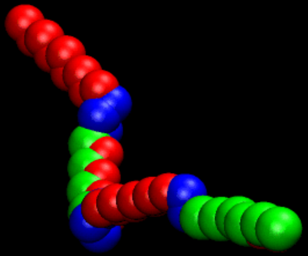
New developments in enhanced conformational sampling using molecular dynamics



Mark E. Tuckerman

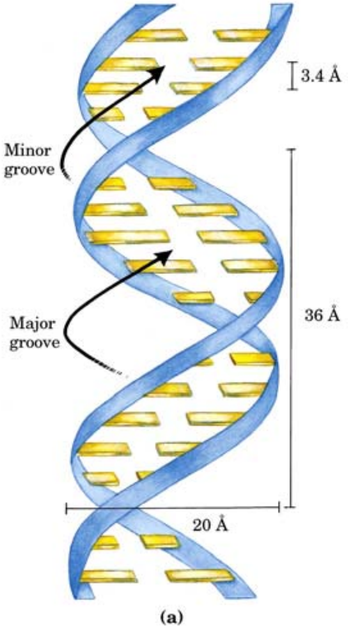
Dept. of Chemistry and Courant Institute of Mathematical Sciences
New York University, 100 Washington Square East, NY 10003

NSF

THE CAMILLE
& HENRY DREYFUS
FOUNDATION, INC.

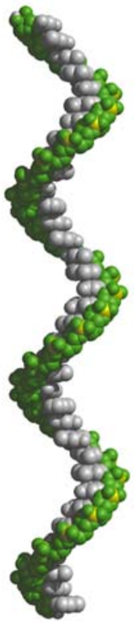
Alexander von Humboldt
Stiftung / Foundation

DNA



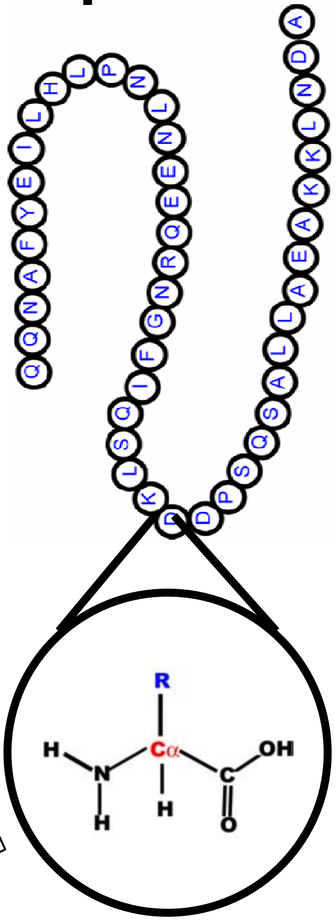
Translation

RNA

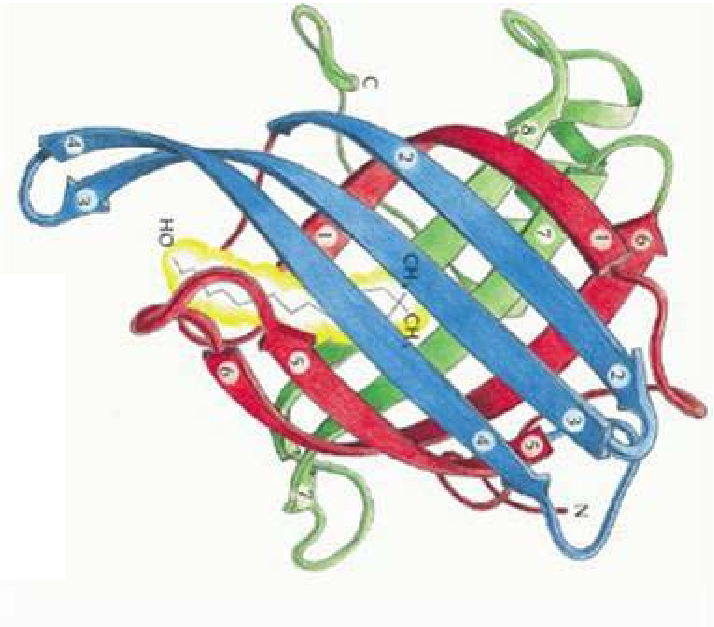


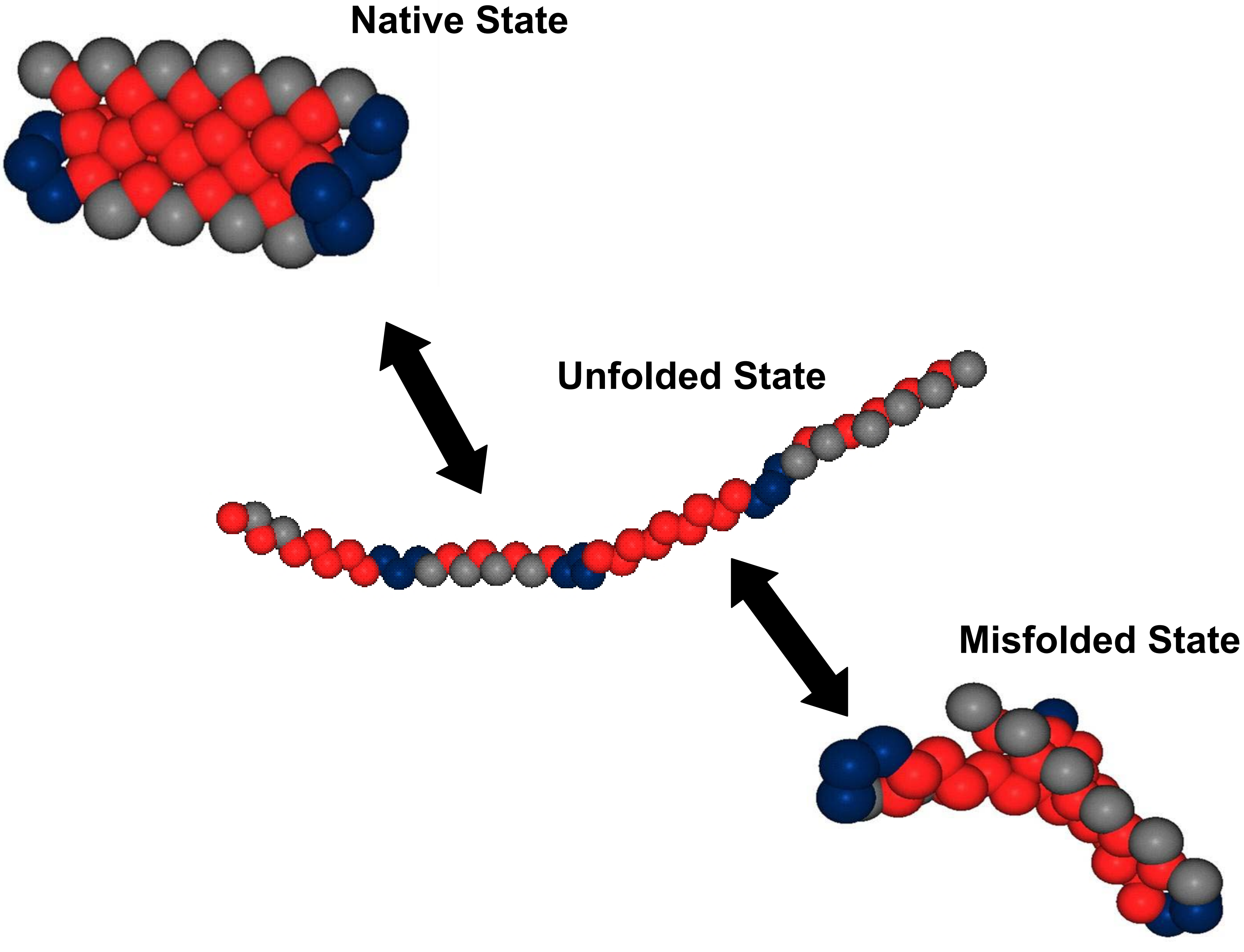
Transcription

Protein Sequence



FOLDING





Native State

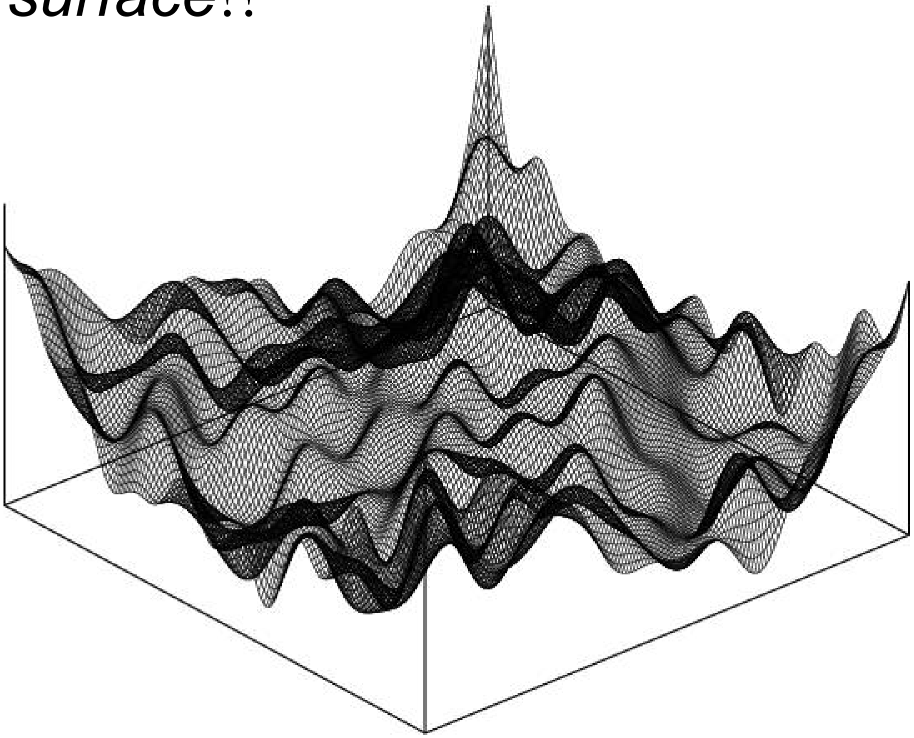
Unfolded State

Misfolded State

Why we cannot sample conformations

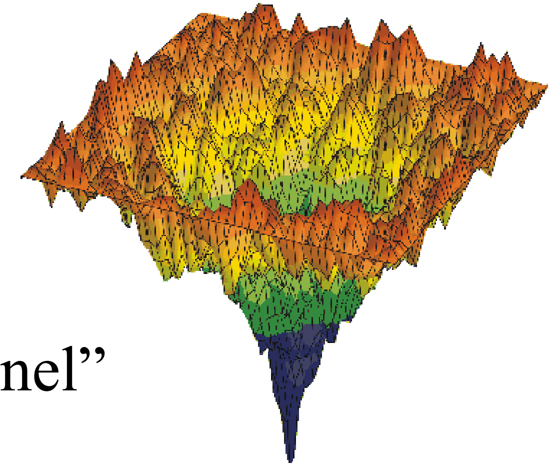
This is our potential energy surface!!

"Rough" energy landscape



Levinthal paradox:

Protein with 100 amino acids has 10^{50} possible conformations!



