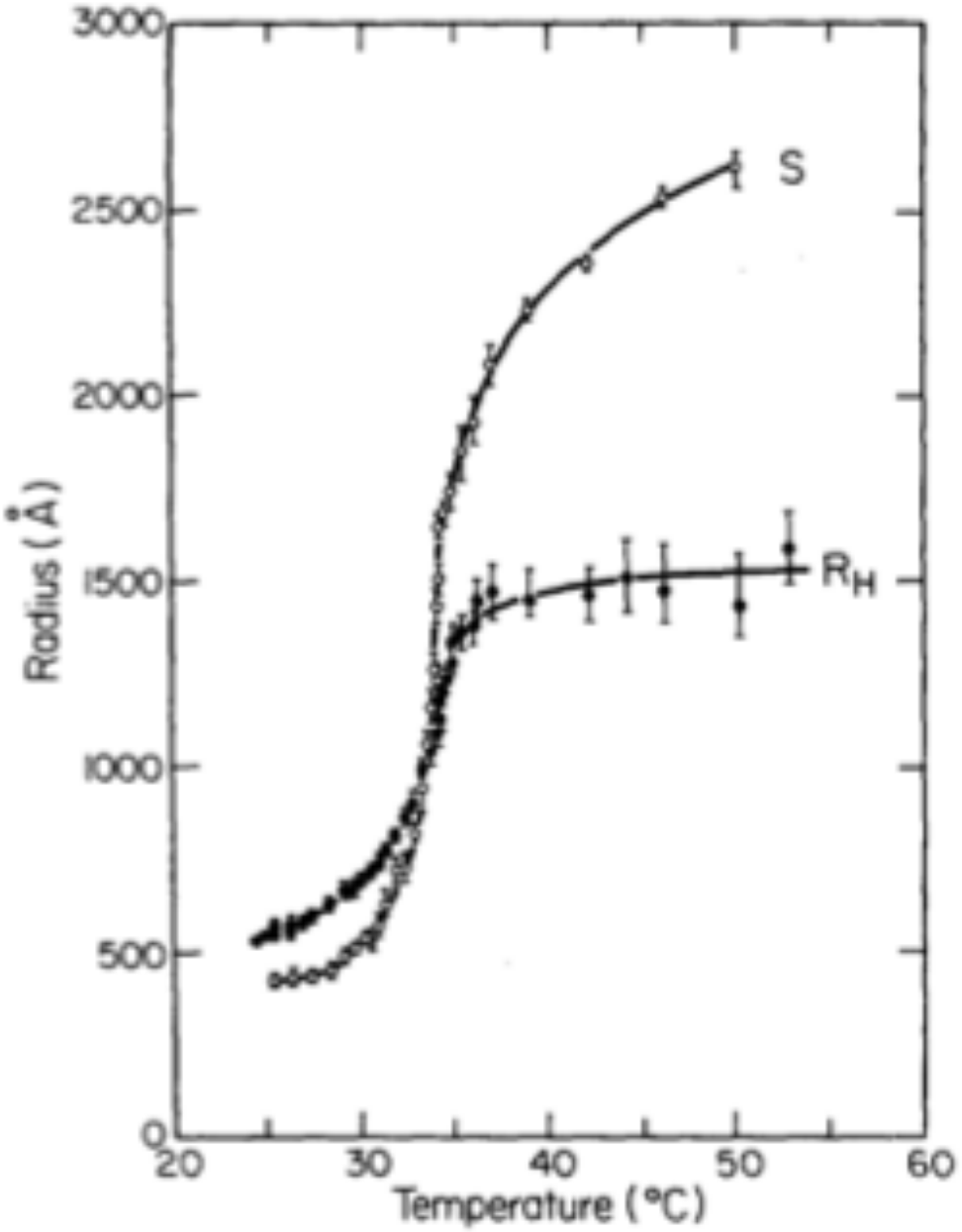
Polymers in Biology KIAS June 29, July 1 2015

- Universal relations in folding
- Stretching Again (Force Spectroscopy)
- Confinement Effects on folding (ribosomes)
- Crowding Effects

Part I : Folding as a finite-size phase transition & consequences

- Polymer size as a function of solvent quality – via Edwards Model
- Order parameters for protein folding & finite size effects
- Scaling of cooperativity in folding with N – number of amino acids

Coil-Globule Transition

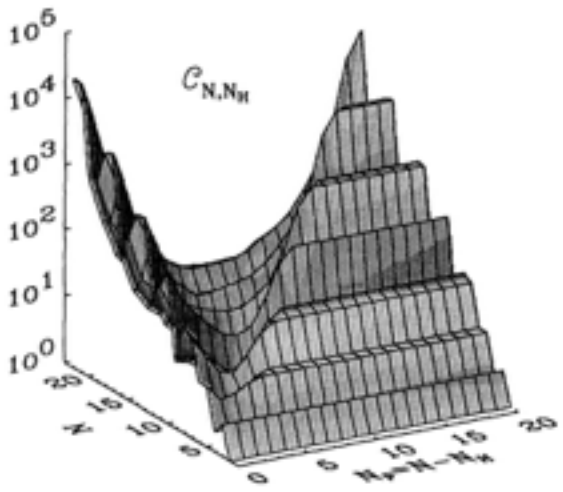


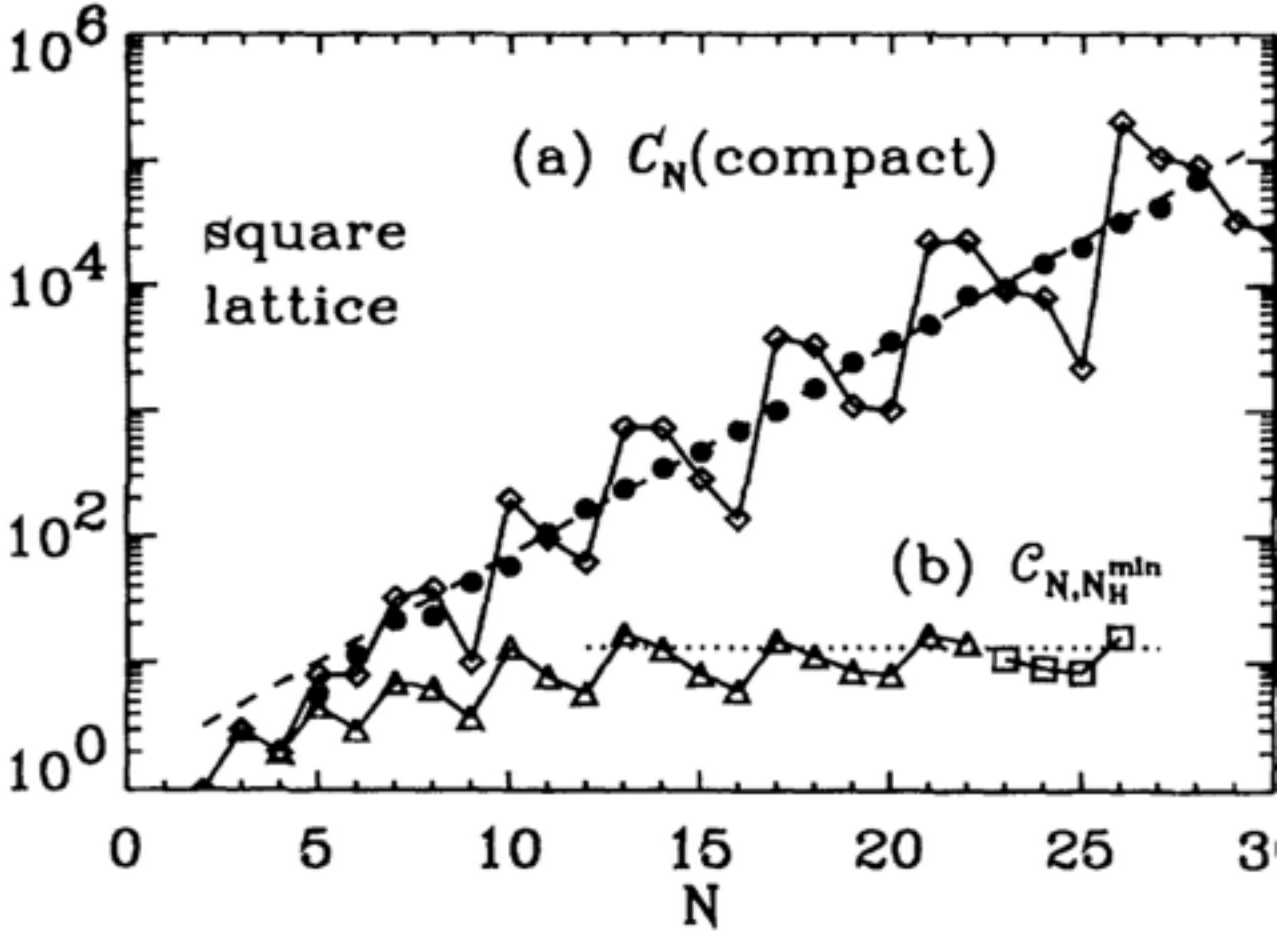
Sun, Nishio, Swislow, Tanaka
J. Chem. Phys. 1995

S = radius; R_H = hydrodymic radius

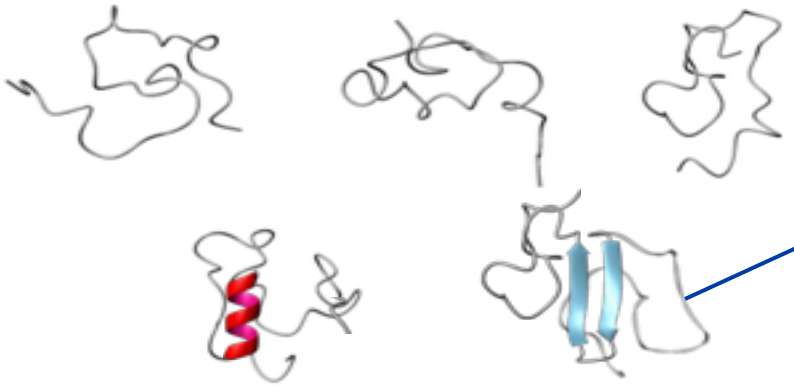
Polystyrene in cyclohexane
(Mol wt = 26 million

Minimum Energy Compact Structures
MECS: Camacho & dt PRL (1993)





Monomeric (spontaneous) folding Mechanisms

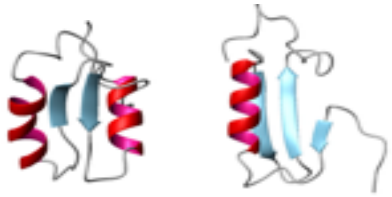


Consequences of finite-size effects & fluctuations on folding thermodynamics

[C]

Or

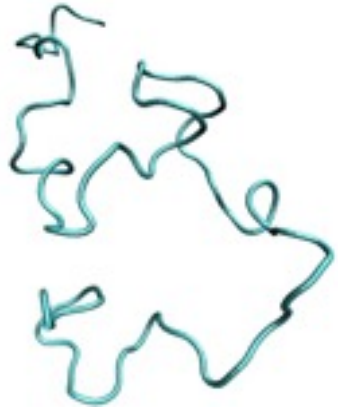
T



dt and C. Hyeon, Biochemistry (2005)

Characteristic Temperatures in Proteins

HIGH T or [C]

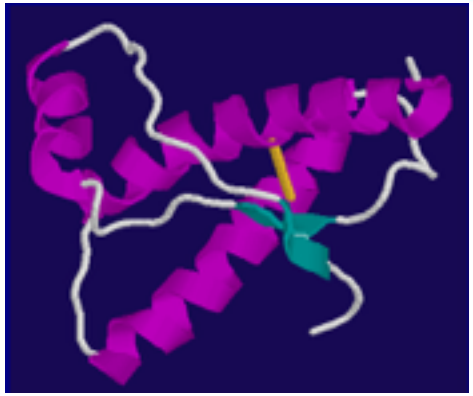


Random Coil (Flory)

Compact

$T \approx T_\theta$ Or $[C_\theta]$

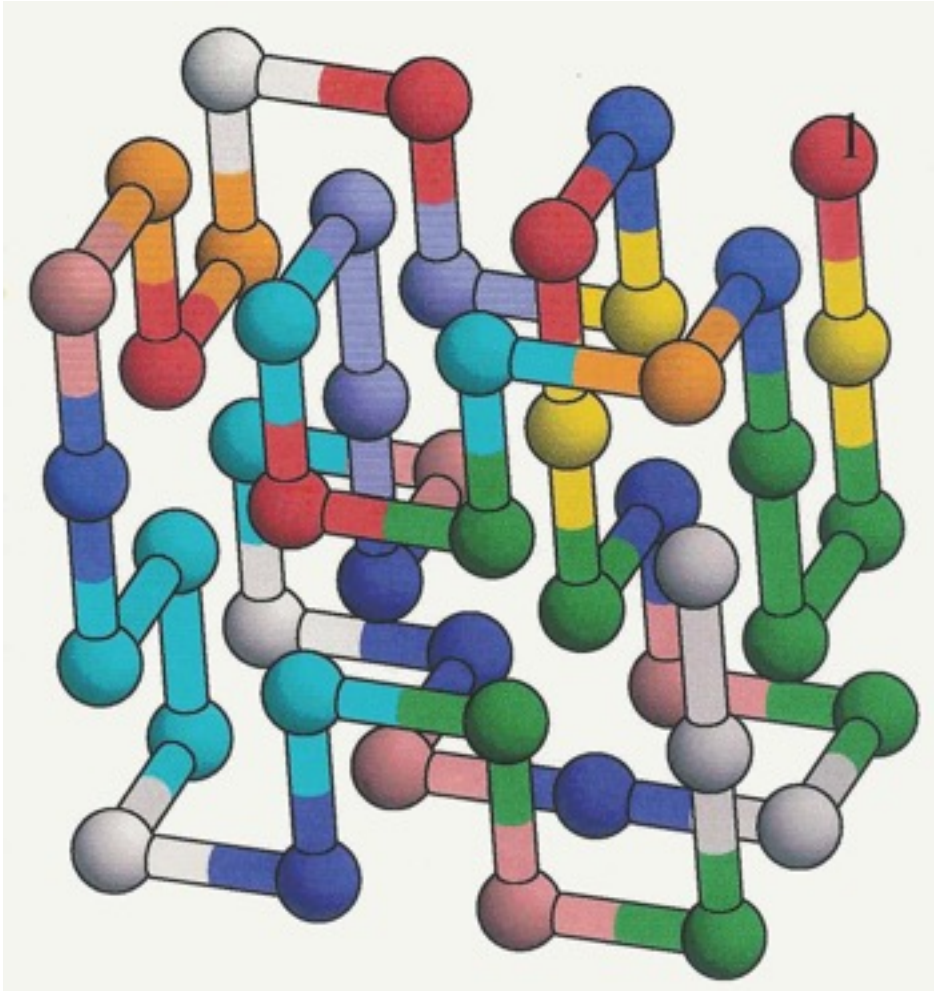
$T \approx T_F$ $[C_F]$



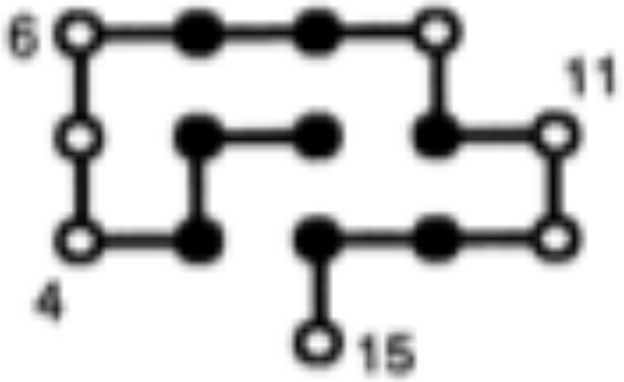
Native State

$\sigma = (T_\theta - T_F)/T_\theta$

Lattice models 3d



Exact enumeration Lattice Models
Camacho & dt PNAS (2003)



Total # of conformations
296,806

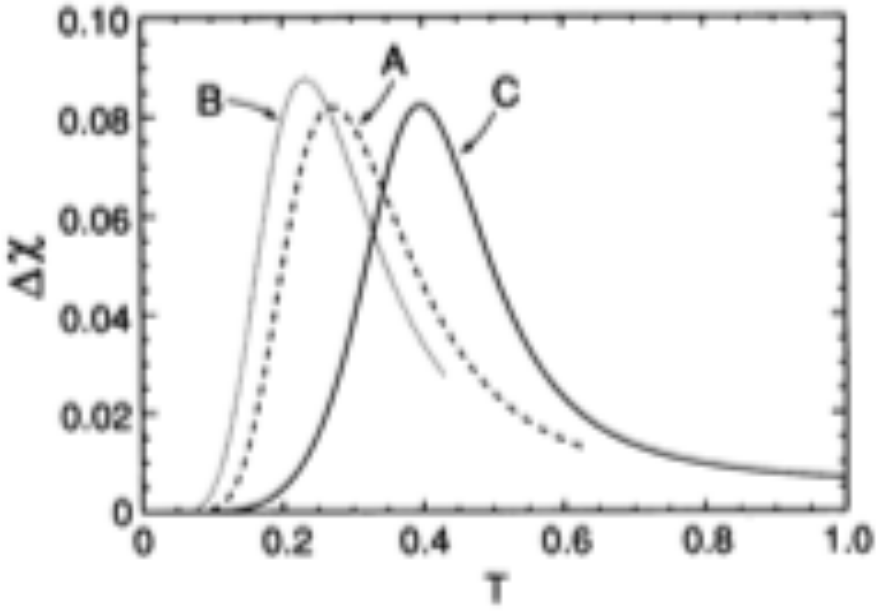


FIG. 3. $\Delta\chi$ as a function of temperature for models A–C.

Size (R_g) of Proteins: Flory Laws

Unfolded state

$R_g \approx a_D N^{\nu}$ "good solvent" $\nu \approx 3/(d + 2)$

Folded and Collapsed Globule

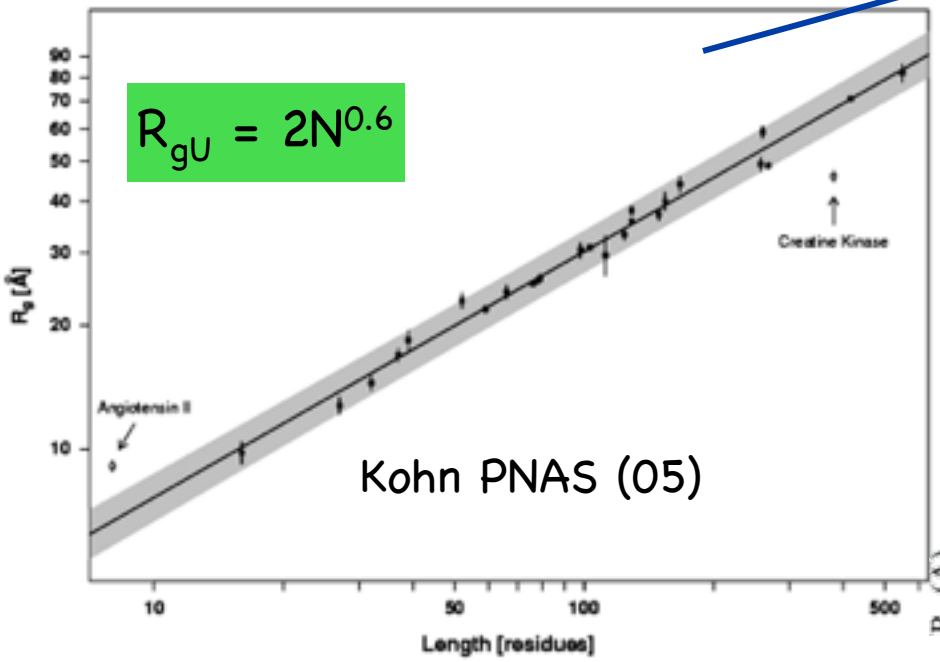
$R_g \approx a_N N^{\nu}$ "poor solvent" $\nu = 1/d$

How good are Flory Laws for proteins?

