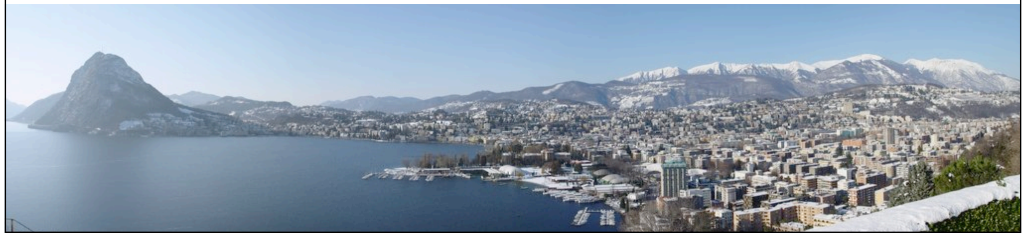
Through mountains and valleys...

Michele Parrinello
Department of Chemistry and Applied Biosciences ETH Zurich
and
ICS, Università della Svizzera Italiana, Lugano, Switzerland



Molecular dynamics

Given a potential energy surface:

$$U\left(\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_N\right)$$

The dynamics can be determined from Newton’s equation:

$$\boldsymbol{M}_I \ddot{\boldsymbol{R}}_I = -\nabla U\left(\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_N\right)$$

Newtonian dynamics and statistical mechanics

From Newtons equations

$$\dot{R} = \frac{P}{M} \qquad \dot{P} = F$$

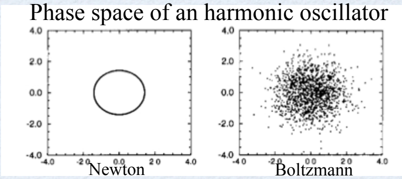
Under the hypothesis of ergodicity

$$\left\langle O(P_I, R_I)\right\rangle = \lim_{\tau\to\infty} \frac{1}{\tau}\int_0^\tau dt\, O(P_I(t), R_I(t)) = \int dP_I dR_I\, O(P_I, R_I) P(P_I, R_I)$$

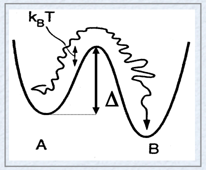
$$P(P_I, R_I) \propto \delta(H(P_I, R_I) - E) \qquad \left\langle K \right\rangle = \frac{3}{2} N k_B T$$

Problems with molecular

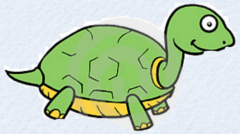
✳ Ergodicity



✳ Activated events



✳ Slow diffusion



Controlling the temperature

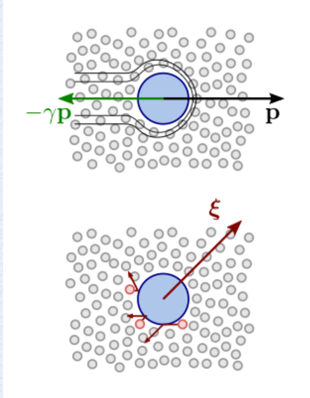
- From (NVE) to (NVT)

$$P(P_I, R_I) \propto \exp(-\beta H(P_I, R_I))$$

- A good thermostat should:
  - Sample the canonical distribution
  - Be ergodic
  - Be tunable
  - Not disturb the dynamics

How about Langevin dynamics?

Brownian particle

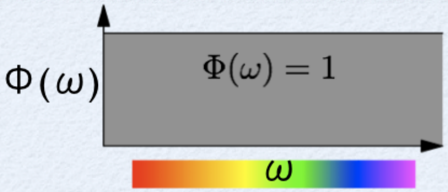


$$\dot{P} = -\gamma P + \xi$$

Langevin equation and statmech

$$\dot{R}_I = \frac{P_I}{M_I}$$

$$\dot{P}_I = F_I - \gamma P_I + \sqrt{2M_I\gamma k_B T}\,\xi_I$$



$$\langle \xi_I(t)\xi_I(t')\rangle = \delta(t-t')$$

Canonical ensemble

$$P(P_I, R_I) \propto \exp\left(-\beta H(P_I, R_I)\right)$$

Generalized Langevin dynamics

$$\dot{R} = \frac{P}{M}$$

Memory function

$$\dot{P} = F - \int_0^t K(t-t')P(t')dt' + \sqrt{M k_B T}\,\zeta(t)$$

$$\langle \zeta(t)\zeta(t') \rangle = K(t-t') \longrightarrow \text{Colored noise}$$

