

**Dynamics and thermodynamics of large conformational rearrangements in confined polymers**

# Lei Huang

\$\$\$:

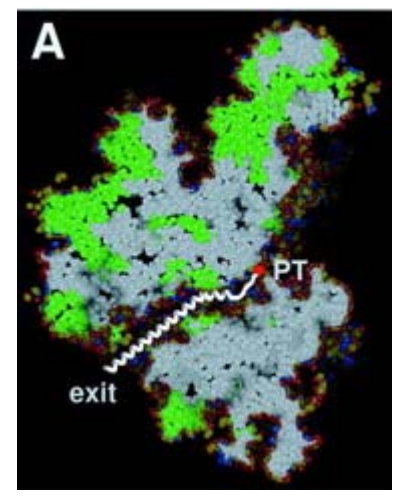
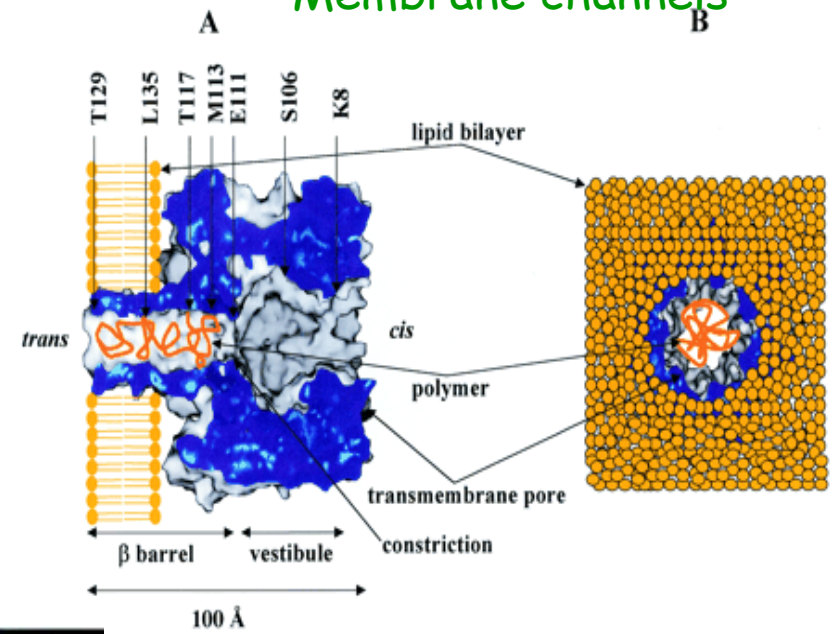
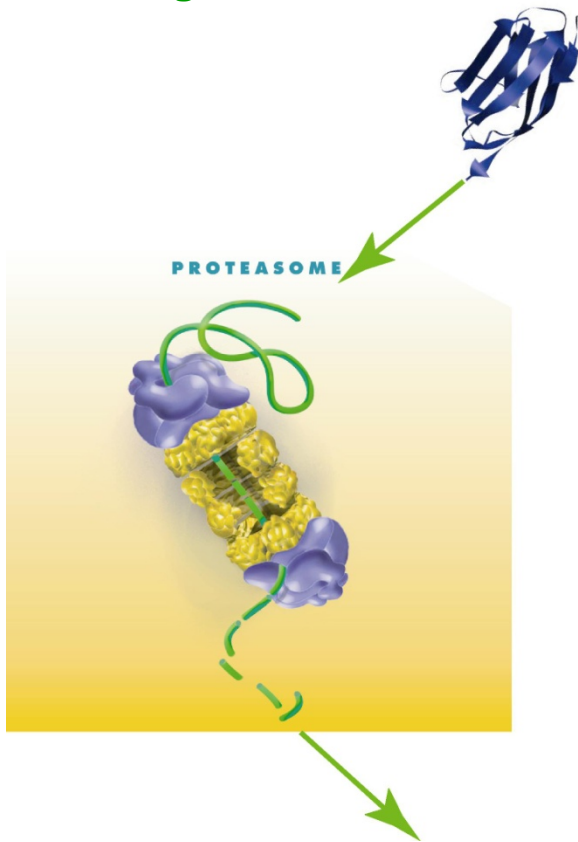
NSF, Welch Foundation,

Texas Advanced Computing Center (CPU)

# Examples of protein translocation:

## Membrane channels

### Protein degradation\*

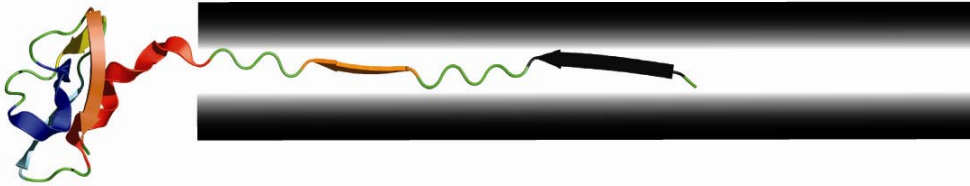


### Protein synthesis\*\*

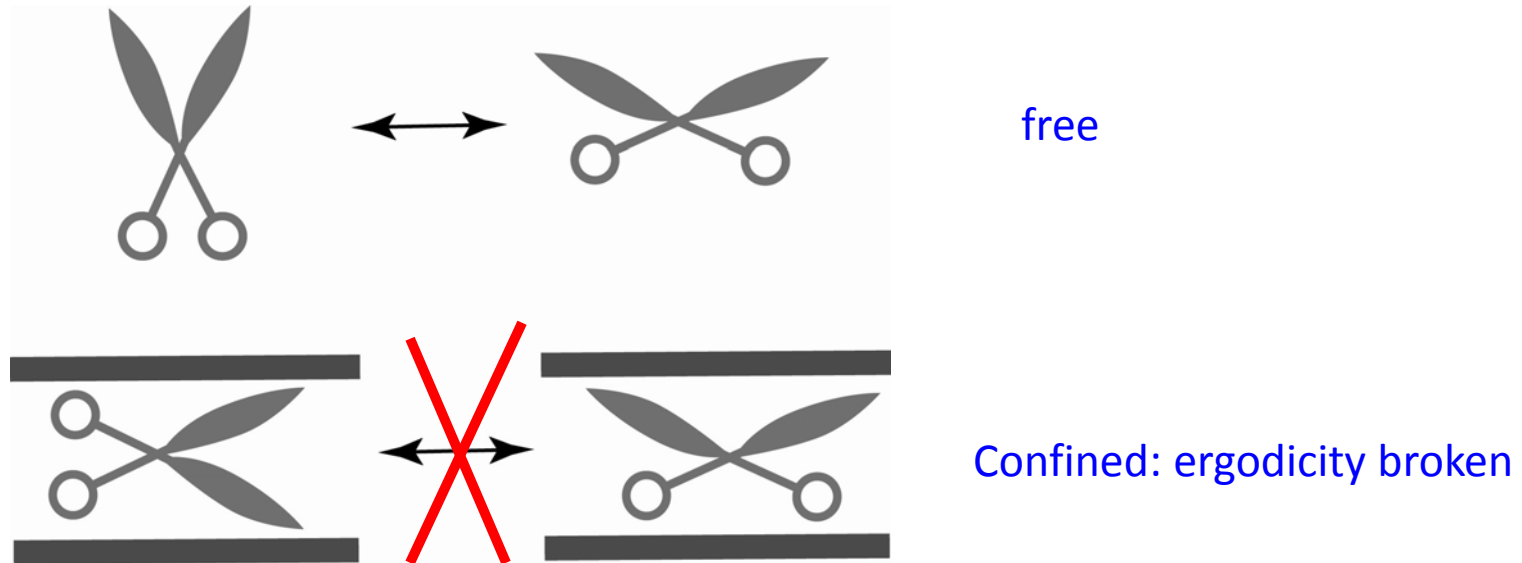
\* image credit: U.S. Department of Energy Genomics:GTL Program, <http://doegenomestolive.org>.