Protein Collapse : R_g follows Flory law

"Unfolded"

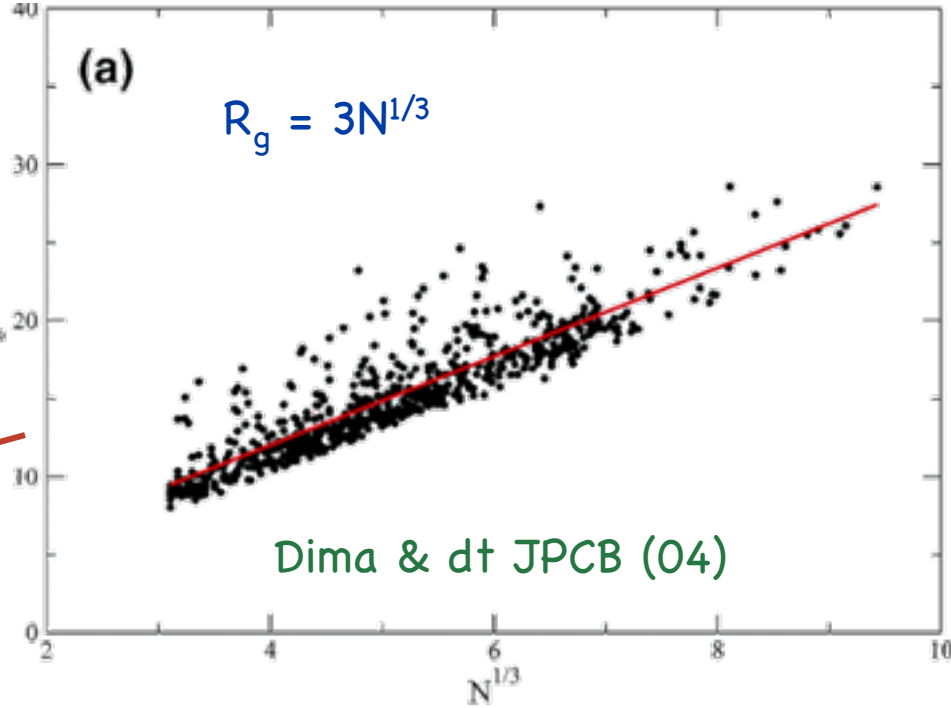
$R_{gU} = 2N^{0.6}$



Kohn PNAS (05)

Fyl's wrath

Folded



Dima & dt JPCB (04)

Folding reaction as a phase transition: A rationale N = number of amino acids

Order Parameter Description

$\rho = N/R_g^3$; χ = Overlap with NBA (0 for NBA)

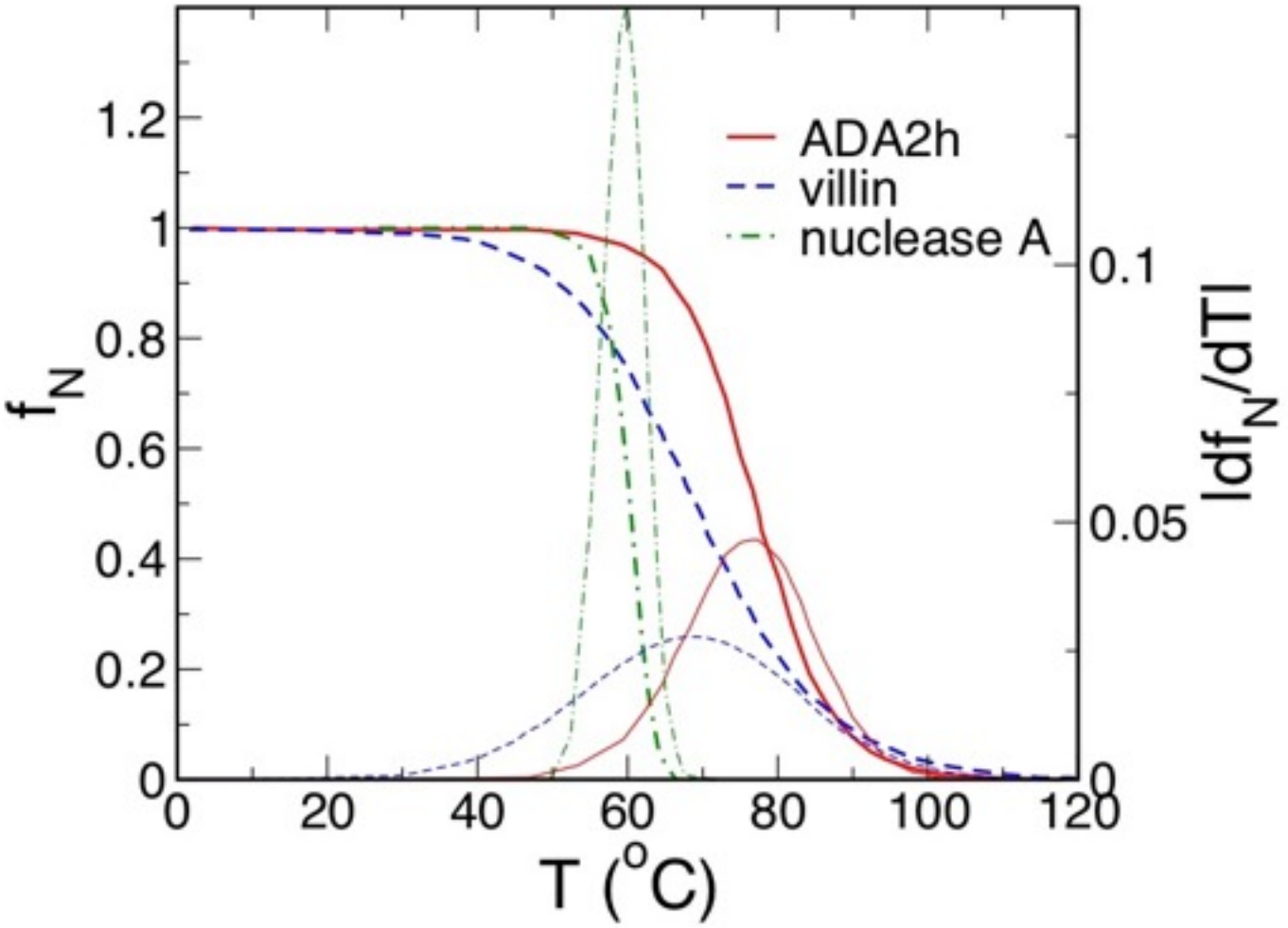
Unfolded (U), Collapsed Globules (CG);
Folded (NBA)

U: ρ (small), χ Large ("vapor")

CG: $\rho \approx O(1)$, χ Large (Dense no order "Liquid")

NBA: $\rho \approx O(1)$, χ Small (Dense order "Solid")

Finite Size Effects on Folding
Order parameters matter



Scaling of Ω_C with N (number of aa)

Two points:

1) T_F = max in $\Delta\chi$ (suceptibility)

$\Delta\chi = T(d\langle\chi\rangle/dh$; h = ordering field (analogy to mag system)
$\Delta\chi$ is dimensionless $\Rightarrow$ h ~ T (in proteins or [C])

2) Efficient folding $T_F \approx T_\theta$ (collapse Temp; Camacho & dt PNAS (1993)) $\Rightarrow \Omega_C$ controlled by protein DSE at $T \approx T_F \approx T_\theta$

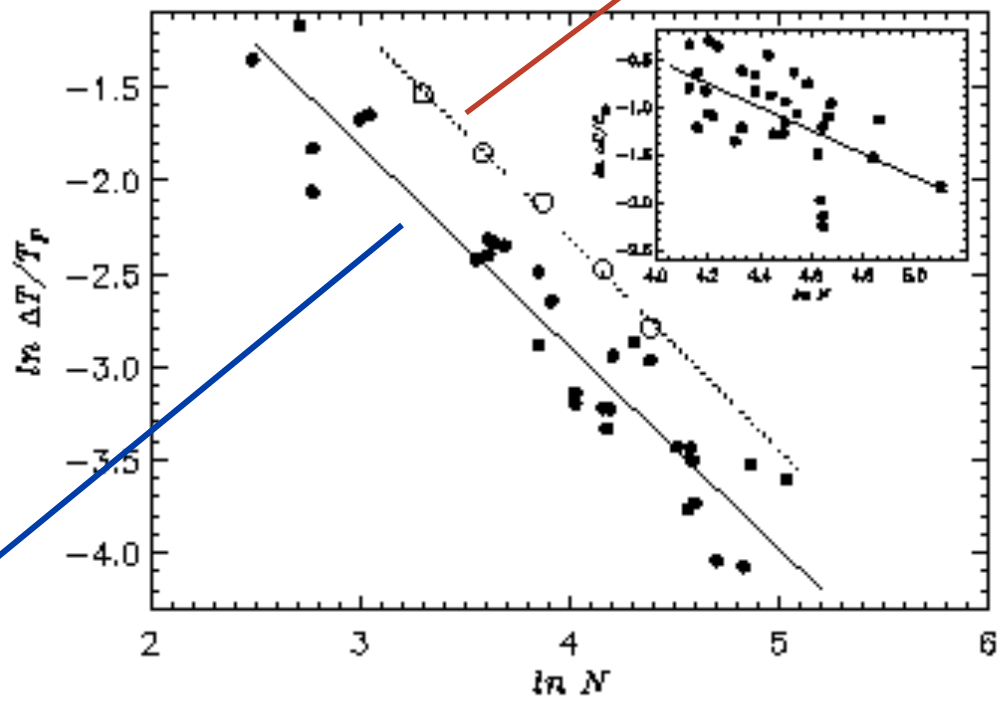
$R_g \sim (\Delta T/T_F)^{-\nu} \sim N^\nu$ (DSE a SAW & manget analogy)

$\Delta T/T_F \sim 1/N$ (Result I)

Finite-size effects on T_F

Lattice models
Side Chains

$\Delta T/T_F \sim 1/N$



Li, Klimov & DT Phys. Rev. Lett. (04)

Experiments

Scaling of Ω_c with N

Magnet-Polymer analogy

$$\Omega_c = (T_F/\Delta T)\,[T_F(d\langle\chi\rangle/dT)]$$

"disp in T_F" X "suspectibility"

$T \approx T_F \approx T_\theta$

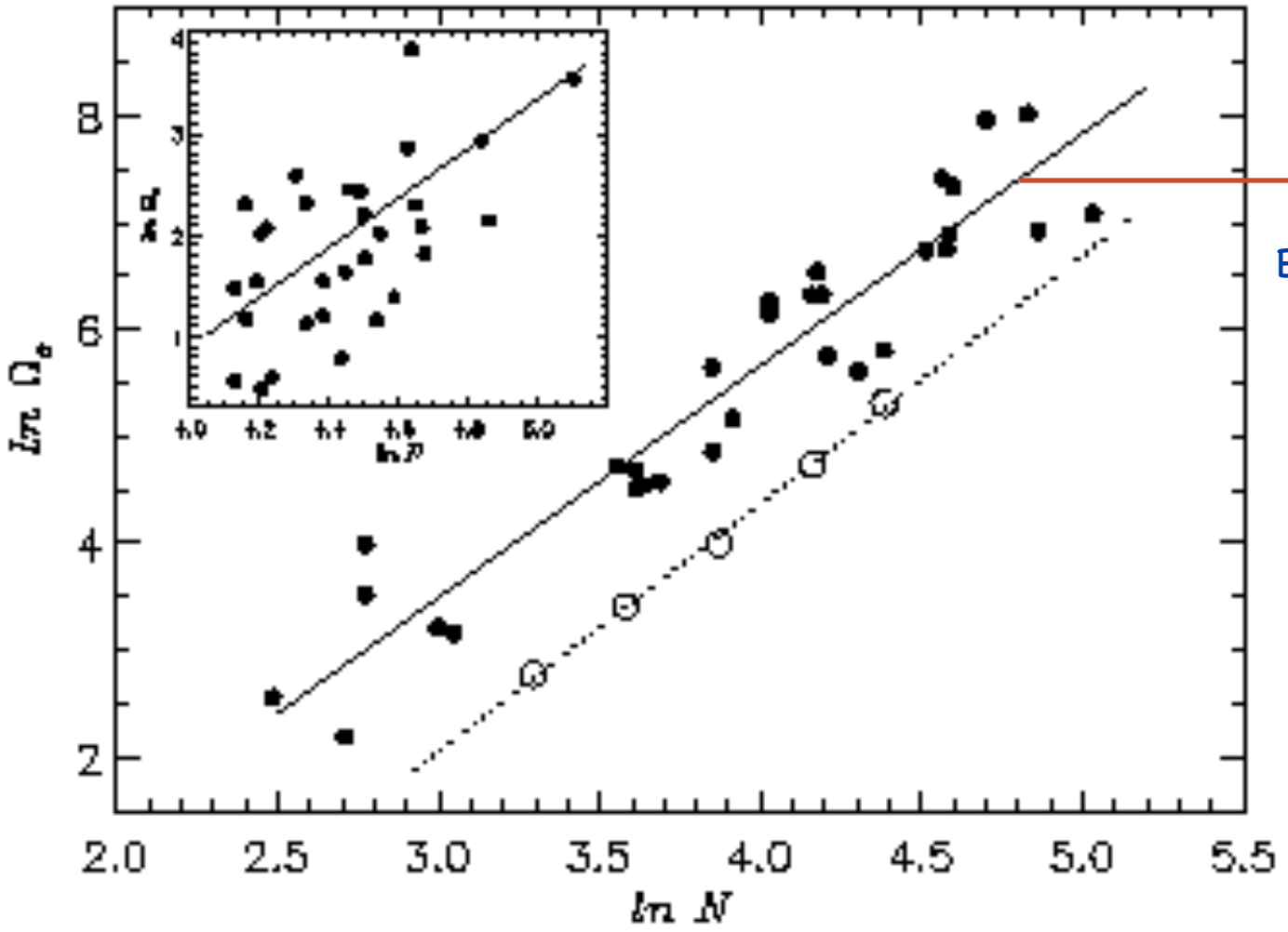
$\Delta\chi \approx N^{\gamma}$

$\Omega_C \approx N^{\zeta}$; $\zeta = 1 + \gamma$(Universal); $\gamma \approx 1.2$

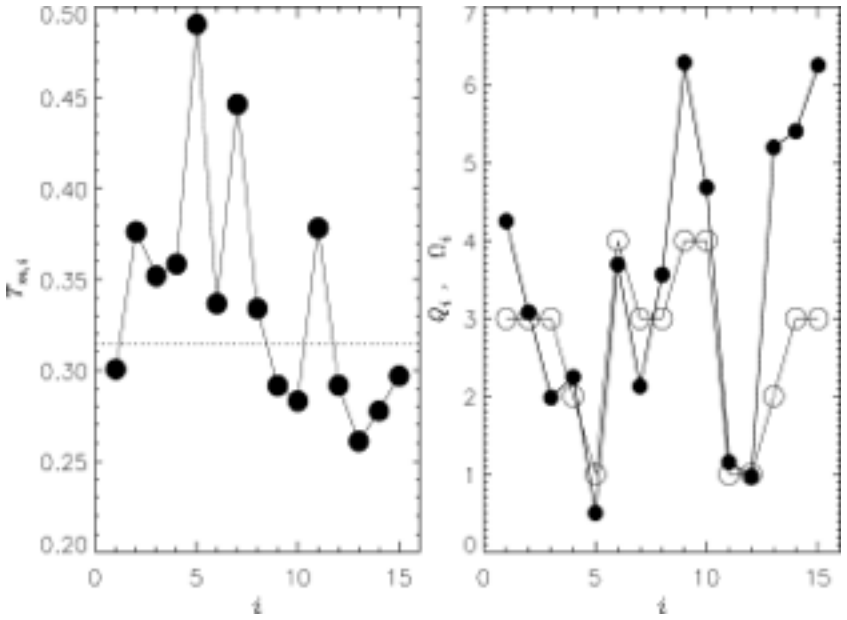
Result II

Universality in Cooperativity

Data from many labs: could be done more precisely



Residue-dependent melting T_m-Holtzer Effect

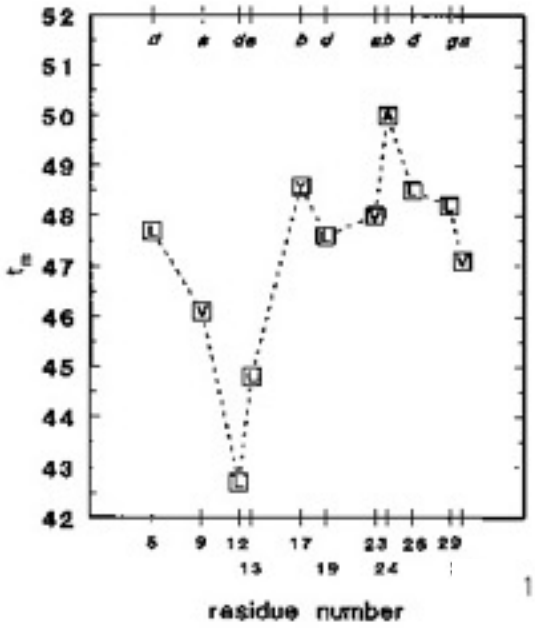


$$f_m(T_{mi}) = 0.5$$

Lattice Models
Side Chains

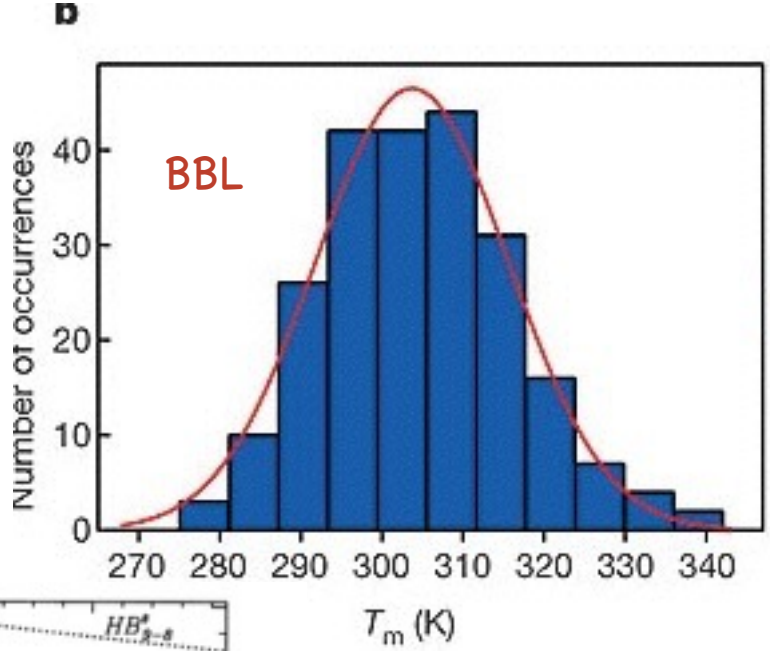
Klimov & dt J. Comp. Chemistry (2002)

Is the melting temperature Unique? Finite-size effects!

