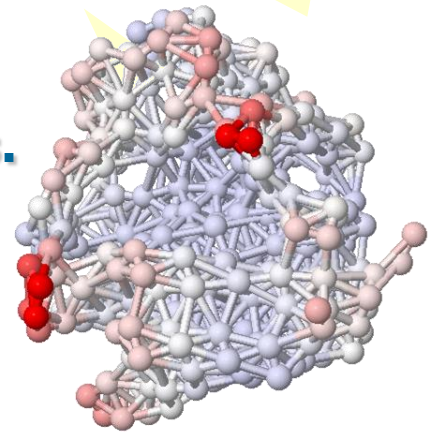
Protein Dynamics and Stability: Universality *vs.* Specificity

Rony Granek

The Stella and Avram Goren-Goldstein Dept. of Biotechnology Engineering
Ben-Gurion University of The Negev

- **Motivation.**
- **Fractal nature of proteins. Fractons – the vibrational normal modes of a fractal (scalar elasticity).**
- **Protein marginal stability – universal "equation of state".**
- **Dynamics: Vibrational & Random Walk MSD's, return probabilities, etc.**
- **Tensorial elasticity models.**
- **Dynamic Structure factor – preliminary results.**
- **Force induced unfolding – preliminary results.**
- **Conclusions.**



Coworkers

Tel-Aviv University

Shlomi Reuveni

Joseph Klafter

Ben-Gurion University

Marina de Leeuw

Roee Ben-Halevi

Amit Srivastava

RG

Natural Proteins

Long sequence of amino acids (20 types).

- Thousands of different proteins.
- Differ by sequence and length.
- Fold in different ways to give different 3-D fold structure.

Conflicting requirements:

- Specific folding – leads to a specific function (lock and key…).
- Large internal motion is needed to allow for biochemical function (enzymatic activity, antibody function, capturing and releasing ions, etc.).

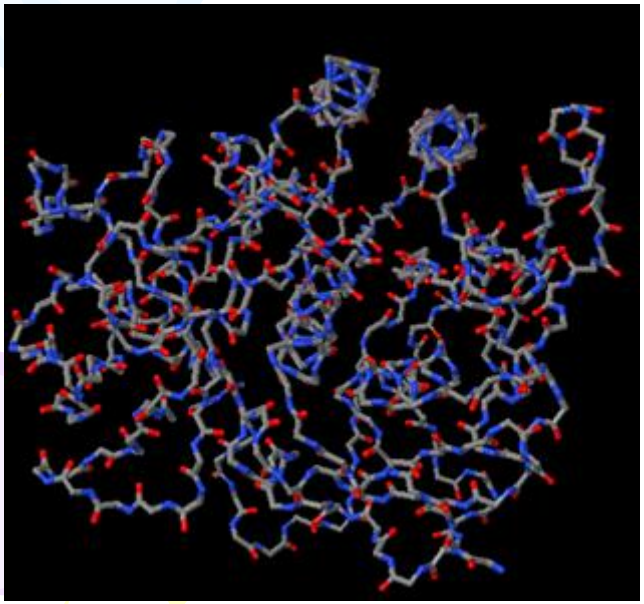
Problem –
a folded protein has less internal motion than an unfolded Protein.

Sequence

3D Structure

Dynamics ?

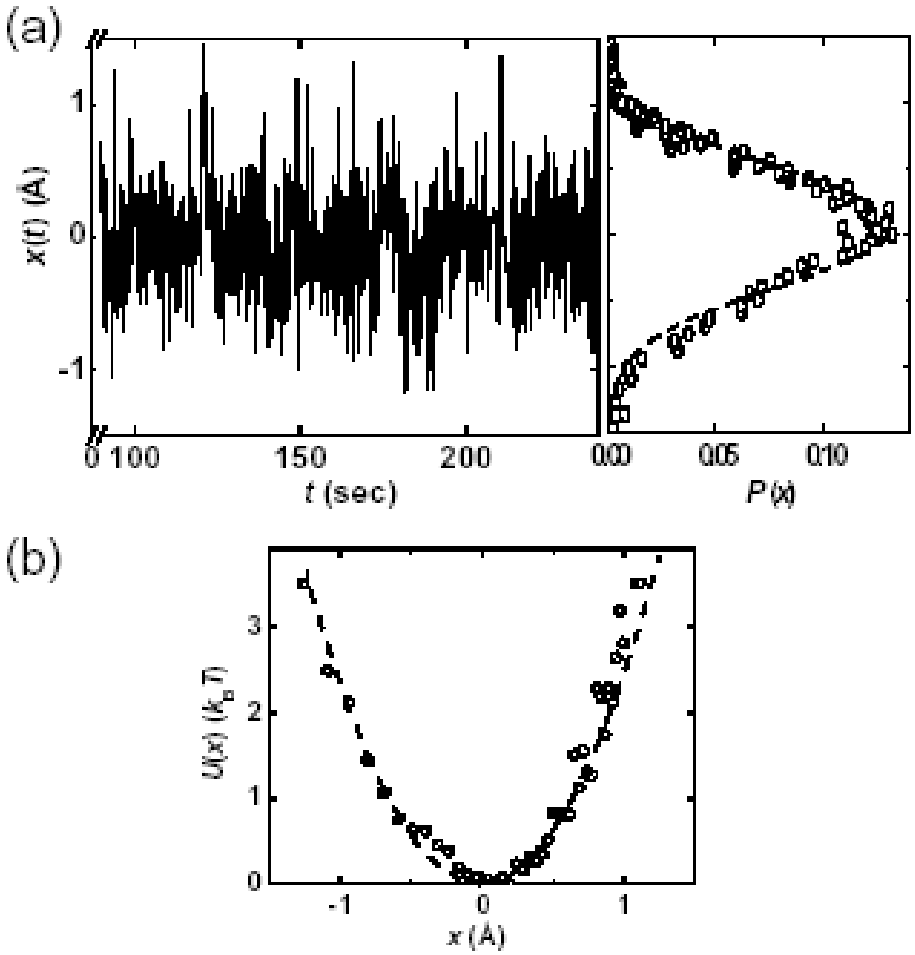
Function ?



www.rcsb.org 2gko

Single molecule experiments in proteins

- **FRET: Fluorescence resonant energy transfer (tens of angstroms).**
- **Photo-induced electron transfer (a few angstroms)**



