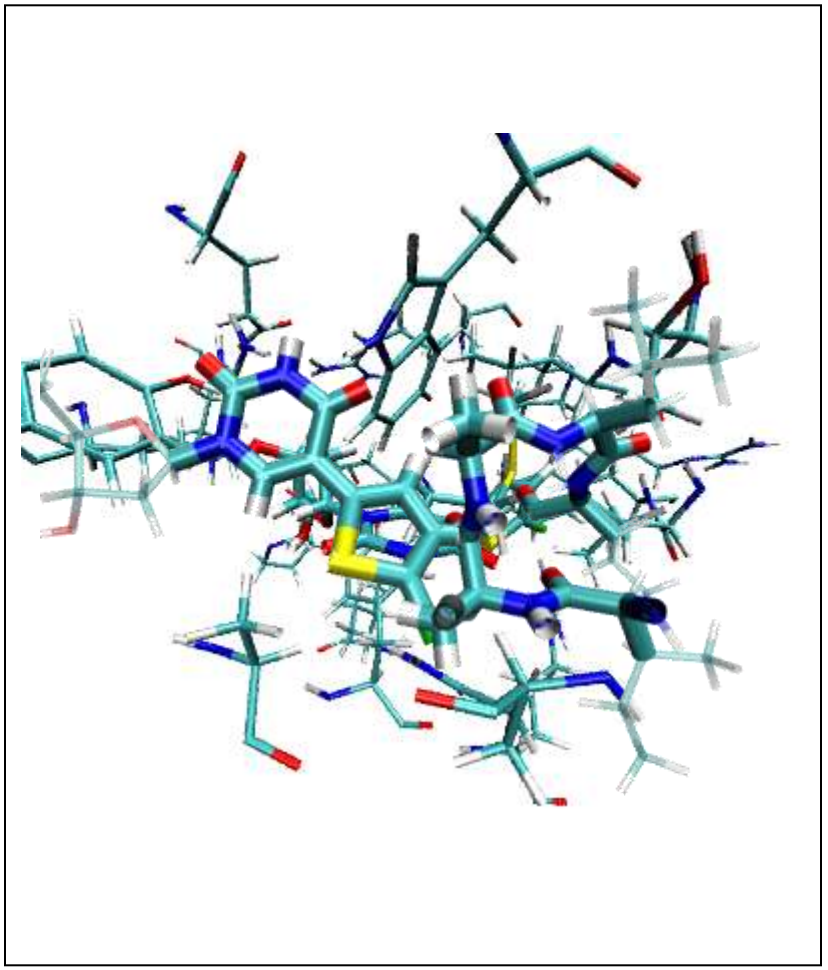
Borrelli KW, Vitalis A, Alcantara R, Guallar V. PELE: Protein Energy Landscape Exploration. A Novel Monte Carlo Based Technique. Journal of Chemical Theory and Computation. 1 (6): NOV 2005

Steered the perturbation



Ligand motion

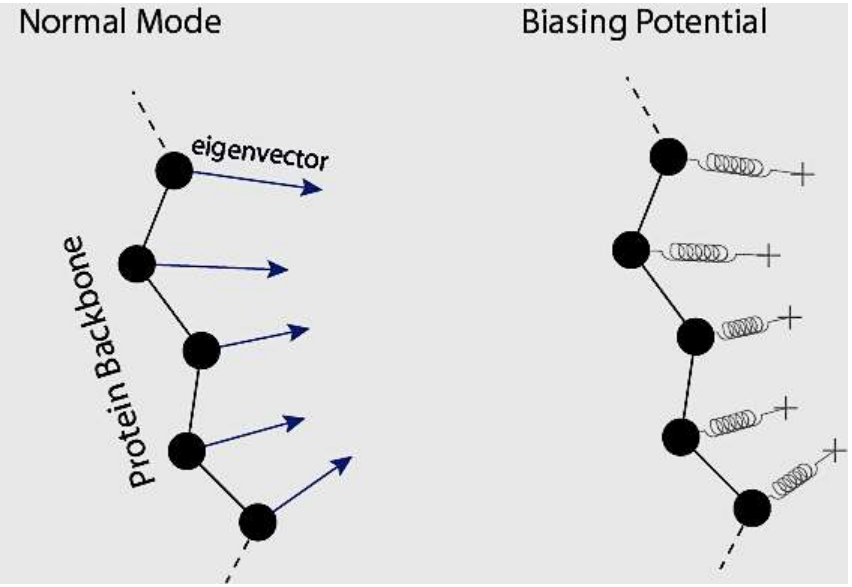




- Eliminate grid-points which always produce steric clashes (0.65% of VdW radius)
  - That create a steric clash between the ligand core and the protein sidechain
  - For which all the rotamers of a single residue or ligand region cause a steric clash with the backbone or ligand core
  - For which a single combination of rotamers cannot be found that does not produce a steric clash

Backbone Perturbation: Anisotropic Network Model
Follow a mode in an steered minimization

This ligand perturbation is coupled to backbone motion by means of a gaussian network of alpha carbon contacts



Side Chain Step (Prediction)