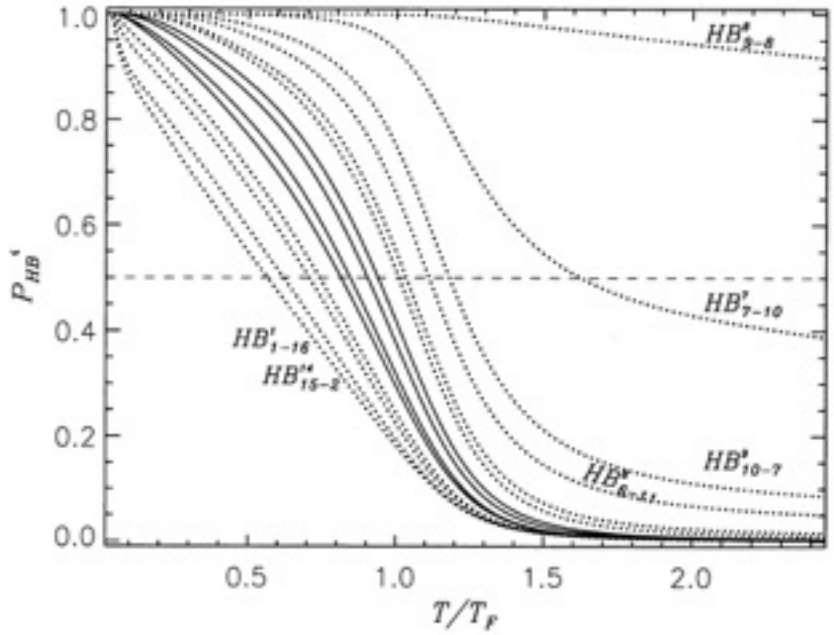


Holtzer
Leucine
Zipper
Biophys J
1997

ΔT large



Munoz
Nature
2006

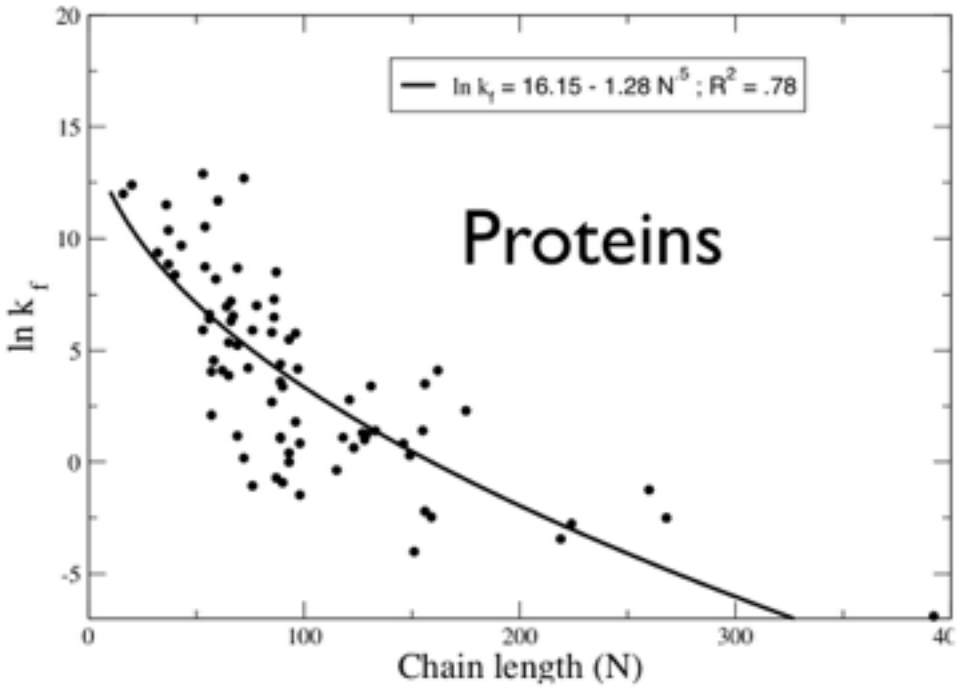


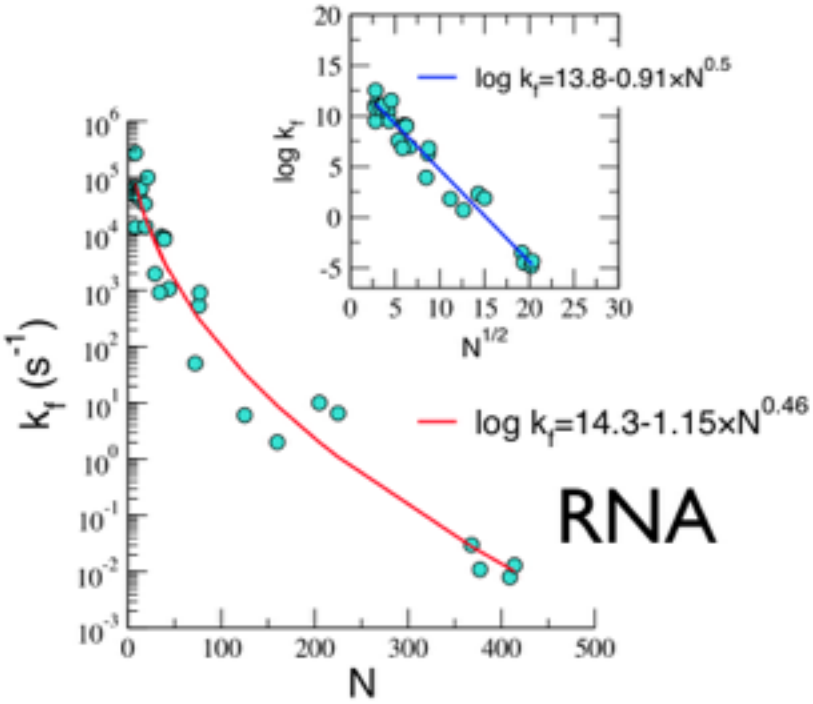
Udgaonkar
Barstar Monnelin

β-hairpin
PNAS 2000
Klimov & dt

Role of N in folding time estimates

$k_F \approx k_0 \exp(-\alpha N^{0.5})$ dt J. de Physique I (1995)
$1/k_0 \approx 1\mu s$; $\alpha \approx O(1)$





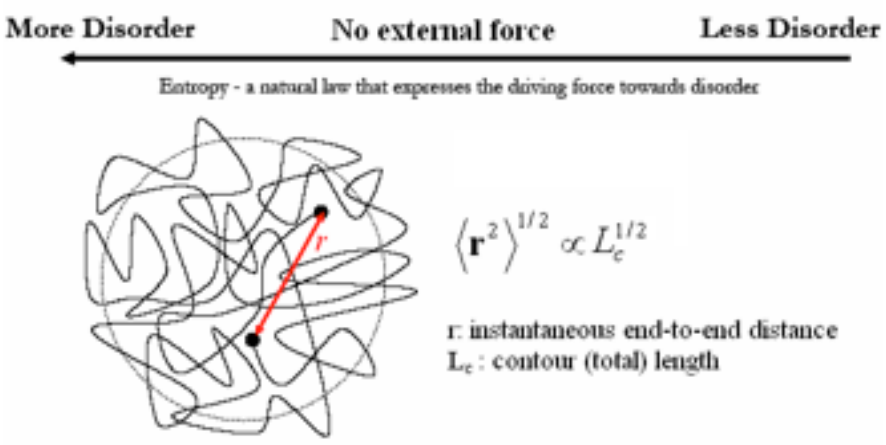
Dill PNAS (2011)

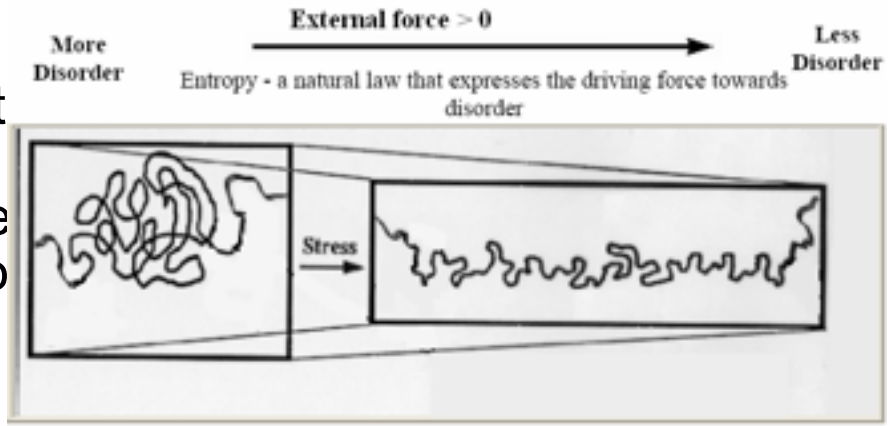
Part II – Stretching Again

- Polymer elasticity WLC and FJC
- Crossover from WLC to FJC
- Forcing globules to unfold (hydrophobic effects)
- Introduction to constant force pulling

Entropic elasticity

- A random coil favors states with high entropy or more disorder.
- When no external force: (unstretched/equilibrium state), the average end-to-end distance is like $L_{con}^{1/2}$, where L_{con} *is* the total length ➔ quite compact.
- To stretch it, an external force must be applied so that the work done gained by the coil compensates the loss of entropy (to stretch means to make the coil more ordered).



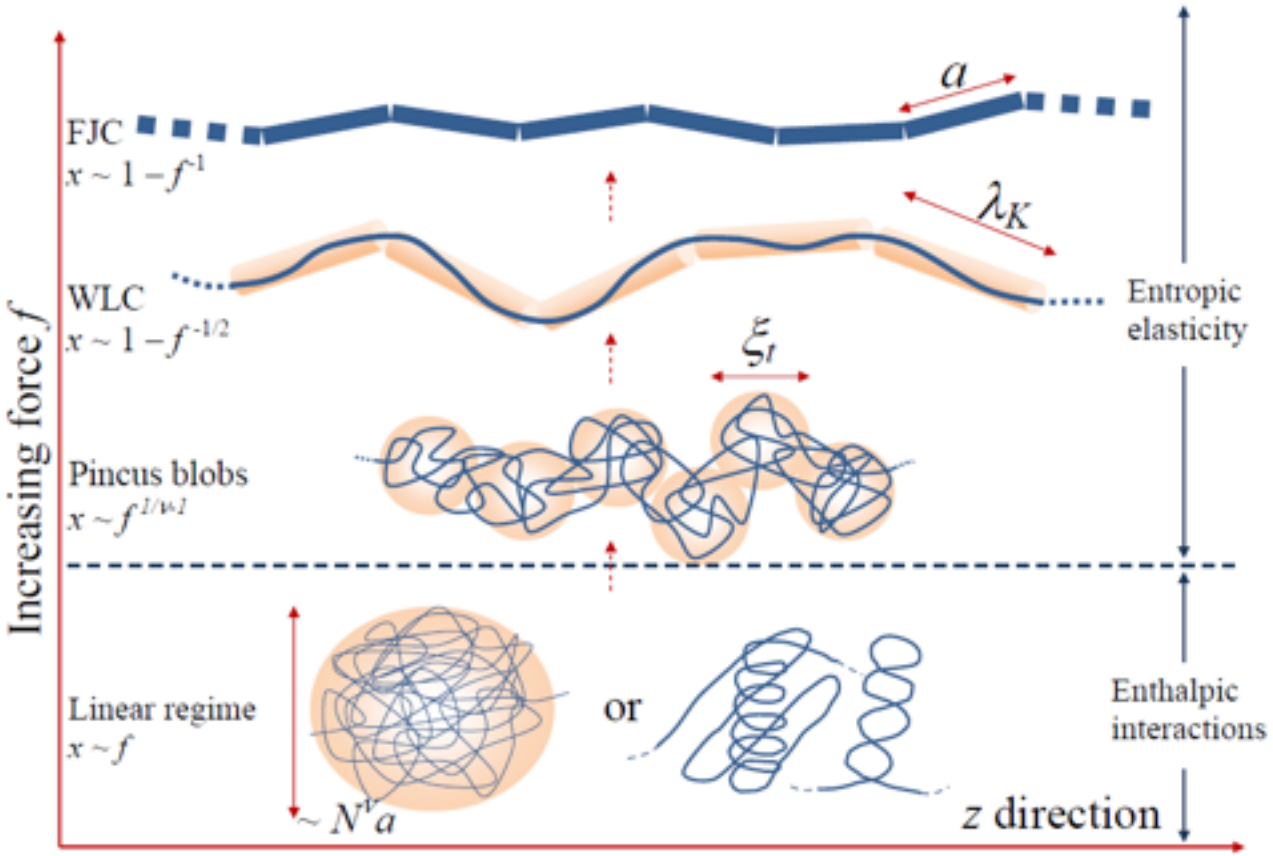


Various stretching regimes
Toan & dt Macromolecules (2010)

$$\mathcal{F}_h = c\frac{k_B T \xi_p}{a^2}$$

$$\mathcal{F}_l \approx \frac{k_B T}{\xi_p}$$

$\lambda_k = (\xi_t \xi_p)^{0.5}$



Stretching DNA (WLC)
Ha & dt JCP 97

$$z = f\,\frac{R_0^2}{3}$$

$$\frac{z}{L} \approx \left(1 - \frac{3}{4}\sqrt{\frac{1}{l_p f}}\right).$$

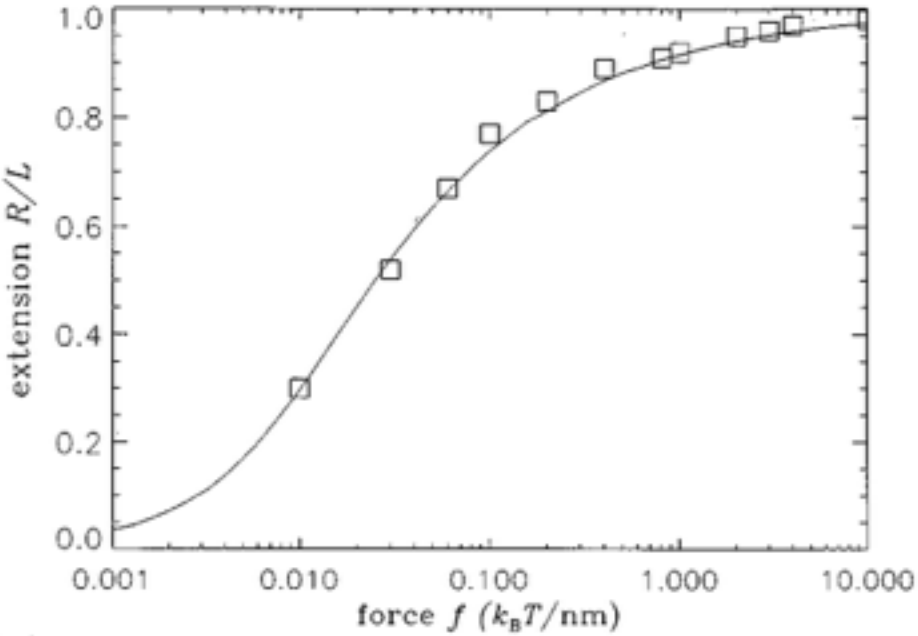
$$F = -k_B T \ln Z,$$

where

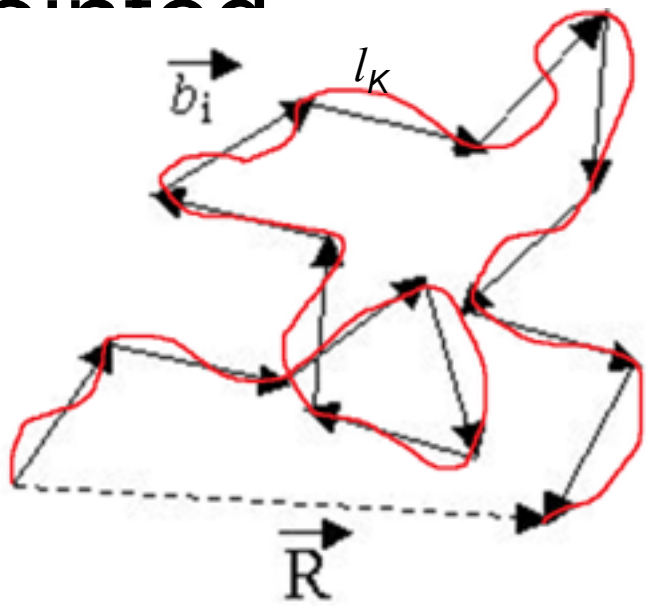
$$Z = \int \mathscr{D}[\mathbf{u}(s)]\,\delta(\mathbf{u}^2(s) - 1)\,e^{-\mathscr{H}/k_B T}$$

with

$$\frac{\mathscr{H}}{k_B T} = \frac{l_p}{2}\int \left(\frac{\partial \mathbf{u}(s)}{\partial s}\right)^2 ds - \int \mathbf{f}(s)\cdot\mathbf{u}(s)\,ds.$$



Freely-Jointed chain m

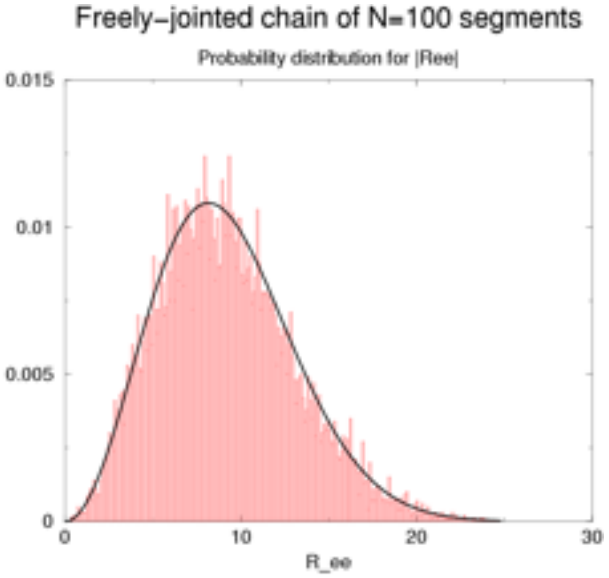


- FJC: Succession of rigid bonds with random orientations
- How it is:
  - *N* rigid bonds (fixed length l_K)
  - Non-self-avoiding
  - End-to-end distance