\*\* Nissen, P.; Hansen, H.; Ban, N.; Moore, P. B.; Steitz, T. *Science* 2000, 289, 920

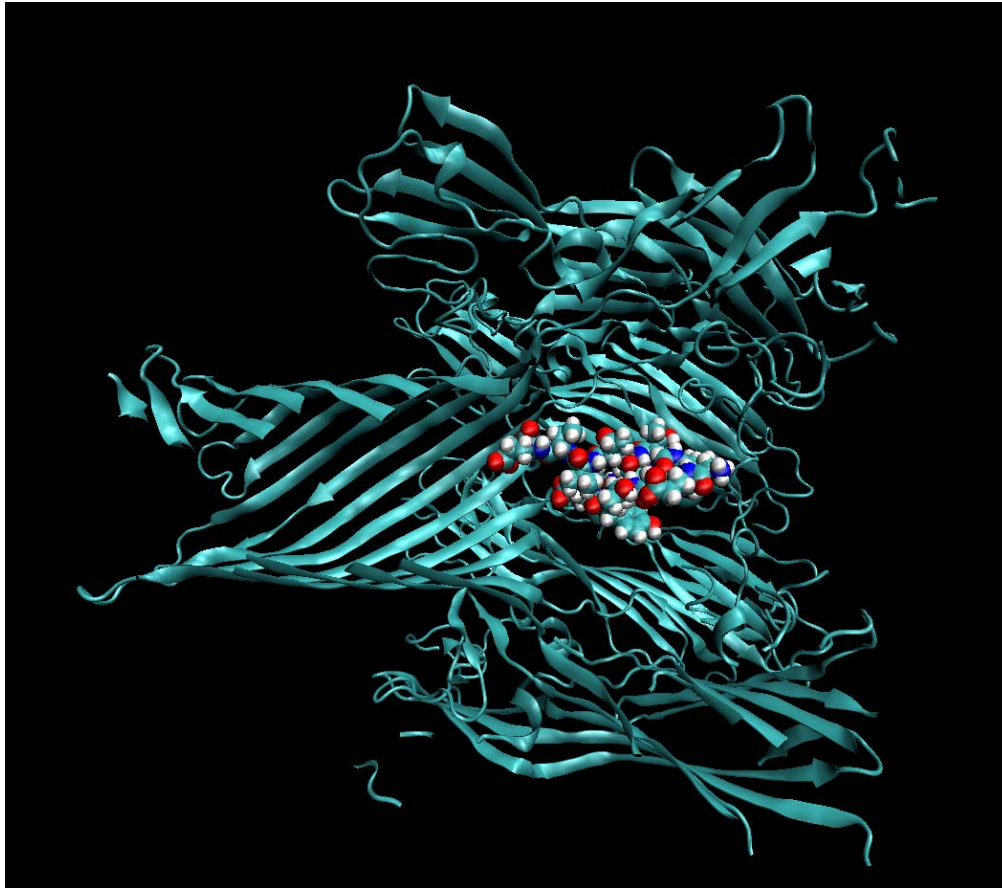
Confinement in pores introduces multiple metastable states separated by large barriers (e.g., proteins unfold in pores via mechanisms as opposed to all-or-none unfolding).



- Tian P, Andricioaei I: *J Mol Biol* 2005, **350**:1017-1034
- Huang L, Kirmizialtin S, Makarov DE: *J. Chem. Phys.* 2005, **123**:124903
- Dmitrii E. Makarov, *Accounts of Chemical Research* 42 (2009) 281-289.



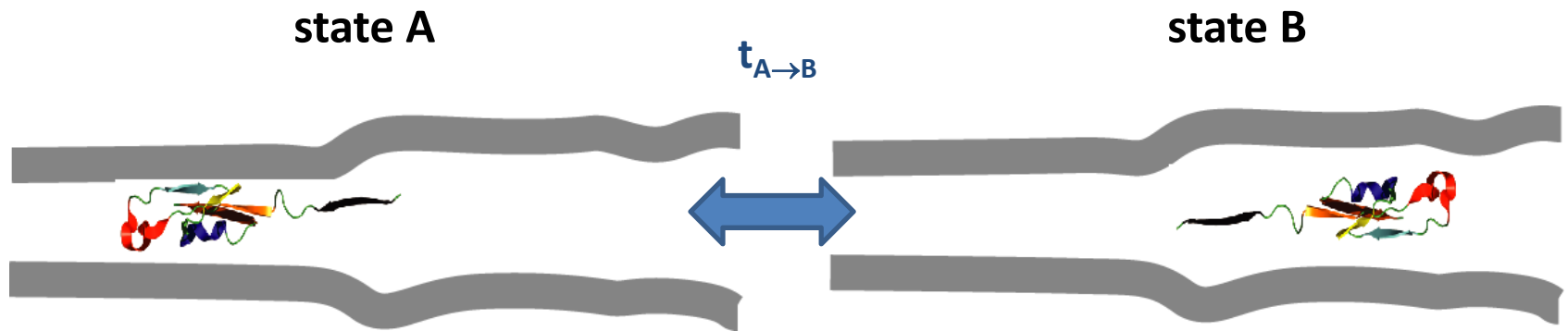
Confinement may thus break ergodicity in protein dynamics, depending on the relative time scales of conformational rearrangements and of the passage through the pore



Kijeong Kwac



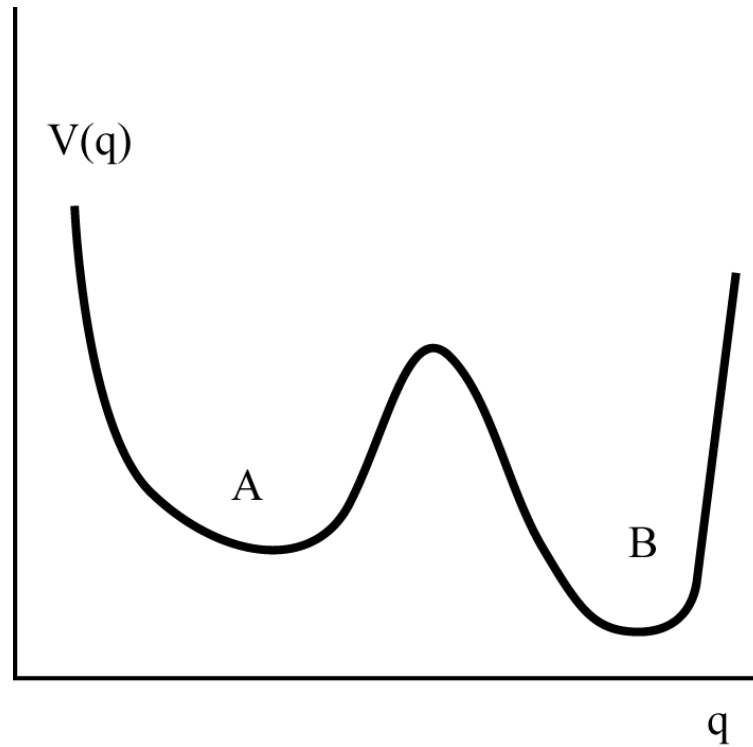
**Does this protein tumble as it goes through the pore?**



## Questions:

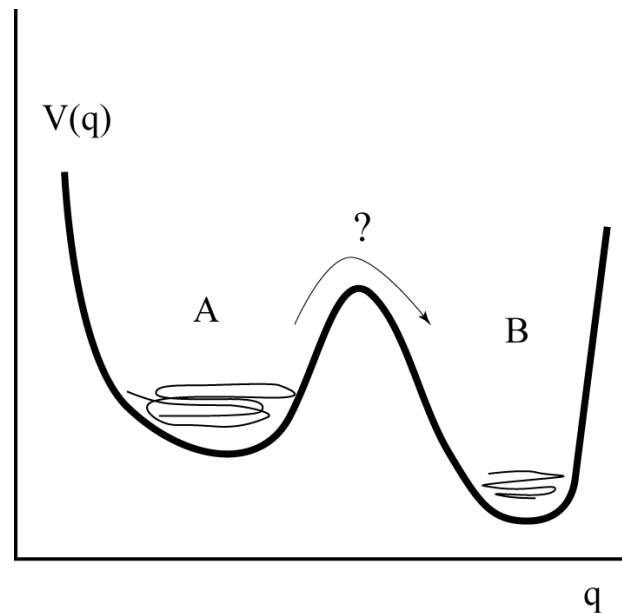
1. What are the populations of different conformational states,  $p_A$  and  $p_B$ ? (i.e., what is the equilibrium constant  $K=p_B/p_A$ ?)
2. What is the time scale of conformational rearrangement  $t_{A \rightarrow B}$ ?