Mori-Zwanzig

A small dictionary

The Canonical Distribution $P(P,R) \propto e^{-\beta\frac{P^2}{2}} e^{-\beta U(R)}$

The Boltzmann Distribution $P(R) = \frac{1}{Z} e^{-\beta U(R)}$ $U(R) = -\frac{1}{\beta}\log P(R)$

The Partition Function $Z = \int e^{-\beta U(R)}\, dR$

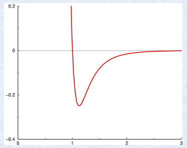
$\beta = \frac{1}{T}$

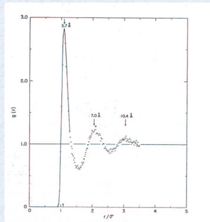
The Free Energy $F = -\frac{1}{\beta}\log Z$

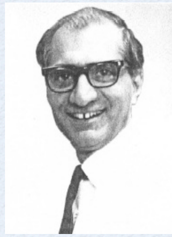
The beginnings

Correlations in the Motion of Atoms in Liquid Argon*

A. Rahman
Argonne National Laboratory, Argonne, Illinois
(Received 6 May 1964)

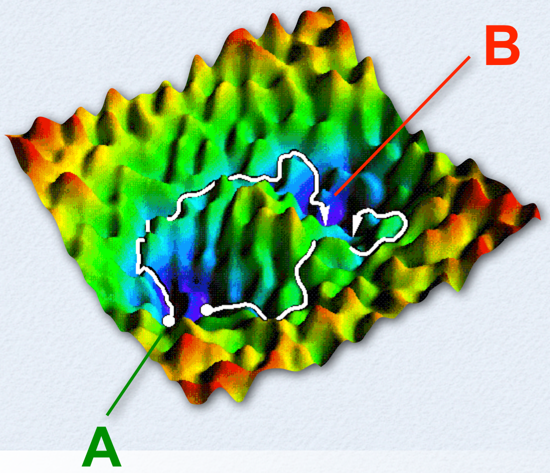


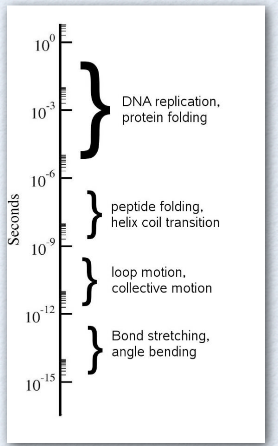




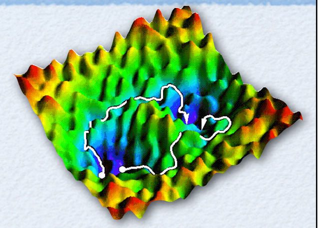
The potential energy surface is rough

Life is complicated

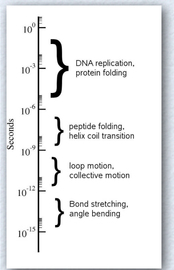




Proposed solutions to the sampling problem



- **Enhanced sampling**:Parallel Tempering, Replica Exchange, Simulated Tempering
- **Trajectory-based schemes**: NEB, Transition Path Sampling, Forward Flux, Finite Temperature String Method



- **Bias potential**: Umbrella Sampling, Local Elevation, Conformational Flooding, Adaptive Force Bias, Self-Healing Umbrella Sampling, Multicanonical MD, Wang-Landau, **Metadynamics**

Focusing on the relevant degrees of freedom

Describe the system using a small set of collective variables S(R)

Explore the free energy surface rather than the rougher potential energy surface

Laio and Parrinello, PNAS 99, 12562 (2002)

Examples of collective variables

- Distances
- Angles: bending and torsional
- Coordination numbers: between individual atoms or between different species
- Contact map
- Solvation energy
- Electric field
- Reaction path