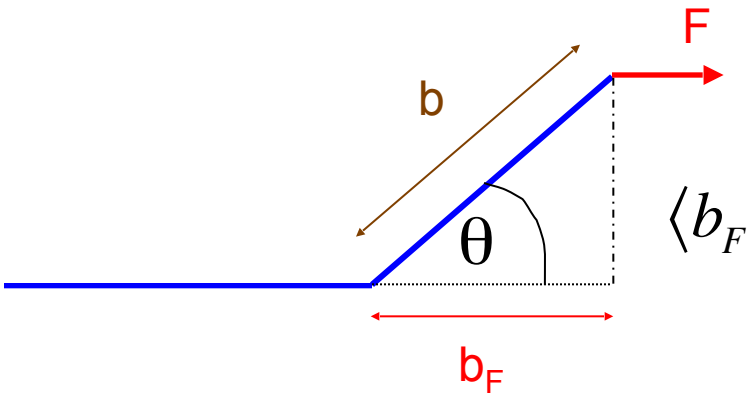
$$\vec{R} = \sum_i \vec{b}_i$$

  - Typical size (~ random walk)

$$\sqrt{\vec{R}^2} = \sqrt{N} l_K$$



The Freely Jointed Chain Model



$$\langle b_F \rangle = \frac{\int_0^{\pi} b\cos(\theta) 2\pi b^2 \sin(\theta) d\theta \; e^{\frac{Fb\cos(\theta)}{k_B T}}}{\int_0^{\pi} 2\pi b^2 \sin(\theta) d\theta \; e^{\frac{Fb\cos(\theta)}{k_B T}}}$$

Where -F x b cos(θ) is the potential energy acquired by a segment aligned along the direction θ with an external force F. Integration leads to:

$$\langle b_F \rangle = b\left[\coth\left(\frac{Fb}{k_B T}\right) - \frac{k_B T}{Fb}\right] = b\Im\left(\frac{Fb}{k_B T}\right)$$ *Langevin Function*

And for a polymer made up of N statistical segments its ave. end-to-end distance is:

$$\langle R(F) \rangle = Nb\Im\left(\frac{Fb}{k_B T}\right) = L\Im\left(\frac{Fb}{k_B T}\right)$$

The Freely Jointed Chain Model

At low forces:

$$x/L = \coth\left(\frac{Fb}{k_B T}\right) - \frac{k_B T}{Fb}, \qquad \text{for } Fb << k_B T:$$

$$\coth(a) = \frac{1}{a} + \frac{a}{3} + \ldots, \text{ therefore:}$$

$$\mathrm{x/L} = \frac{1}{a} + \frac{a}{3} - \frac{1}{a} + \ldots. = \frac{Fb}{3k_B T}$$

$$F = \frac{3k_B T}{b}\left(\frac{x}{L}\right)$$

Thus, at low forces, the chain behaves as a hookian spring with a spring constant, $\kappa = \partial F / \partial x = 3k_B T/b$

Topics for Today

WLC to FJC crossover; Non-linear elasticity?

Confinement Effects on folding (Folding in the ribosome tunnel)

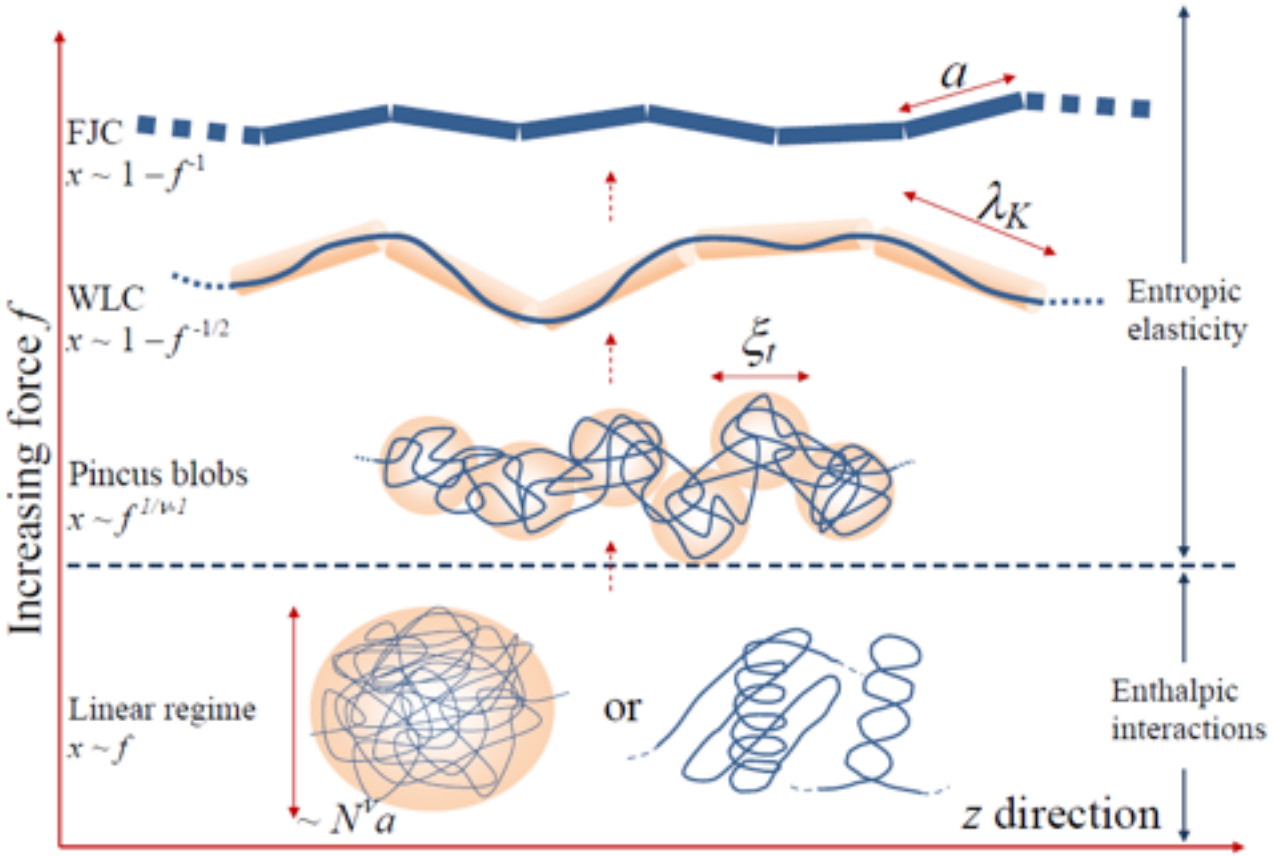
Basic ideas of SMFS: Polymer theory in data analysis

Various stretching regimes
Toan & dt Macromolecules (2010)

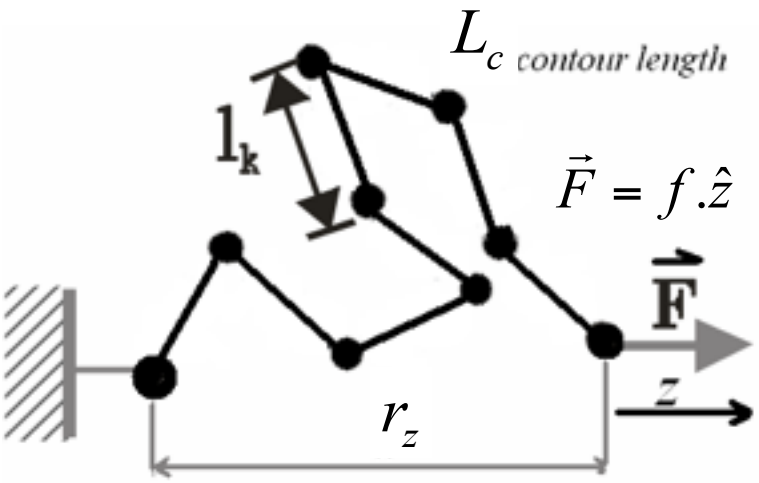
$$\mathcal{F}_h = c\frac{k_B T \xi_p}{a^2}$$

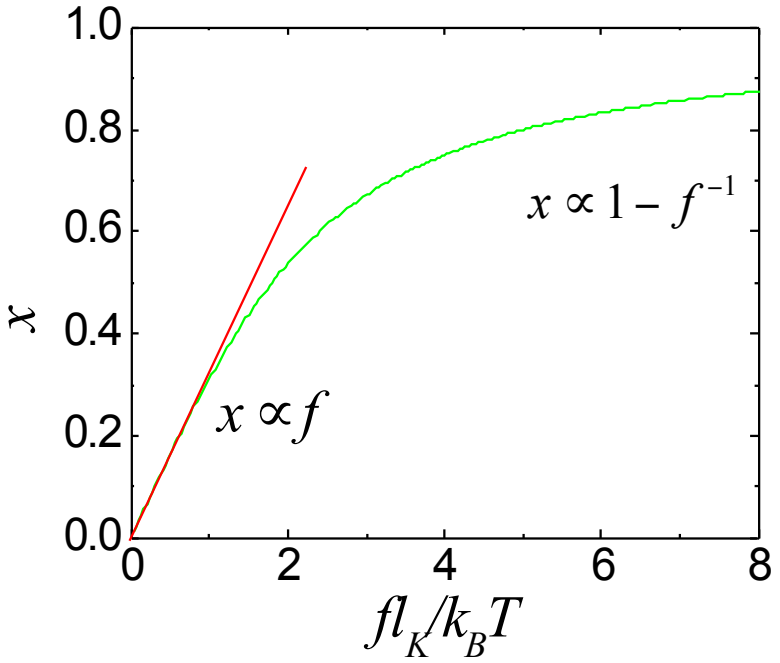
$$\mathcal{F}_l \approx \frac{k_B T}{\xi_p}$$

$\lambda_k = (\xi_t \xi_p)^{0.5}$



FJC under stretch





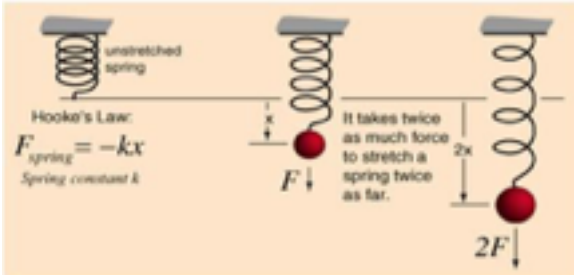
Exact force-extension relation:

$$x \equiv \frac{r_z}{L_c} = \coth\left(\frac{f\,l_K}{k_B T}\right) - \frac{k_B T}{f\,l_K}$$

Linear response at low force: ("entropic" Hooke's law)

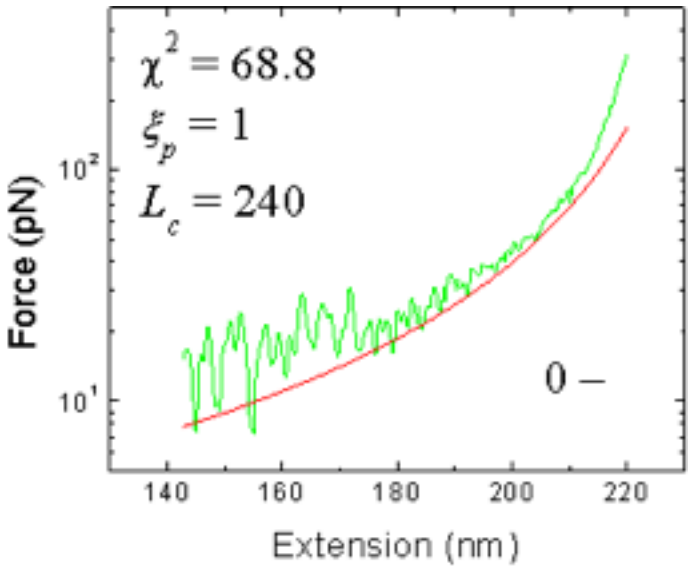
$$x \propto f$$

At high force: $x \propto 1 - f^{-1}$



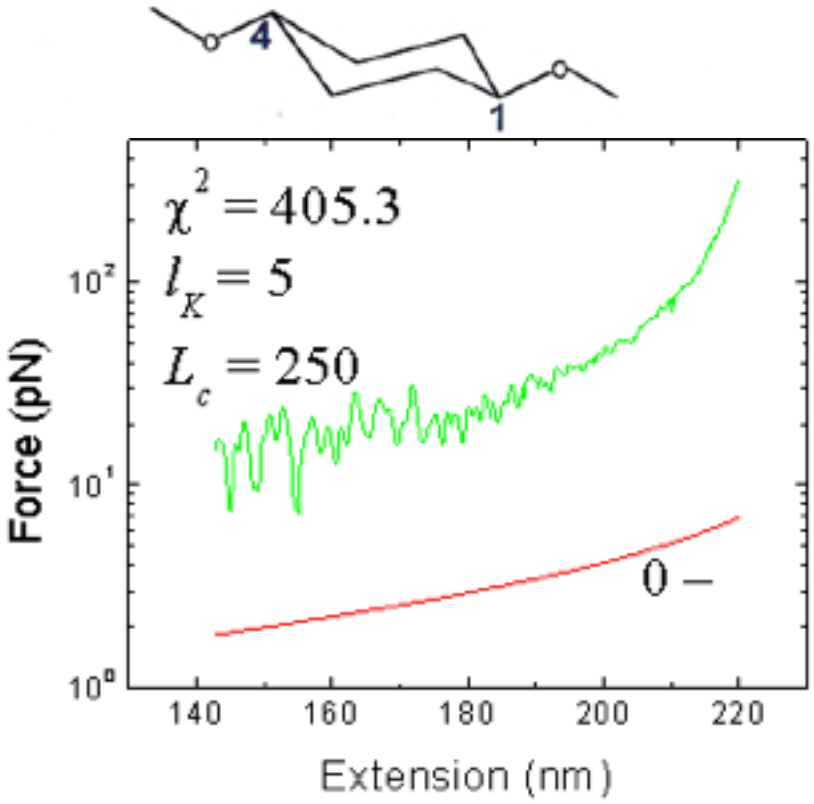
WLC for cellulose?

- Best-fit values:
  - $\chi^2 \approx 5$ (FJC: 1.25)
  - ξ_p = 1.39 nm
- WLC is not appropriate to model the stretching response of cellulose!



Cellulose as a FJC?

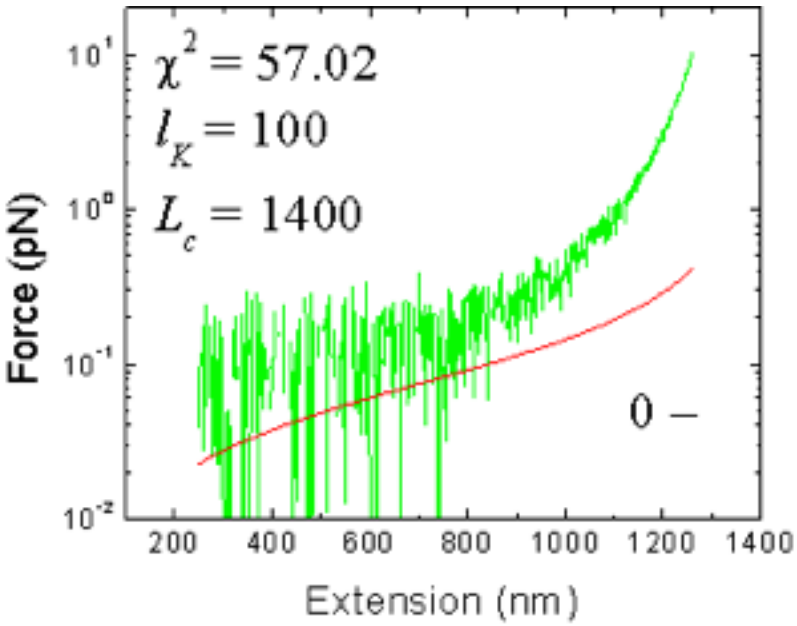
- Initial guess:
  - l_K = 5 nm
  - L_c = 250 nm
  - χ^2= 405.3
- Best-fit values:
  - l_K = 0.849 nm
  - L_c = 223.3 nm
  - χ^2= 1.25
- Good χ^2 and l_K (~ 0.6 nm by *ab initio* (B3LYP/ 6–31G) calculations) (Marszalek, PNAS 1999)



Data: Marszalek *et al.*, Nature (2001)

dsDNA as a FJC?

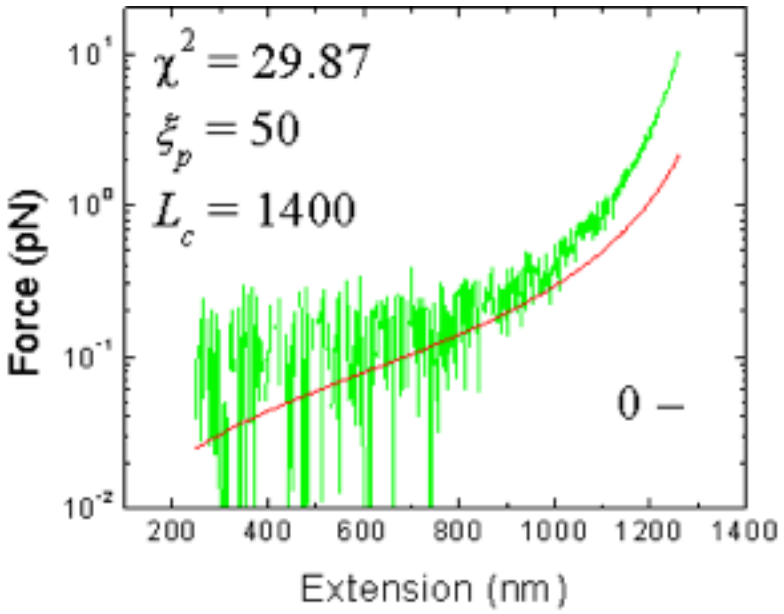
- FJC is not working to account for the stretching behavior of dsDNA
  - Bad fit quality (χ^2)
  - l_K much larger than base-pair separation (0.34 nm).



Data: Wang et al., Biophys. J. 1995

dsDNA as a WLC?

- Best-fit values
  - ξ_p = 40.7 nm
  - L_c= 1324 nm
  - χ^2= 1.13
- Good fit quality!