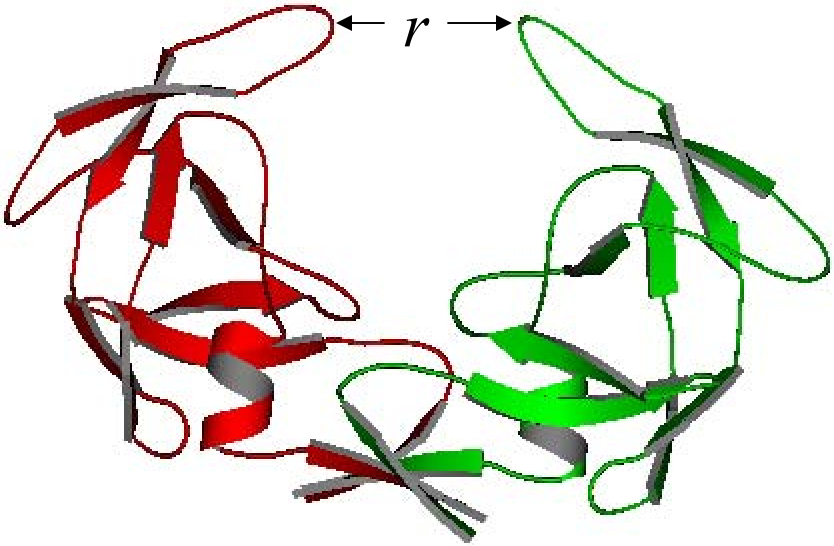
"Folding funnel"

Adiabatic free-energy dynamics (AFED)

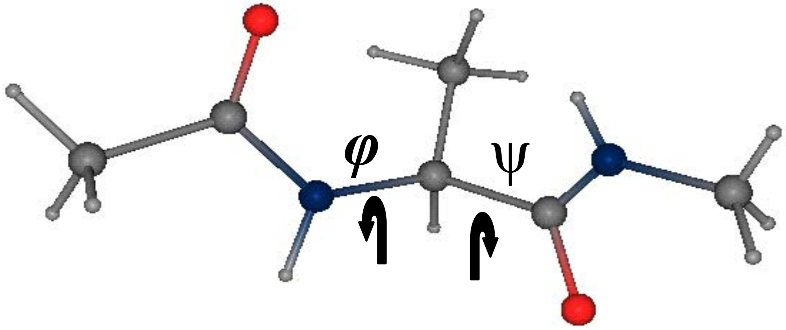
With: **Lula Rosso** (Imperial College) and **Jerry B. Abrams** (BU)

Small subset of collective variables of interest

Distances



Dihedral angles:



Denote these coordinates as

$$q_\alpha(\mathbf{r}) \qquad \alpha = 1,\ldots,n$$

Adiabatic free-energy dynamics

L. Rosso and MET *Mol. Simulat.* **28**, 91 (2002); L. Rosso, P. Minary, Z. Zhu, MET *J. Chem. Phys.* **116**, 4389 (2002)

In a transformation to generalized coordinates:

$$q_\alpha = q_\alpha(\mathbf{r}_1,\ldots,\mathbf{r}_N) \qquad \alpha = 1,\ldots,3N \quad ; \qquad \mathbf{r}_i = \mathbf{r}_i(q_1,\ldots,q_{3N})$$

Suppose first *n* are of particular interest.

$$P(s_1,\ldots,s_n) = \int d^N\mathbf{p}\,d^N\mathbf{r}\; e^{-\beta H(\mathbf{p},\mathbf{r})}\prod_{\alpha=1}^{n}\delta(q_\alpha(\mathbf{r})-s_\alpha)$$

$$= \int d^N\mathbf{p}\,d^{3N}q\; e^{-\beta\tilde{H}(\mathbf{p},q)}\prod_{\alpha=1}^{n}\delta(q_\alpha-s_\alpha)$$

$$H = \frac{1}{2}\mathbf{p}^T M^{-1}\mathbf{p} + V(\mathbf{r}) \qquad \tilde{H} = \frac{1}{2}\mathbf{p}^T M^{-1}\mathbf{p} + V(\mathbf{r}(q)) - kT\ln J(q)$$

Adiabatic and temperature conditions: $m_{1,\ldots,n} \gg m_{n+1,\ldots,3N}$ $\qquad T_s \gg T$

Free-energy surface: $$A(s_1,\ldots,s_n,T) = -kT_s\ln P_{\text{adb}}(s_1,\ldots,s_n)$$

Adiabatic Free-Energy Dynamics

L. Rosso and MET *Mol. Simulat*. **28**, 91 (2002); L. Rosso, P. Minary, Z. Zhu, MET *J. Chem. Phys*. **116**, 4389 (2002)

Equations of motion:

$\alpha = 1,\ldots,n$	$\alpha = n+1,\ldots.,3N$
$m_\alpha \ddot{q}_\alpha = F_\alpha(q) + \text{ heat bath}(T_s)$	$m_\alpha \ddot{q}_\alpha = F_\alpha(q) + \text{ heat bath}(T)$

Analysis using the Liouville operator:

S. Samuelson and G. J. Martyna, *J. Chim. Phys*. **94**, 1503 (1997)
D. Marx, MET, G. J. Martyna, *Comput. Phys. Comm*. **118**, 166 (1999)
L. Rosso, P. Minary, Z. Zhu, MET *J. Chem. Phys*. **116**, 4389 (2002)
L. Rosso and MET *Mol. Simulat*. **28**, 91 (2002)
J. B. Abrams, L. Rosso and MET *J. Chem. Phys*. **125**, 074115 (2006) (solvation free energy)
J. B. Abrams and MET *J. Phys. Chem. B* **112**, 15742 (2008)

Adiabatic Free-Energy Dynamics

Free energy surface at temperature *T*:

$$e^{-\beta A(q_1,\ldots,q_n;\beta)} = \int d^{3N-n}q \; e^{-\beta\tilde{V}(q_1,\ldots,q_{3N})}$$

Free energy surface at temperature T_s:

$$e^{-\beta_s A(q_1,\ldots,q_n;\beta)} = \left[e^{-\beta A(q_1,\ldots,q_n;\beta)} \right]^{\beta_s/\beta}$$

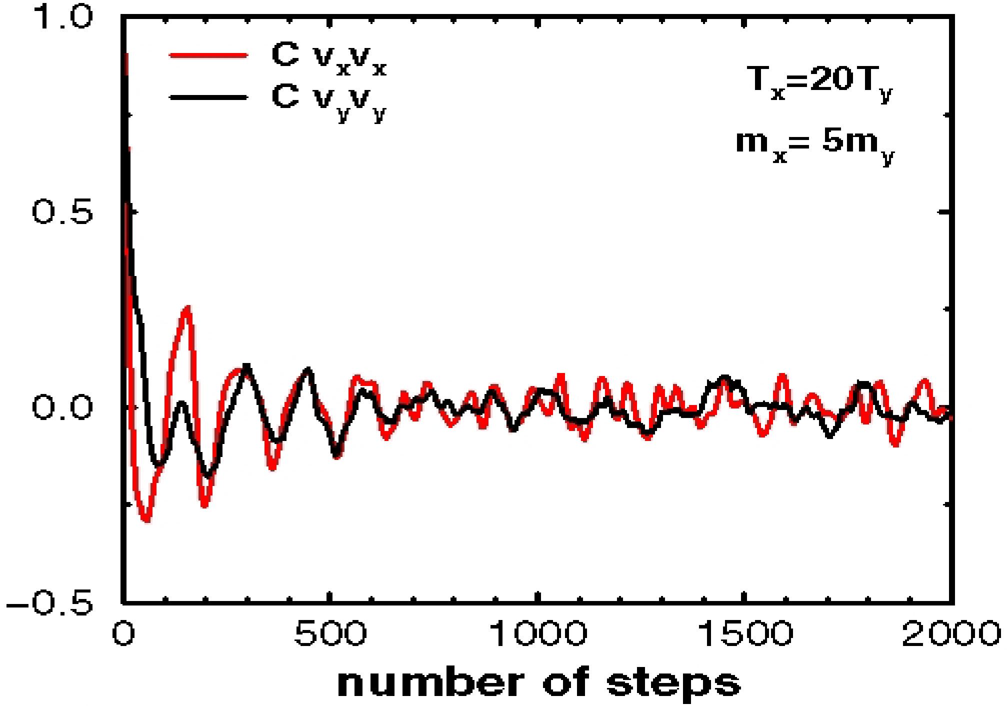
$$= \left[\int d^{3N-n}q \; e^{-\beta\tilde{V}(q_1,\ldots,q_{3N})} \right]^{\beta_s/\beta}$$

$$= P_{\text{adb}}(q_1,\ldots,q_n;\beta,\beta_s)$$

Free energy profile given by:

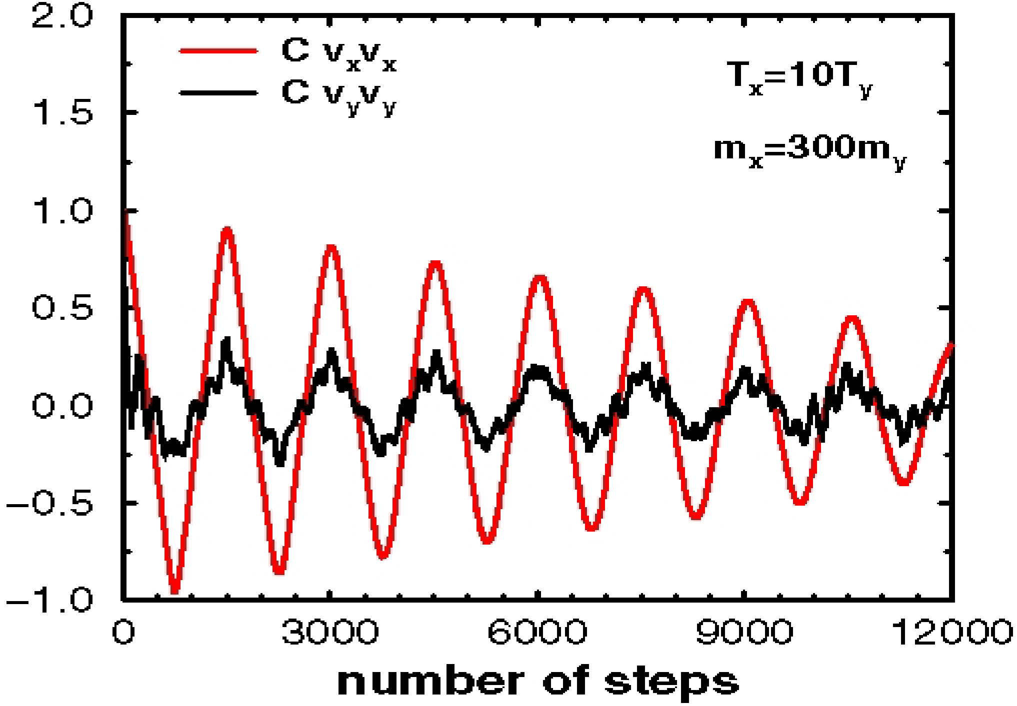
$$\boxed{A(q_1,\ldots,q_n;\beta) = -kT_s \ln P_{\text{adb}}(q_1,\ldots,q_n;\beta,\beta_s)}$$

L. Rosso, P. Minary, Z. Zhu, MET
J. Chem. Phys. **116**, 4389 (2002)



Toy model:

$$V(x,y)=D_0\left(x^2-a^2\right)^2 +\frac{1}{2}ky^2+\lambda xy$$

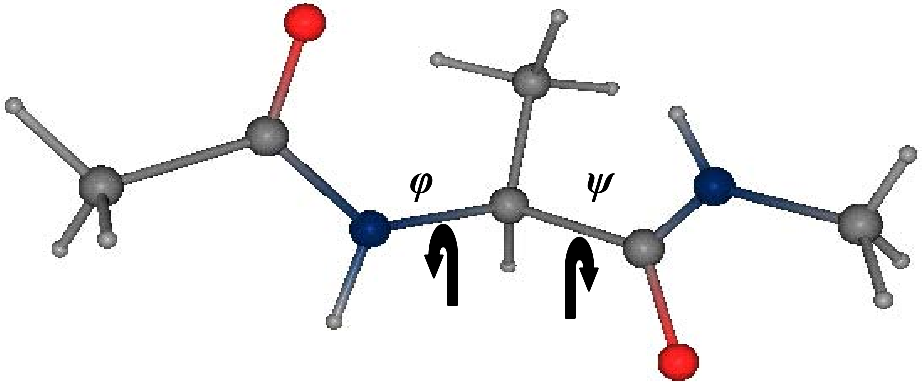


$$C_{v_xv_x}(t)=\left\langle \dot{x}(0)\dot{x}(t)\right\rangle$$

$$C_{v_yv_y}(t)=\left\langle \dot{y}(0)\dot{y}(t)\right\rangle$$

Conformational sampling of the solvated alanine dipeptide

[L Rosso, J. B. Abrams and MET J. Phys. Chem. B **109**, 2099 (2005)]