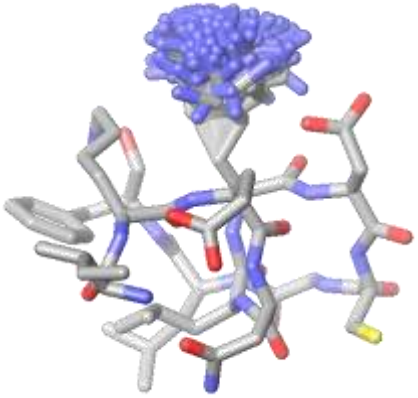
Rotamer Libraries (Barry Honig 2001) Define All Possible Residue Movements

Steric Filtering Removes Unfeasible Ones

Resulting Rotamers are Clustered

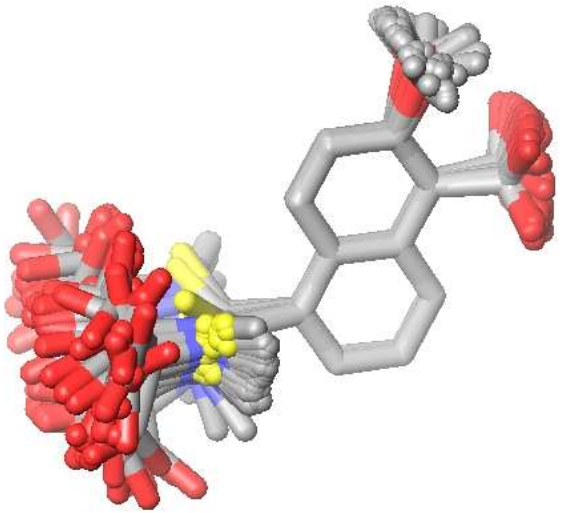
Clusters are Minimized

Lowest Energy Cluster is Selected



The Ligand also has rotamers associated!!

- Choosing a "core" region
  - The frozen region (no rotatable bonds within the region) that minimizes the number of bond lengths from this region to the farthest atom
- Populate the Library
  - Using Macromodel: eliminating all those that create a steric clash (two atoms closer than 75% of their combined VdW distance)



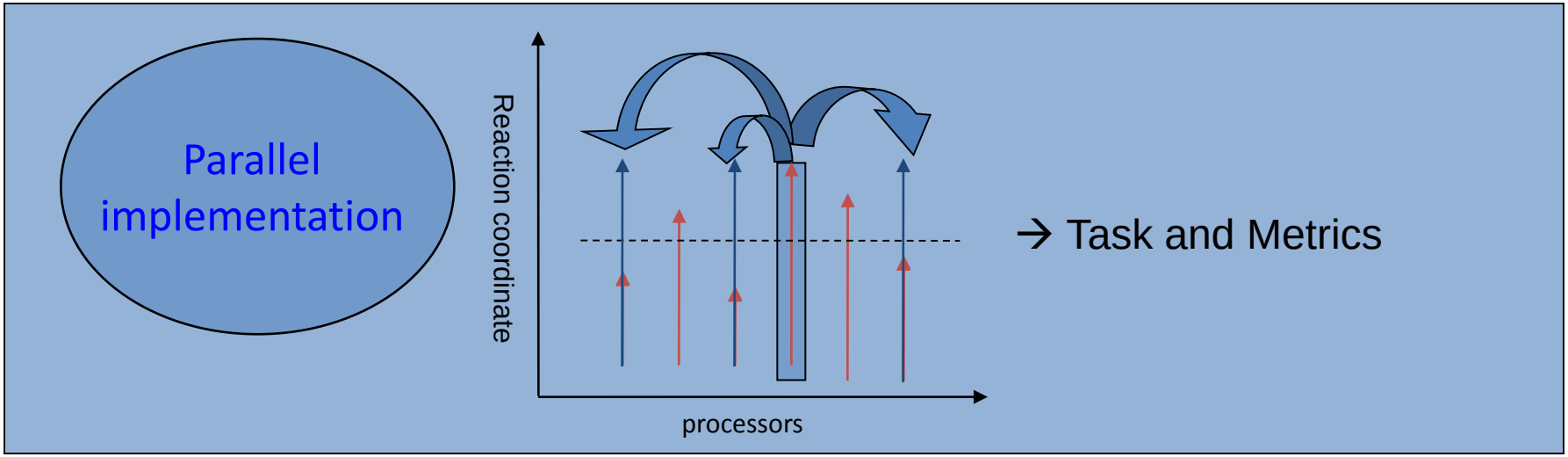
3 independent rotamers

Last step: minimization

Truncated Newton with long/short range force update and numerical born alpha derivatives