Wang et al., Biophys. J. 1995

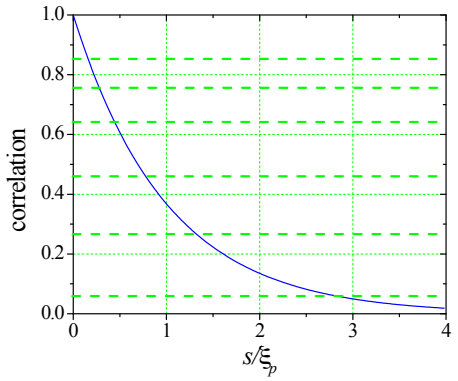
Persistence length ξ_p

- Tangent-tangent correlation of a non-self-avoiding polymer

$$\langle \cos(\theta(s)) \rangle \sim \exp\left(-\frac{s}{\xi_p}\right)$$

- ξ_p is related to the stiffness of polymer
- $\xi_p = ½\, l_K$

$$\sqrt{\vec{R}^2} = \sqrt{N} l_K$$



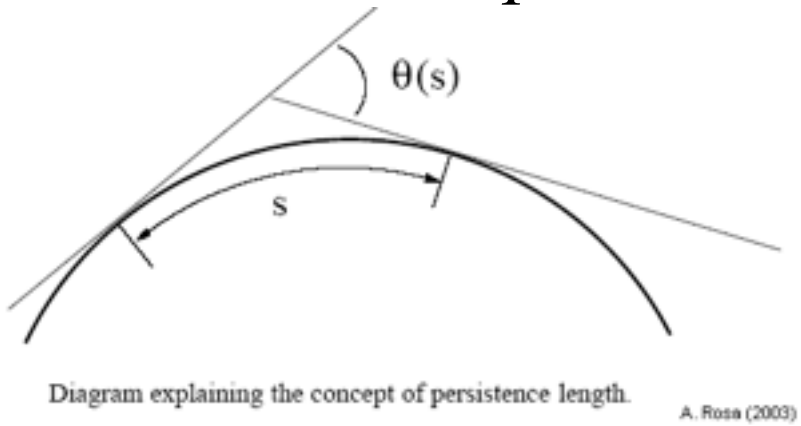


Diagram explaining the concept of persistence length. A. Rosa (2003)

Biopolymers	ξ_p *(nm)*
dsDNA	~ 50
ssDNA/RNA	~ 1.0
Protein	~ 0.7

Persistence length and end-to-end distance

$$\left\langle \vec{R}^2 \right\rangle = \left\langle \vec{R} \cdot \vec{R} \right\rangle$$

$$= \left\langle \int_0^{L_c} \hat{t}(s)ds \cdot \int_0^{L_c} \hat{t}(s')ds' \right\rangle$$

$$= \int_0^{L_c} ds \int_0^{L_c} \left\langle \hat{t}(s) \cdot \hat{t}(s') \right\rangle ds'$$

$$= \int_0^{L_c} ds \int_0^{L_c} e^{-|s-s'|/\xi_p} ds'$$

$$= 2L_c\xi_p \left[1 - \frac{\xi_p}{L_c}\left(1 - e^{-L_c/\xi_p}\right) \right]$$

$$\left\langle \vec{R}^2 \right\rangle = 2L_c\xi_p \left[1 - \frac{\xi_p}{L_c}\left(1 - e^{-L_c/\xi_p}\right) \right]$$

$$\approx 2L_c\xi_p$$

when $L_c >> \xi_p$

FJC regime is universal

in all discrete chains at high forces (Toan & dt Macro, (2012)

- The crossover force to FJC regime

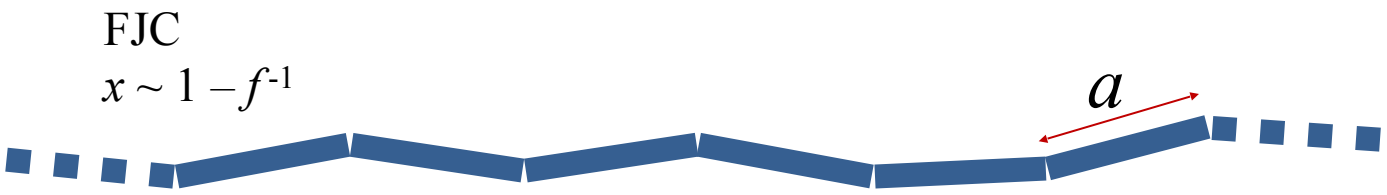
$$\mathcal{F}_h = c\frac{k_B T \xi_p}{a^2},$$

c in the range from $\frac{4+\pi}{2}$ to 4

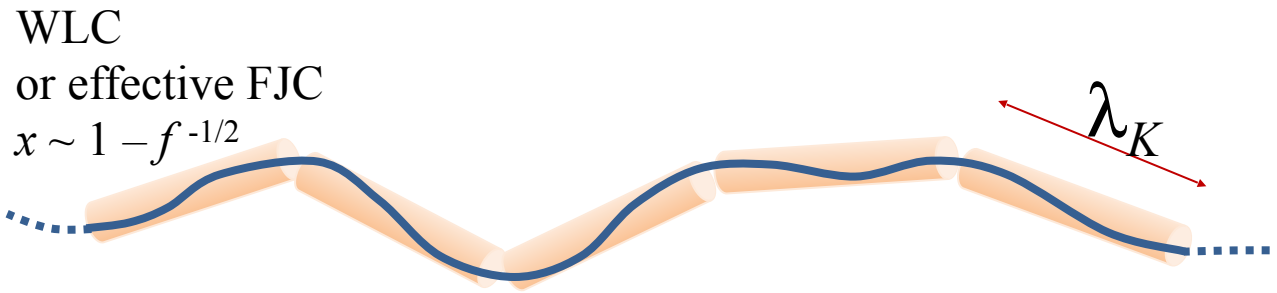
As long as $a > 0,$ regardless of intrachain interaction, the chain behaves as a FJC under high force and the stretching equation is that of the FJC at high force:

$$x = 1 - \frac{\xi_t}{a}.$$

$$\xi_t \equiv k_B T / f.$$



The WLC regime or the effective FJC regime



WLC can also be *statistically* viewed as a FJC with discreteness and free bending, in certain force range

The view is true for any semiflexible chain.

Lower bound of force for the WLC regime

$$\lambda = l_K \Rightarrow \xi_t \approx \xi_p \Rightarrow \mathcal{F}_l \approx \frac{k_B T}{\xi_p}$$

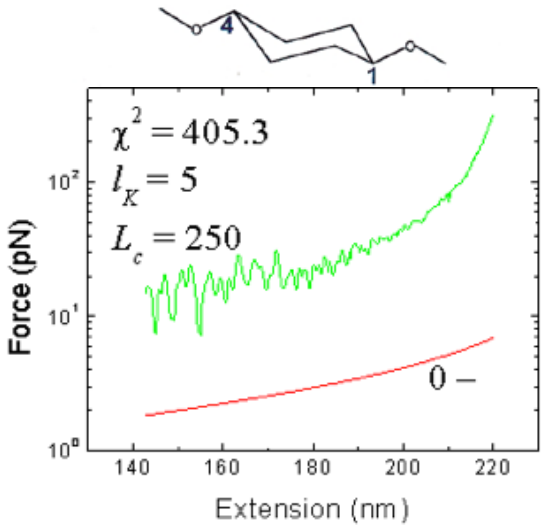
Why the varying performances of the models: Cellulose/Polysaccharides

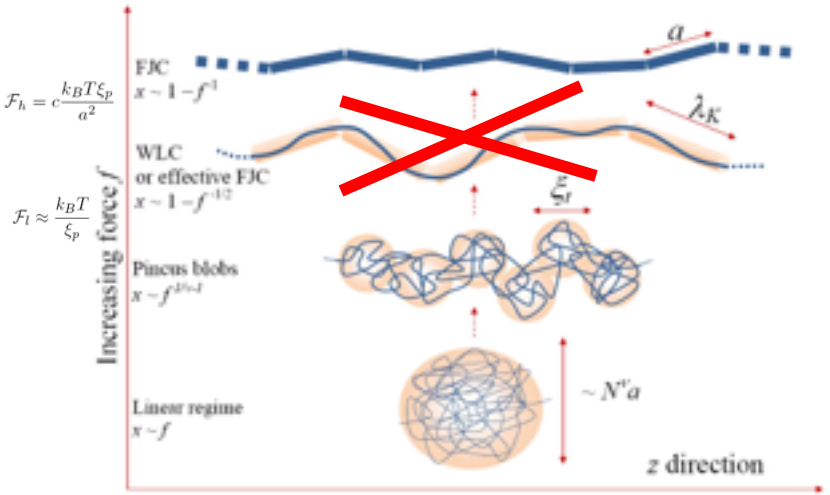
- Range of force!
- Persistence length on the order of or smaller than monomer length

$$l_K \approx a \approx 0.6\,\text{nm}$$

$$\Rightarrow F_l \equiv 2\frac{k_B T}{l_K} \approx F_h \equiv c\frac{k_B T l_K}{2a^2} \approx 15\,pN$$

  - ➔No WLC regime!





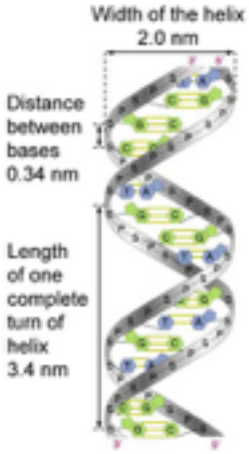
Why the varying performances of the models: dsDNA

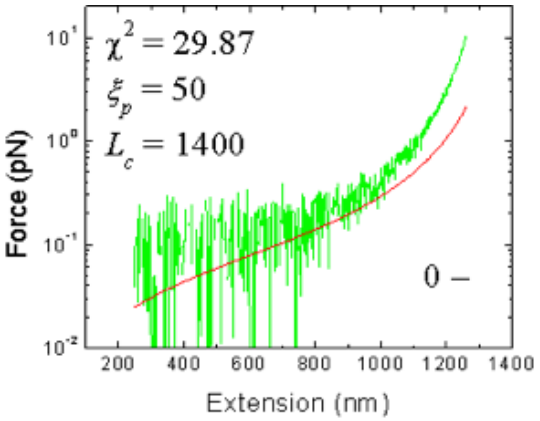
- Range of force:

$a = 0.34\text{nm}, \xi_p = 50\text{nm}$

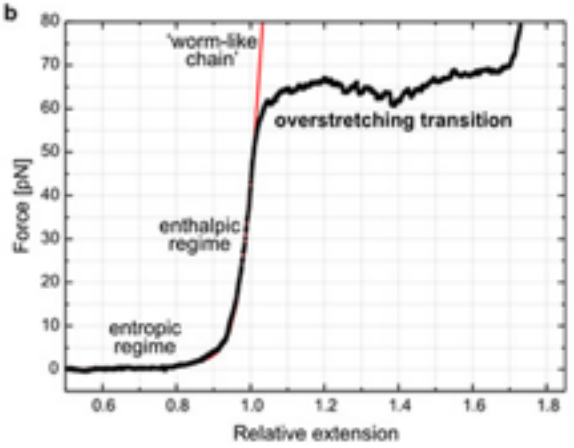
$\Rightarrow F_l \approx 10^{-1}\text{pN}$ & $F_h \approx 6.5 \times 10^3\text{pN}$

Even if $a = 3.4\text{nm}, F_h \approx 65\text{pN}$





  - But dsDNA undergoes overstretching transition at 65 pN➔Not possible to observe the FJC regime.
  - ➔Entire portion of the FEC lays comfortably within the WLC regime and that's why WLC model itself is very good for the dsDNA stretching data.



Why the varying performances of the models:

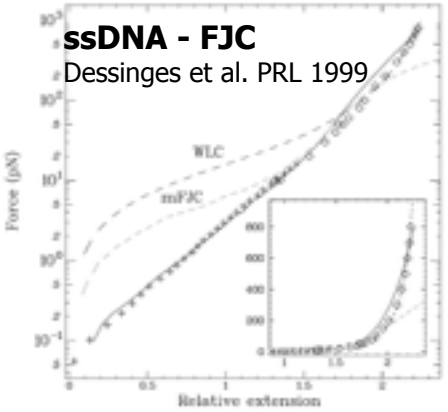
ssDNA & RNA

$$a = 0.5\text{nm}, \xi_p = 1\text{nm}$$

$$\Rightarrow F_l \approx 4\text{pN} \;\&\; F_h \approx 50\text{pN}$$

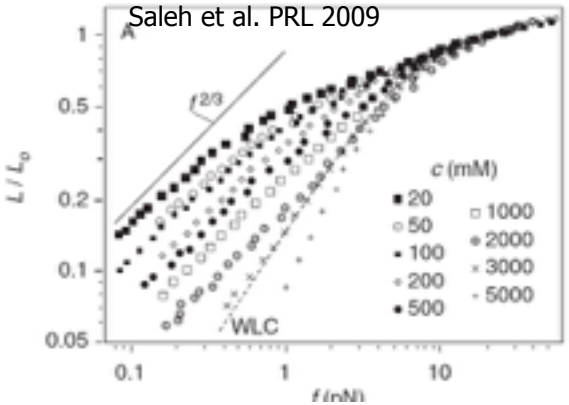
ssDNA - FJC

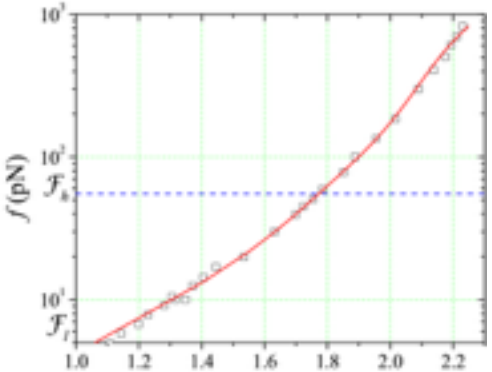
Dessinges et al. PRL 1999



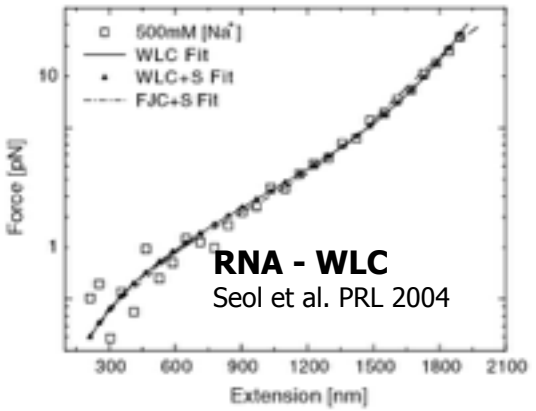
ssDNA - WLC

Saleh et al. PRL 2009





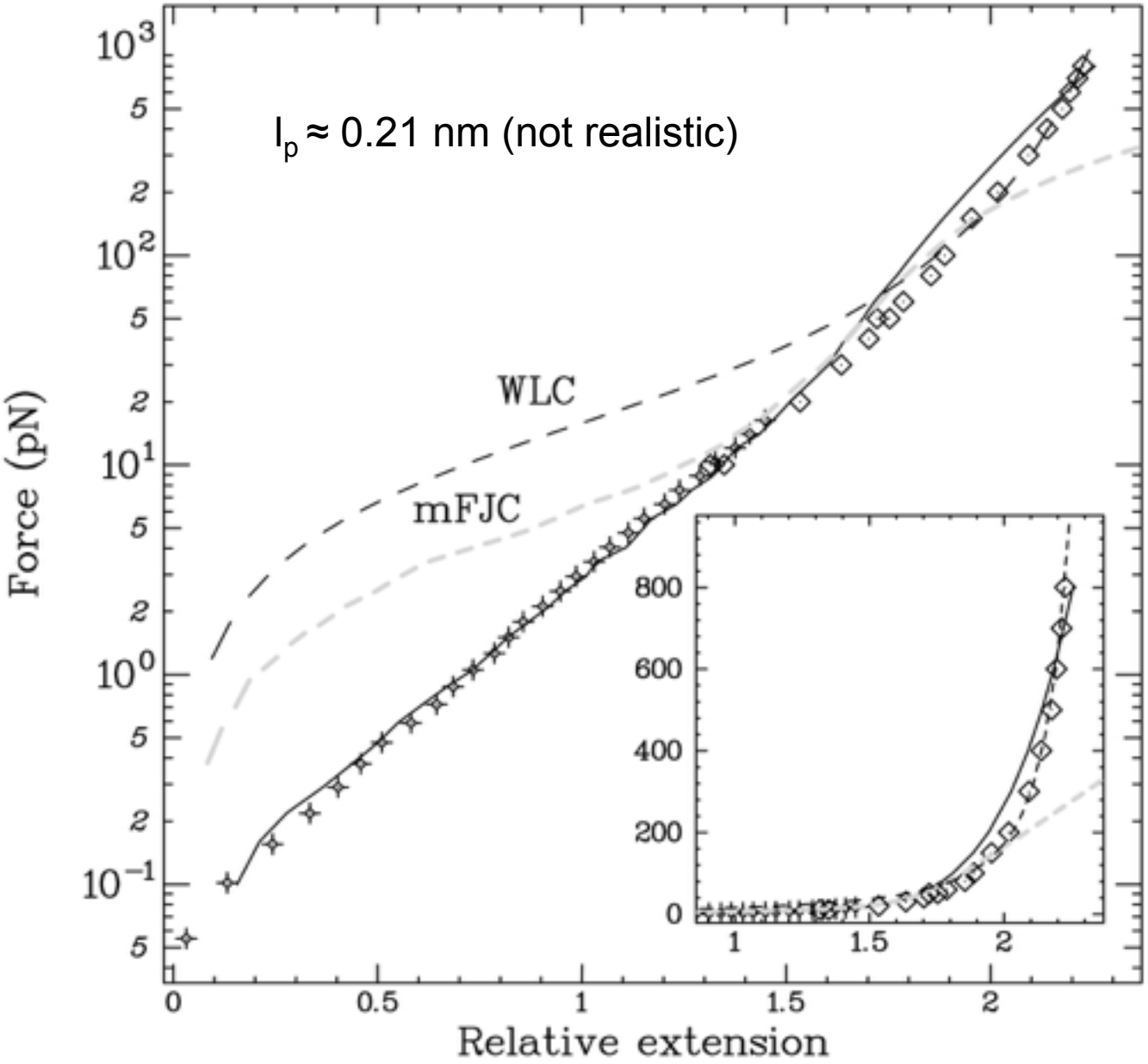
$$x \approx 1 - \frac{\xi_p}{(\lambda_K^{\alpha} + a^{\alpha})^{\frac{1}{\alpha}}}$$



RNA - WLC

Seol et al. PRL 2004

Dessinges Bensimon,Croquette PRL (2002)



Non-linear elasticity at high F?

WLC models

TABLE 1 DNA elasticity models

Model	Formula	Comments
Marko-Siggia (1995) WLC	$F = \left(\frac{k_B T}{L_p}\right)\left[\frac{1}{4(1 - x/L_o)^2} - \frac{1}{4} + \frac{x}{L_o}\right]$	Entropic theory. Interpolation formula of exact solution when [salt] ≥ 10 mM. Applicable when $F \ll \frac{1}{4}\,(k_B T K_o^2/L_p)^{1/3}$. Differs from exact solution by up to ~10% near $F \approx 0.1$ pN. Approaches exact solution at lower and higher forces.
Odijk (1995) WLC	$x = L_o\left[1 - \frac{1}{2}\left(\frac{k_B T}{F L_p}\right)^{1/2} + \frac{F}{K}\right]$	Entropic/enthalpic theory. Applicable for $\lvert x - L_o\rvert/L_o \ll 1$ (high-force regime).
Smith et al. (1995) FJC	$x = L_o\left[\coth\left(\frac{2FL_p}{k_B T}\right) - \frac{k_B T}{2FL_p}\right]\left(1 + \frac{F}{K_o}\right)$	Entropic/enthalpic theory. Applicable to polymers that approximate a FJC. Note that the Kuhn length = $2L_p$, and Langevin function $L(\alpha) = \coth(\alpha) - 1/\alpha$.
Modified Marko-Siggia WLC	$F = \left(\frac{k_B T}{L_p}\right)\left[\frac{1}{4(1 - x/L_o + F/K_o)^2} - \frac{1}{4} + \frac{x}{L_o} - \frac{F}{K_o}\right]$	Entropic/enthalpic theory. Modification of Marko-siggia formula to incorporate enthalpic stretching. Has limitations similar to Marko-Siggia near $F \approx 0.1$ pN.