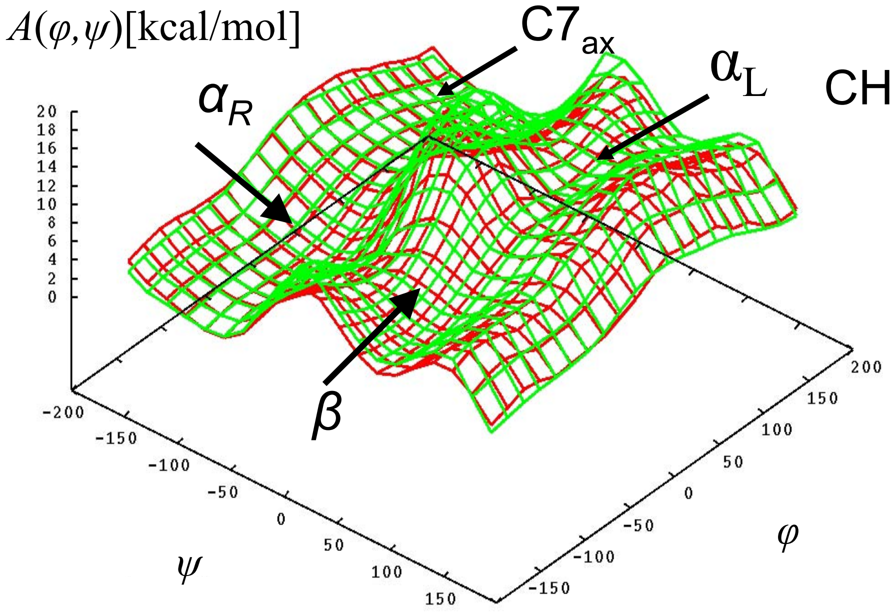


	Time	FF	β	α_R	$C7_{ax}$	α_L
Meta[1]	5 ns	CHm27	1.0	0.0	4.8	7.4
US	400 ns	CHm22	0.3	0.0	3.9	4.4
AFED	4.7 ns	CHm22	0.2	0.0	4.6	8.2

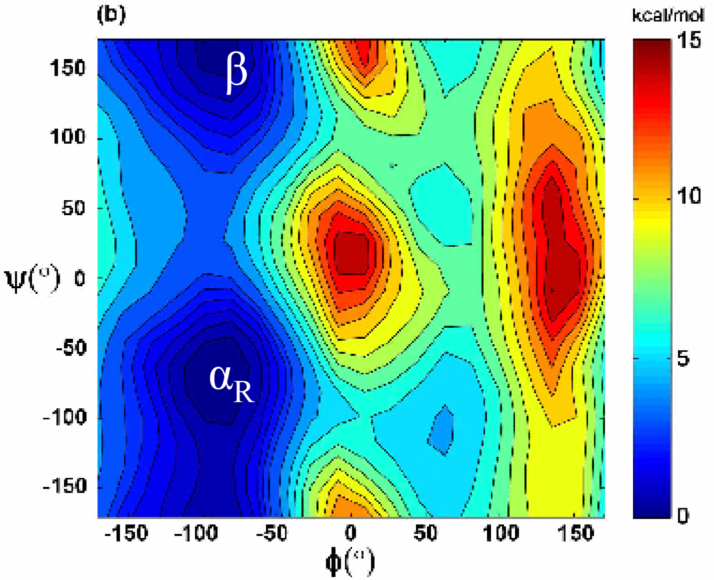
🟥 AFED $T_{\varphi,\psi} = 5T$, $M_{\varphi,\psi} = 50M_C$ 4.7 ns

🟩 Umbrella Sampling 50 ns

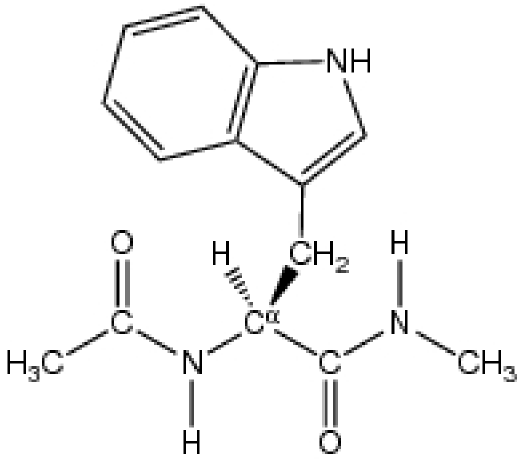
1. Laio and Parrinelli, PNAS (2002)
   Ensing, et al. ACR **39**, 73 (2005)

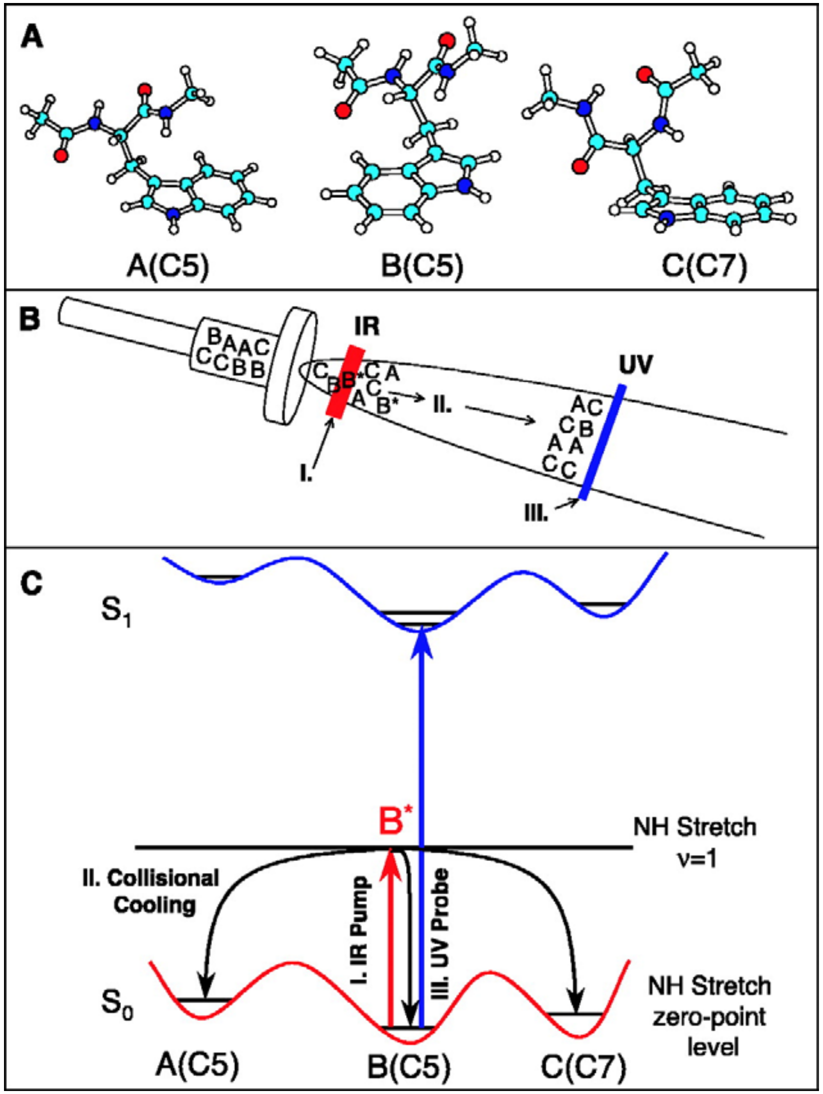


CHARMm22



N-acetyl-tryptophan-methylamide (NATMA)





- Selective IR excitation of molecules from jet gives molecules excess energy.
- Excess energy causes isomerization.
- Collisions lead to cooling.
- UV Probe collision-free cooled populations.

Dian, *et al*. *Science* **296**, 2369 (2002)
Dian, *et al*. *JCP* **117**, 10688 (2002)
Dian, *et al*. *JCP* **120**, 133 (2004)
Dian, *et al*. *JCP* **120**, 9033 (2004)

Ramachandran Surfaces

J. B. Abrams and MET (in preparation)

Simulation Details

$T = 423$ K

$T_{\phi,\psi} = 5T$

$M_{\phi,\psi} = 50M_C$

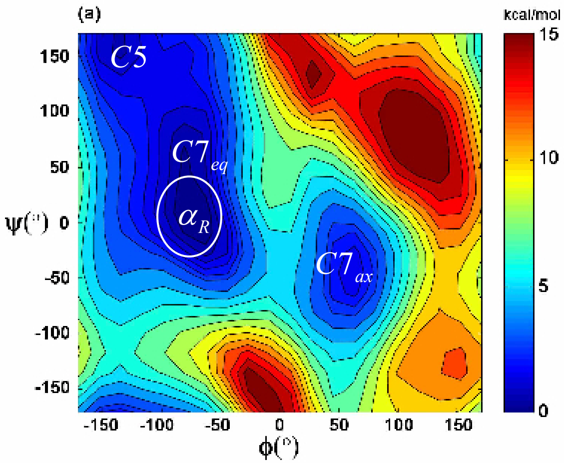
$\Delta t = 2.0$ fs with r-RESPA

Simulation time = 5 ns

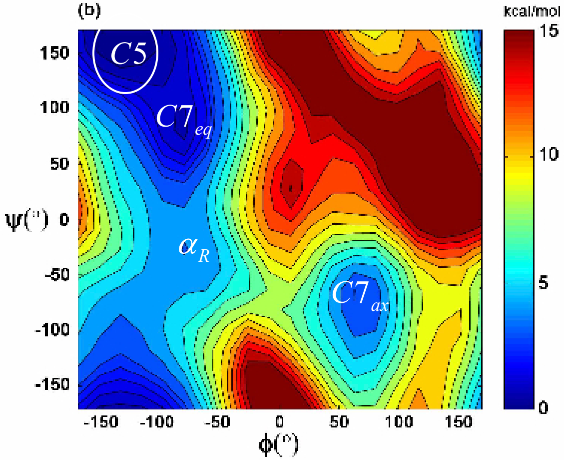
FFs: CHARMM22, AMBER95, OPLS-AA

GGMT Thermostats

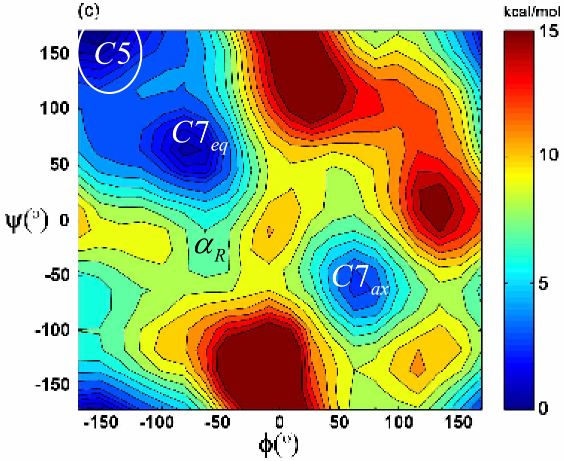
[Liu, et al. JCP **112**, 1685 (2000)]



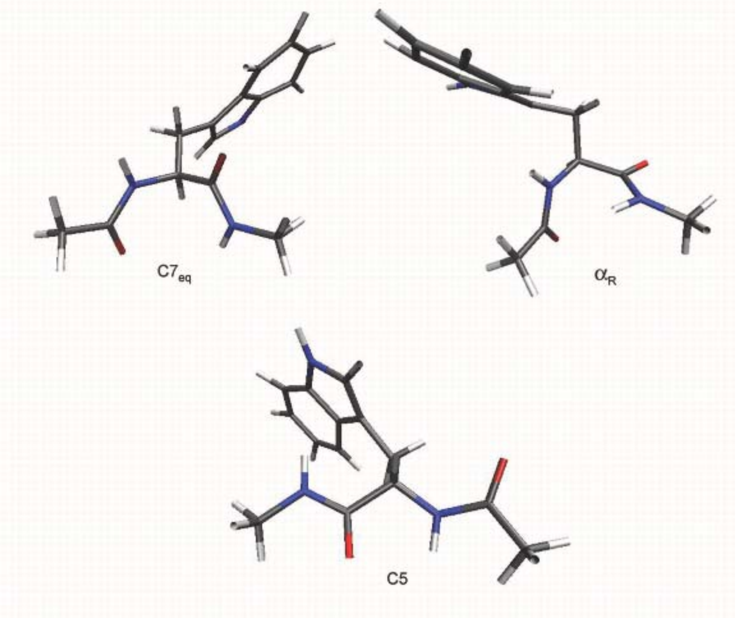
AMBER95



CHARM22



OPLS-AA



Expt. C5 (0.0), C7eq (0.73 kcal/mol)

AFED without transformations/TAMD

Margliano and Vanden-Eijnden, *Chem. Phys. Lett.* **426**, 168 (2006)
J. B. Abrams and MET, *J. Phys. Chem. B* **112**, 14752 (2008)

$$P(s_1,\ldots,s_n) = \int d^N\mathbf{p}\,d^N\mathbf{r}\;e^{-\beta H(\mathbf{p},\mathbf{r})}\prod_{\alpha=1}^{n}\delta(q_\alpha(\mathbf{r})-s_\alpha)$$