Alpha carbon can/should be constrained to the outcome of ANM

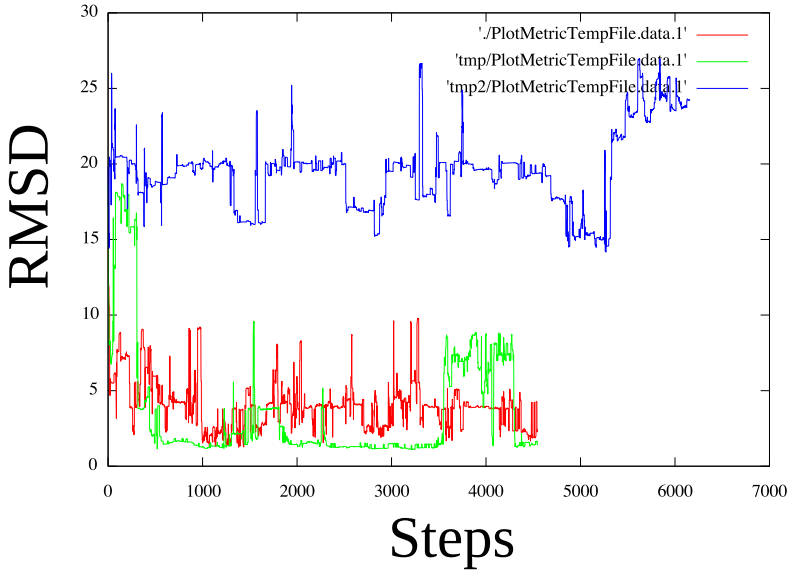
All rap up in MPI (parallel implementation: one node controlling and n-1 running)

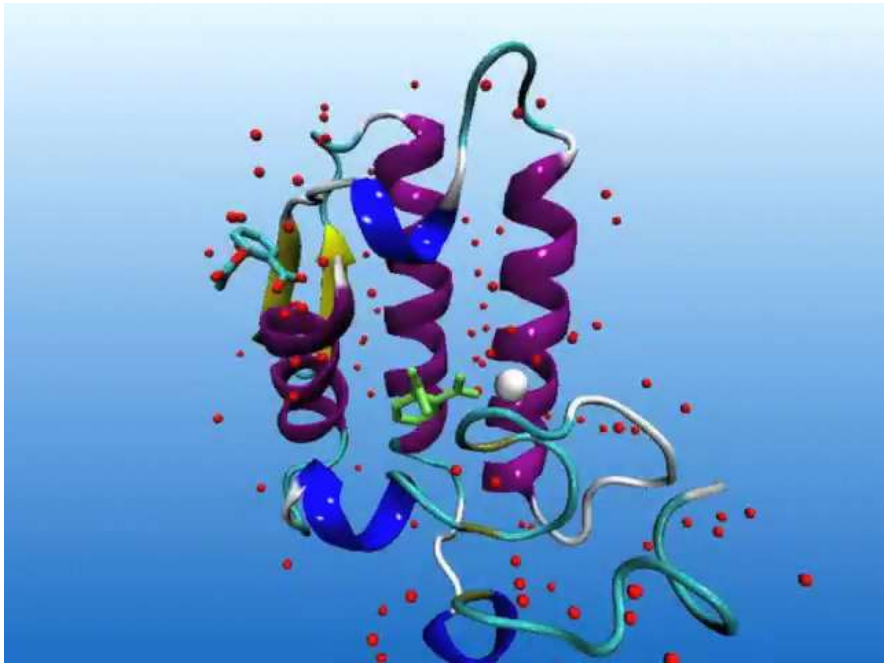


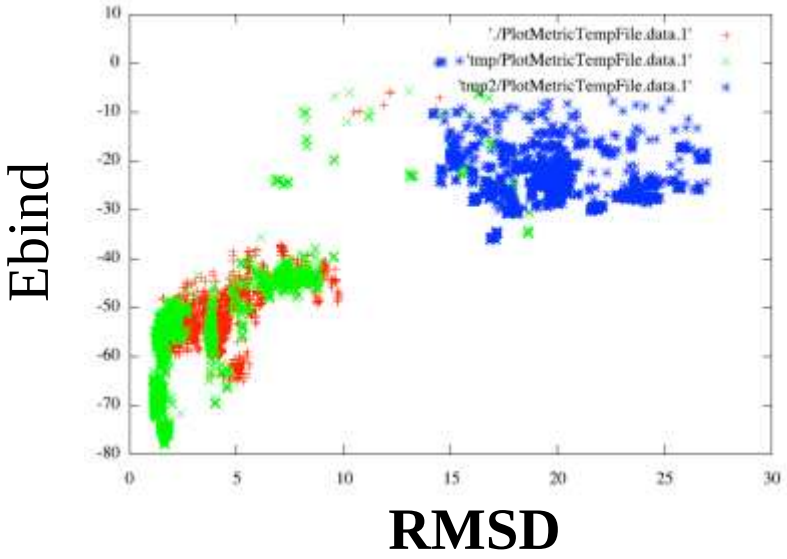
Tasks: lower energy, binding energy, move ligands away/to, move domains, change initial rmsd, etc.