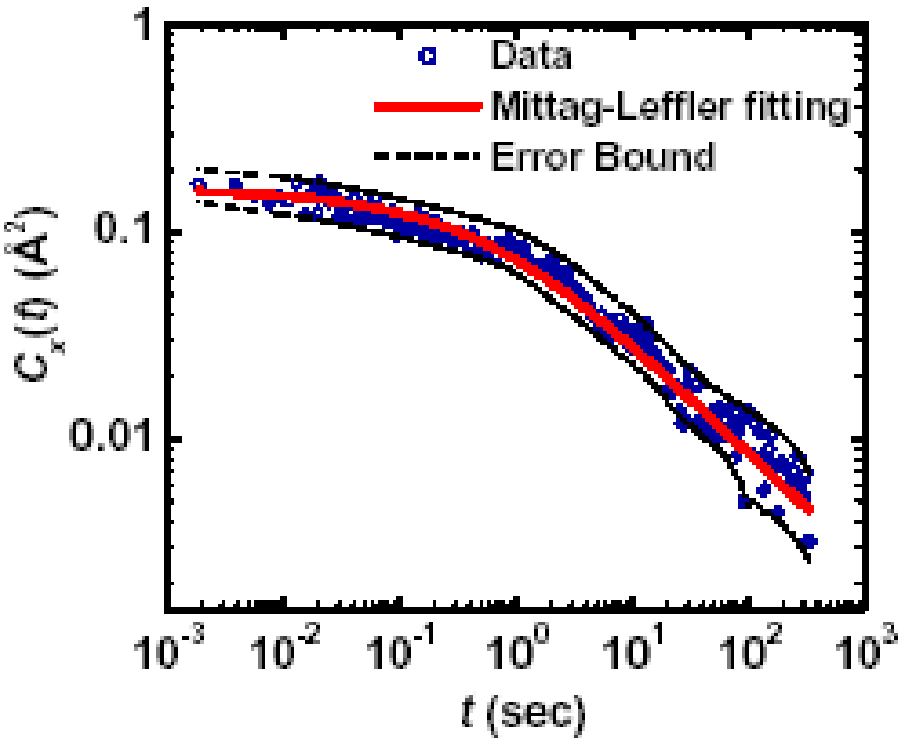
FL (fluorescein) and anti-FL complex
Tyr (tyrosine) – donor, FL – acceptor

$$x(t) = X(t) - X_{eq}$$

W. Min et al., PRL (2005)

Autocorrelation function

$$C_x(t) = \langle x(t)\,x(0) \rangle$$



$$C_x(t) \sim \begin{cases} 1 - \textbf{const.} & t^{1/2} \qquad t << 1\,s \\ t^{-1/2} & \qquad\quad t >> 1\,s \end{cases}$$

- **Small scale motion – VIBRATIONS?**

MSD of dihedral angles

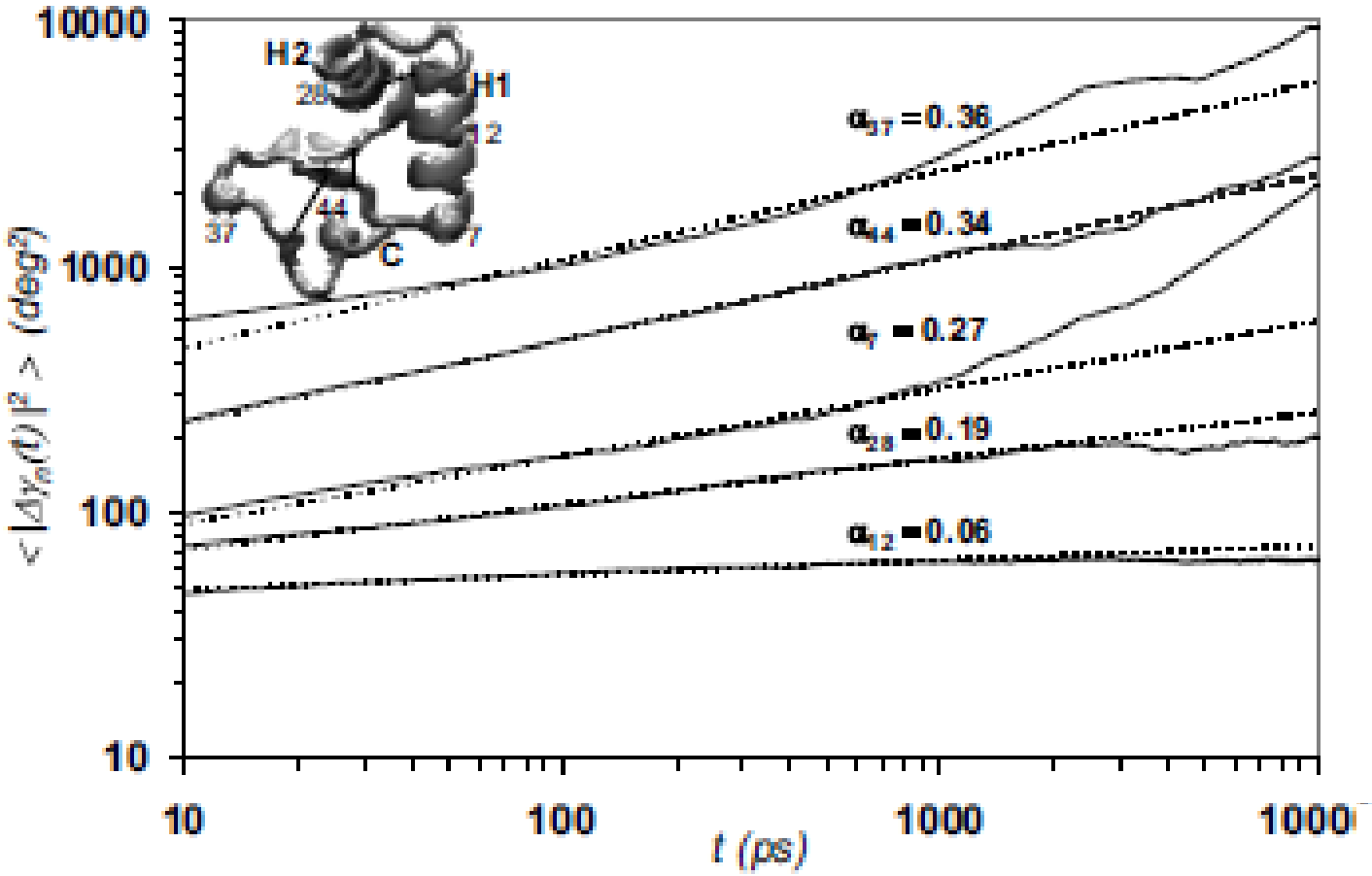
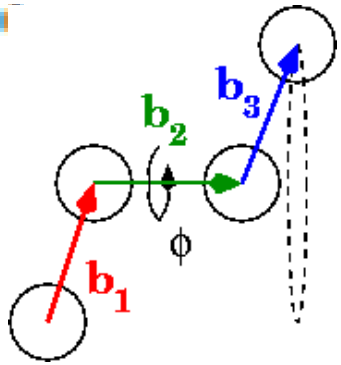
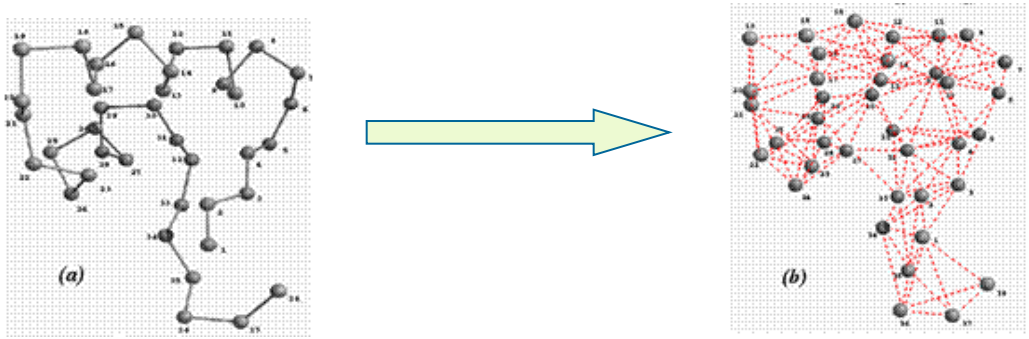