Write δ-functions as product of Gaussians:

$$P(s_1,\ldots,s_n) = \lim_{\kappa\to\infty}\left(\frac{\beta\kappa}{2\pi}\right)^{n/2}\int d^N\mathbf{p}\,d^N\mathbf{r}\;e^{-\beta H(\mathbf{p},\mathbf{r})}\prod_{\alpha=1}^{n}\exp\left[-\frac{\beta\kappa}{2}\left(q_\alpha(\mathbf{r})-s_\alpha\right)^2\right]$$

Introduce uncoupled Gaussian integrations:

$$P_\kappa(s_1,\ldots,s_n) = C_\kappa\int d^N\mathbf{p}\,d^N\mathbf{r}\;d^n p_s\,\exp\left\{-\beta\left[H(\mathbf{p},\mathbf{r})+\sum_{\alpha=1}^{n}\frac{p_{s_\alpha}^2}{2m_\alpha}+\frac{\kappa}{2}\sum_{\alpha=1}^{n}\left(q_\alpha(\mathbf{r})-s_\alpha\right)^2\right]\right\}$$

AFED without transformations/TAMD

Now impose adiabatic conditions on $\{s, p_s\}$: $T_s \gg T, \quad m_\alpha >> m$

Distribution generated:

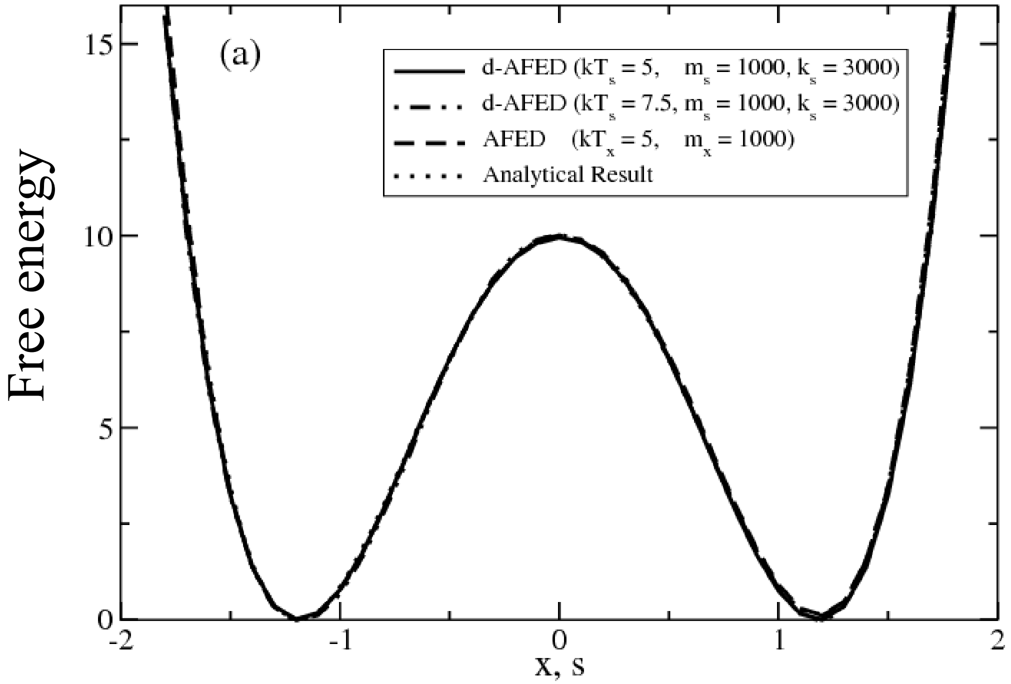
$$P_\kappa^{(\mathrm{adb})}(s_1,\ldots,s_n) \propto \int d^n p_s \exp\left[-\beta_s \sum_{\alpha=1}^{n} \frac{p_{s_\alpha}^2}{2m_\alpha}\right]\left[Z_\kappa(s_1,\ldots,s_n)\right]^{\beta_s/\beta}$$

$$Z_\kappa(s_1,\ldots,s_n) = \int d^N\mathbf{p}\,d^N\mathbf{r}\ \exp\left\{-\beta\left[H(\mathbf{p},\mathbf{r}) + \sum_{\alpha=1}^{n}\frac{p_{s_\alpha}^2}{2m_\alpha} + \frac{\kappa}{2}\sum_{\alpha=1}^{n}\left(q_\alpha(\mathbf{r}) - s_\alpha\right)^2\right]\right\}$$

Free energy surface:

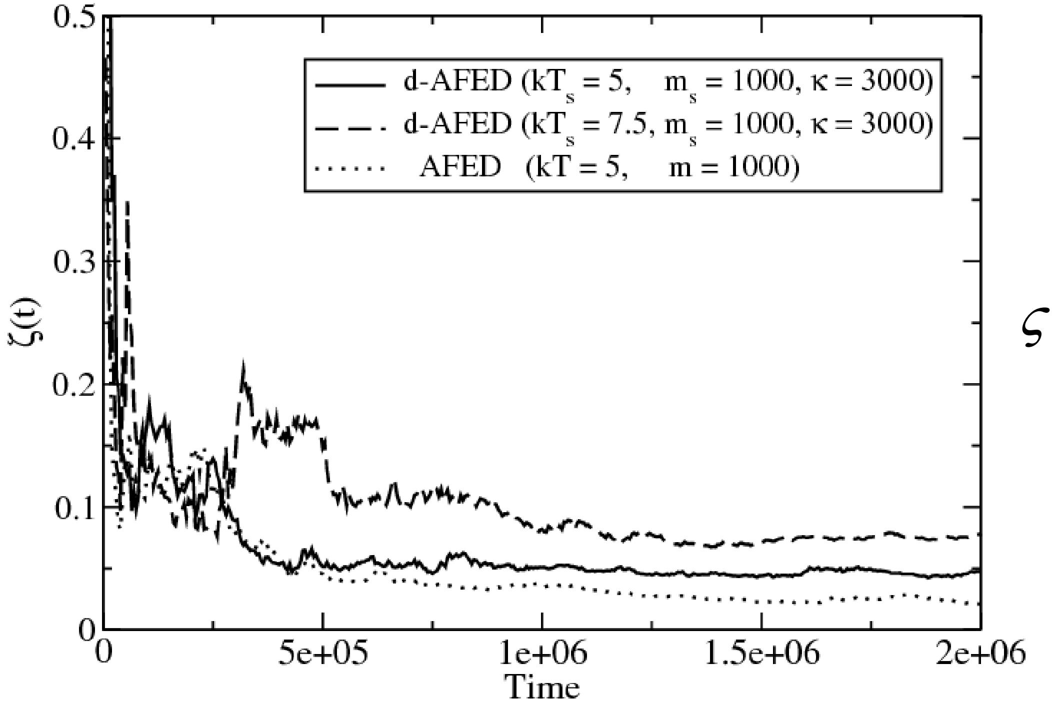
$$A_\kappa(s_1,\ldots,s_n) = -kT_s \ln P_\kappa^{(\mathrm{adb})}(s_1,\ldots,s_n)$$

$$A(s_1,\ldots,s_n) = \lim_{\kappa\to\infty} A_\kappa(s_1,\ldots,s_n)$$



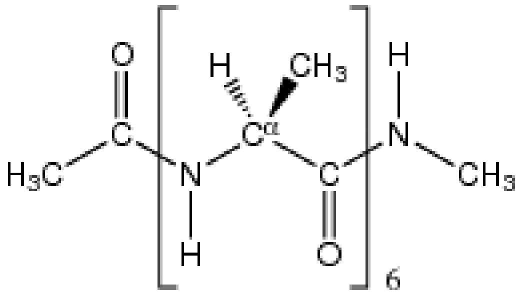
Toy model:

$$V(x,y) = D_0\left(x^2 - a^2\right)^2 + \frac{1}{2}ky^2 + \lambda xy$$



$$\varsigma(t) = \frac{1}{N_{\text{bins}}}\sum_{i=1}^{N_{\text{bins}}}\left[P(x_i;t) - P_{\text{exact}}(x_i)\right]$$

Preliminary results for N-acetyl-(AAAAAA)-methylamide



Simulation Details

$T = 300$ K, $T_s = 600$ K

$m_s = 150 m_C$

$\kappa = 5.4\times10^6\,\mathrm{K}\bullet(\mathrm{A})^{-2}$

(5 r-RESPA steps used for harmonic coupling)

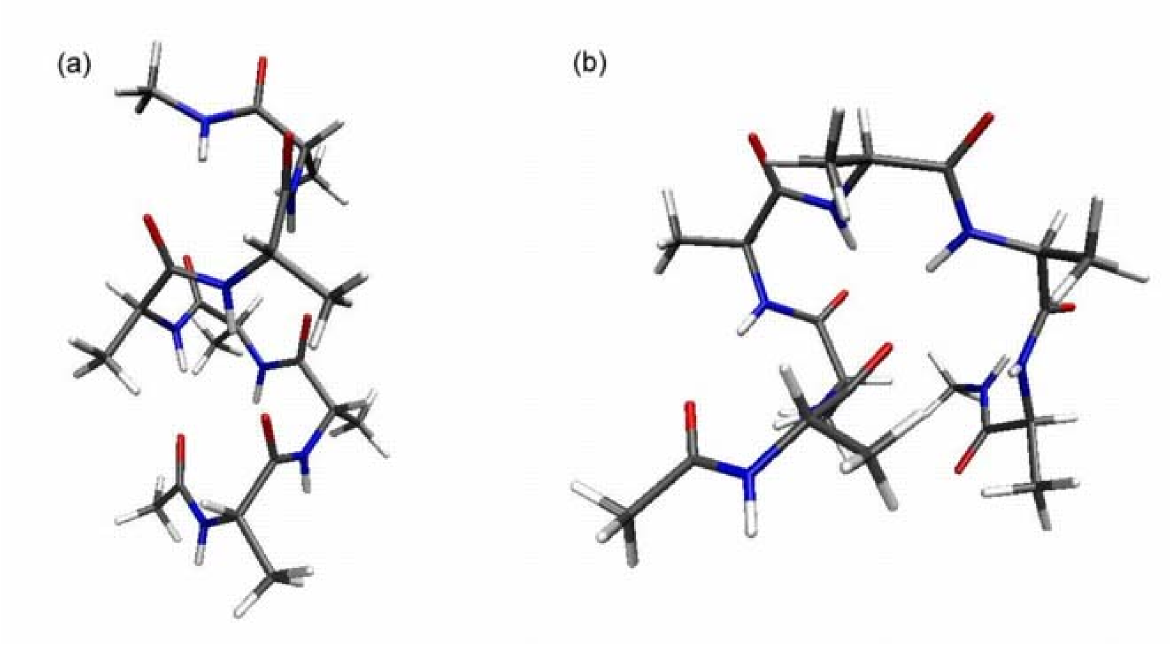
$\Delta t = 6.0$ fs

FF: AMBER95, 698 TIP3P waters

GGMT Thermostats

Run length = 5 ns

Minimum: α_R



Collective variables:

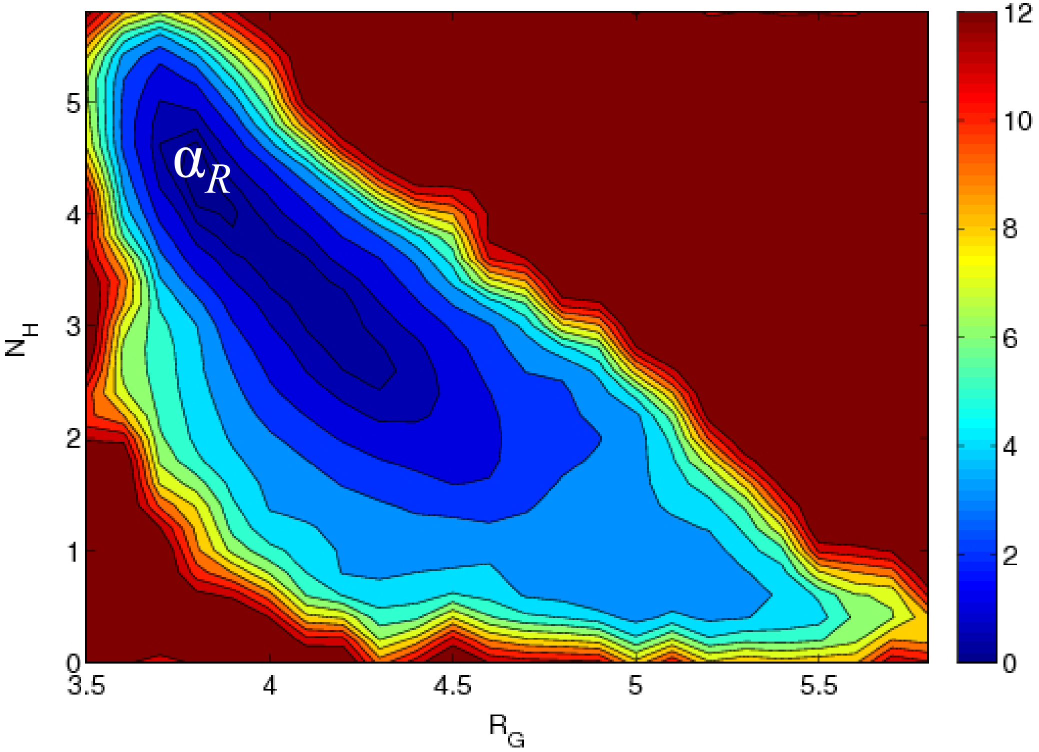
Radius of gyration of backbone:

$$R_G = \sqrt{\frac{1}{N_b}\sum_{i=1}^{N_b}\left(\mathbf{r}_i - \frac{1}{N_b}\sum_{j=1}^{N_b}\mathbf{r}_j\right)^2}$$