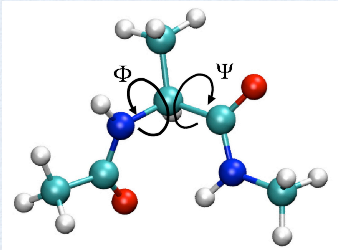
Non-linear functions of atomic coordinates S(R)

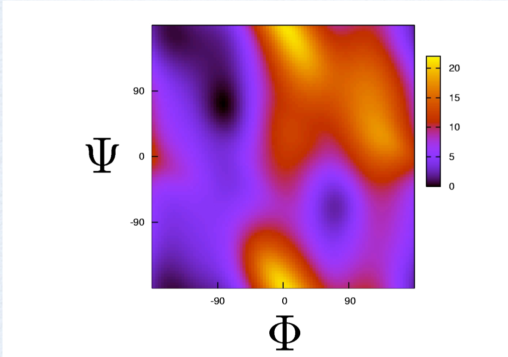
Free energy surface

$$P(\mathbf{S}) = \frac{\int d\mathbf{R}\,\delta\left(\mathbf{S} - \mathbf{S}(\mathbf{R})\right) e^{-\beta V(\mathbf{R})}}{\int d\mathbf{R}\, e^{-\beta V(\mathbf{R})}}$$

$$F(\mathbf{S}) = -k_B T \ln P(\mathbf{S})$$

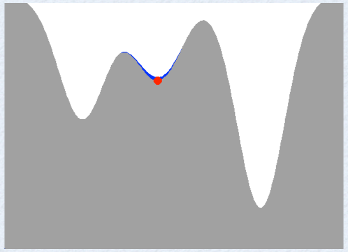
A simple example: alanine dipeptide





Metadynamics

- Wherever you go put a "small" Gaussian
- Always move in the direction that minimizes the sum of F(s) and all the Gaussian potential



$$V(S) = -F(S)$$

Laio and Parrinello, PNAS 99, 12562 (2002)

Well tempered metadynamics

$$M\ddot{R} = F(R) - \frac{\partial V\left(s(R),t\right)}{\partial R}$$

s(R) is a collective coordinate

V(s,t) is a time dependent bias potential

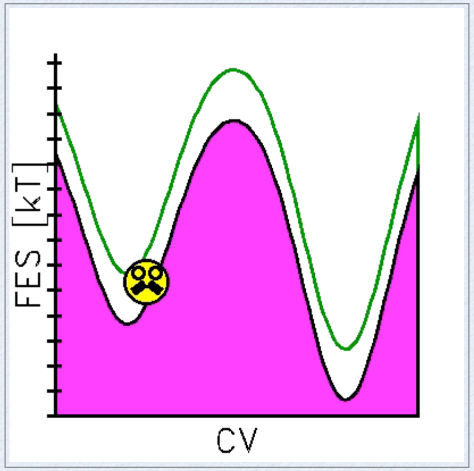
$$\dot{V}\left(s,t\right) = \omega e^{-\beta V(s,t)/(\gamma-1)}\delta_{s,s(t)}$$

ω measures the strength of the bias potential

γ is a boosting factor

In the practice the δ is replaced by a Gaussian

Useful properties of the tempered metadynamics



It fills partially the free energy surface

$$V(s,t\rightarrow\infty)=-(1-\frac{1}{\gamma})F(s)$$

It enhances the s fluctuations

$$P_{WTE}(s)\propto\left(P_{CE}(s)\right)^{\frac{1}{\gamma}}$$

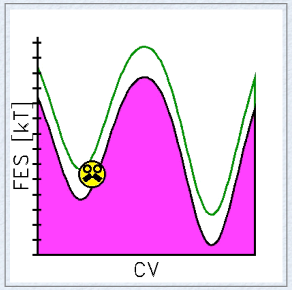
Barducci, Bussi and Parrinello PRL (2008)

Well tempered metadynamics

The initial Gaussian height ω_0 is rescaled during the simulation:

$$\omega = \omega_0 e^{-V(s,t)/\Delta T}$$

$T + \Delta T$ is a fictitious CV temperature



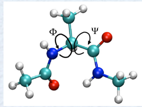
$\gamma = \frac{T + \Delta T}{T}$ is a boosting factor