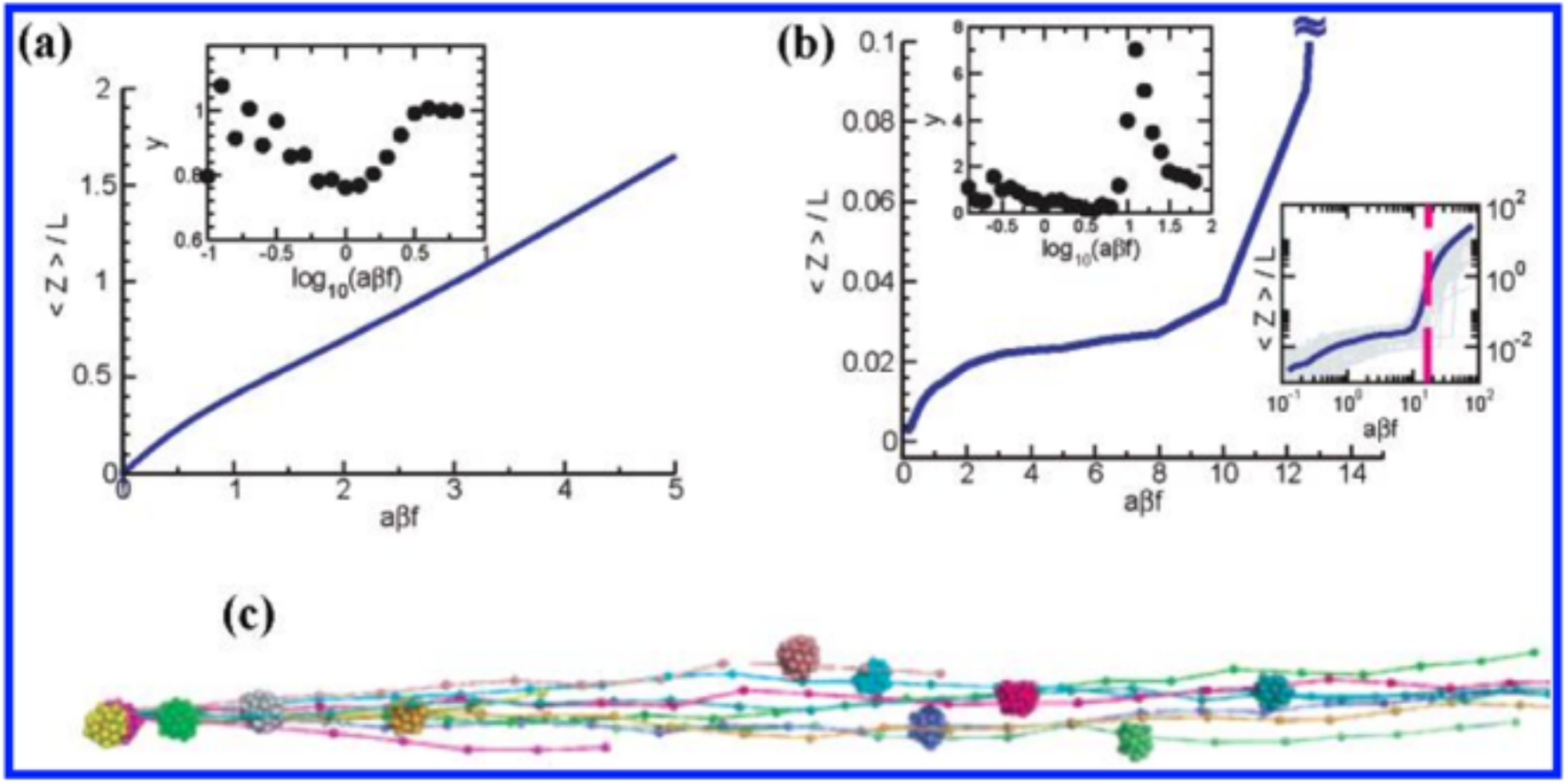
WLC, Wormlike chain; FJC, freely jointed chain; F, force; x, extension (end-to-end distance); L_p, persistence length; L_o, contour length; K_o, elastic modulus; $k_B T$, Boltzmann's constant times absolute temperature.

Biophysical Journal Volume 72 March 1997 1335-1346

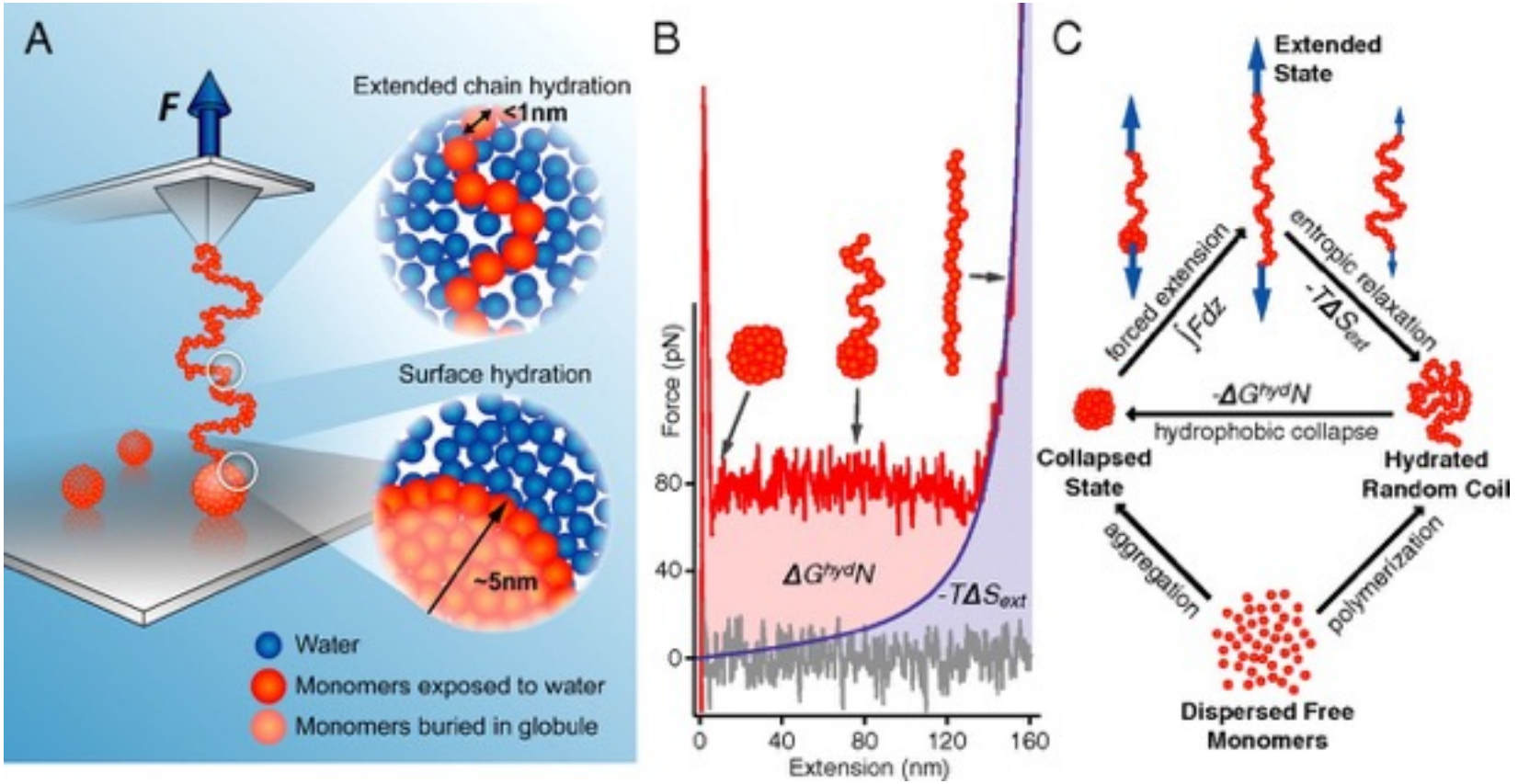
Stretching Polymer in Poor Solvent: Force Plateau

Morrison, Hyeon, Toan, Ha, dt

Macromolecules (2007)

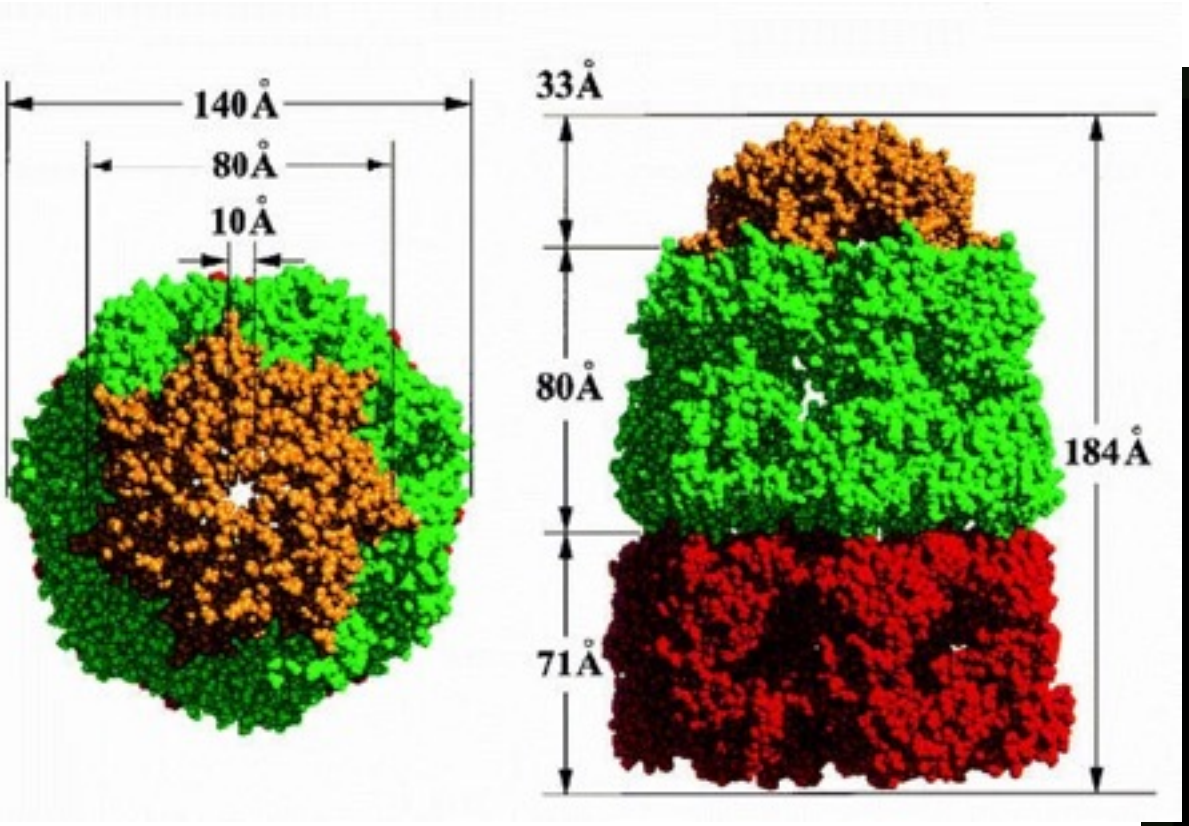


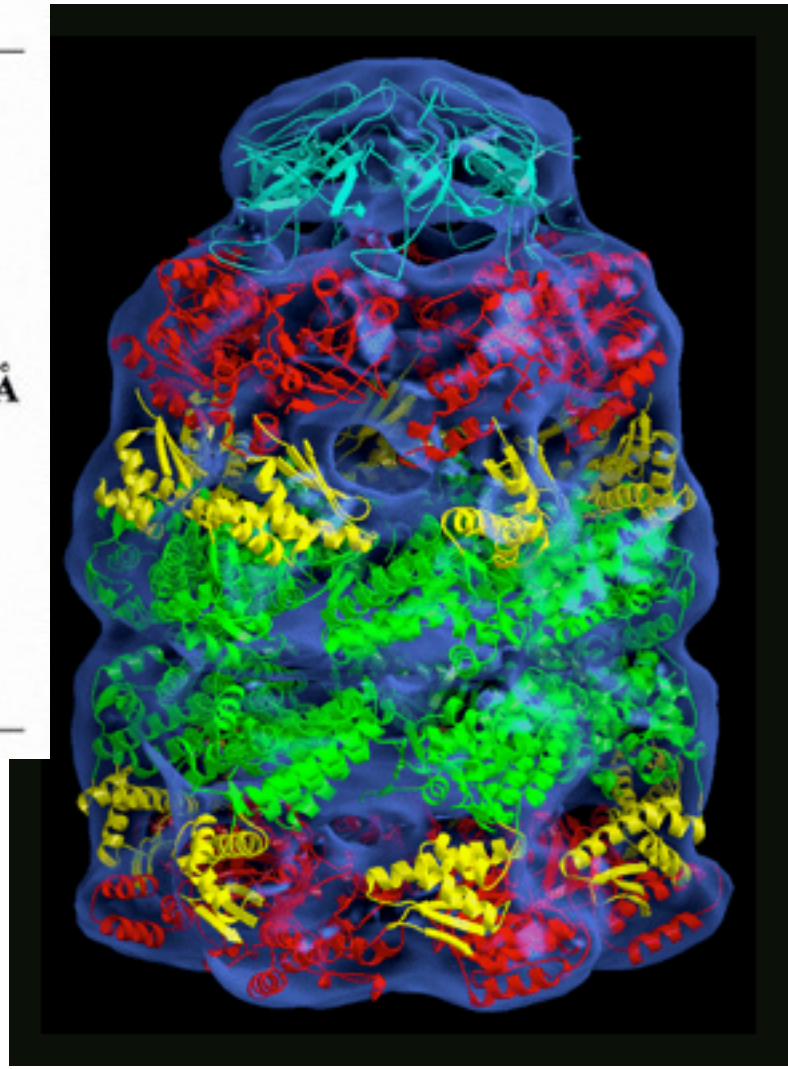
Single molecule pulling experiments are used to obtain the hydration free energy of a hydrophobic polymer as it transitions from a collapsed to an extended state.



Isaac T. S. Li, and Gilbert C. Walker PNAS 2011;108:16527-16532

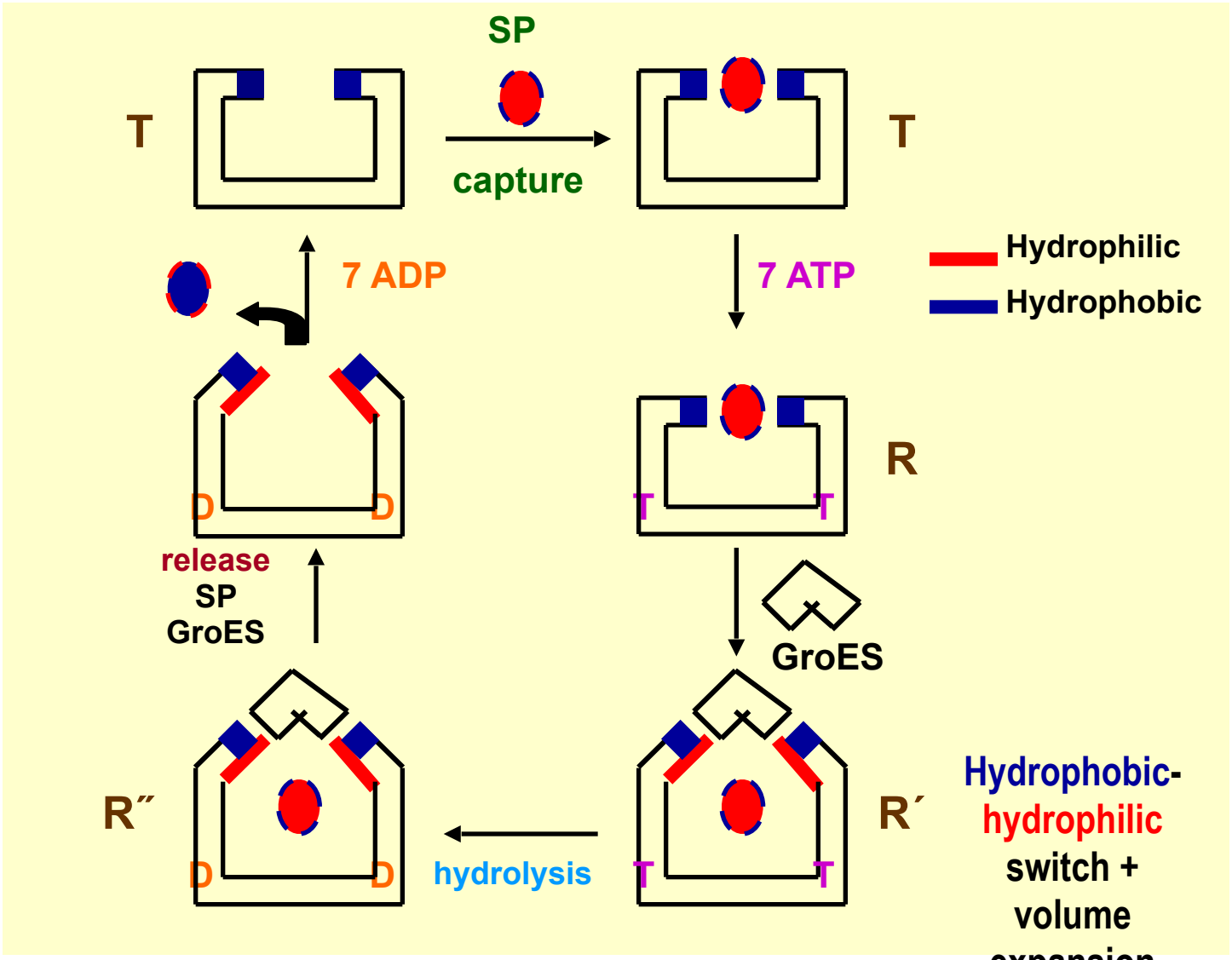
Confinement Effects : Motivation from Biology





R'' (ATP hydrolyzed) state (machine)