Number of hydrogen bonds:

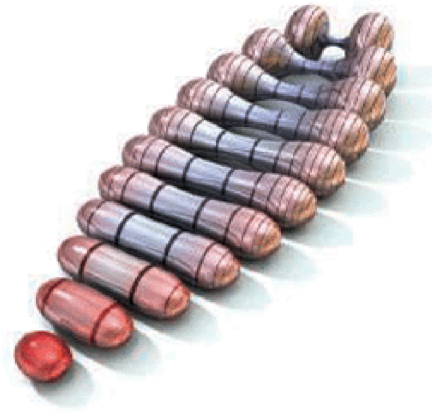
$$N_H = \sum_{i=1}^{N_O}\sum_{j=1}^{N_N}\frac{1-\left(\left|\mathbf{r}_i - \mathbf{r}_j\right|/d_0\right)^6}{1-\left(\left|\mathbf{r}_i - \mathbf{r}_j\right|/d_0\right)^{12}}$$

Free-energy surface of alanine hexamer

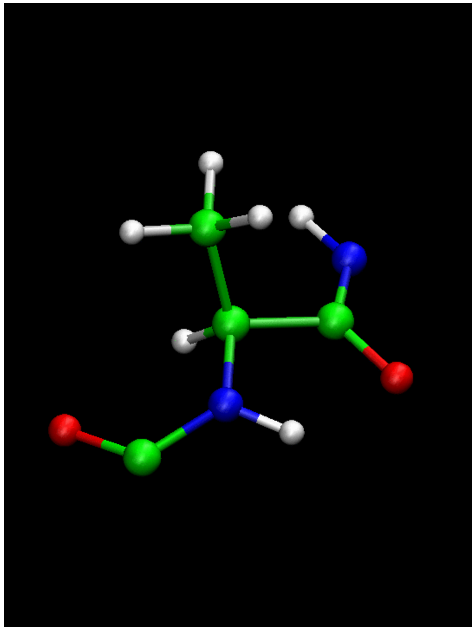


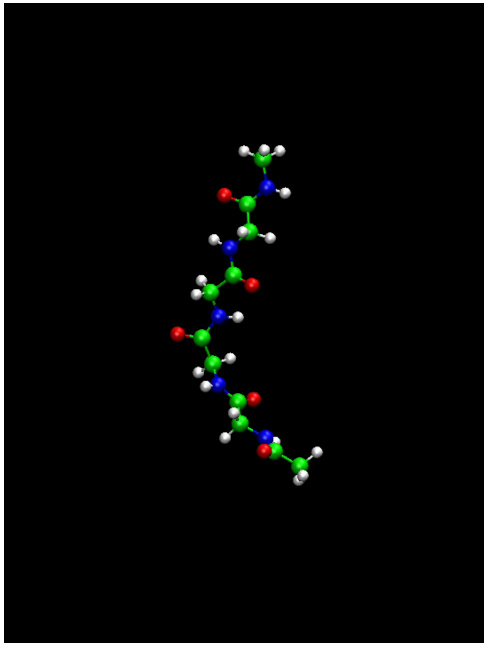
Spatial-Warping Transformations

With: **Glenn J. Martyna** (IBM), **Zhongwei Zhu (**Wall Street**)**, **Peter Minary** (Stanford)

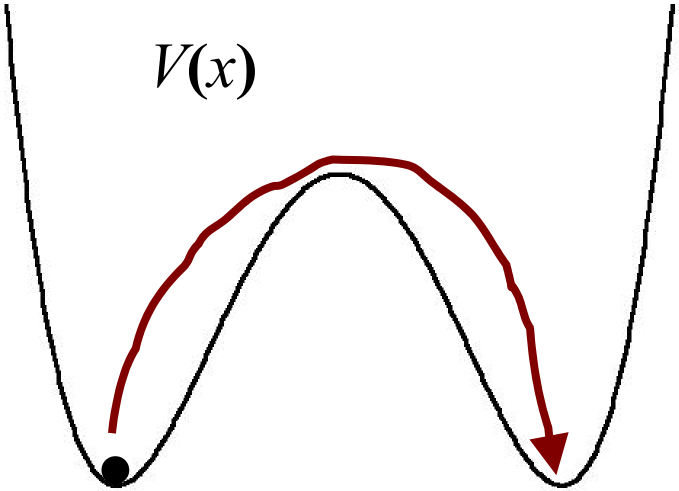


Science, December 22, 2006



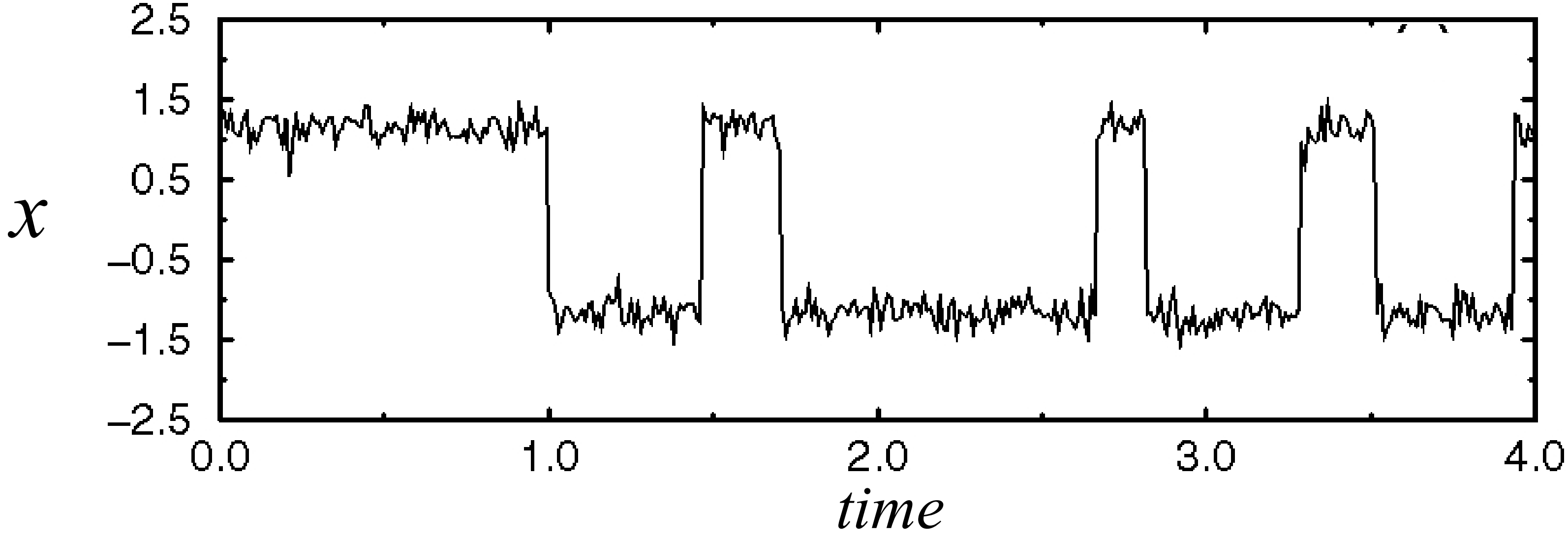


Molecular dynamics in a double well



$$\dot{x} = \frac{p}{m}$$

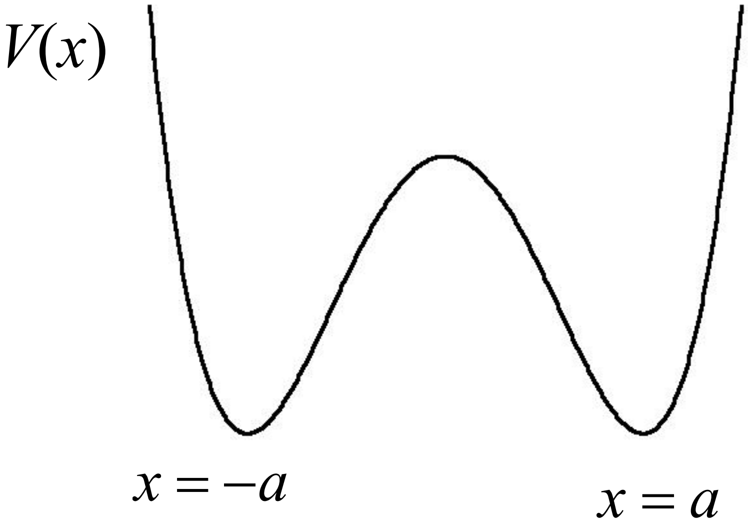
$$\dot{p} = -\frac{dV}{dx} = F(x) + \text{heat bath}$$



Crossing the barrier is a rare event!

Spatial-warping transformations

Z. Zhu, MET, S. O. Samuelson, G. J. Martyna, *Phys. Rev. Lett*. **88** 100201 (2002)



Hamiltonian:

$$H(x,p)=\frac{p^2}{2m}+V(x)$$

Canonical partition function:

$$Q=\int dp\ dx\ \ \exp\left\{-\beta\left[\frac{p^2}{2m}+V(x)\right]\right\}$$

Change the integration variables:

$$u=f(x)=-a+\int_{-a}^{x}dy\ e^{-\beta\tilde{V}(y)}$$

$$dx=e^{\beta\tilde{V}(f^{-1}(u))}du$$

Transformed partition function:

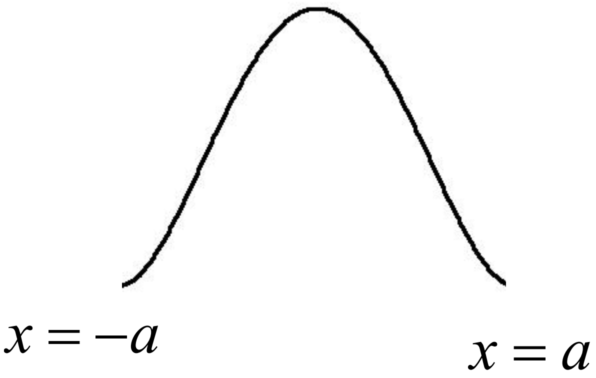
$$Q=\int dp\ du\ \ \exp\left\{-\beta\left[\frac{p^2}{2m}+V\left(f^{-1}(u)\right)-\tilde{V}\left(f^{-1}(u)\right)\right]\right\}$$

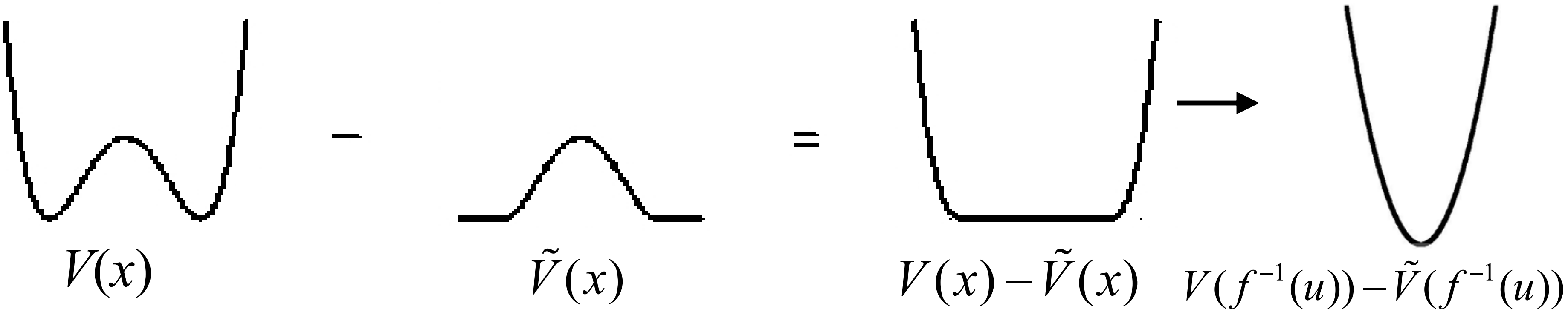
Spatial-warping transformations (cont'd)

$\tilde{V}(x)$

Choosing the transformation:

$$\tilde{V}(x) = \begin{cases} V(x) & -a < x < a \\ 0 & \text{Otherwise} \end{cases}$$





MD: $\dot{u} = \frac{p}{m}$ $\dot{p} = F(u) + \text{heat bath}$

Reference potential spatial warping (REPSWA)

Difference from bias potential methods

Bias potential needs the unbiasing factor:

$$Q = \int dp\ dx\ \exp\left\{-\beta\left[\frac{p^2}{2m} + V\left(x\right) - \tilde{V}\left(x\right)\right]\right\}\exp[-\beta\tilde{V}(x)]$$

Spatial-warping transformations:

$$Q = \int dp\ du\ \exp\left\{-\beta\left[\frac{p^2}{2m} + V\left(f^{-1}(u)\right) - \tilde{V}\left(f^{-1}(u)\right)\right]\right\}$$

System *naturally* moves on the smoother surface. No unbiasing needed.