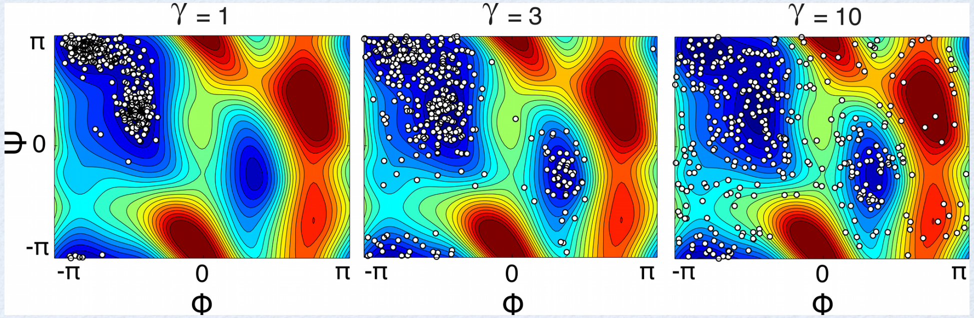
$$V(s,t \to \infty) = -(1 - \frac{1}{\gamma})F(s)$$

$$P_{WTE}(s) \propto \left(P_{CE}(s)\right)^{\frac{1}{\gamma}}$$

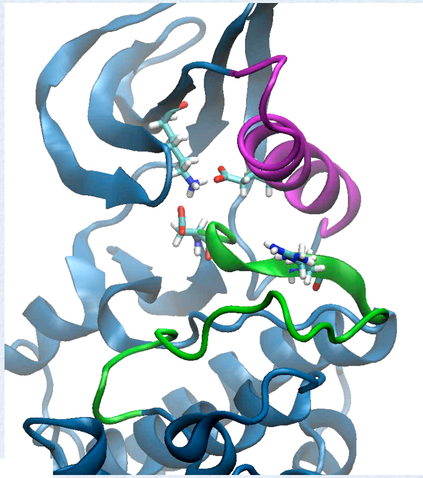
Barducci, Bussi and Parrinello PRL (2008)

Enhancing the fluctuations





Large scale motions in kinases



J|A|C|S ARTICLES

Published on Web 12/09/2008

Protein Conformational Transitions: The Closure Mechanism of a Kinase Explored by Atomistic Simulations

Anna Berteotti,[†] Andrea Cavalli,[*,‡,§] Davide Branduardi,[||]
Francesco Luigi Gervasio,[*,||] Maurizio Recanatini,[‡] and Michele Parrinello[||]

Addressing the CV problem

Combine metadynamics with other sampling methods:

- Parallel tempering, bias exchange, multiple walkers

Find smart collective variables

- Path variables, energy, conjugate variables ...

Discover your own CV

- Reconnaissance metadynamics

Addressing the CV problem

Combine metadynamics with other sampling methods:

- Parallel tempering, bias exchange, multiple walkers

Find smart collective variables

- Path variables, **energy**, ...

Find out your CV

- Reconnaissance metadynamics

The energy as collective variable

$$M\ddot{R} = -\frac{\partial U(\boldsymbol{R})}{\partial \boldsymbol{R}} - \frac{\partial V(U(\boldsymbol{R}), t)}{\partial \boldsymbol{R}},$$