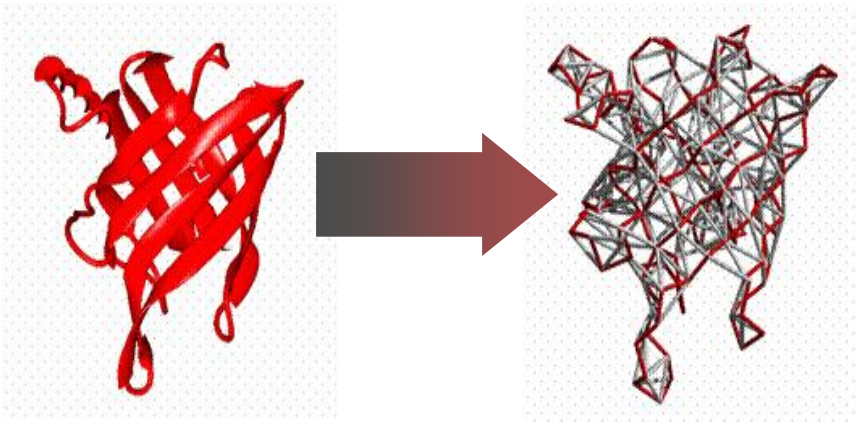
Fig. 1. Typical results for the mean-square displacement (MSD) of dihedral angles γ_n in crambin computed from MD on a double-log scale. For each dihedral angle γ_n, the MSD (solid lines) is compared with a power law (dotted lines) with an exponent α_n. *Inset* shows the main structural elements of crambin. The spheres represent the location of residues 7, 12, 28, 37, and 44, and the thick lines indicate the positions of the disulfide bridges.

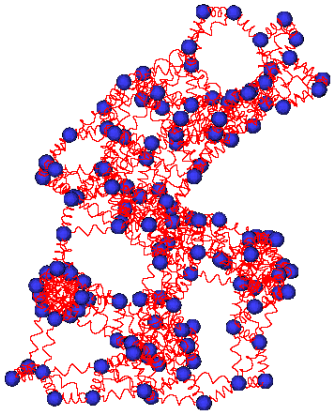


Patrick Senet[a,b], Gia G. Maisuradze[a], Colette Foulie[b], Patrice Delarue[b], and Harold A. Scheraga[a,1]

The Gaussian Network Model (GNM)