# Calculating equilibrium between A and B



$$K = \frac{p_B}{p_A} = \frac{Q_B}{Q_A} = \frac{\int_B dpdx \exp(-\beta H(p, x))}{\int_A dpdx \exp(-\beta H(p, x))} = \exp\left[-\frac{F_B - F_A}{k_B T}\right] = \exp\left[\frac{S_B - S_A}{k_B}\right] \exp\left[-\frac{\langle E_B \rangle - \langle E_A \rangle}{k_B T}\right]$$

MD or MC



Is it possible to find  $Q_{A(B)}$  from a simulation of A(B) alone? (In other words, we want to post process the simulation data rather than change the way we do simulations)

Method proposed independently in:

- F.M. Ytreberg and D. Zuckerman, JCP 124, 104105 (2006)
- L. Huang and D.E. Makarov, JCP 124, 064108 (2006)

MC or MD sampling is good to compute averages of physical quantities

$$\langle A \rangle = \frac{\int dpdx e^{-\beta H} A(p, x)}{\int dpdx e^{-\beta H}}$$

BUT...

$$Q = \int dpdx e^{-\beta H} \quad \text{or} \quad F = -\beta^{-1} \ln Q$$

are not averages of the above form. Hence the common wisdom:  
one cannot compute absolute  $F$ , only a difference in  $F$   
(thermodynamic integration, Bennett's method, Jarzynski etc.)

How to calculate absolute free energies:

$$\rho(p, x) = Q^{-1} \exp[-H(p, x) / k_B T]$$

$$\int dp dx \rho(p, x) = 1$$

$\rho(p, x)$  is just the histogram of  $p$  and  $x$  obtained from sampling

$$Q^{-1} = e^{F/k_B T} = \rho(p, x) \exp[H(p, x) / k_B T]$$

## Single-point estimate of the partition function

$$Q^{-1} = \exp(F / k_B T) = \rho(p, x) \exp[H(p, x) / k_B T]$$

Bad because the accuracy of the estimated phase-space density is poor in many dimensions and because we are not using all the information available

We can improve this estimate by averaging over phase space points

Window function:  $w(p, x)$ :

$$\int w(p, x) dp dx = 1$$

$$Q^{-1} = \int dp dx w(p, x) \rho(p, x) \exp[H(p, x) / k_B T]$$

$$Q^{-1} = \left\langle w e^{H/k_B T} \right\rangle_{\rho}$$

$$\langle \dots \rangle_{\rho} = \int dp dx \rho(p, x) (\dots)$$

True for any normalized  $w(p, x)$

## Choosing $w(p,x)$ :

$w(x) \sim \delta(x-x_0)$  - reduces to a single point estimate – **BAD**,  
since lots of information is thrown away (Veith et al made  
it work to some extent by choosing  $x_0$  wisely. See **M. Veith, A. Kolinski, and J.  
Skolnick, J. Chem. Phys. 102, 6189 1995)**)

$w(x)$  is highly delocalized – BAD because  $Q^{-1} = \langle w \exp(\beta H) \rangle$   
will be dominated by regions where  $H$  is large (i.e., tails of the  
distribution)

$w(x)$  should be large where  $\rho(x)$  is large.

In fact, if we choose  $w(x) = \rho(x)$  then statistical errors will  
disappear (i.e. even a single sampling point will give the exact result)

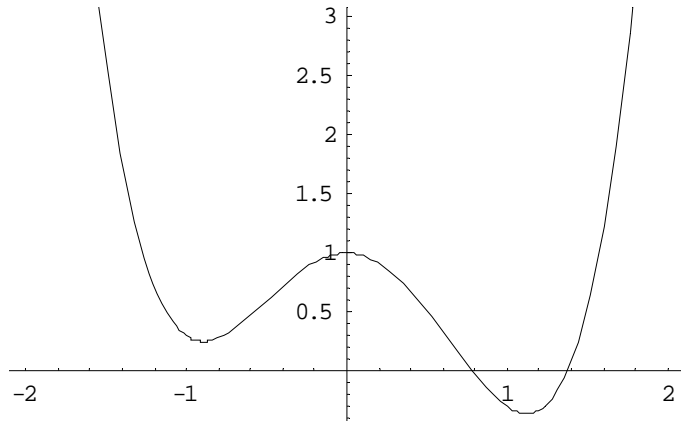
Relationship to the reference state method (see, e.g., F.M. Ytreberg and D. Zuckerman, JCP 124, 104105 (2006) :

If

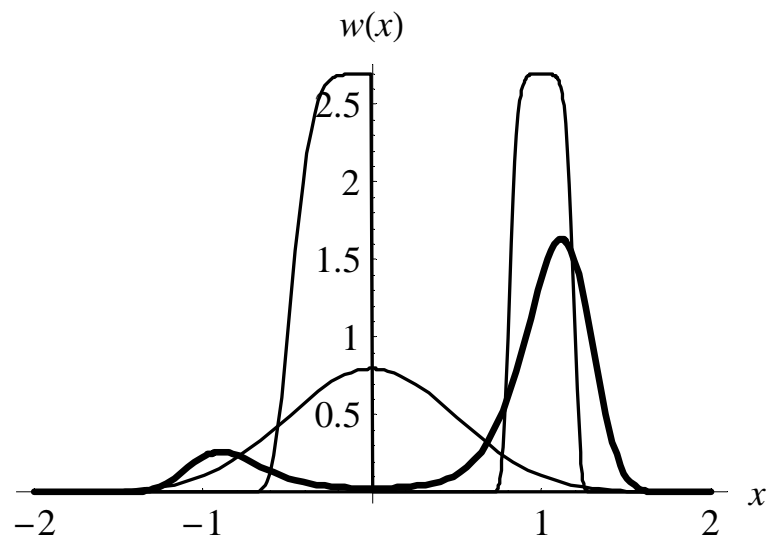
$$w(p,x) = \exp(-H_{\text{ref}}(p,x)/k_B T) / Q_{\text{ref}}$$

then effectively we are computing the free energy difference  $F - F_{\text{ref}}$