$$\dot{V}(U,t) = -\omega e^{-\frac{V(U,t)}{k_B \Delta T}} \delta_{U,U(t)},$$

$$V(U, t \to \infty) = -\left(1 - \frac{1}{\gamma}\right) F(U)$$

$$\gamma = \frac{T+\Delta T}{T} \geq 1$$

$$F(U) = U - \frac{1}{\beta} \ln N(U)$$

$$N(U) = \int \delta(U - U(R))\, dR$$

Density of states

$$N(U) = \int d\boldsymbol{R}\, \delta(U - U(\boldsymbol{R}))$$

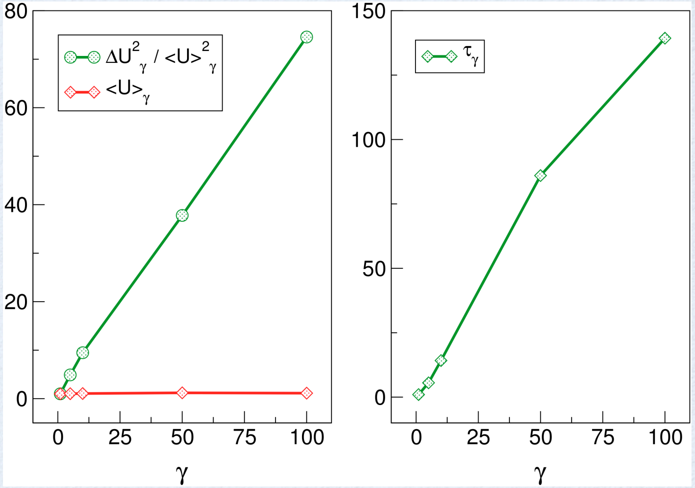
Well-tempered ensemble

$$\mathcal{Z}_\gamma = \int dU\, P(U)^{\frac{1}{\gamma}}$$

$$P(U) = e^{-\beta U} N(U)$$

From the canonical (γ=1) to multicanonical (γ=∞) ensemble

Quasi critical



Why?

$$Z = \int dU e^{-\beta U} N(U) = \int dU P(U)$$

$$P(U) = e^{-\beta U} N(U) \approx C e^{-\frac{\left(U-\bar{U}\right)^2}{2\Delta U^2}}$$

$$Z_\gamma = \int dU P(U)^{\frac{1}{\gamma}} \approx C^{\frac{1}{\gamma}} \int dU e^{-\frac{\left(U-\bar{U}\right)^2}{2\gamma\Delta U^2}}$$

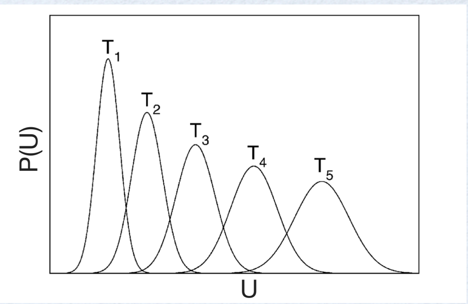
SAME AVERAGES BUT ENHANCED FLUCTUATIONS

Parallel tempering

- N replicas at different T_i

- Avoid colder replica to get trapped in local energy minima

- Exchange configurations with a Metropolis criteria:

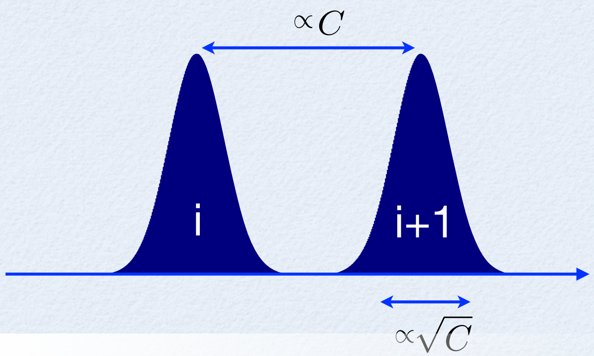


$$p(i \rightarrow j) = \min\left\{1, e^{\Delta}\right\}$$

$$\Delta = \left(\beta_i - \beta_j\right)\left(U\left(R_i\right) - U\left(R_j\right)\right)$$

Overlap necessary

$$U_i = \langle U \rangle_{\beta_i} \pm \sqrt{\langle \Delta U^2 \rangle_{\beta_i}} = \langle U \rangle_{\beta_i} \pm \sqrt{-\frac{\partial \langle U \rangle_\beta}{\partial \beta}}$$



- Number of replicas grows as the <u>square root of the specific heat</u>

Combining PT and WT

- Enhance sampling in the conjugate variables: T and E

Zheng, Chen, Yang PNAS (2009)

- The acceptance should be modified to take into account the bias potential:

Bussi, Gervasio, Laio, Parrinello JACS (2006)

which gives:

$$\Delta = \frac{(\beta_i - \beta_j)(U(\boldsymbol{R}_i) - U(\boldsymbol{R}_j))}{\gamma}$$

Amplified fluctuations

γ-reduction

Round-trip time as a measure of efficiency

Addressing the CV problem

Combine metadynamics with other sampling methods:

- Parallel tempering, bias exchange, multiple walkers

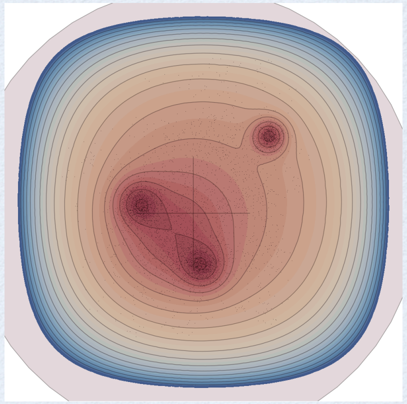
Find smart collective variables

- Path variables, energy, ...

Find out your CV

- **Reconnaissance metadynamics**

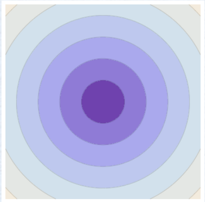
Gaussian mixtures



- M. E. Tipping and C. M. Bishop "Mixtures of probabilistic principal component analysers" Neural Computation 11 pp 443-482 (1999)
- P. Meinicke and H. Ritter "Resolution-Based Complexity Control for Gaussian Mixture Models" Neural Computation 13 pp 453-475 (2001)

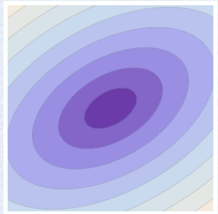
G. Tribello, M. Ceriotti, and M. P., PNAS (2010)

CV's from partitions



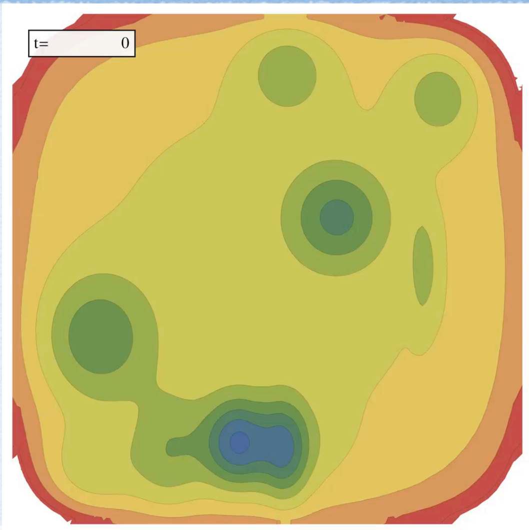
$$R(s) = |s - \mu|$$

$$R(s) = \sqrt{(s-\mu)C^{-1}(s-\mu)}$$



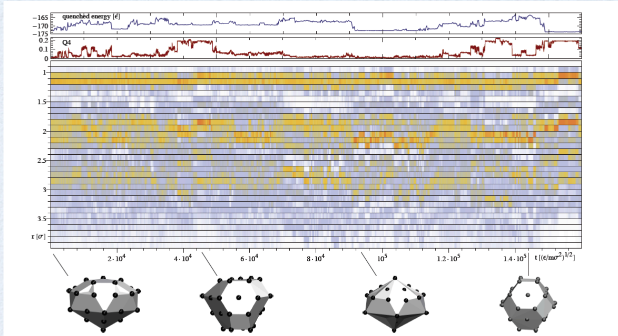
G. Tribello, M. Ceriotti, and M. P., PNAS (2010)

Reconnaissance metadynamics



G. Tribello, M. Ceriotti, and M. P., PNAS (2010)

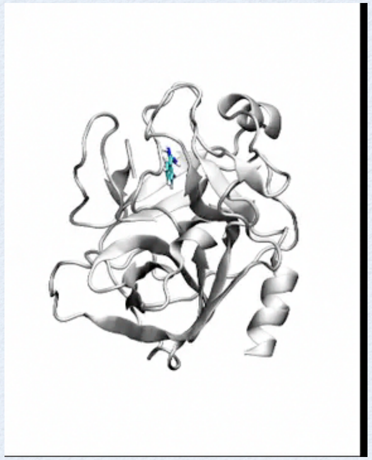
A Lennard-Jones cluster (38 atoms)