Borrelli et al.. PELE: Protein Energy Landscape Exploration. A Novel Monte Carlo Based Technique. J Chem Theo Comp, 1:1304-11 (2005)

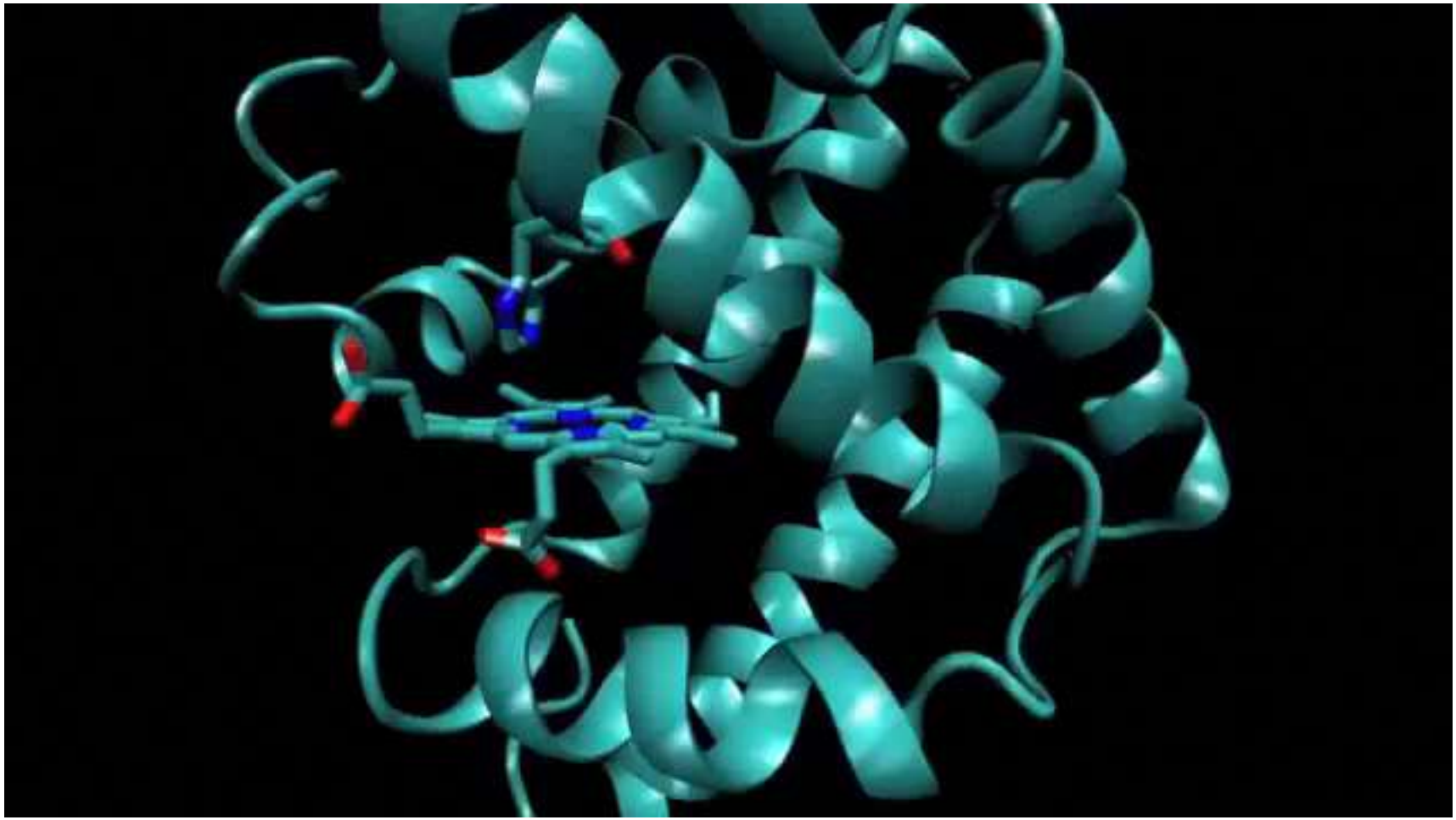
Free search with PELE



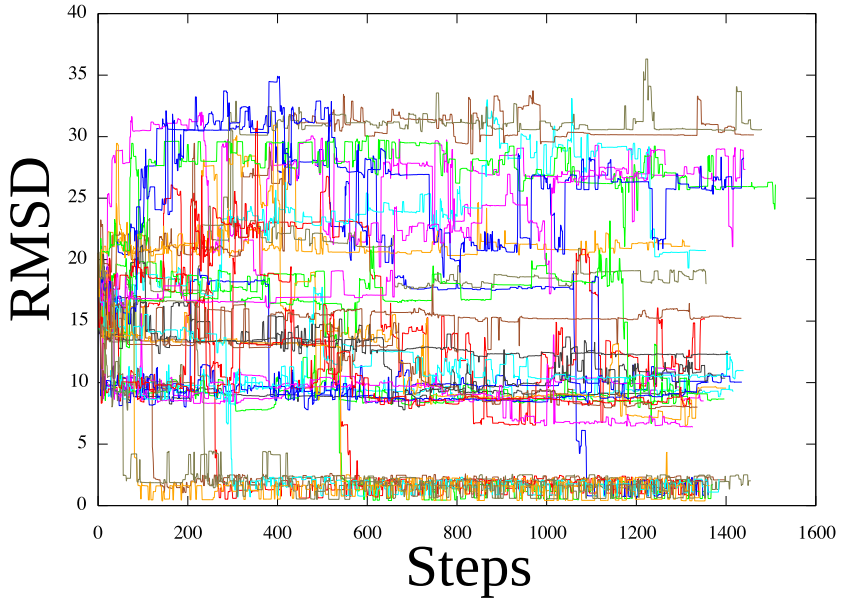