In low-dimensional cases even crazy window functions work



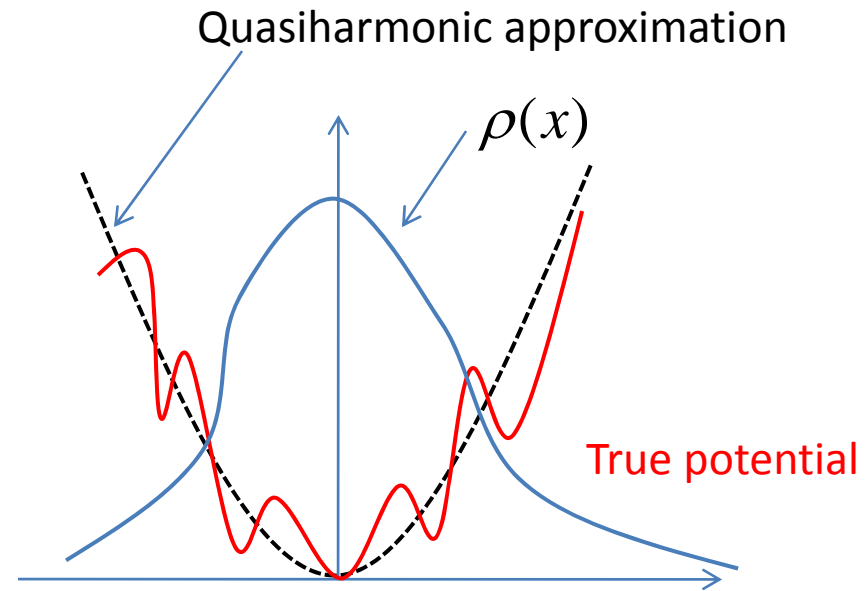
$$U(x) = 3(x^2 - 1)^2 - 0.9x$$



**Gaussian window:**

$$w(x) = \frac{1}{\sqrt{2\pi \langle \Delta x^2 \rangle}} \exp \left[ -\frac{(x - \langle x \rangle)^2}{2 \langle \Delta x^2 \rangle} \right]$$

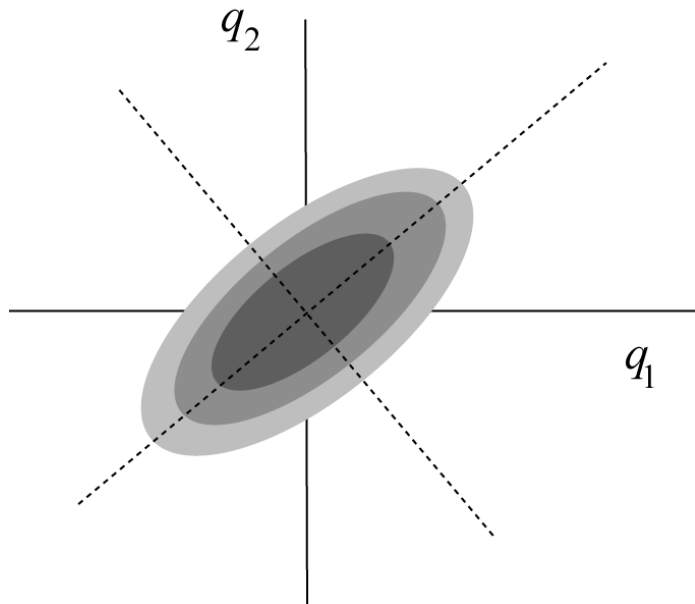
$$\beta m \omega_{\text{eff}}^2 = 1 / \langle \Delta x^2 \rangle$$



only requires knowledge of averages,  $\langle x \rangle$  and  $\langle x^2 \rangle$

## Quasiharmonic approximation in for $D>1$

contour map of  $p(q)$



The best set of coordinates is obtained by diagonalizing:

$$\boldsymbol{\sigma} = \begin{pmatrix} \langle \Delta q_1^2 \rangle & \langle \Delta q_1 \Delta q_2 \rangle & \langle \Delta q_1 \Delta q_3 \rangle & \dots \\ \langle \Delta q_2 \Delta q_1 \rangle & \langle \Delta q_2^2 \rangle & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}$$

Eigenvalues of  $\boldsymbol{\sigma}$  are related to frequencies of effective harmonic modes

$$\exp(S / k_B) \propto (\det \boldsymbol{\sigma})^{-1/2}$$

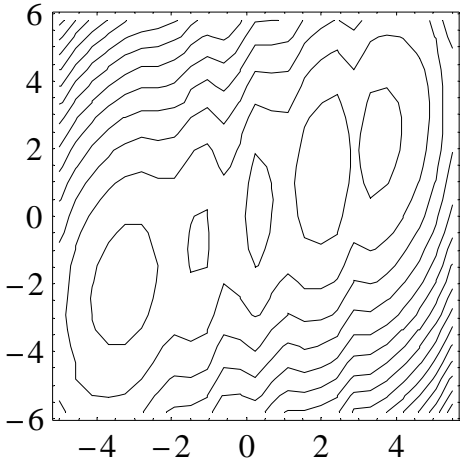
See, e.g., M. Kaplus and J. N. Kushick, *Macromolecules* **14**, 325 1981  
I. Andricioaei and M. Karplus, *J. Chem. Phys.* **115**, 6289 2001.

## Models:

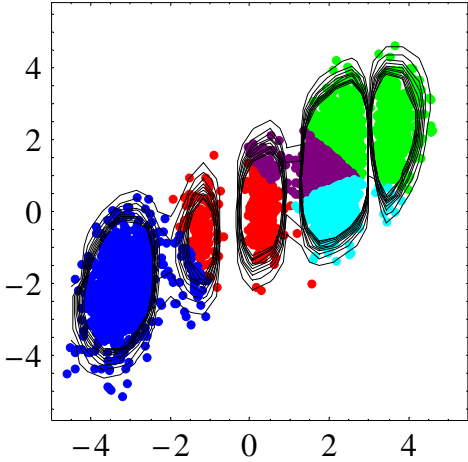
- Ideal gas in 3D consisting of 15 atoms
- 1D potential coupled to 30 harmonic oscillators
- 30D nonlinear oscillator

In each case it was relatively easy to get the exact Q to within an order of magnitude using the Gaussian window, while quasiharmonic approximation itself was 1 – 7 orders of magnitude off

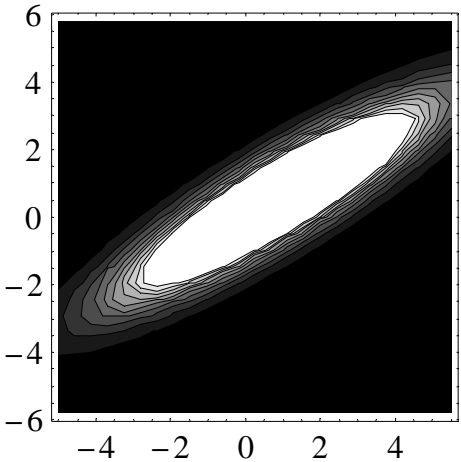
Better window functions: Avoiding poorly sampled regions by identifying clusters of data points



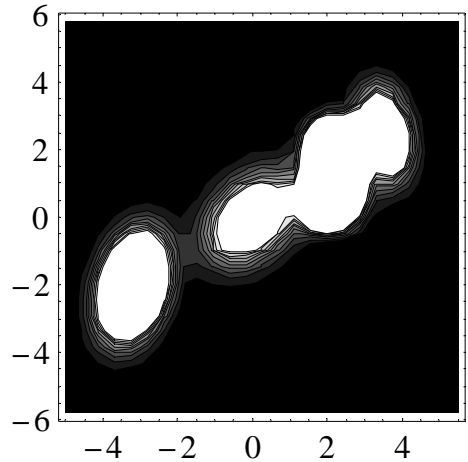
A rough energy landscape



Clustering algorithm breaks up simulated data points into clusters



Optimal single Gaussian window does not exclude poorly sampled, high energy regions



Multi-gaussian window (1 Gaussian per cluster)