Two funnels
one large and entropic
one narrow and enthalpic

G. Tribello, J. Cuny, H. Eshet, and M. P.(2011)

A ligand finds its target



Dimensional reduction

In order to perform metadyamics we have to project the CV in a space of lower dimension

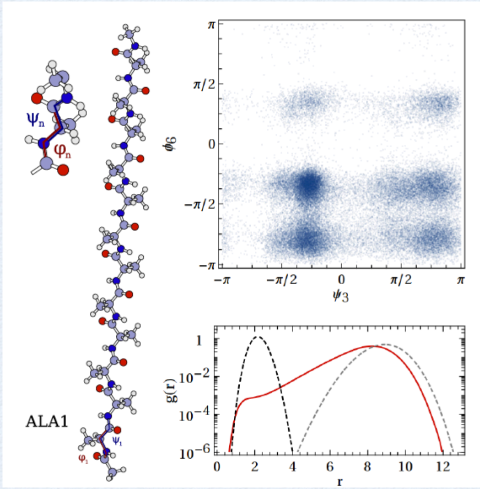
$$\mathbf{X}_n \longrightarrow \mathbf{x}_n$$

$$\min_{\forall x_n} \sum_{i=2}^{N} \sum_{j=1}^{i} (D_{ij} - d_{ij})^2$$

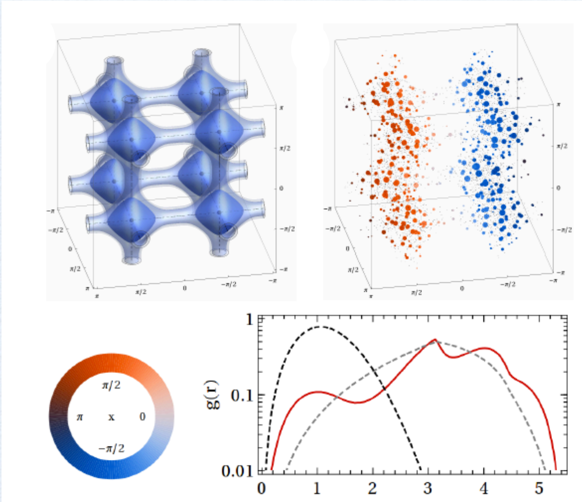
$$\min_{\forall x_n} \sum_{i=2}^{N} \sum_{j=1}^{i} (f(D_{ij}) - f(d_{ij}))^2$$

Stress functions

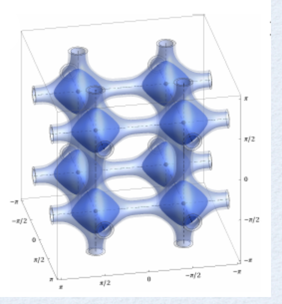
Which dimensional reduction?

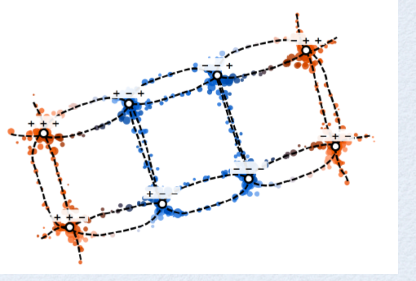


A simple model



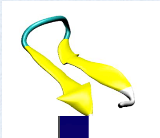
The sketch map projection

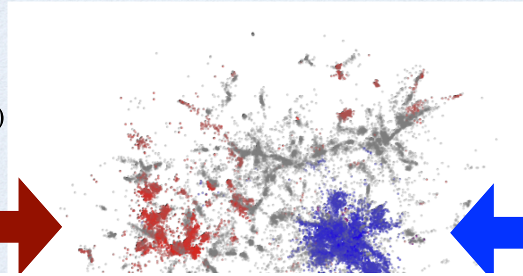




Dimensional reduction the isolucine case

The multidimensional CVs are mapped into a two dimensional manifold (Sketch Map)