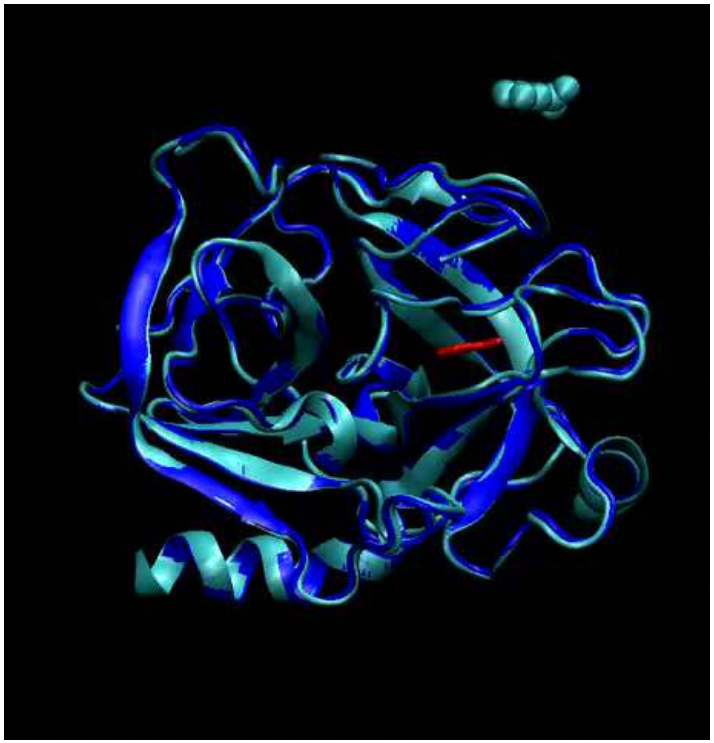


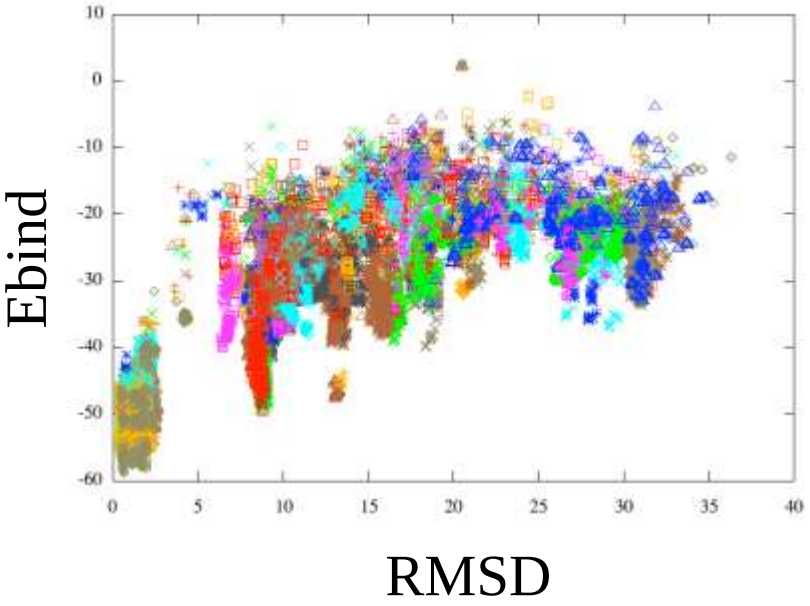
Want to see carbon monoxide getting inside Mb?



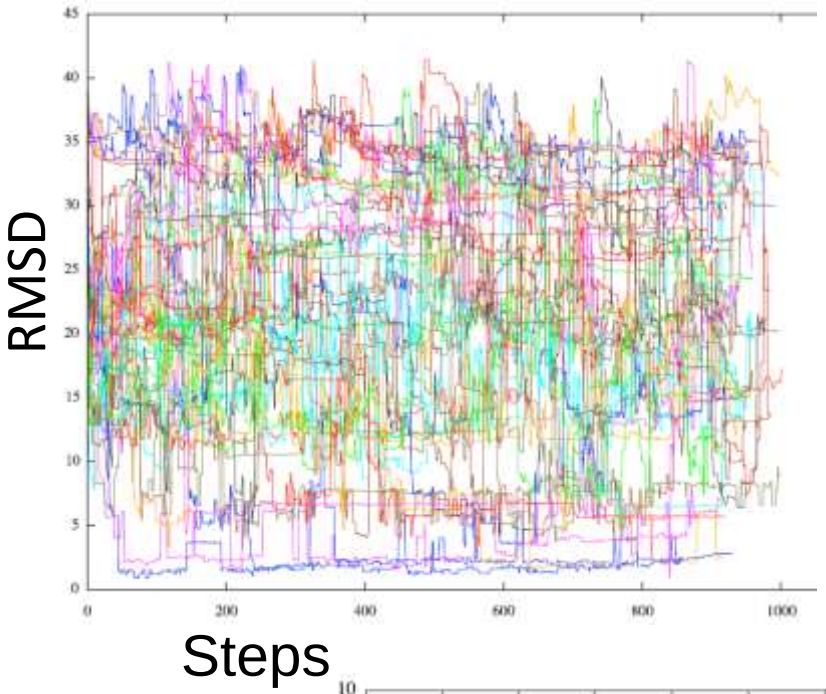
Full ligand random exploration: Benzamidine binding to the bovine trypsin.