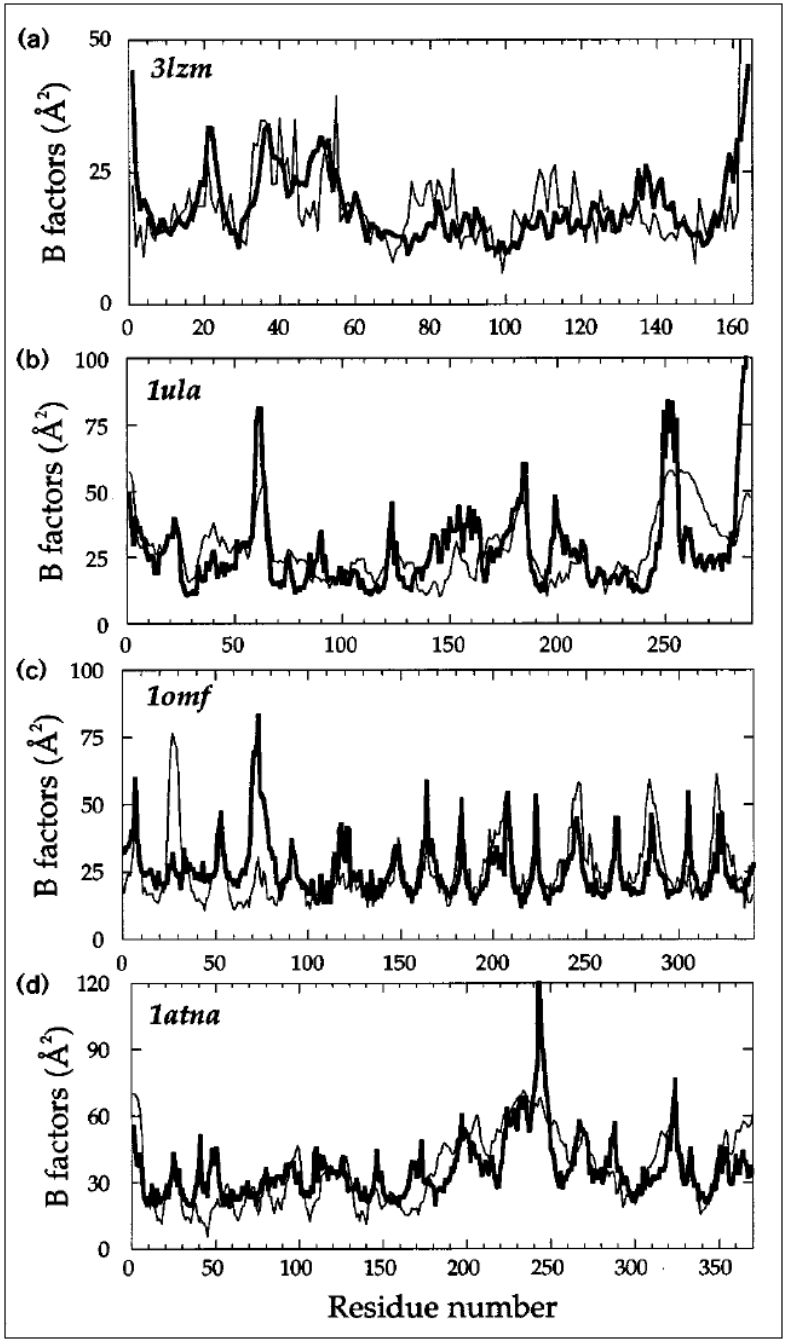


- Scalar elasticity.
- Springs exist only below a cutoff distance R_c.
- All springs have equal spring constant.

I. Bahar and coworkers

The Vibrational Mean Square Displacement (B – Factors)

$$B_i = \frac{8\pi^2}{3}\left\langle u_i^2 \right\rangle$$



I. Bahar and coworkers

Vibrations – normal mode analysis

Equations of motion

$$m \frac{d^2}{dt^2} \vec{u}(\vec{l}, t) = m \omega_o^2 \sum_{\vec{l}' \in \vec{l}} \left(\vec{u}(\vec{l}', t) - \vec{u}(\vec{l}, t) \right)$$

m mass

ω_o Spring natural frequency

$\vec{u}$ displacement

$\vec{l}$ "name" of an Alpha-carbon

Normal modes (eigenmodes, eigenstates)

$$u(\vec{l}, t) = \Psi_\alpha(\vec{l}) e^{i \omega_\alpha t}$$

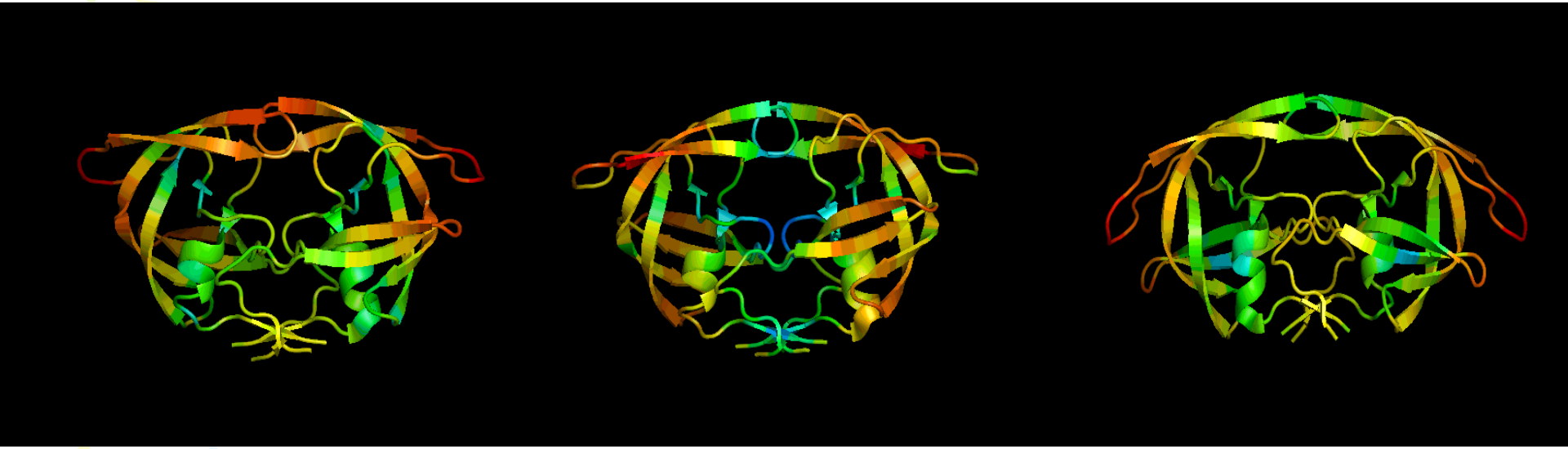
$$- \omega_\alpha^2 \Psi_\alpha(\vec{l}) = \omega_o^2 \sum_{\vec{l}' \in \vec{l}} \left(\Psi_\alpha(\vec{l}') - \Psi_\alpha(\vec{l}) \right)$$

Density of states (DOS, density of normal modes): $g(\omega)$

Cumulated DOS:

$$G(\omega) = \int_0^\omega g(\omega) d\omega$$

Modes of Motion – HIV Protease

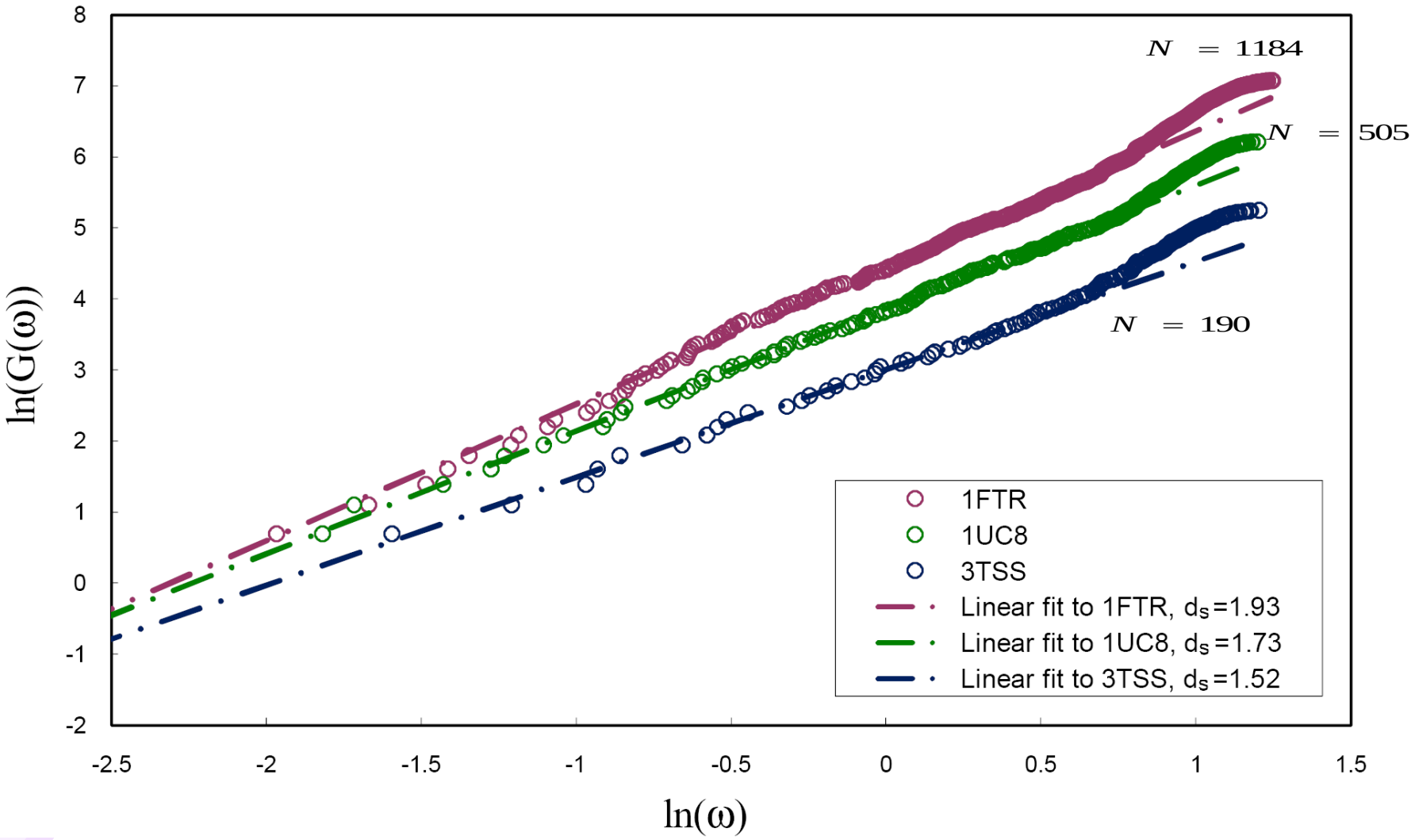


Mode 1

Mode 2

Mode 3

Calculation of cumulated DOS using the GNM



Powerlaws!

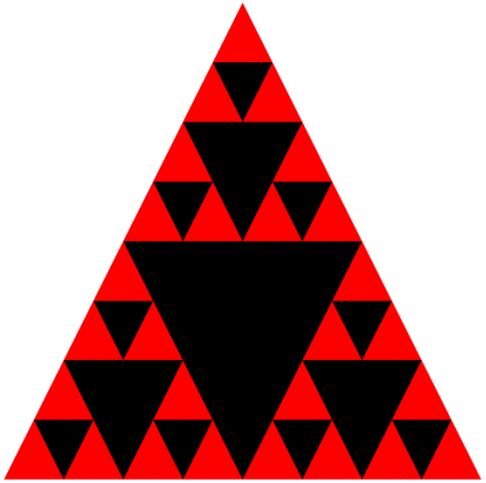
Reuveni et al., PRL (2007)

Fractals

- Definition –

  “a rough or fragmented geometric shape that can be split into parts, each of which is (at least approximately) a reduced-size copy of the whole”

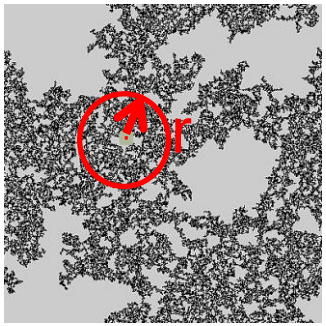
  Benoît Mandelbrot (1975)



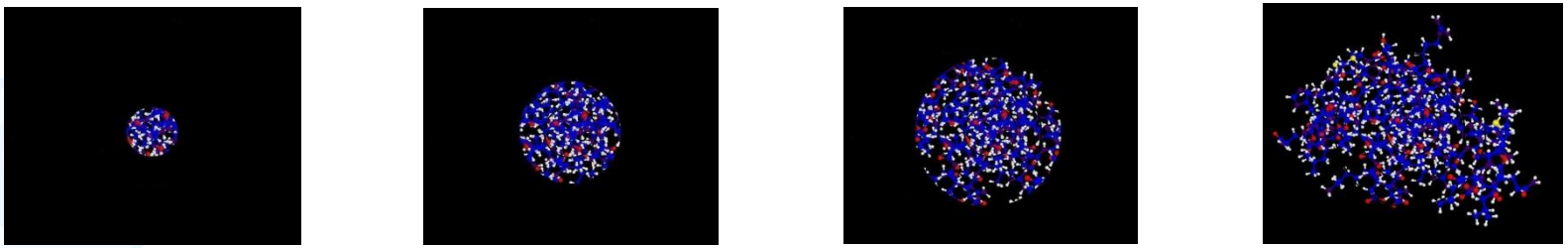
Fractal nature of proteins

Mass fractal dimension d_f :