$V^{\ddagger}$

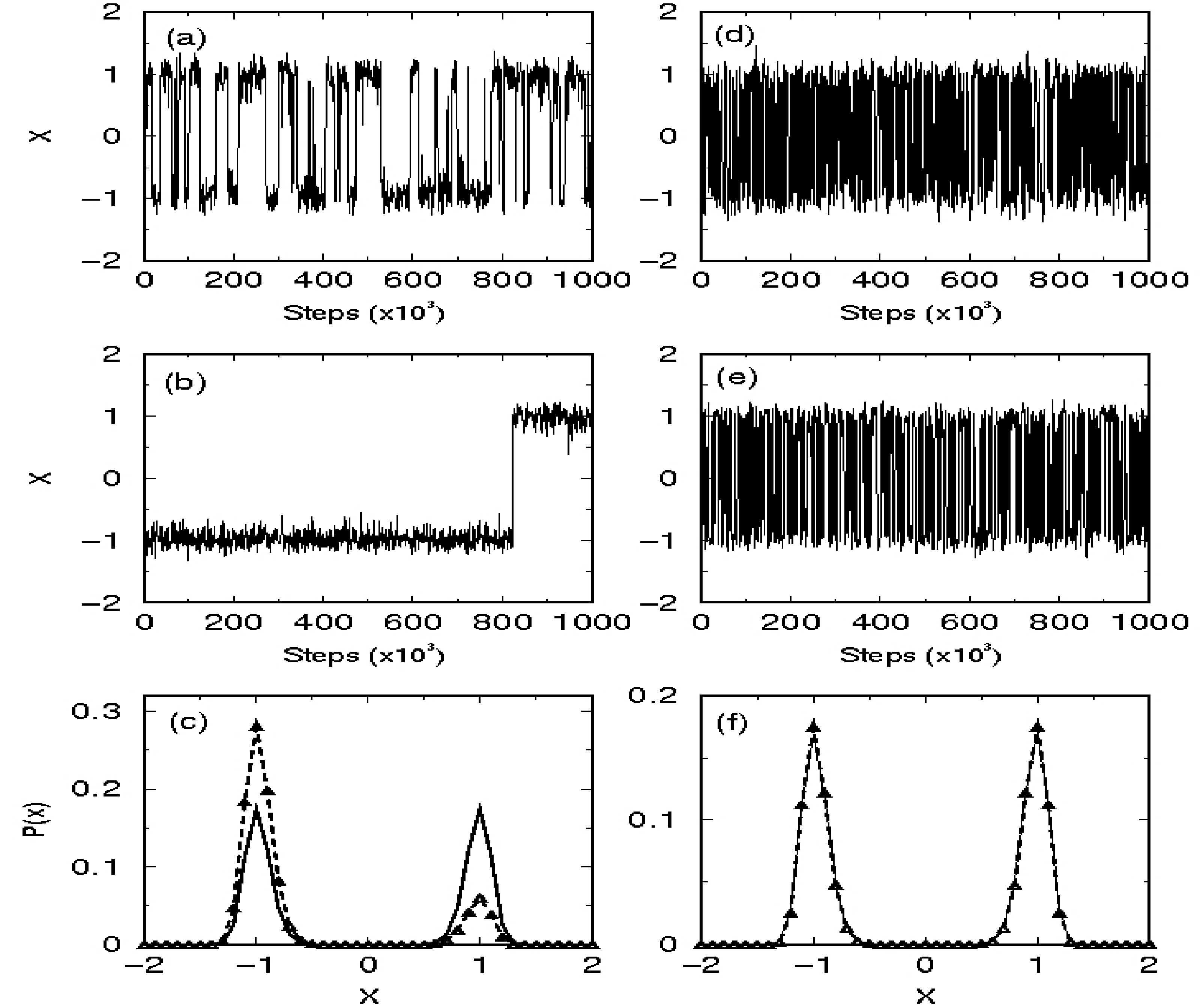
No Transformation

Transformation

5kT

10kT

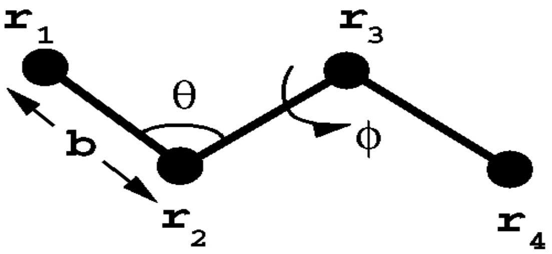
10kT

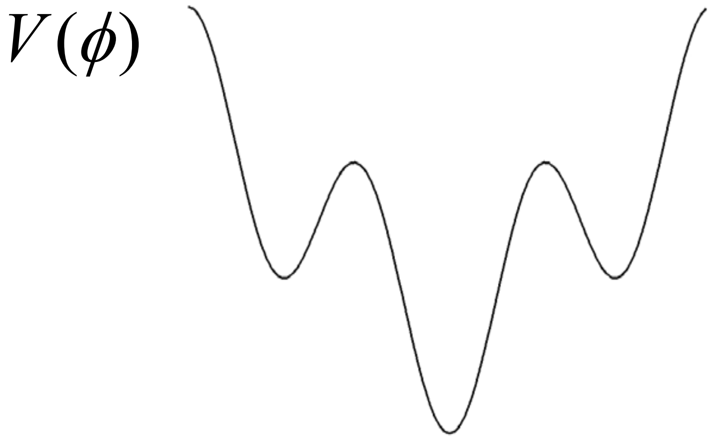


How to do biomolecules and polymers

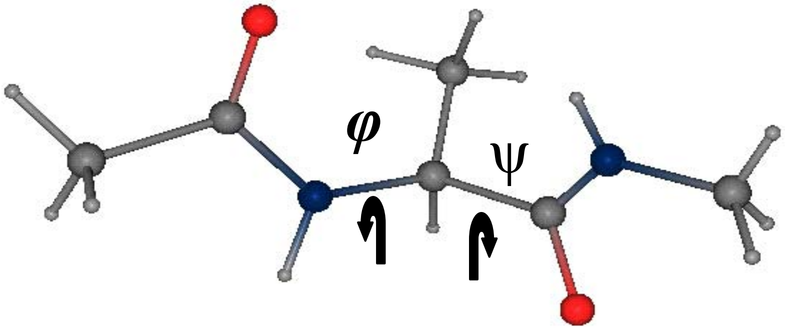
P. Minary, MET, G. J. Martyna *SIAM J. Sci. Comput.* **30**, 2055 (2007)

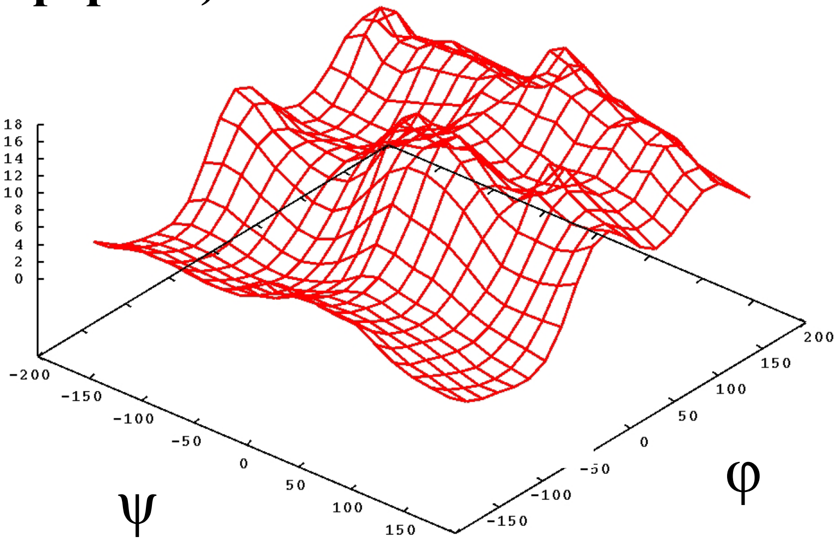
Sampling hindered by barriers in the dihedral angles:



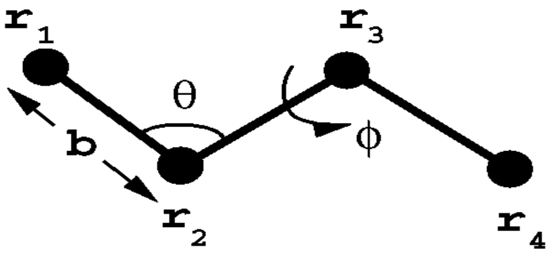


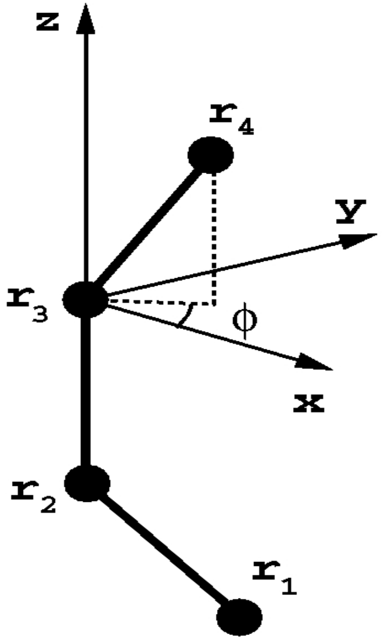
Ramachandran dihedral angles (alanine dipeptide):





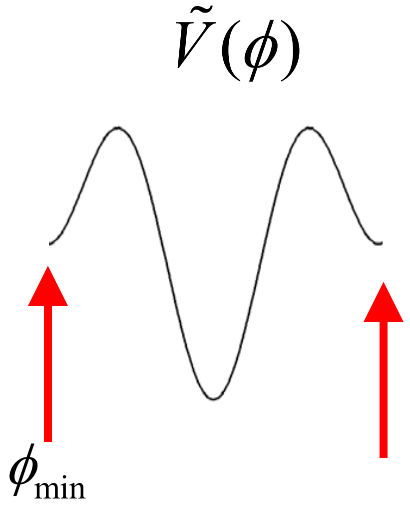
Dihedral angle transformations





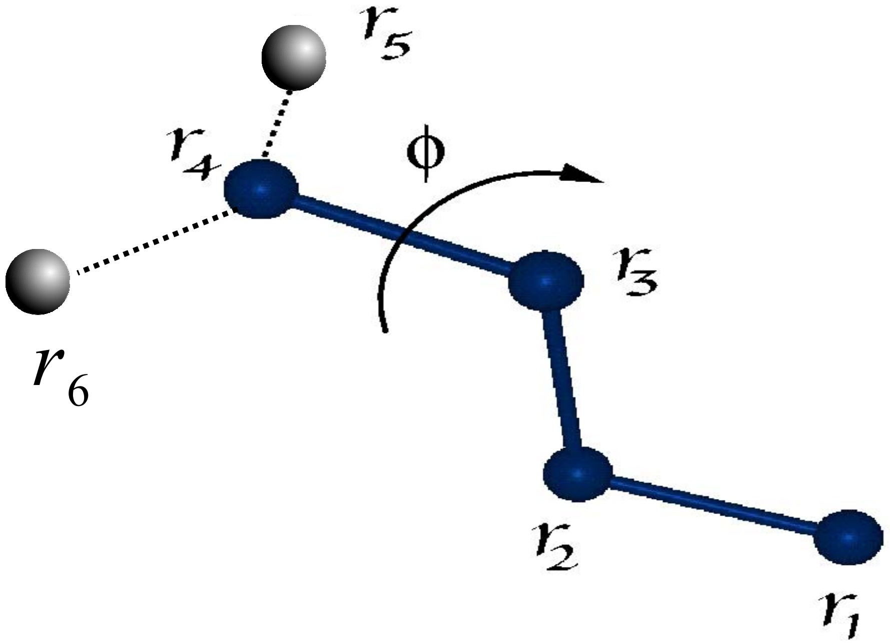
Transformed angle:

$$\phi_u = \phi_{\min} + \int_{\phi_{\min}}^{\phi} d\chi \; e^{-\beta \tilde{V}(\chi)}$$