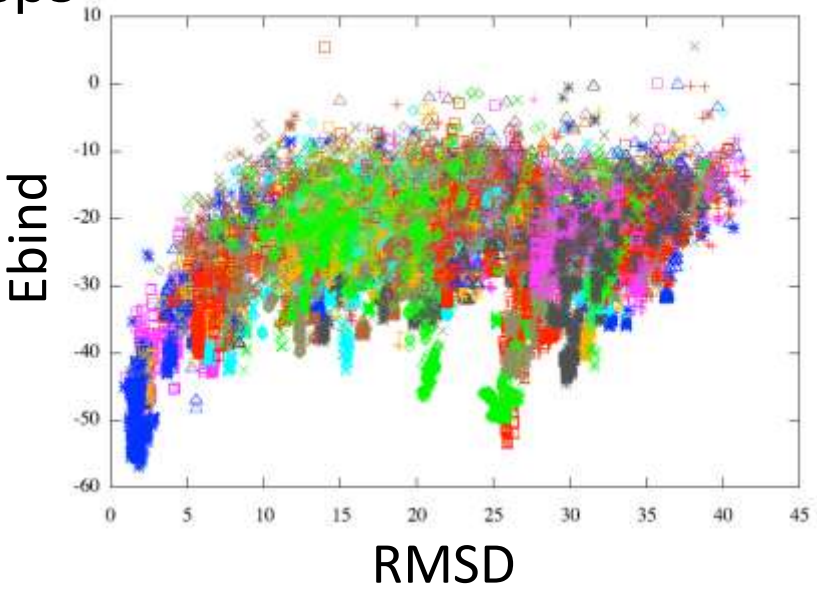


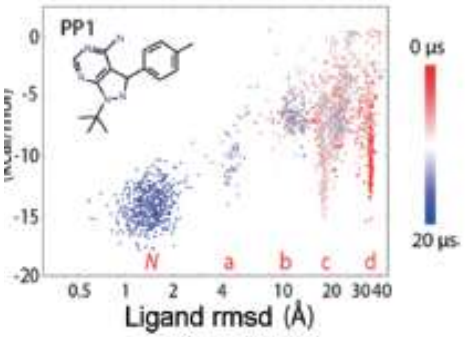


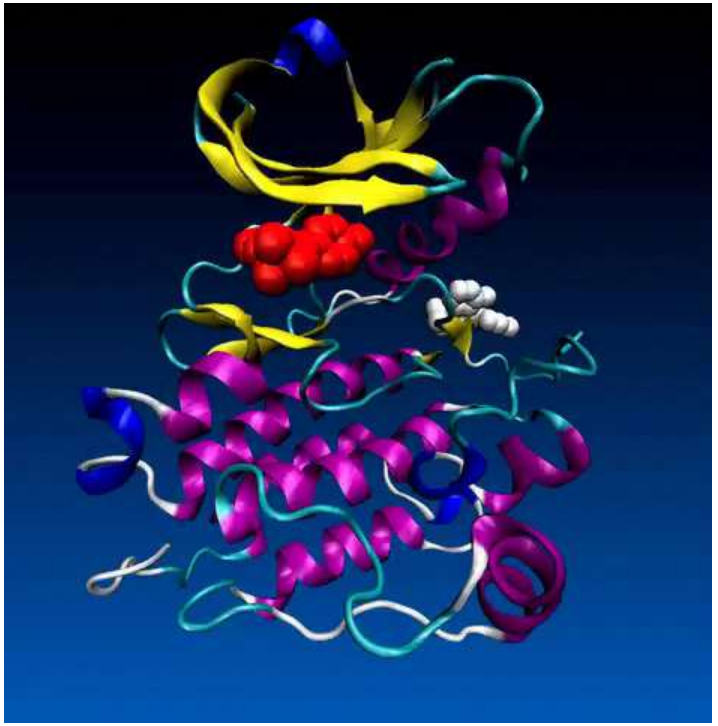


Can we do a difficult case like the Src Kinase?