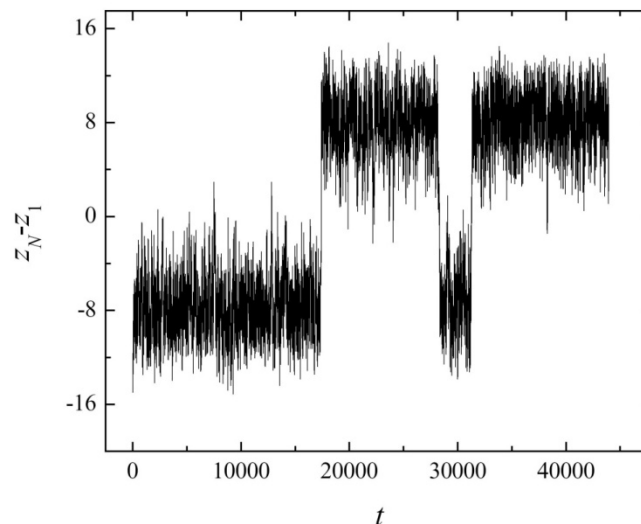
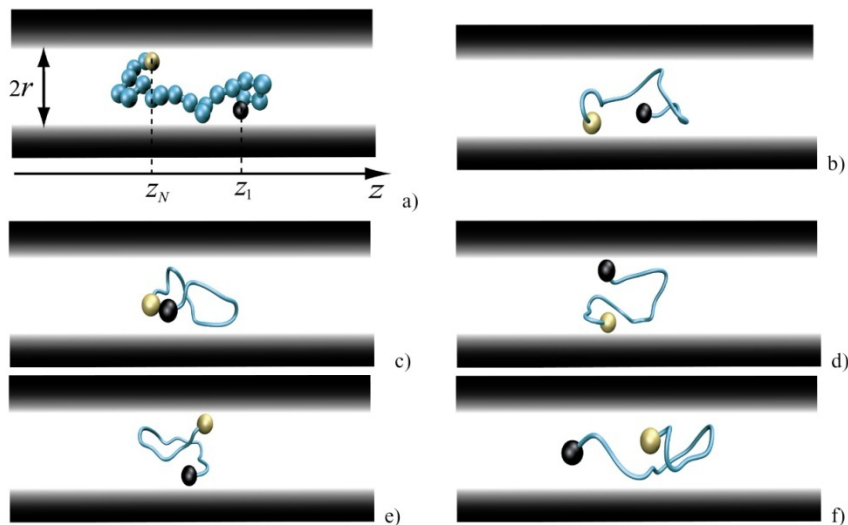
## NOTES

- Clustering algorithms are cheap and don't care much about dimensionality (they are based on pairwise distances between points)
- No free lunch (method is as good as the sampling is). But the method addresses the question how to construct the best free energy estimate given available information

Time scales of rearrangement between confined states

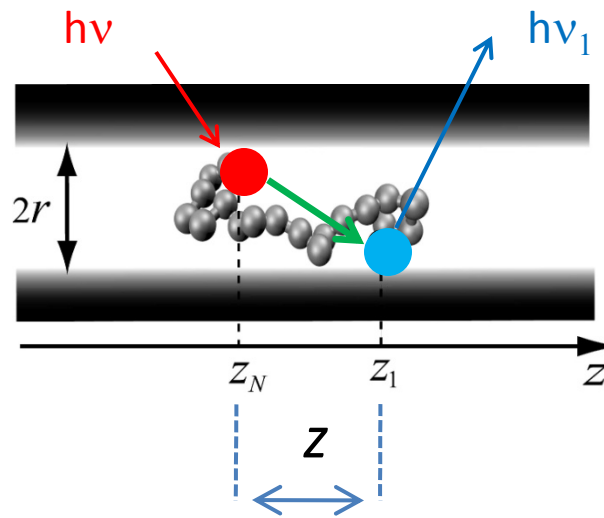
## A toy model: Polymer reversal inside a tube:

Lei Huang and Dmitrii E. Makarov, J. Chem. Phys. 128 (2008) 114903



### “complex” transition:

- Reactant and products are not well defined structures
- no single energy barrier/pathway dominates - some kind of path sampling is needed
- no obvious simple approximate solution (such as harmonic TST)
- Barrier is entropic
- No obviously successful reaction coordinate is anticipated



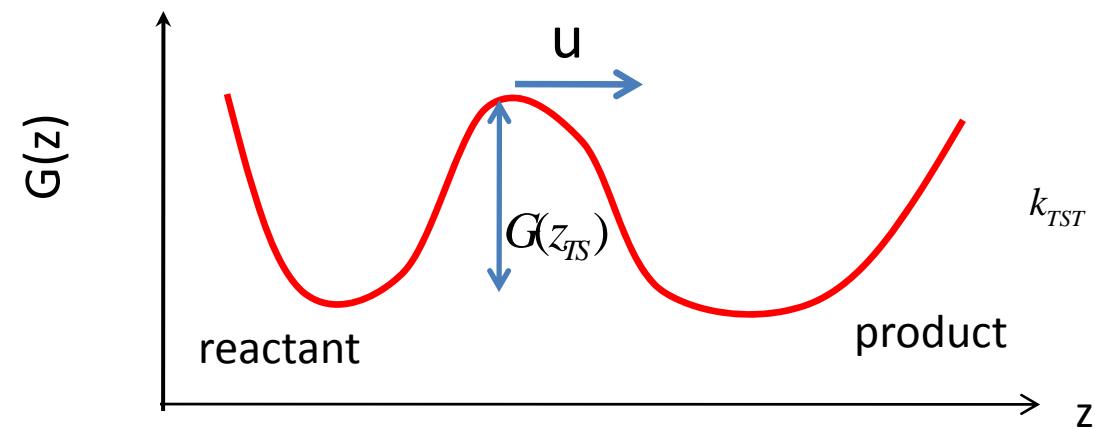
## QUESTIONS

1. How does the reversal rate depend on the polymer length  $N$  and the pore radius  $r$ ?
2. Can we interpret translocation dynamics as 1D motion in an effective potential of mean force  $G(z)$ ?

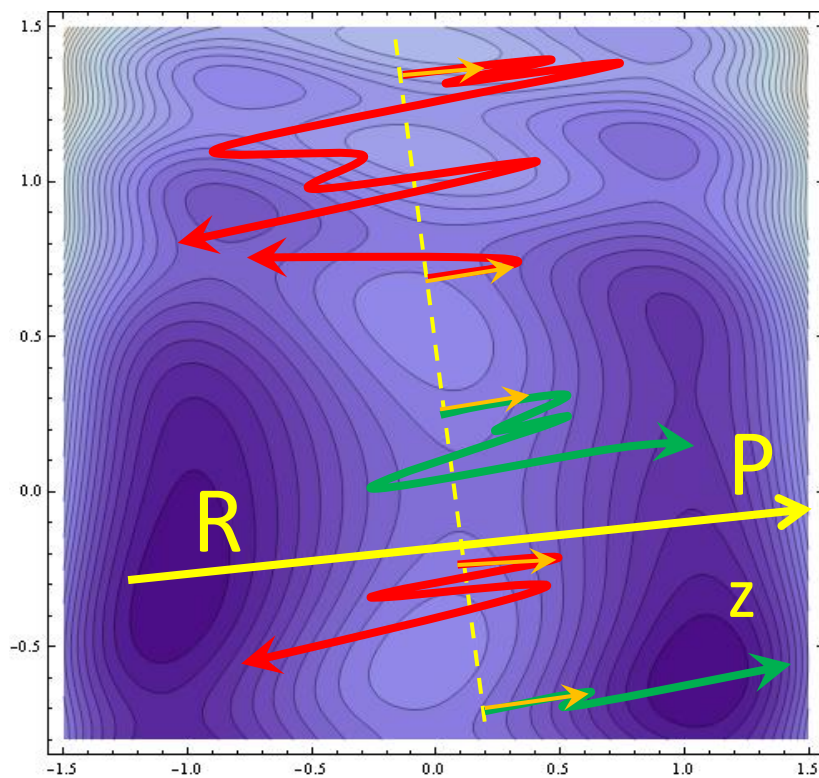
Why care about potentials of mean force and the 1D picture?