Dynamic Transformations

P. Minary, G. J. Martyna and MET *SIAM J. Sci. Comp.* **30**, 2055 (2008)



Reference potential:

$$\tilde{V}(\phi,\{\mathbf{r}\}) = V_{\text{tors}}(\phi)S_1(\phi) + \alpha \sum_{i\in \text{nn}} V_{\text{n.b.}}(|\,\mathbf{r}_4(\phi,\{\mathbf{r}\}) - \mathbf{r}_i\,|)S_2(|\,\mathbf{r}_4(\phi,\{\mathbf{r}\}) - \mathbf{r}_i\,|)$$

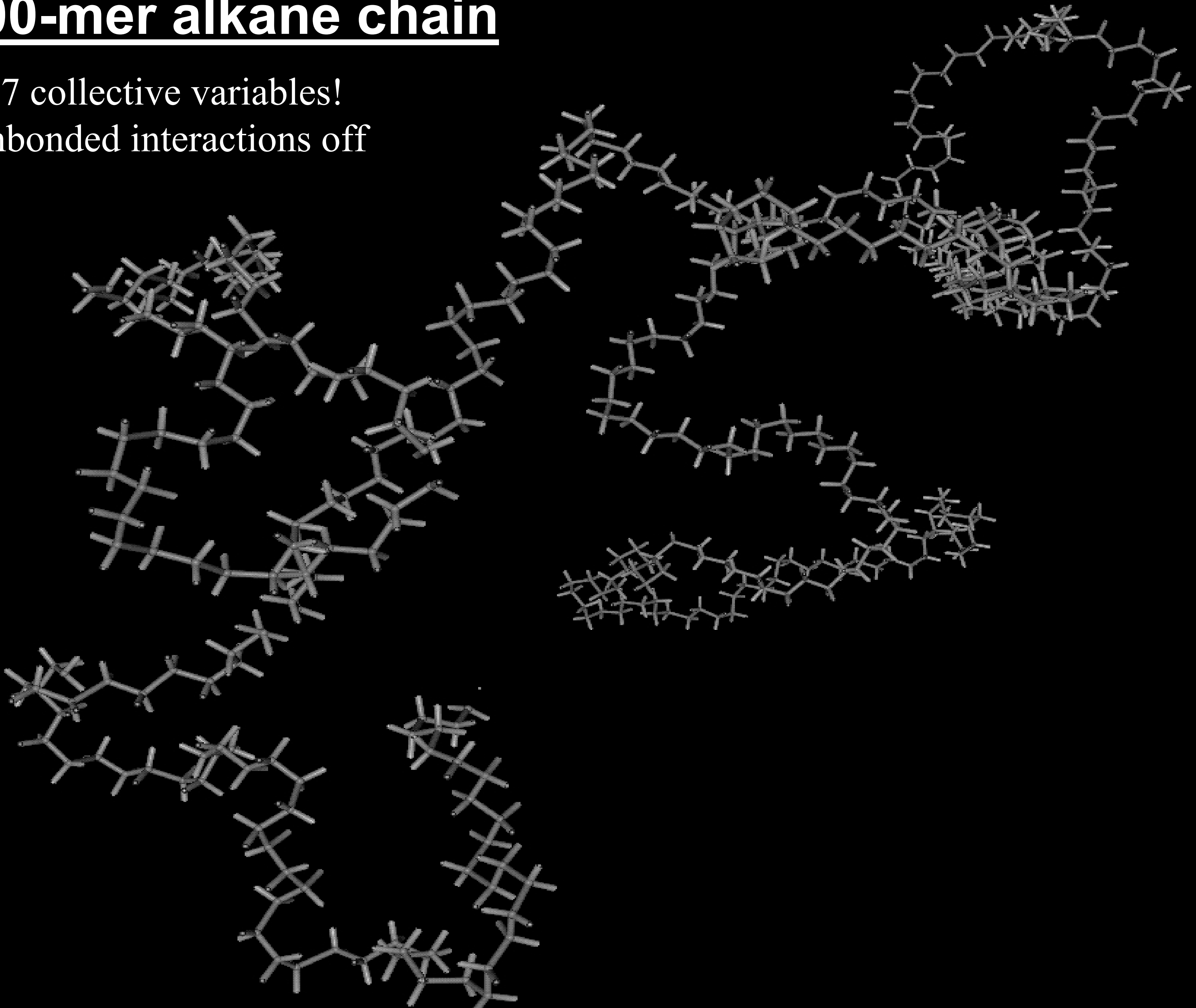
Transformation

$$\phi_u(\{\mathbf{r}\}) = \phi_{\min}(\{\mathbf{r}\}) + c(\{\mathbf{r}\}) \int_{\phi_{\min}(\{\mathbf{r}\})}^{\phi(\{\mathbf{r}\})} d\phi' \; e^{-\beta V_{\text{ref}}(\phi',\{\mathbf{r}\})}$$

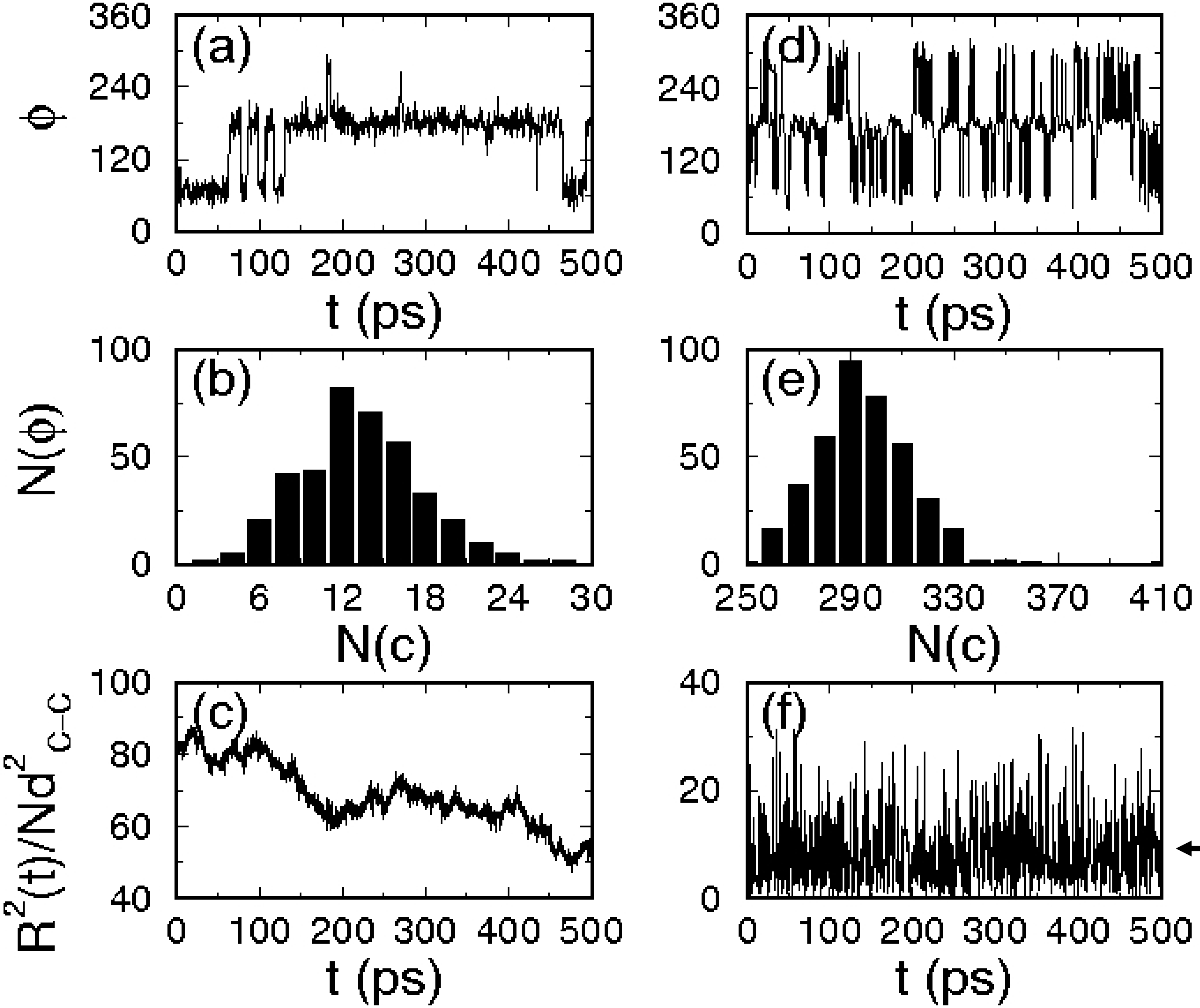
Additional force contributions from $\partial\phi_u/\partial\mathbf{r}_i$.

A 400-mer alkane chain

397 collective variables!
Nonbonded interactions off



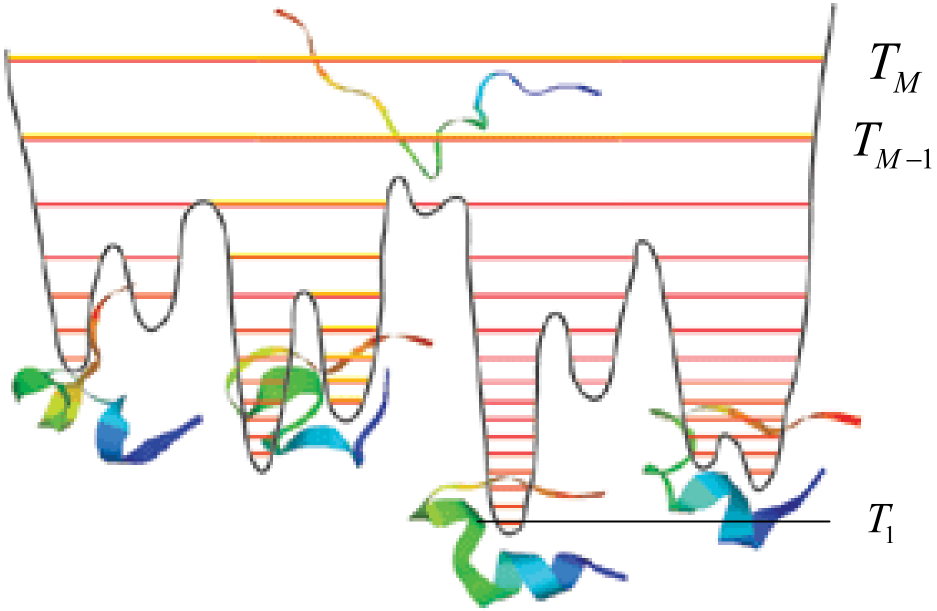
No Transformation **REPSWA Transformations**



RIS Model value: 10

Replica-exchange Monte Carlo (Parallel Tempering)

Swendsen and Wang, *Phys. Rev. Lett.* **57**, 2607 (1986)



Basic Algorithm (Parisi, Whittington, Nemoto)

A. Create M replicas of a system with temperatures $T_1 < T_2 < \cdots T_M$

B. Run molecular dynamics on each replica and periodically attempt neighbor exchanges:

$$A(\mathbf{r}^{k+1}, \mathbf{r}^{k} \mid \mathbf{r}^{k}, \mathbf{r}^{k+1}) = \min\left[1, e^{-\Delta_{k,k+1}}\right]$$

$$\Delta_{k,k+1} = \left(\beta_k - \beta_{k+1}\right)\left[V(\mathbf{r}^{k}) - V(\mathbf{r}^{k+1)})\right]$$