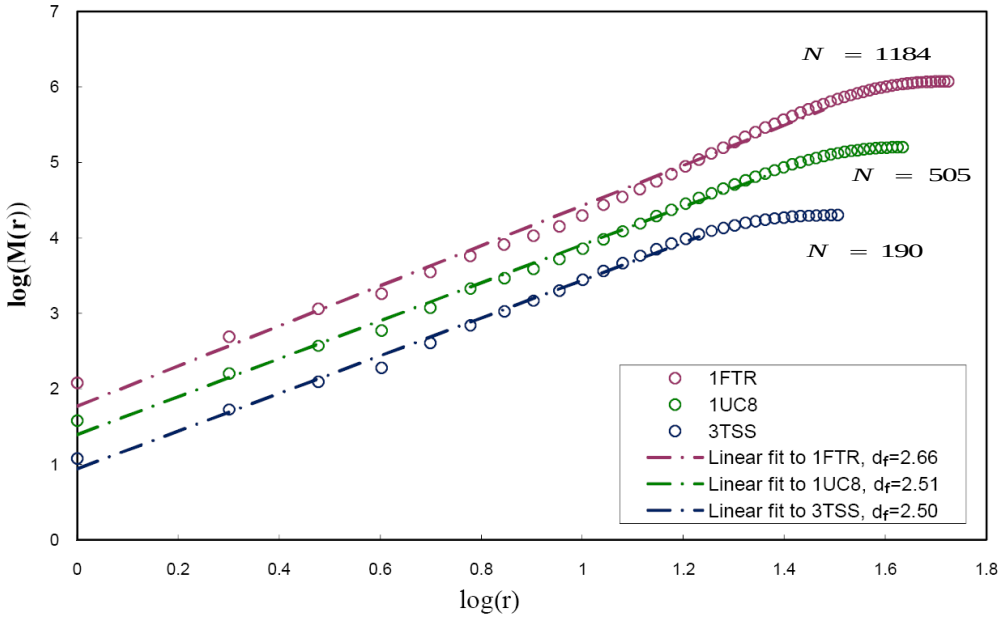
$$M \sim r^{d_f}$$



The atoms enclosed in spheres of different radii (pdb: 1OCP)



D. M. Leitner
and coworkers



S. Alexander and R. Orbach (1982)

Fractons

Vibrations of the fractal

$$m \frac{d^2}{dt^2} \vec{u}(\vec{l}, t) = m \omega_o^2 \sum_{\vec{l}' \in \vec{l}} \left(\vec{u}(\vec{l}', t) - \vec{u}(\vec{l}, t) \right)$$

m mass

ω_o Spring natural frequency

$\vec{u}$ displacement

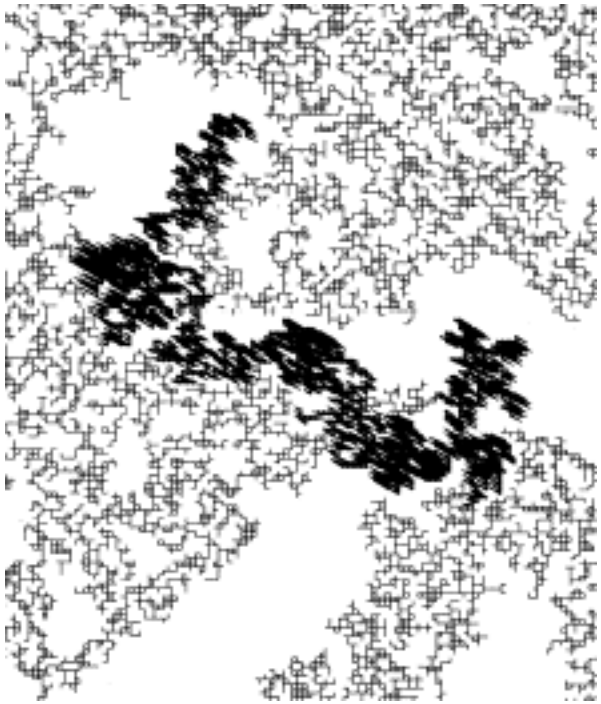
$\vec{l}$ "name" of a point mass

Normal modes (eigenmodes, eigenstates) – Fractons:

$$u(\vec{l}, t) = \Psi_\alpha(\vec{l}) e^{i \omega_\alpha t}$$

- **Strongly localized eigenstates** $\Psi_\omega(\vec{l})$
- **Density of states** $g(\omega) \sim \omega^{d_s - 1}$