1. Blame it on Jarzynski
2. Experiments measure 1D projection of dynamics (e.g., FRET), we need to understand them
3. Hope to have simple approximations and to avoid expensive simulations

# 1D view: TST and beyond



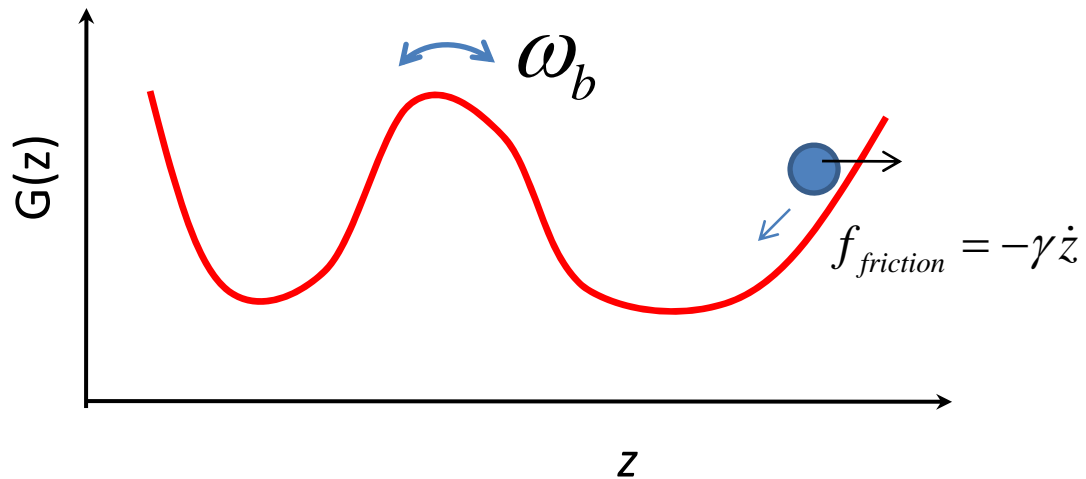
$$k_{TST} = \frac{\langle |u| \rangle}{2} \left\{ \int_{\text{reactants}} \exp[-G(z)/k_B T] \right\}^{-1} \exp\left[-\frac{G(z_{TS})}{k_B T}\right]$$



$$k = \kappa k_{TST}$$

transmission factor

Simple (but approximate) estimates of the transmission factor exist:  
E.g. Kramers' theory



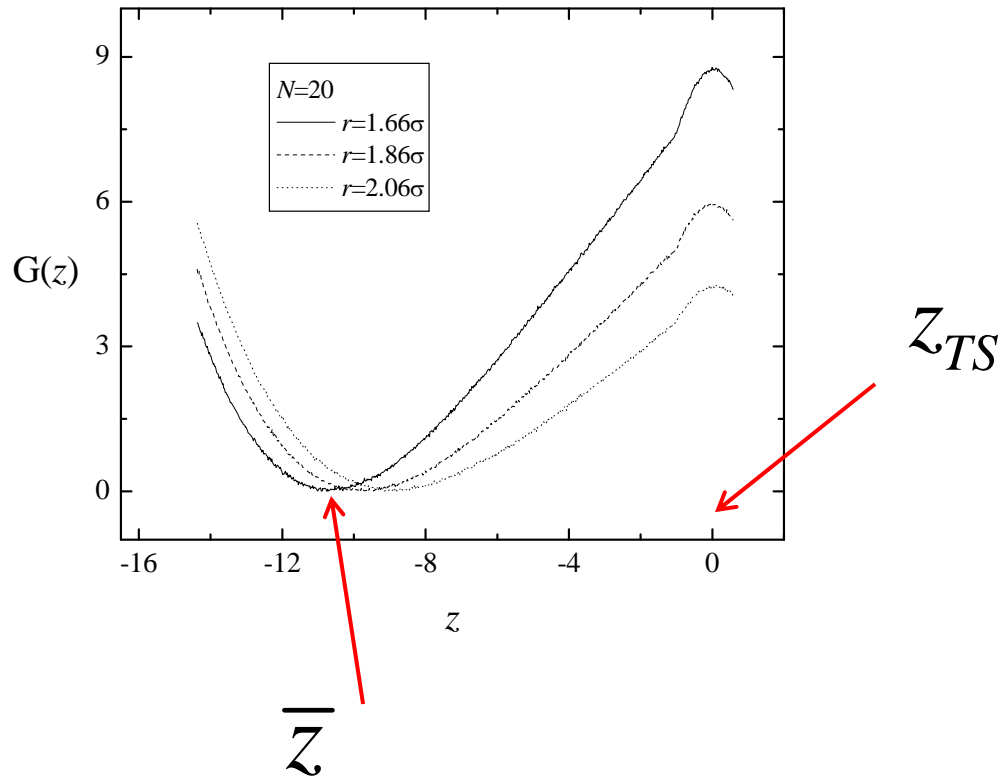
$$k_{Kramers} \approx \frac{\omega_b}{\gamma} k_{TST}$$