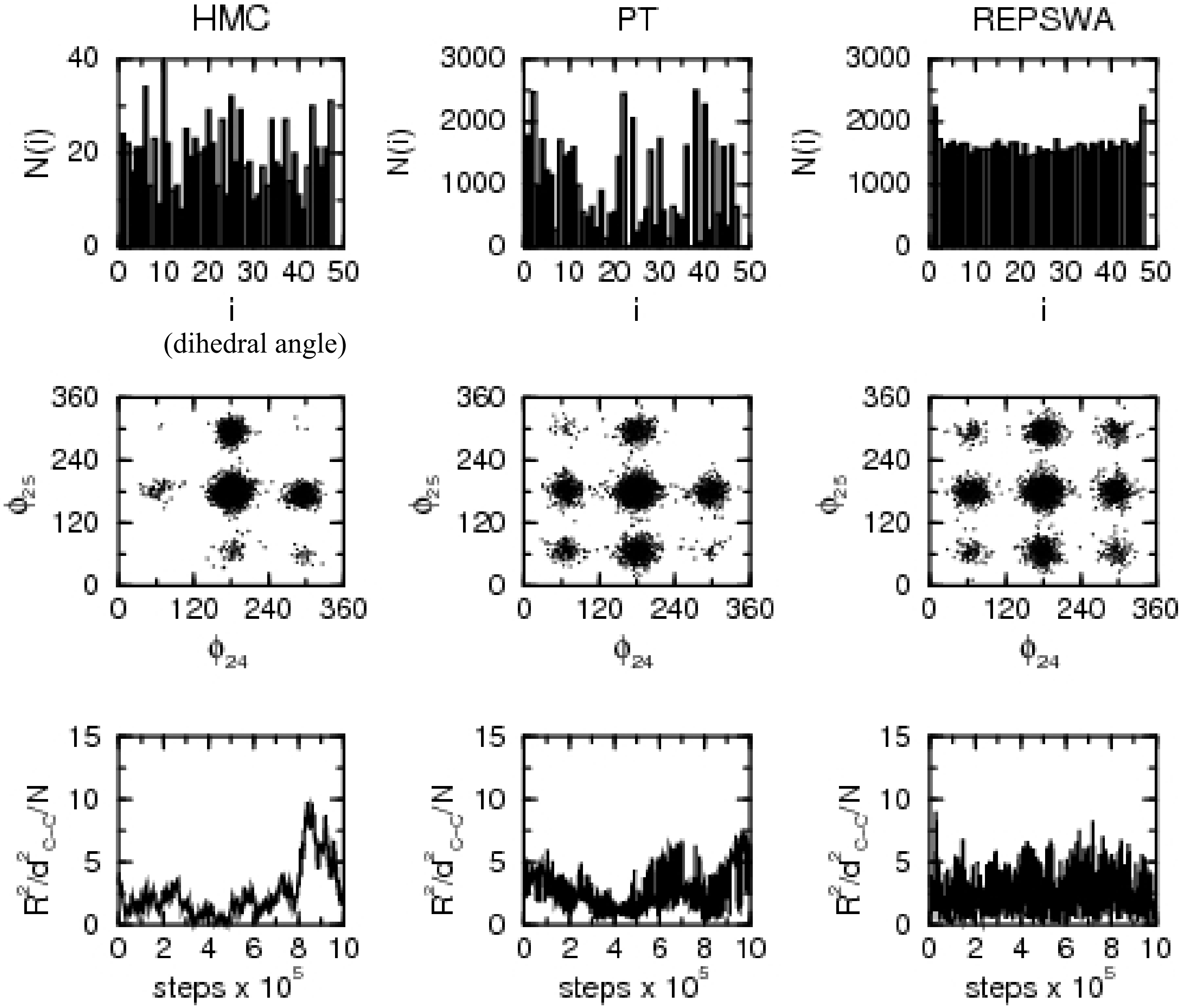
10-20 Replicas
5%-20% Exchange prob.

Comparison for a 50-mer chain using TraPPE

Martin and Siepmann, *J. Phys. Chem. B* **102**, 2569 (1998)

PT replicas = 10; 300 < T <1000 PT exchange prob. = 5%

Nonbonded interactions off

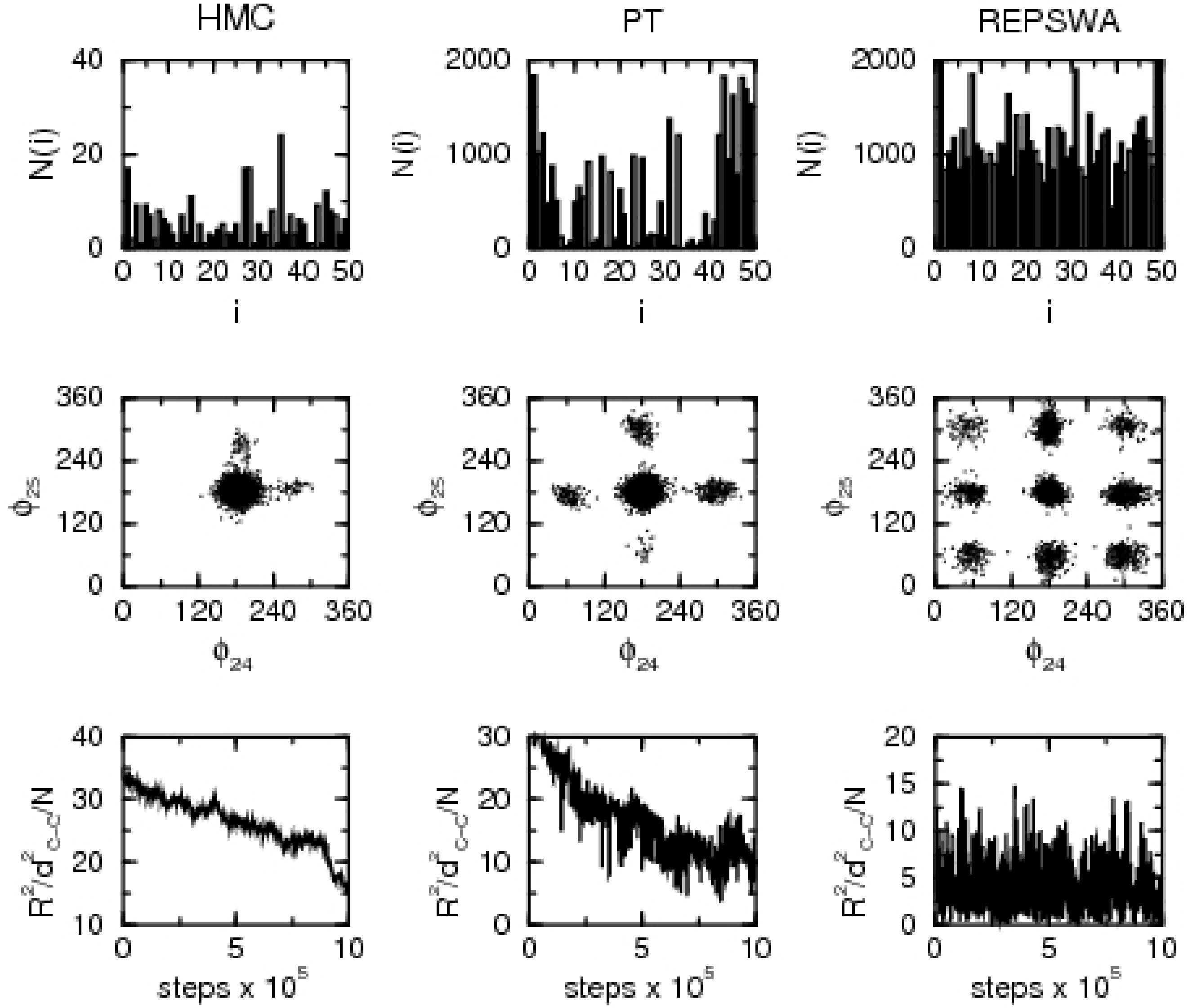


(dihedral angle)

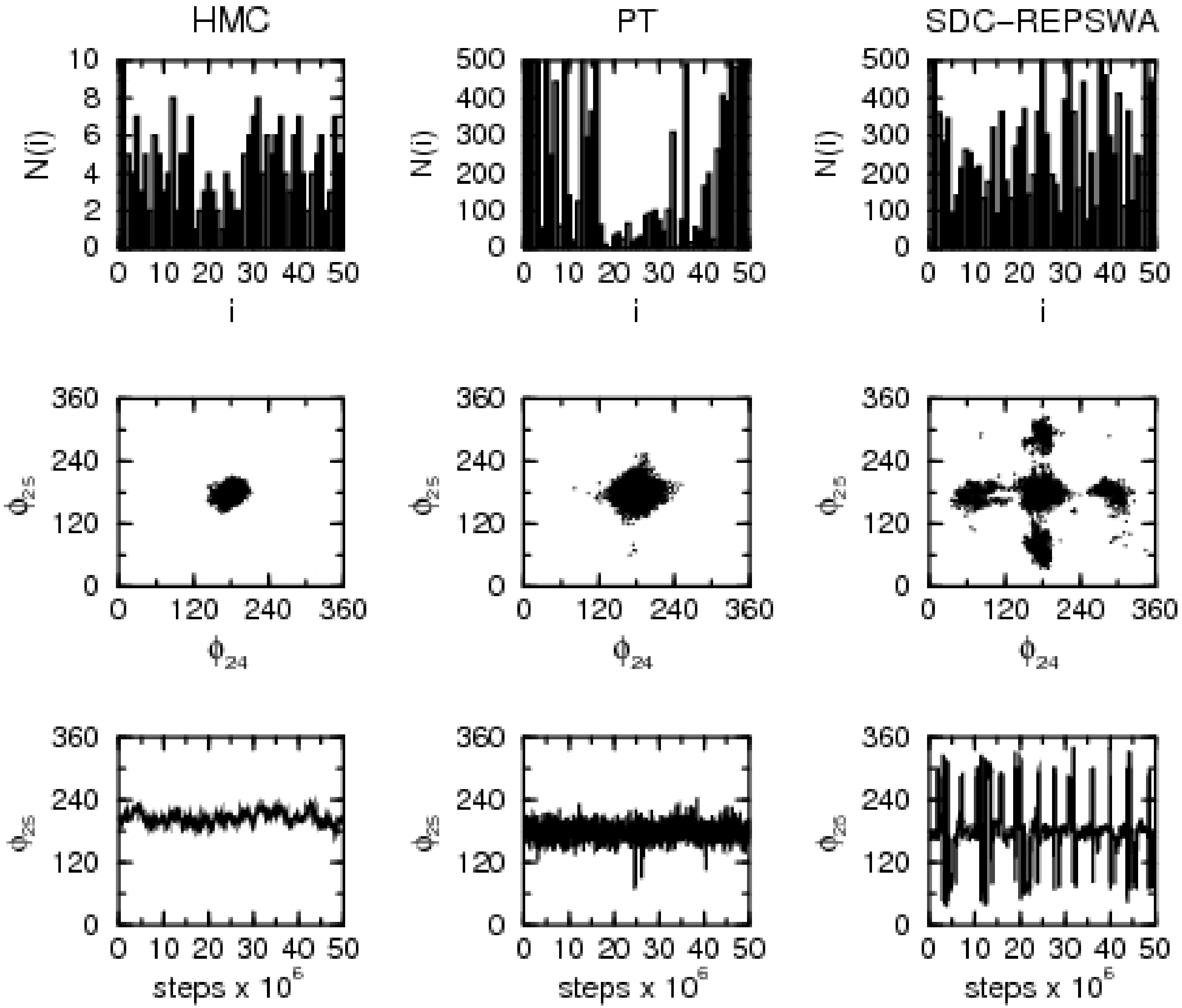
Comparison for a 50-mer chain using CHARMM22

REMC/PT replicas = 10; 300 < T < 1000

Nonbonded interactions off



Comparison for 50-mer using CHARMM22 all interactions



Conclusions and Questions

1. Spatial-warping transformations lead to substantial gains in conformational sampling over ordinary MD and REMC.

   Future: Combine two approaches to gain advantages of both
   Still need transformations for hydrogen bonds.

2. Adiabatic dynamics can be used to explore and generate free-energy hypersurfaces efficiently.

3. Use of external variables [TAMD (Maragliano and Vanden-Eijnden)/ d-AFED (Abrams and Tuckerman)] eliminates the need for explicit variable transformations.

4. Simple mass scaling leads to some enhancement of conformational sampling

Future: Combine REMC with REPSWA – advantages of both (I-Chun Lin).
How does light water help in all of these methods?