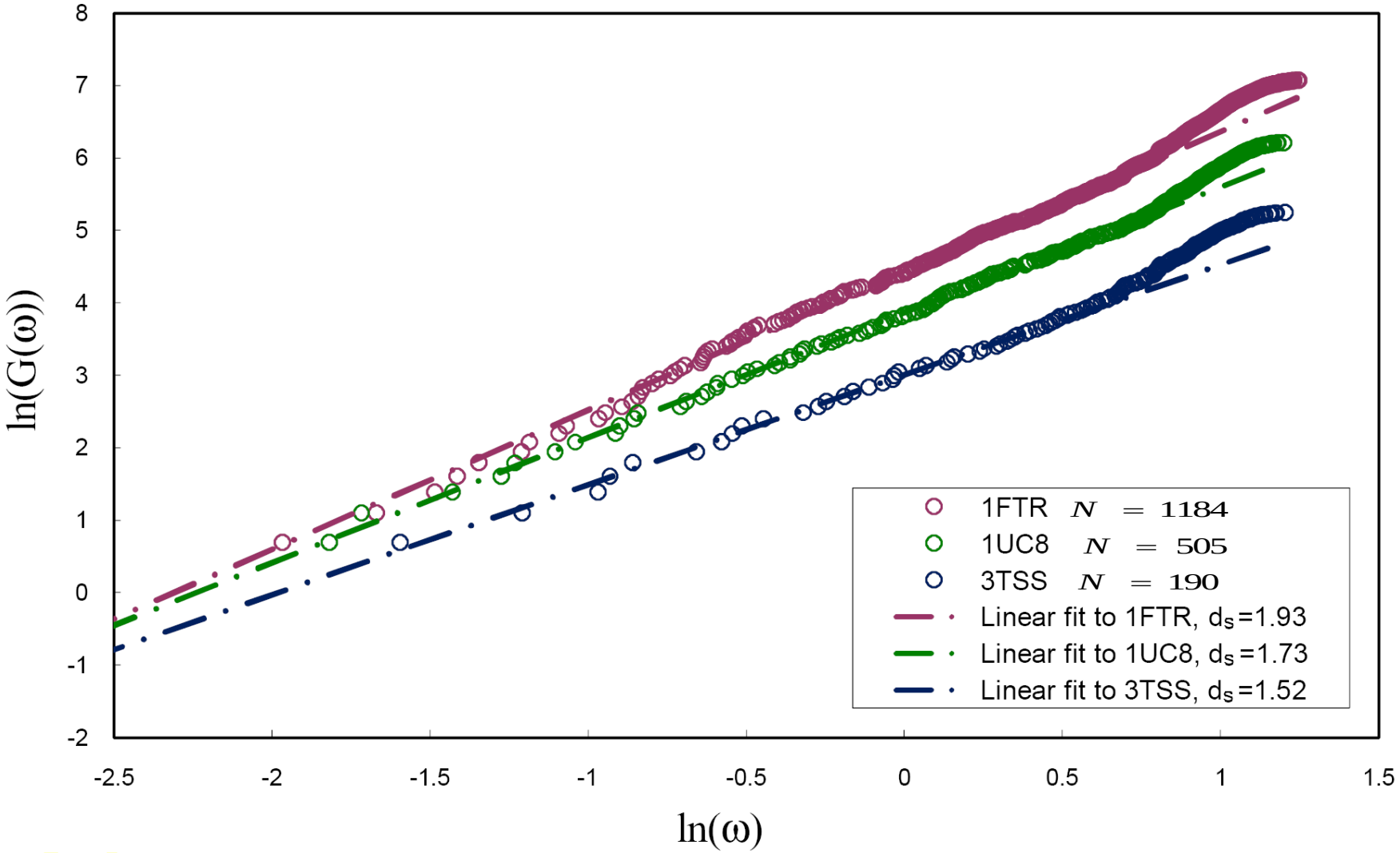
d_s – Spectral dimension



Yakubo and Nakayama (1989)

Calculation of cumulated DOS using the GNM

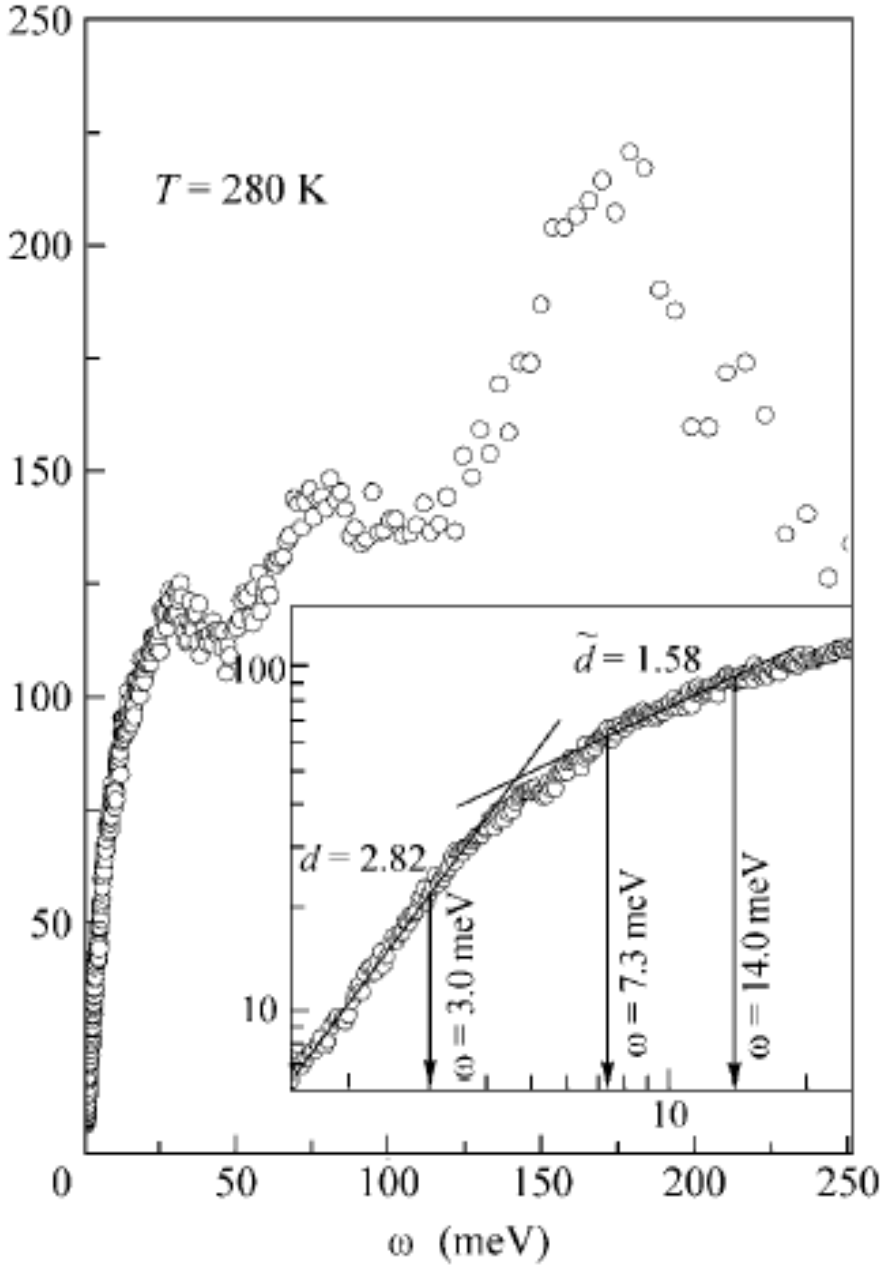
$$G(\omega) = \int_0^{\omega} g(\omega)\, d\omega \sim \omega^{d_s}$$



Vibrational Density of States of Hen Egg White Lysozyme

S. G. Lushnikov[1], A. V. Svanidze[1], and I. L. Sashin[2]

$g(\omega)$



JETP Letters, Vol. 82, No. 1, 2005, pp. 30–33.