Transmission factor

Effective friction force along the reaction coordinate

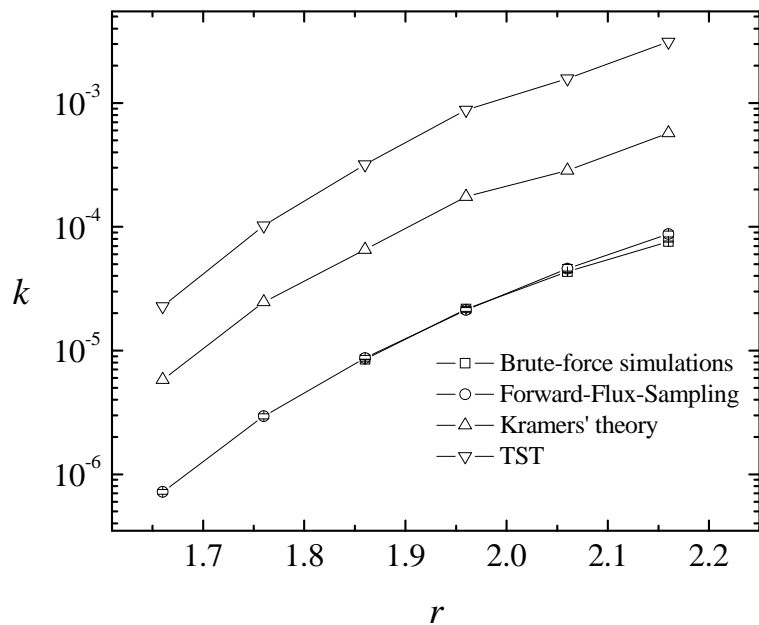
Reaction coordinate:  $z = z_N - z_1$

Potential of mean force for a polymer in a pore

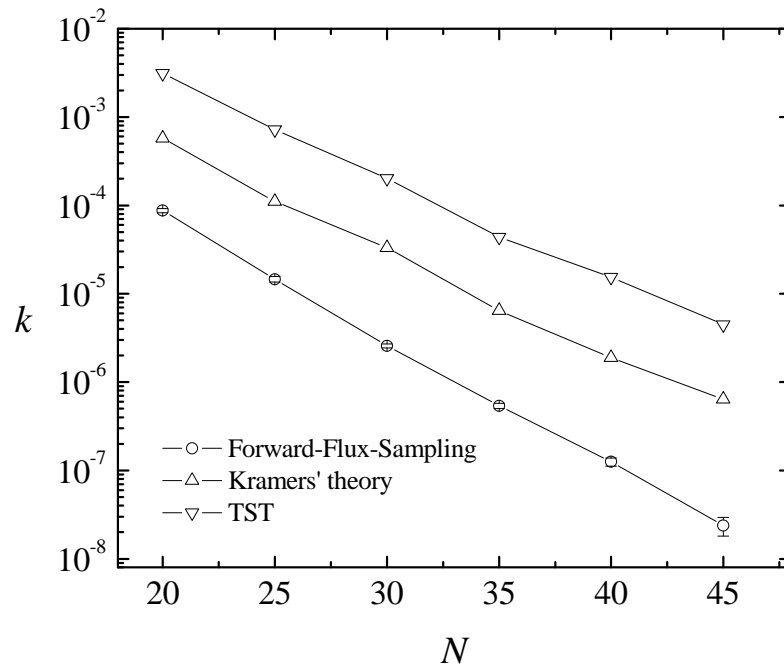


Interestingly, the barrier exists only when excluded effects are present

## Reversal rate dependence on the pore radius



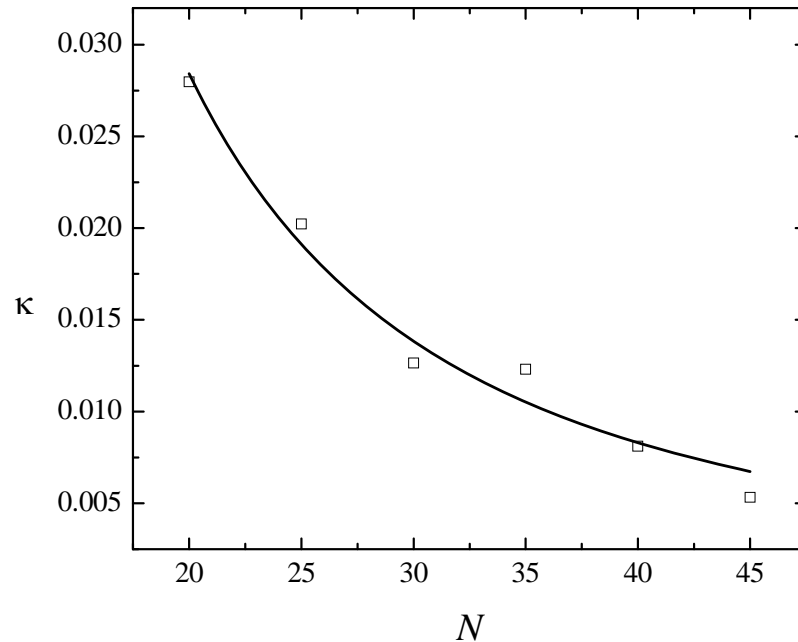
## Reversal rate dependence on the polymer length



$$\mathcal{K} = \frac{k}{k_{TST}} \ll \mathcal{K}_{Kramers}$$

$$k_{exact} = \kappa k_{TST}$$

## Scaling of the transmission factor $\kappa$

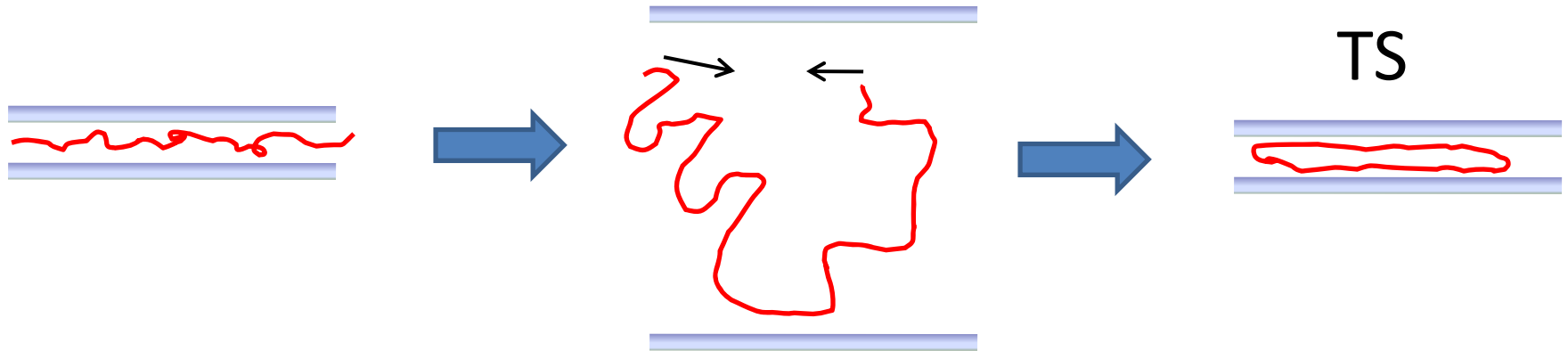


$$\kappa \propto N^{-1.8}$$

Kramers' theory predicts

$$\kappa \approx \omega_b / \gamma \propto N^{-1}$$

# The scaling of the free energy barrier:



$$\Delta G_{TS} \approx G_{ring} + \Delta G_{loop} - G_{linear} \sim G_{ring} - G_{linear}$$

$\uparrow$   
 $\propto \ln N$

$\swarrow \quad \searrow$   
 $\propto k_B T N / r^{5/3}$

De Gennes and other classics

# Summary

$$\begin{array}{ccc} \propto N^{-1.8} & & \propto N / r^{5/3} \\ \searrow & & \downarrow \\ k = \kappa(N, r) \times v_{TST}(N, r) \times \exp\left[-\frac{\Delta G_{TS}(N, r)}{k_B T}\right] & & \\ \nearrow & & \\ \propto N^{-0.37} r^{-1.11} & & \end{array}$$

- 1D TST/Kramers' theory account for the exponentially strong dependence of  $k(N, r)$ . The prefactor has a weaker (power law) dependence on  $N$  and  $r$ .
- Kramers' theory with constant friction does not provide a good description of barrier recrossings/transmission factor

Finally, a challenge:

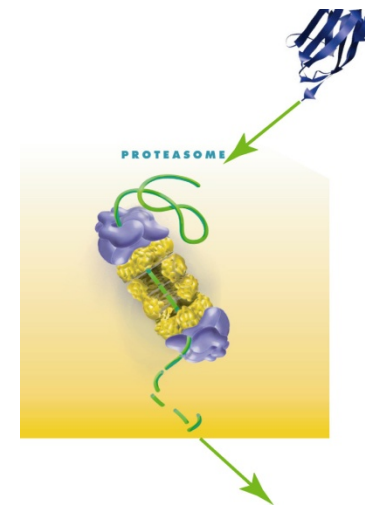
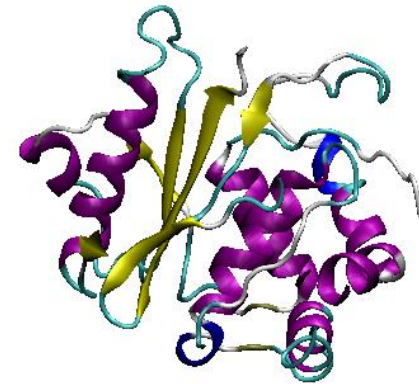
# Translocation of knotted proteins

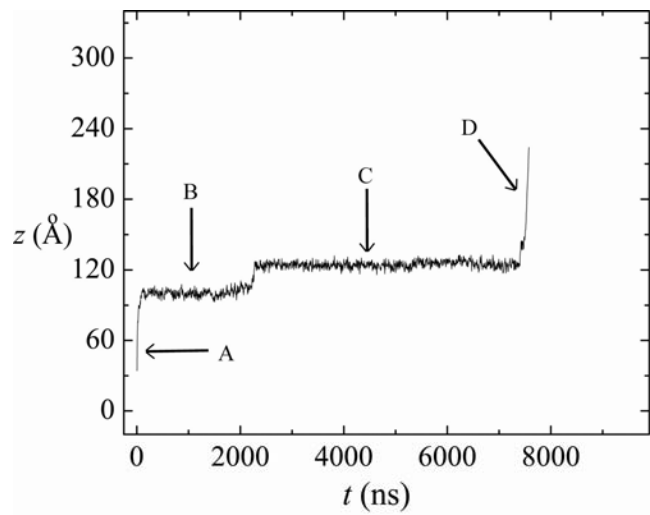
- Very few proteins (~36) are knotted
- Most are enzymes
- Whether knots are biologically important is unknown