PNAS

COMMENTARY

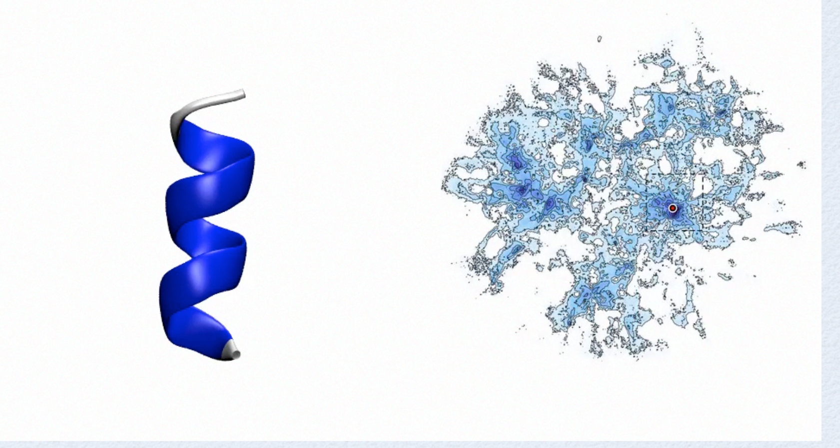
The social network (of protein conformations)

John D. Chodera[a,1] and Vijay S. Pande[b]
[a]California Institute for Quantitative Biosciences (QB3), University of California, Berkeley, CA 94720; and [b]Department of Chemistry, Stanford University, Stanford, CA 94305

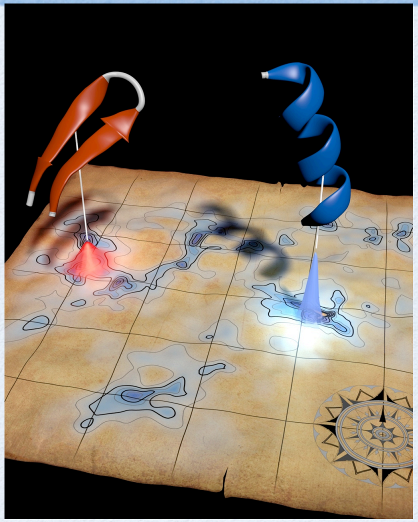
Proteins possess thousands of degrees of freedom, and yet, they can rapidly and reliably find their way into well-defined folded configurations (1). In the cell, these folded proteins carry out highly specific motions

mensionality of a dataset (the location of large land masses on the Earth) is reduced from three (their location in 3D space) to two (coordinates on a flat, rectangular sheet of paper). In constructing his projection in 1569, Gerardus Mercator

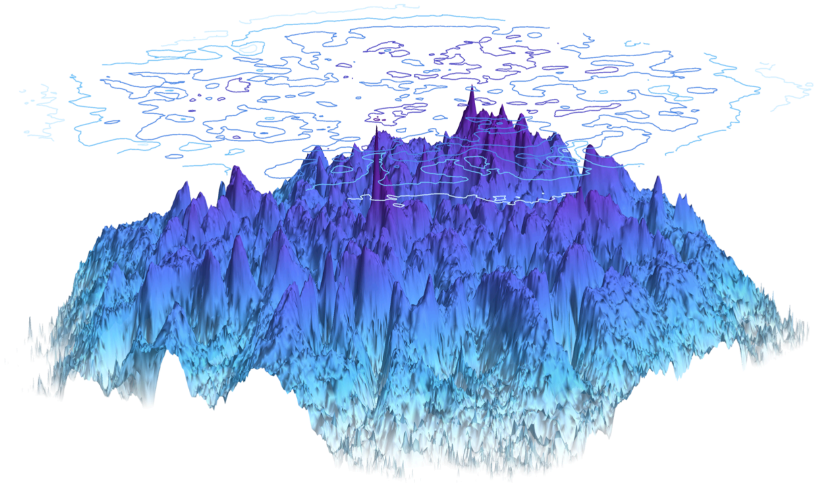
Exploring a rough landscape



The free-energy surface



Mountains and valleys ...as promised



Recognizing the contribution of the reconnaissance people

Michele Ceriotti

Gareth Tribello

For the brave and the curious

PLUMED 1.2

- Metadynamics
- Umbrella Sampling
- Steered MD

portable to

- GROMACS
- NAMD
- AMBER
- DL_POLY
- LAMMPS
- ACEMD
- QUANTUM ESPRESSO
- CPMD

http://merlino.mi.infn.it/plumed

The end

Thank you for your attention