The PELE Server

https://pele.bsc.es/pele.wt/about

0.9.6.1

Barcelona Supercomputing Center
Centro Nacional de Supercomputación

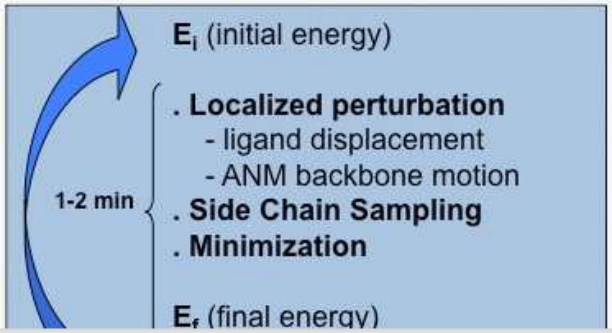
PELE

- About
- News
- Help
- Register
- Login

About PELE

Using technological advances in protein structure prediction, we have developed ***PELE*** (an acronym for Protein Energy Landscape Exploration), a novel method to perform protein energy landscape explorations. ***PELE*** combines a Monte Carlo stochastic approach with protein structure prediction algorithms and is capable of accurately reproducing long time scale processes in only few hours of CPU.

Figure below shows the heuristic algorithm for the landscape exploration method which is based on three main steps: *an initial perturbation*, *a side chain sampling* and a *final minimization*.



First thing: Ligand and protein check!

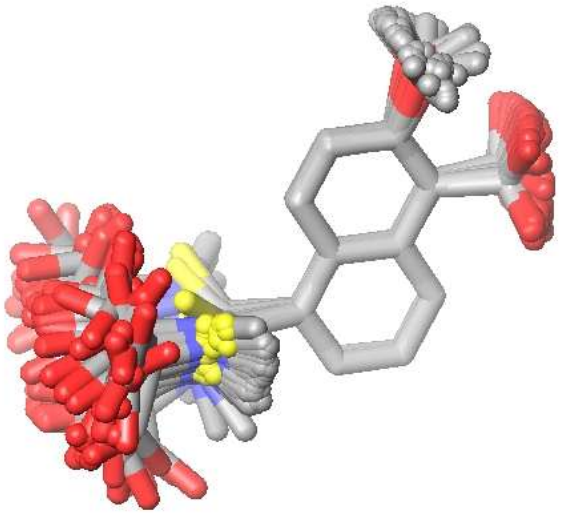
- Check Protein, missing loops, side chain...
- Ligand need to have unique atom names!
- Check protonation state, **read literature**!
- It needs to be one ligand (not composition of different ones...) → unique ligand number, name..

- Run the Protein Wizard of Maestro but check interactively each hydrogen bond, ligand, etc!!!

The **lig** file

```
* LIGAND DATABASE FILE (OPLS2005)
*
INH     57    60    117    205     362
   1     0 M   CM   _C1_     1   46.079180  104.111990   50.326750
   2     1 M   CM   _C3_     3    1.338190   85.268590  104.753090
   3     1 S   CO4  _C5_     5    1.468930   37.845760 -103.947950
.....
11    16   8   4  10  18   3  18   2  10   1  13   4   3  19   4   3
12    3   4   5   1   2   1   1   2   1   1   2   1   1   8   6   4
.....
NBON
    1   3.5500   0.0760  -0.115000   2.0020   1.7750   0.023028004  -0.852763146
    2   3.5500   0.0760  -0.115000   2.0020   1.7750   0.023028004  -0.852763146
.....
BOND
    1     2   549.000  1.340
    1     3   320.000  1.462
.....
THET
    1     2     8    70.00000  124.00000
    1     2     9    35.00000  120.00000
....
PHI
    1     2     8     6  -0.28300  1.0 1.0
    1     2     8     6   0.26700 -1.0 2.0
....
IPHI
    3     4     1     2   1.10000 -1.0 2.0
    3     7     5     6   1.10000 -1.0 2.0
....
END
```

The **LIG.rot.assign** file



```
rot assign res LIG &
  sidelib FREE30 _C16 _C6_ &
  sidelib FREE30 _C6_ _C22 &
  sidelib FREE30 _C22 _O5_ &
    newgrp &
  sidelib FREE10 _C10 _C11_ &
    newgrp &
  sidelib FREE10 _C16 _O4_ &
```

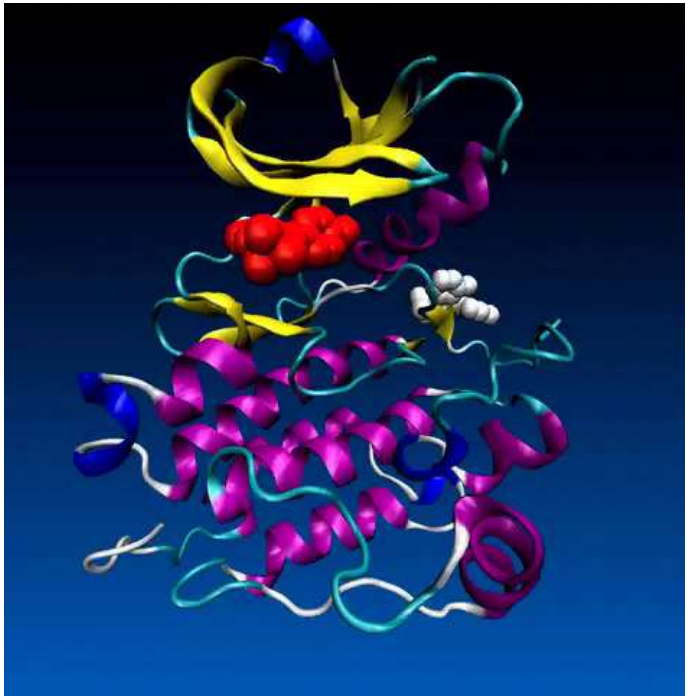
```
rot assign res LIG &
  sidelib FREE30 _C16 _C6_ &
  sidelib FREE30 _C6_ _C22 &
  sidelib FREE30 _C22 _O5_ &
  sidelib FREE30 _C10 _C11_ &
  sidelib FREE30 _C16 _O4_ &
  sidelib FREE30 _C16 _O4_ &
```