The structure of human ubiquitin hydrolase contains a complex knot

Virnau, P.; Mirny, L. A.; Kardar, M. *PLoS Comput Biol* 2006, 2, e122

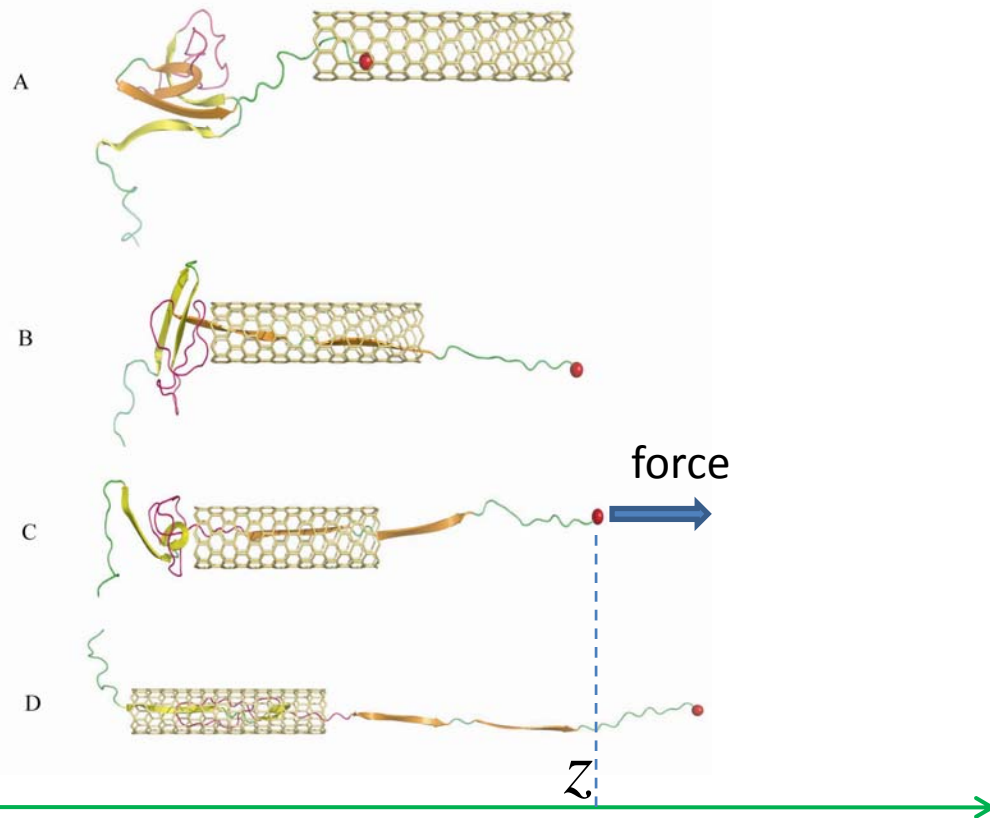
One hypothesis is that knots make proteins resistant to degradation by the proteasome



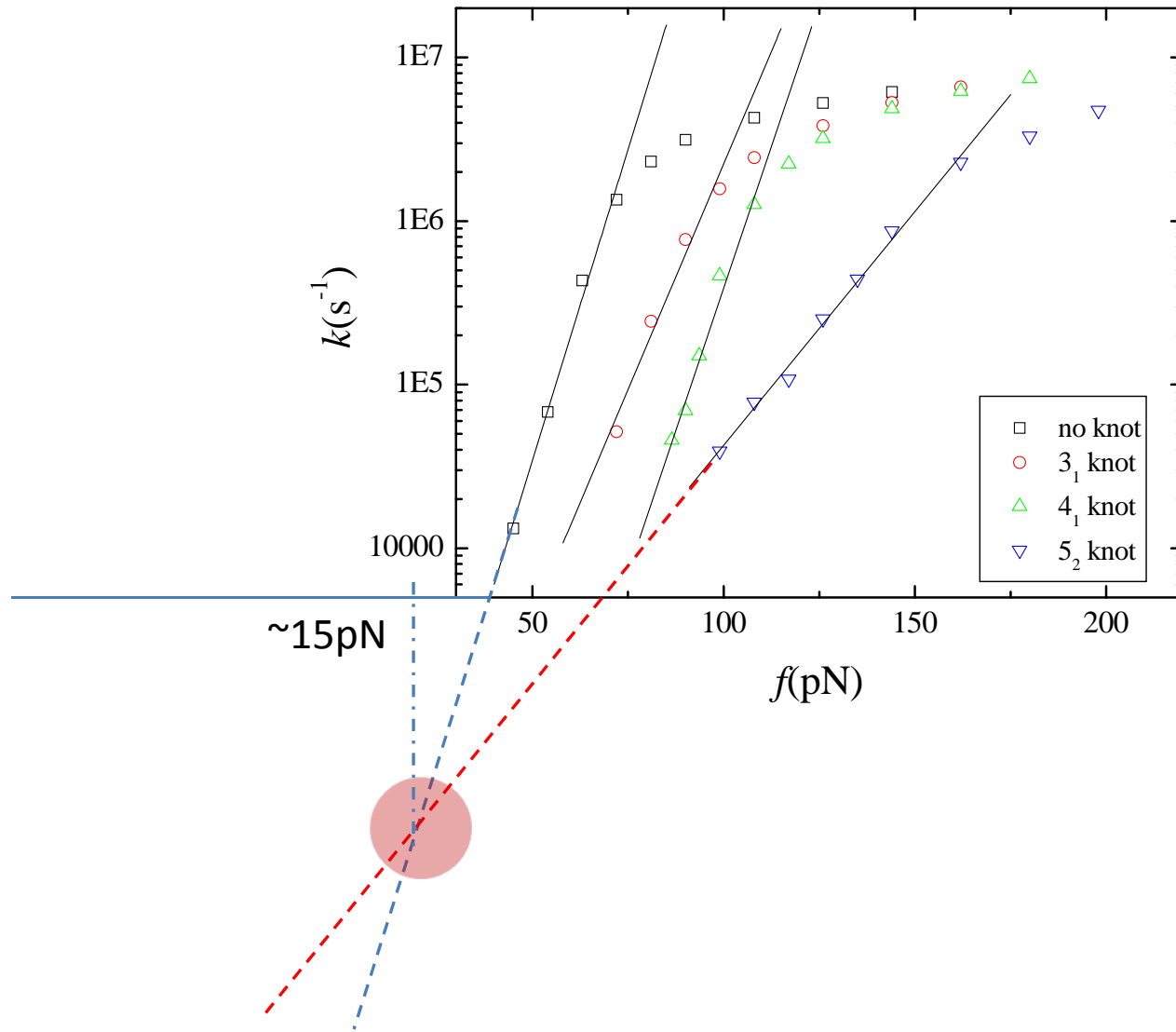


Translocation of knotted proteins is a complex process involving multiple intermediates (knot slippage events), which vary from run to run.

Lei Huang and Dmitrii E. Makarov, *J. Chem. Phys.* 129 (2008), 121107.



# Translocation rate for different types of knots



- Translocation in the low-force regime ( $\sim 10$  pN) must have a different barrier
- This regime is so far inaccessible – Is anyone interested in solving this problem?