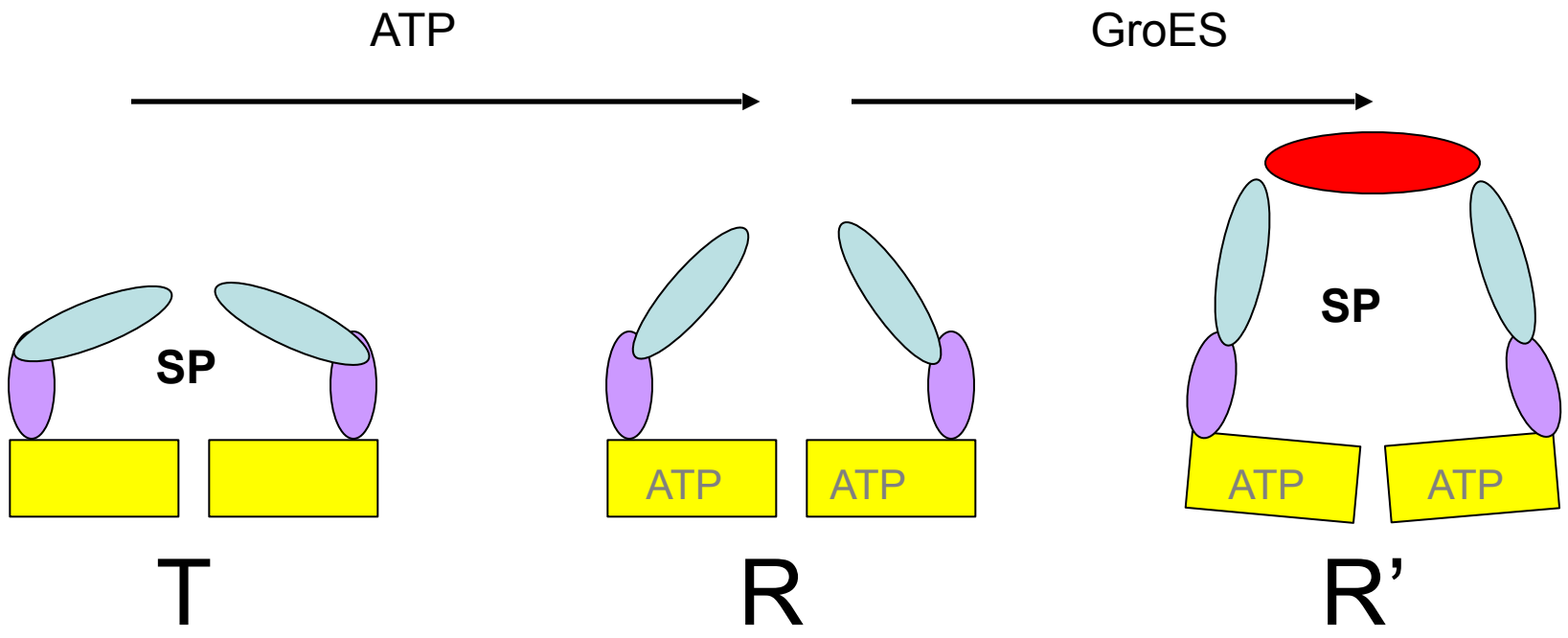
GroEL hemicycle



Rescue of Substrate Proteins by Iterative Annealing Mechanism

- Todd, Lorimer, dt PNAS (1996)
- Hyeon, Lorimer, dt PNAS (2006)
- GroEL allostery Amnon Horovitz
- Review: dt & Lorimer Ann. Rev. Biophys (2001)

Substrate Protein in a GroEL cavity



"Hydrophobic"

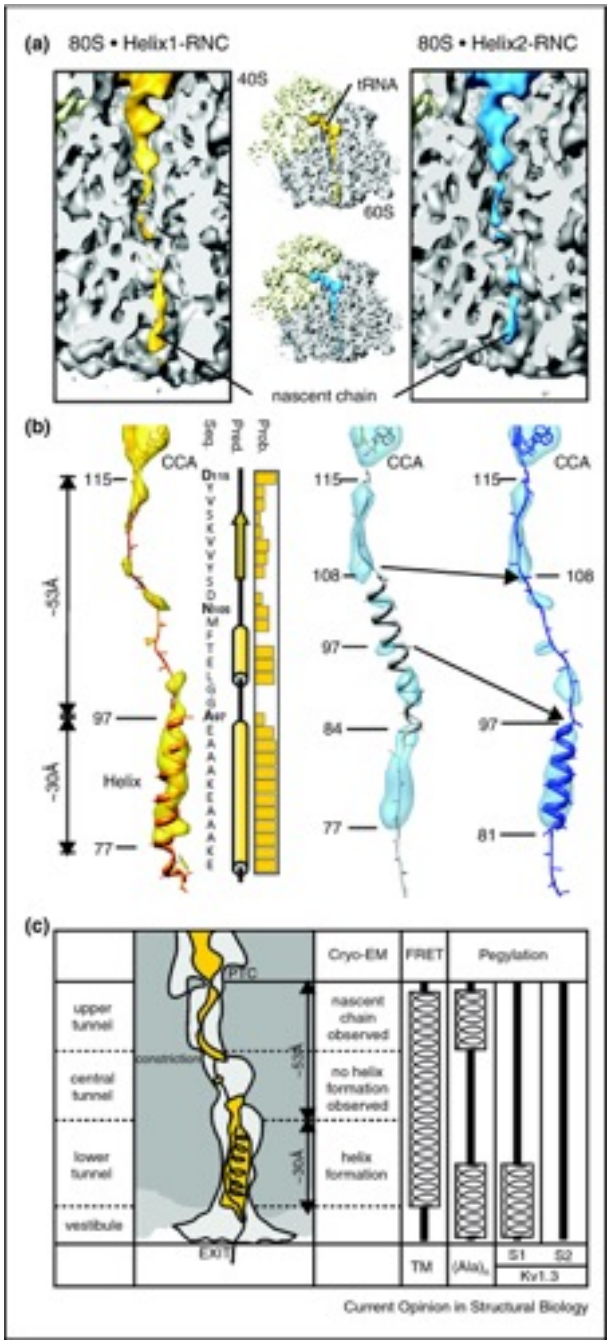
85,000 Å^3

SR1 mutant
GroES dissociates
In 300 mins! 15 times cell
Doubling time

"Hydrophilic"

185,000 Å^3

Ribosome Tunnel accommodates α-Helix (some regions)

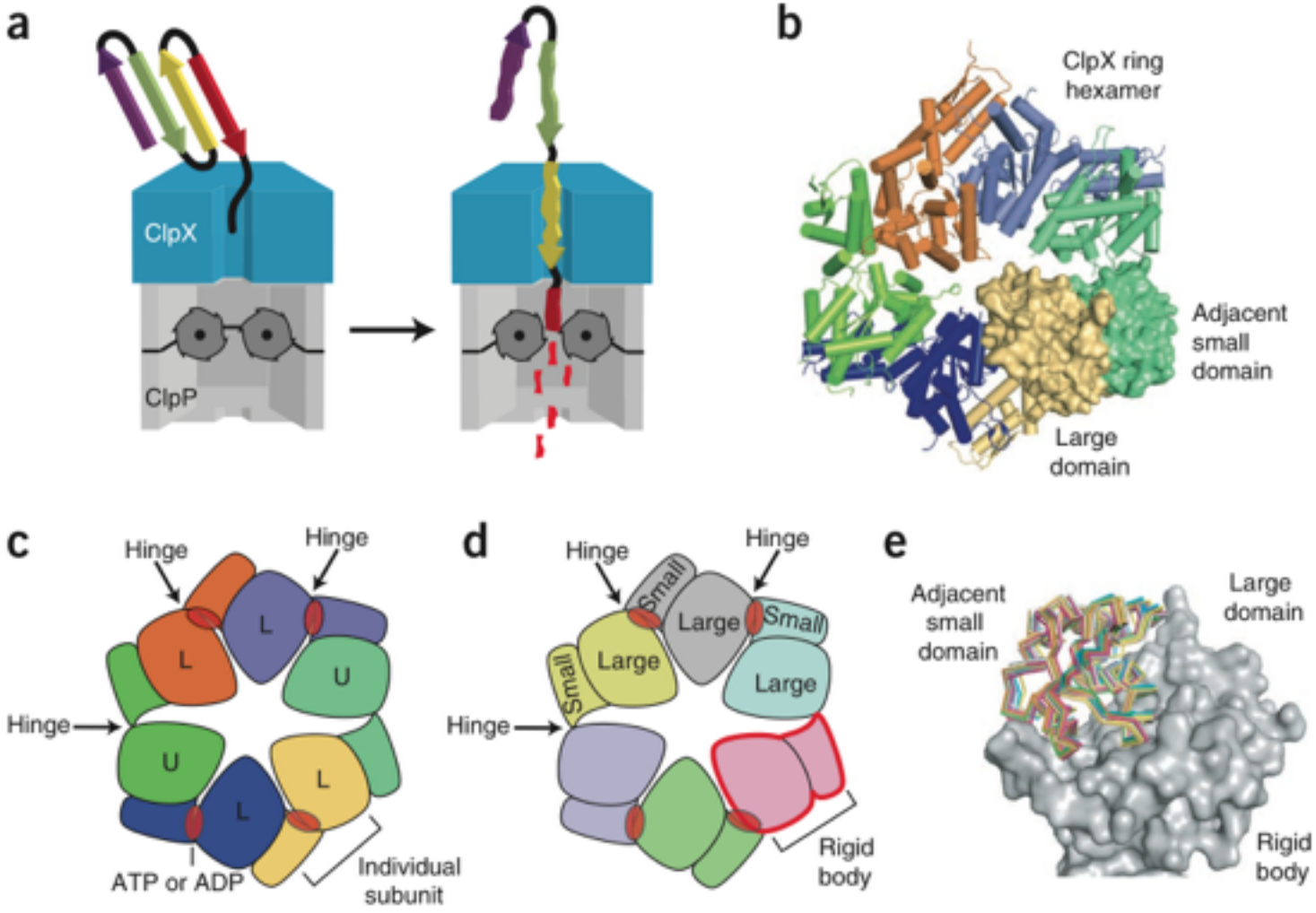


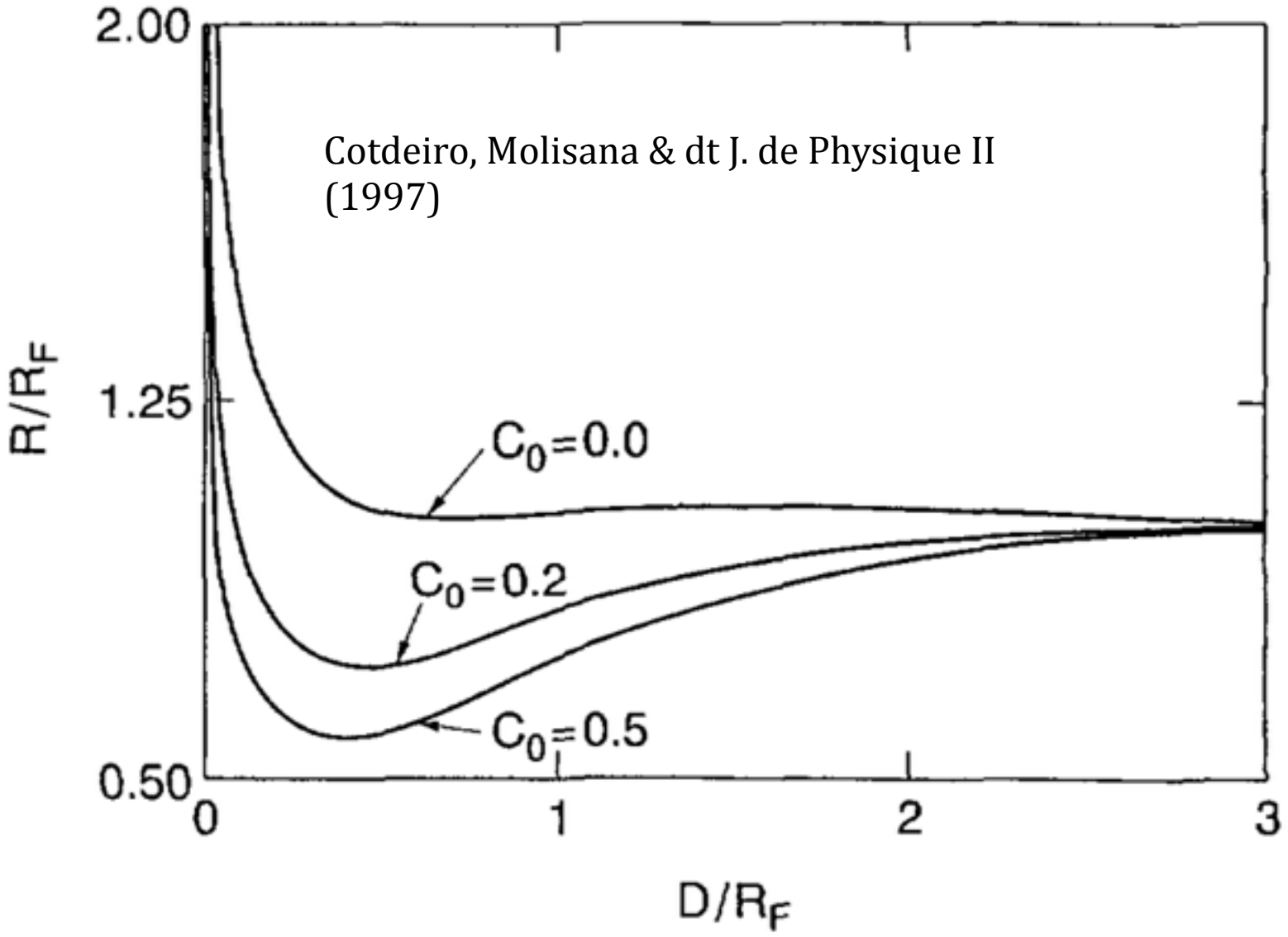
Cheryl Woolhead & A. Johnson *Cell* 2004

Protein exit tunnel:
About 100 Å long,
10-20 Å
in diameter; not uniform

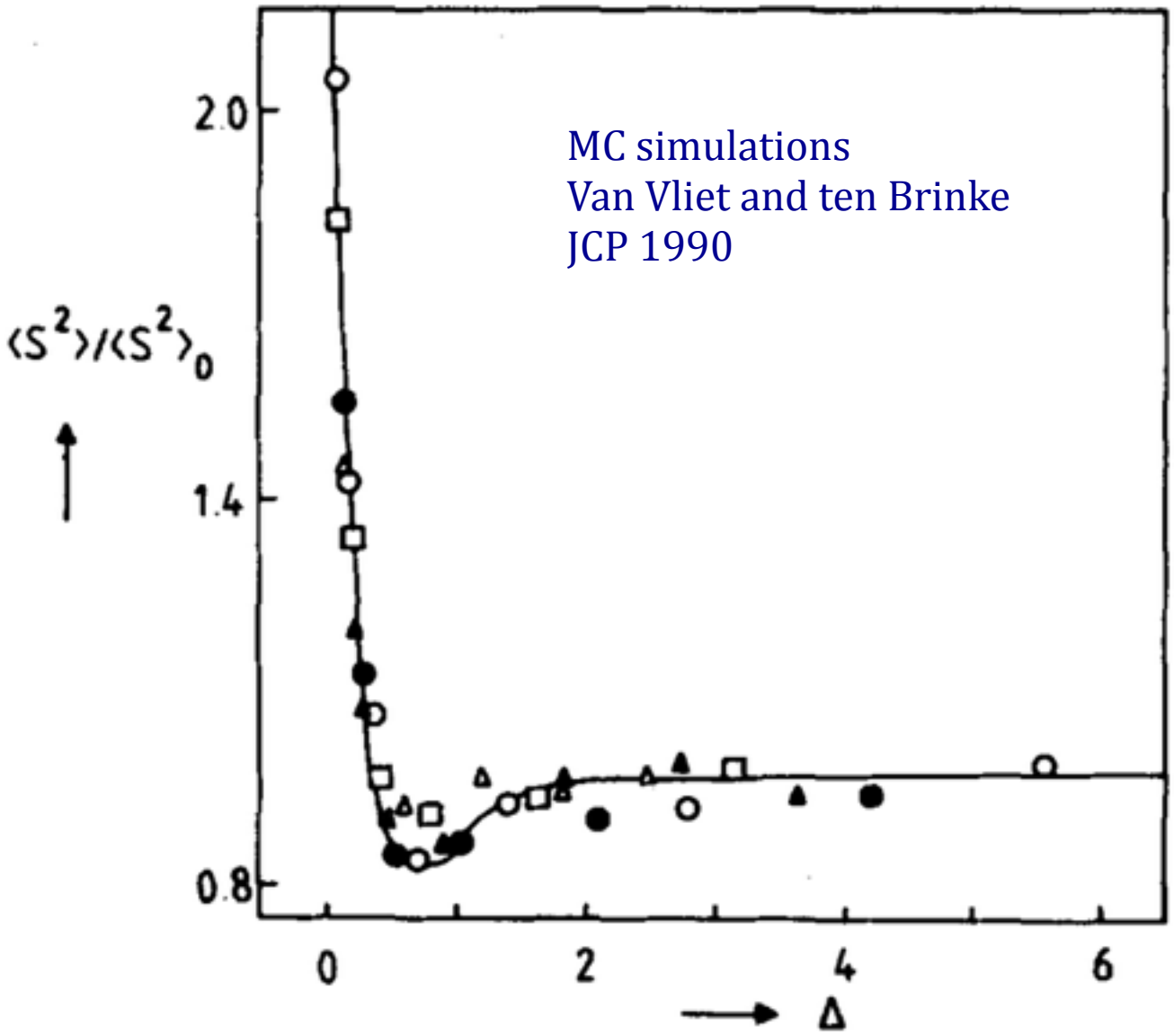
Helix formation in different Sequences possible
Wilson and R. Beckman *Curr .Opin Struct. Biol.* 2011

Proteosome degrades proteins: Nat. Struct. Mol. Biol
Bob Sauer (2012)





Cotdeiro, Molisana & dt J. de Physique II (1997)