Might the ligand dynamics be coupled to protein dynamics??? Global normal modes (INSPECT MODES AND MAYBE SELECT?):

…..
caconst 0.1 *&*
anm_eig_freq 100000 *&*

Specific mode! to open close some domain…
anm_altm_freq 8 &
anm_altm_type 4 &
lanmanm mode 5 &

general random case:
anm_altm_freq 4 &
anm_altm_type 3 &



And... avoid some highly moving parts to dominate the ANM motion!

caconst 0.01 &
rem_bulk_mov 3 &
anm_eig_freq 10000 &
anm_altm_freq 6 &
anm_altm_type 3 &
lanmanm move_ca 0.75 &
lanmanm neig 6 &
lanmanm mix_modes 0.8 &
lanmanm omit_no A:907 A:917 &
lanmanm omit_no A:980 A:983 &

The **control** file

```
file datadir /home/bsc72/bsc72328/plop/data
file log dyn.log


energy params solvent vdgbnp
energy params ionic 0.15


load pdb INIT0001.pdb ions yes waters no het yes


constraint atom A:500:MG__ atom A:217:_OD1 200 1.9 &
atom A:500:MG__ atom A:217:_OD2 200 1.9 &
atom A:500:MG__ atom A:12:_OD1 200 1.82 &
atom A:500:MG__ atom A:209:_OD1 200 1.82 &


load native 3vhv.pdb het yes


pele &
  het L:1 &
```

Focus more on concepts than on sintax!

```
file datadir /home/bsc72/bsc72328/plop/data
file log dyn.log

energy params solvent vdgbnp
energy params ionic 0.15

load pdb INIT0001.pdb ions yes waters no het yes

constraint atom A:500:MG__ atom A:217:_OD1 200 1.9 &
atom A:500:MG__ atom A:217:_OD2 200 1.9 &
atom A:500:MG__ atom A:12:_OD1 200 1.82 &
atom A:500:MG__ atom A:209:_OD1 200 1.82 &

load native 3vhv.pdb het yes

pele &
  het L:1 &
```

This will eliminate the “abstraction of the MG ion by the ligand as we saw before

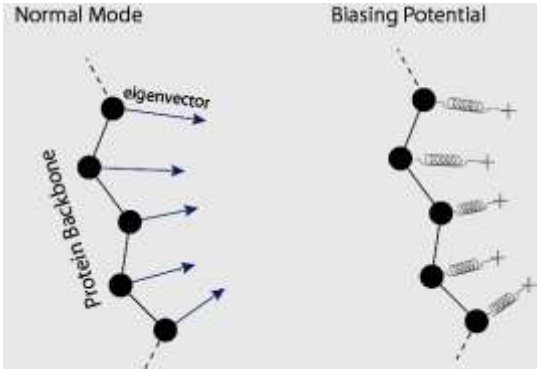
For comparison or guiding!

```
pele &
  het L:1 &
  task &
      show bind_ene 1 &
      show rmsd_lig 1 heavy 0 &
      show atom 1  A:234:_CA_  &
      spawn point 1 -13.5  -3.6  76.8  lt 5.0 &
      if random 1 gt 0.5 then rot_r 0.1 else rot_r 0.25 endif  &
exit steps gt 2000 &
      exit point 1 -13.5  -3.6  76.8  lt  8.0 &
      exit atom 1 A:256:_CA_
  task &
      spawn point 1 -13.5  -3.6  76.8  within 8.0 &
      exit steps gt 500 &
  end_task &
  temp 1000 &
  anmfreq 1 &
  spfreq 1 &
  mifreq 1 &
```

.....
wrfreq 1 &
path traj_ &
rem_bulk_mov 3 &
spradius **6** &
side &
randomize yes &
iter 1 &
sideend &
mirad 100 &
min &
rmsg 0.03 &
nbup yes &
alphaup yes &
minimend &
caconst 0.1 &
.....

Side options

Min options



ANM options

.....
anm_eig_freq 10 &
anm_altm_freq 6 &
anm_altm_type 3 &
lanmanm move_ca 0.75 &
lanmanm neig 6 &
lanmanm mix_modes 0.75 &

ANM-Min options

lanmmin &
mxitn 100 &
iter 1 &
rmsg 0.02 &
nbup yes &
alphaup no &
minimend

Free search

task &

....

if random 1 gt 0.5 then rot_r 0.1 else rot_r 0.25 endif &

if random 2 gt 0.5 then tra_r 6.0 tries 100 else tra_r 1.0 tries 5 endif &

end_task &

steered 1 & ?????

waitfor 3 & ?????

lcom_con 0.1 &