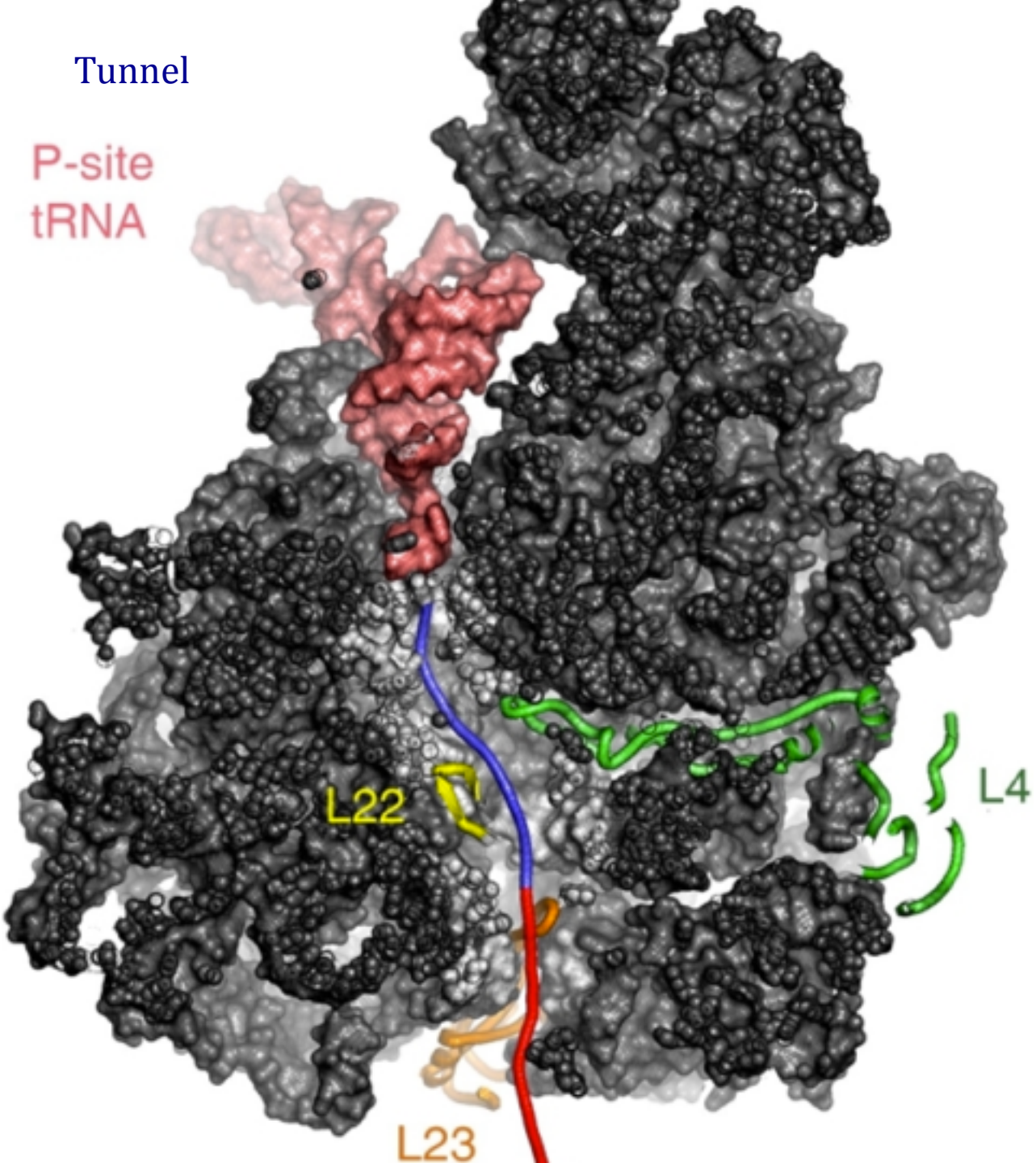

MC simulations
Van Vliet and ten Brinke
JCP 1990
⟨S²⟩/⟨S²⟩₀
2.0
1.4
0.8
0
2
4
6
Δ

Free energy cost of confining SAW in a slit or cylinder

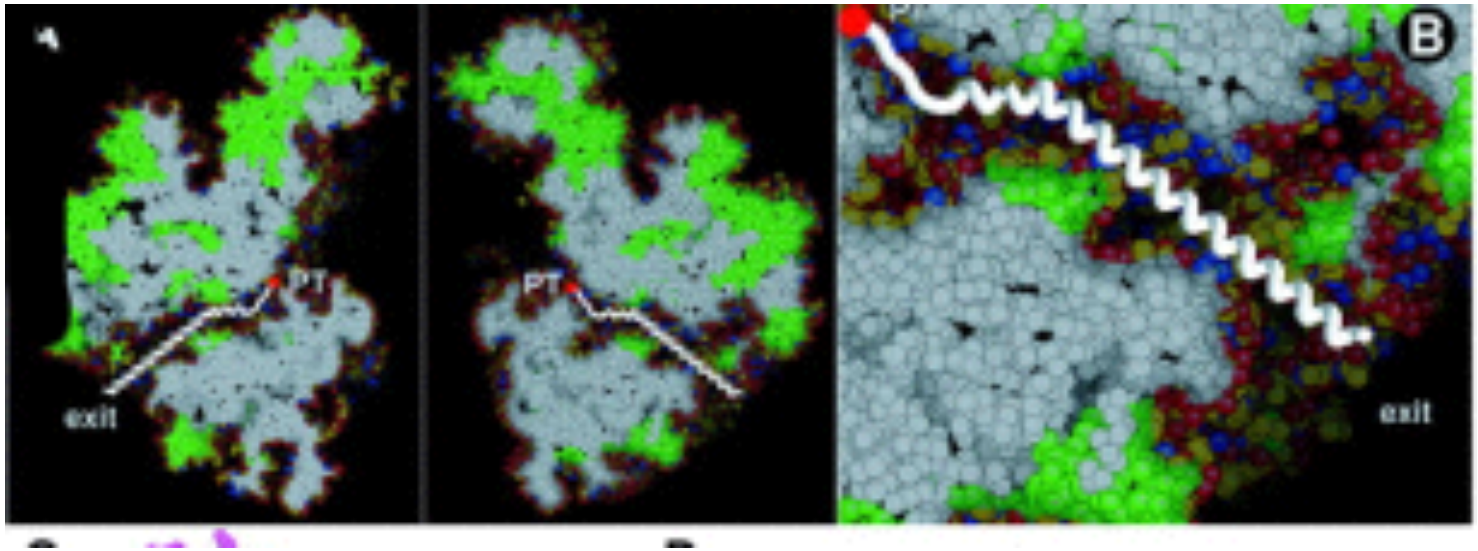
$\Delta F \approx k_B T\,(a/D)^{1/\nu} \approx (\text{const}) D^{-5/3}$

For a long Gaussian chain $\Delta F \approx k_B T\,(a/D)^2$

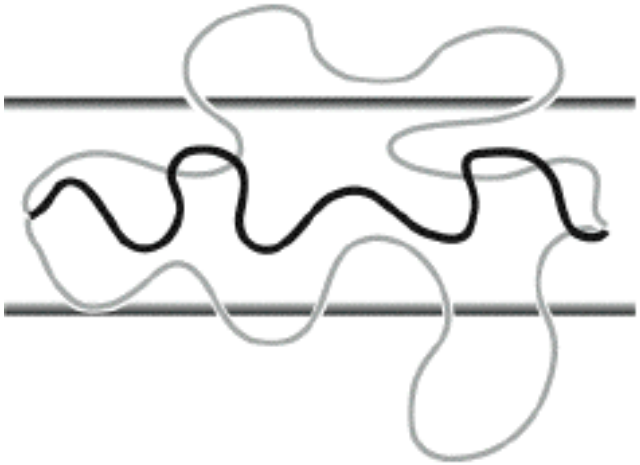
Analogy to particle in a box ground state energy


Tunnel
P-site
tRNA
L22
L4
L23

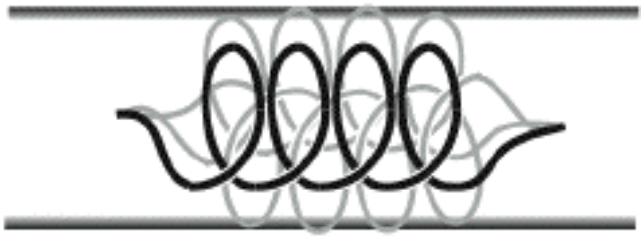
Model of a peptide in the Ribosome exit tunnel



α-helix
in a cylinder



COIL



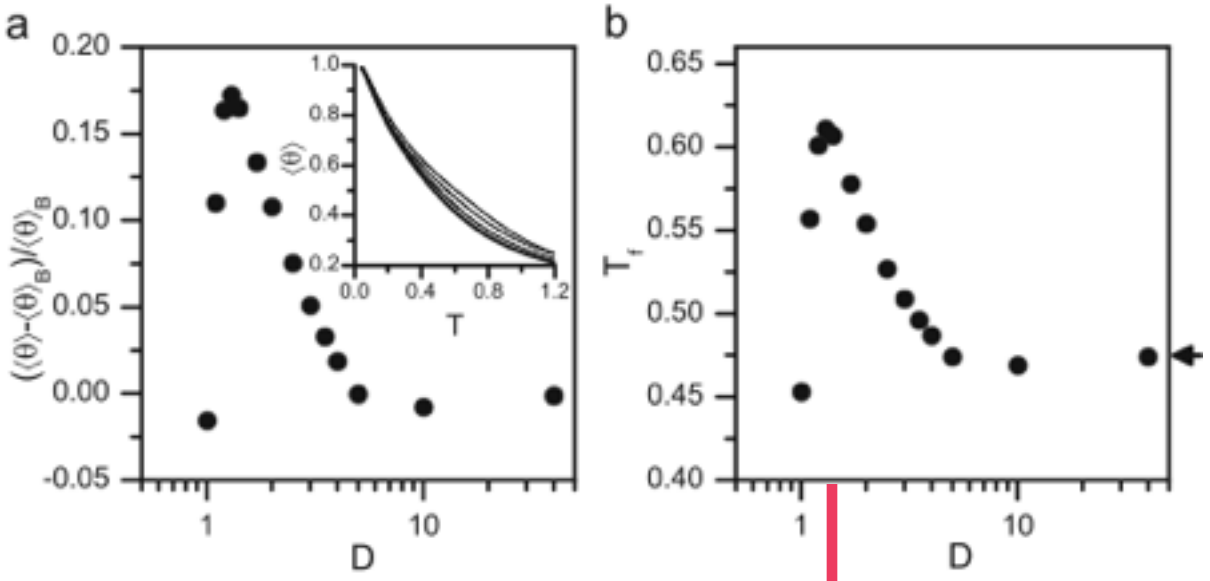
HELIX

Helix stabilization in a cylindrical pore

Guy Ziv, Gilad Haran, & dt PNAS 2005

Helix content

Melting temperature

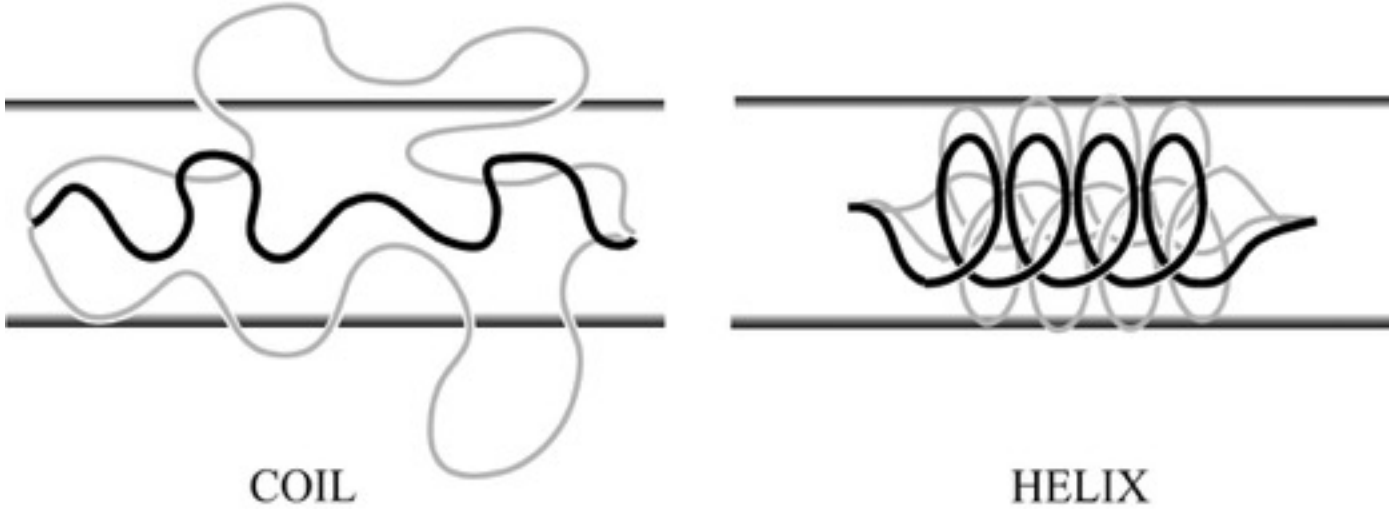


D* balance of conf entropy between folded and unfolded states

D is diameter of cylinder

D* (D < D* Helix not stable)

Confinement effects on conformational entropy of a polypeptide chain

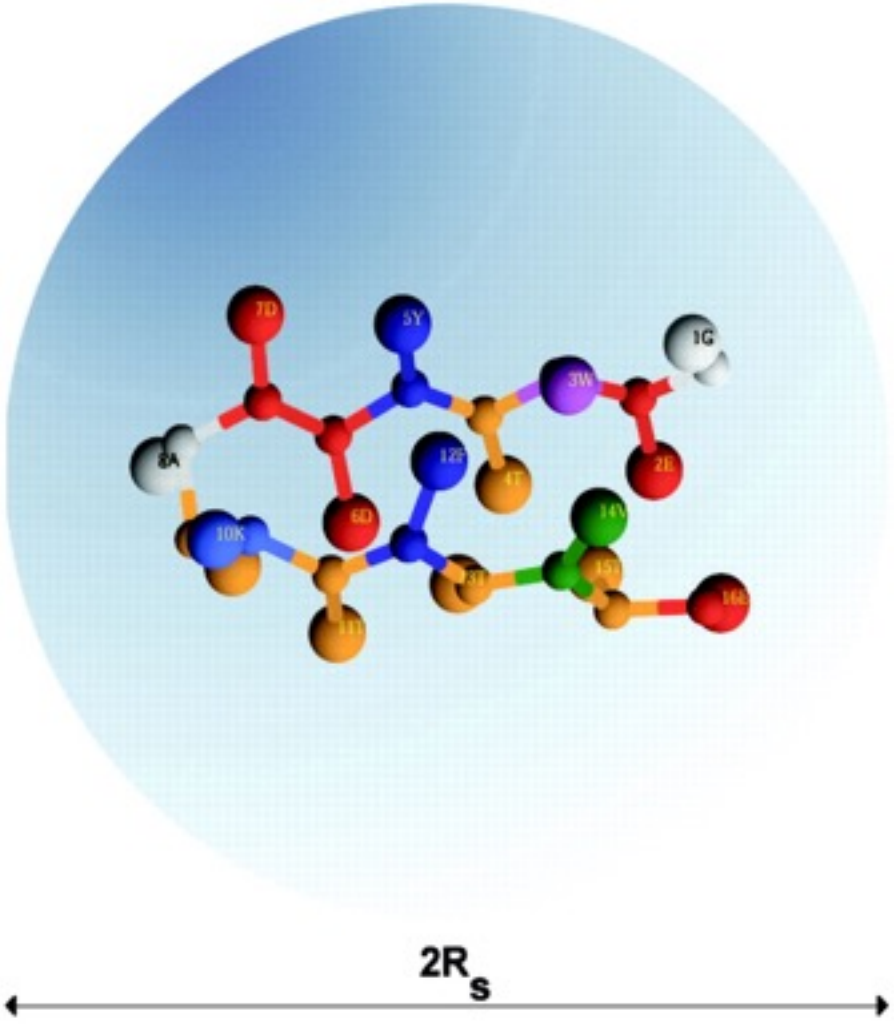


$$\Delta\Delta F \approx a_1 D^{-(2/3)} - a_2 D^{-(5/3)} = \Delta F(\text{helix}) - \Delta F(\text{coil})$$

$D < D^*$ confinement destabilizes α-helices

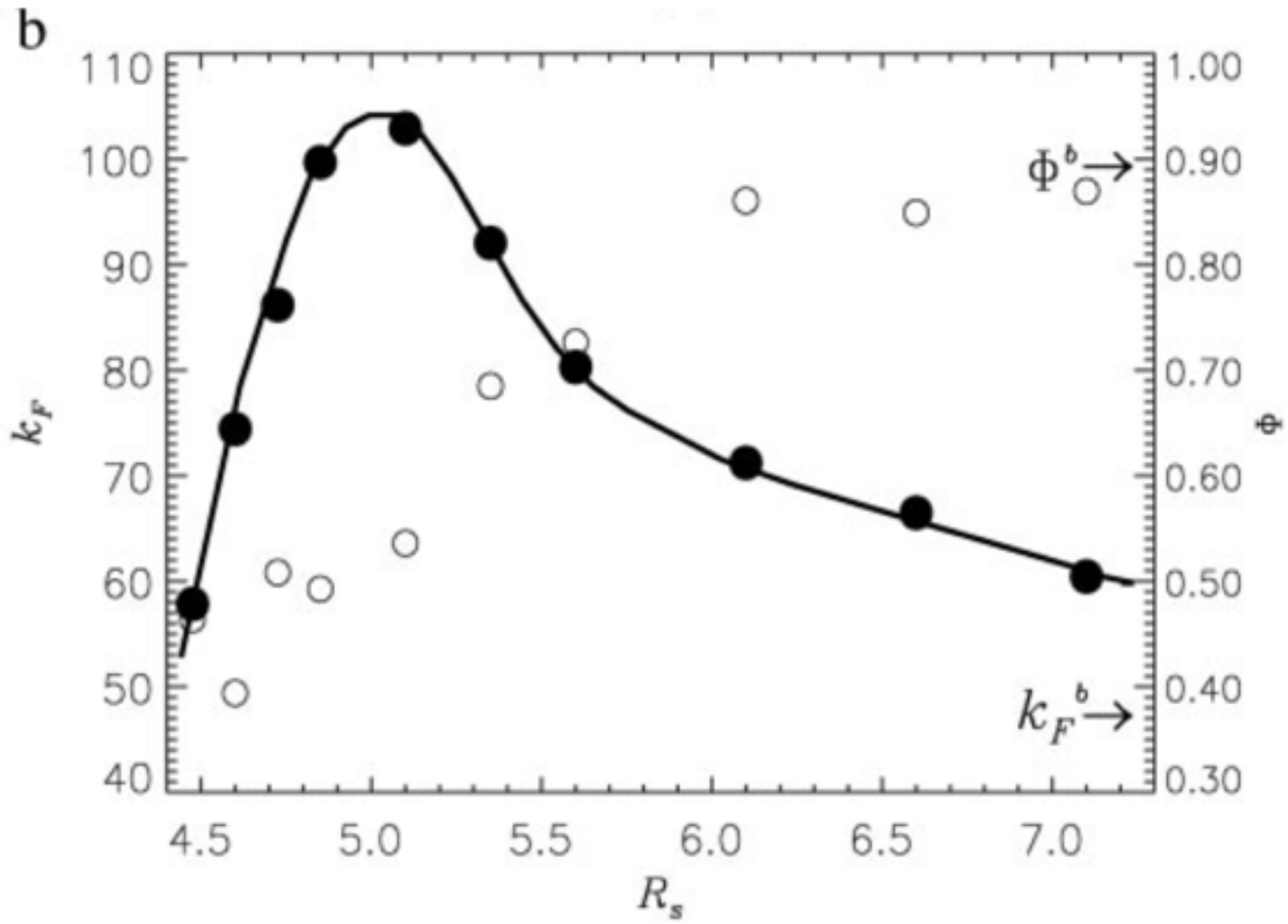
Ziv, Guy et al. (2005) Proc. Natl. Acad. Sci. USA 102, 18956-18961

Encapsulation of 16-mer C-terminal β-hairpin from the protein G in a spherical pore of the radius Rs is sketched.



D. K. Klimov et al. PNAS 2002;99:8019-8024

Folding Rates in Confined Spaces



k_F/k_F(bulk) can be computed analytically. Cheung & dt JMB (2006)