Protein Stability & Unfolding

$\vec{u}_\alpha$ – Amplitude of a normal mode $\Psi_\alpha(\vec{l})$

Equipartition $$\left\langle \vec{u}_\alpha^{\,2} \right\rangle_T = \frac{3k_B T}{m\,\omega_\alpha^{\,2}}$$

Thermal fluctuations of the displacements ($d_s < 2$)

$$\left\langle \vec{u}^{\,2} \right\rangle_T = \sum_\alpha \left\langle \vec{u}_\alpha^{\,2} \right\rangle_T = \int_{\omega_{\min}}^{\omega_o} d\omega\; g(\omega) \left\langle \vec{u}_\omega^{\,2} \right\rangle_T \sim \omega_{\min}^{-(2-d_s)} \sim N^{(2/d_s - 1)}$$

Landau-Peierls Instability

$$\omega_{\min} \sim R_g^{-d_f/d_s} \sim N^{-1/d_s}$$

N – # of amino acids ("polymer index")

If $d_s < 2$, $\left\langle \vec{u}^{\,2} \right\rangle_T$ increases with increasing N !

Large fluctuations may assist enzymatic/biological activity.

But $\left\langle \vec{u}^{\,2} \right\rangle^{1/2}$ should **not** exceed the mean inter-amino acid distance, otherwise protein must **unfold** (or not fold).

- **Marginal stability. To have large amplitude motion but remain folded:**

  Proteins can "live" in the "twilight" zone: Folded-Unfolded !

- **To keep proteins folded, d_s should depend on N :**

  d_s **should approach the value of** 2 **for large proteins.**

Instability threshold: Universal relation between exponents

Cluster melting analog

Unfolding/Melting occurs from the surface inward

Landau-Peierls Instability: $\left\langle \vec{u}^{\,2} \right\rangle_T \sim \frac{k_B T}{m \omega_o^2} N^{(2/d_s - 1)}$

Lindenman-like criterion for unfolding $\left\langle \vec{u}^{\,2} \right\rangle_{surface} \sim R_c^{\,2}$

Assume $\left\langle \vec{u}^{\,2} \right\rangle = p \left\langle \vec{u}^{\,2} \right\rangle_{surface} + (1 - p) \left\langle \vec{u}^{\,2} \right\rangle_{bulk} \sim N^{(2/d_s - 1)}$

Unfolding: $\frac{k_B T}{m \omega_o^2} N^{(2/d_s - 1)} \sim p R_c^{\,2}$

Take: $p \sim \frac{S}{V} \sim \frac{1}{R_g} \sim N^{-(1/d_f)}$

$$\frac{2}{d_s} + \frac{1}{d_f} = 1 + \frac{b}{\ln N}$$

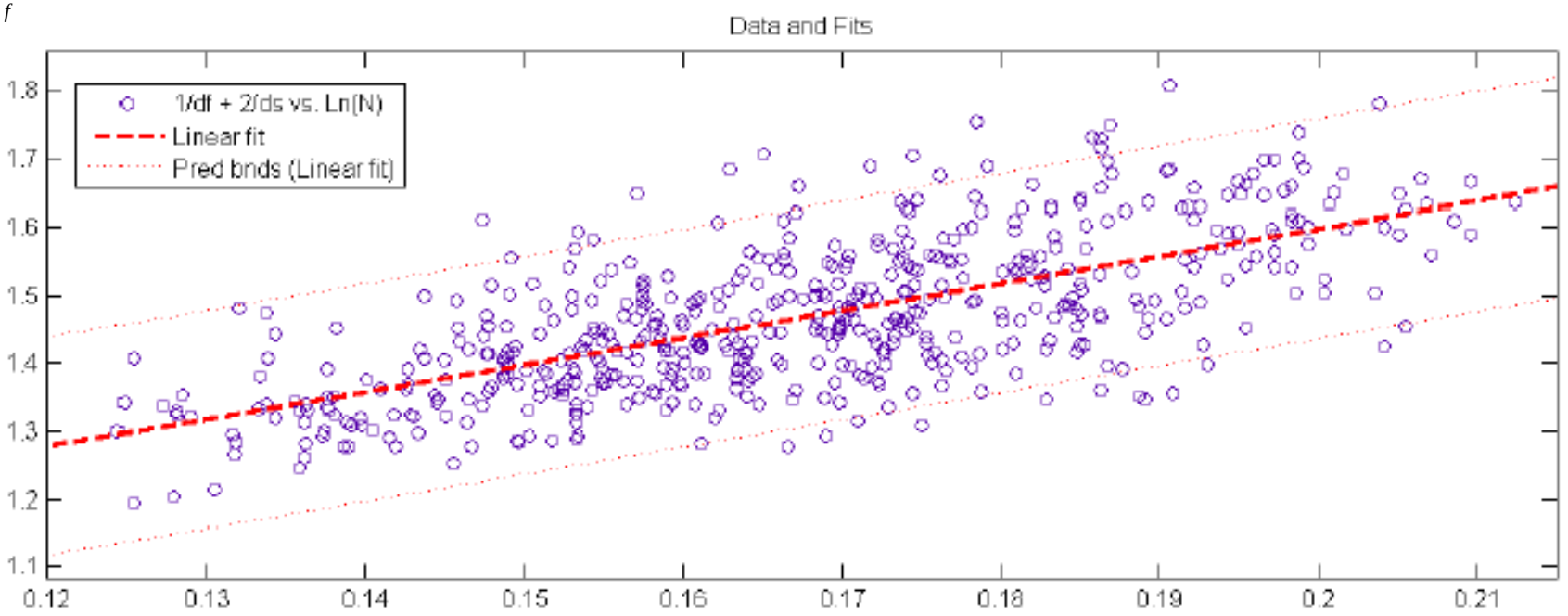
$$b \approx \ln\left(\frac{m \omega_o^2 R_c^2}{k_B T} \right)$$

Reuveni et al., PRL (2007)

Fitting the data of 543 proteins to: $\frac{2}{d_s} + \frac{1}{d_f} = a + \frac{b}{\ln N}$

Best fit: CC=0.55, $a = 0.90$, $b = 4.53$ $\qquad a_{the} = 1,\ b_{the} \approx 4$

$\frac{2}{d_s} + \frac{1}{d_f}$



$1/\ln N$

Reuveni et al., PRL (2007)

GNM cutoff length R_c=6A

Larger set:

Selecting Proteins

Structures in the PDB : 48,638

- No RNA, no DNA
- No ligands
- No more than 95% ID
- No Peptides (less than 100 amino acids)

Files with missing data - removed

High accuracy in determining d_s and d_f,

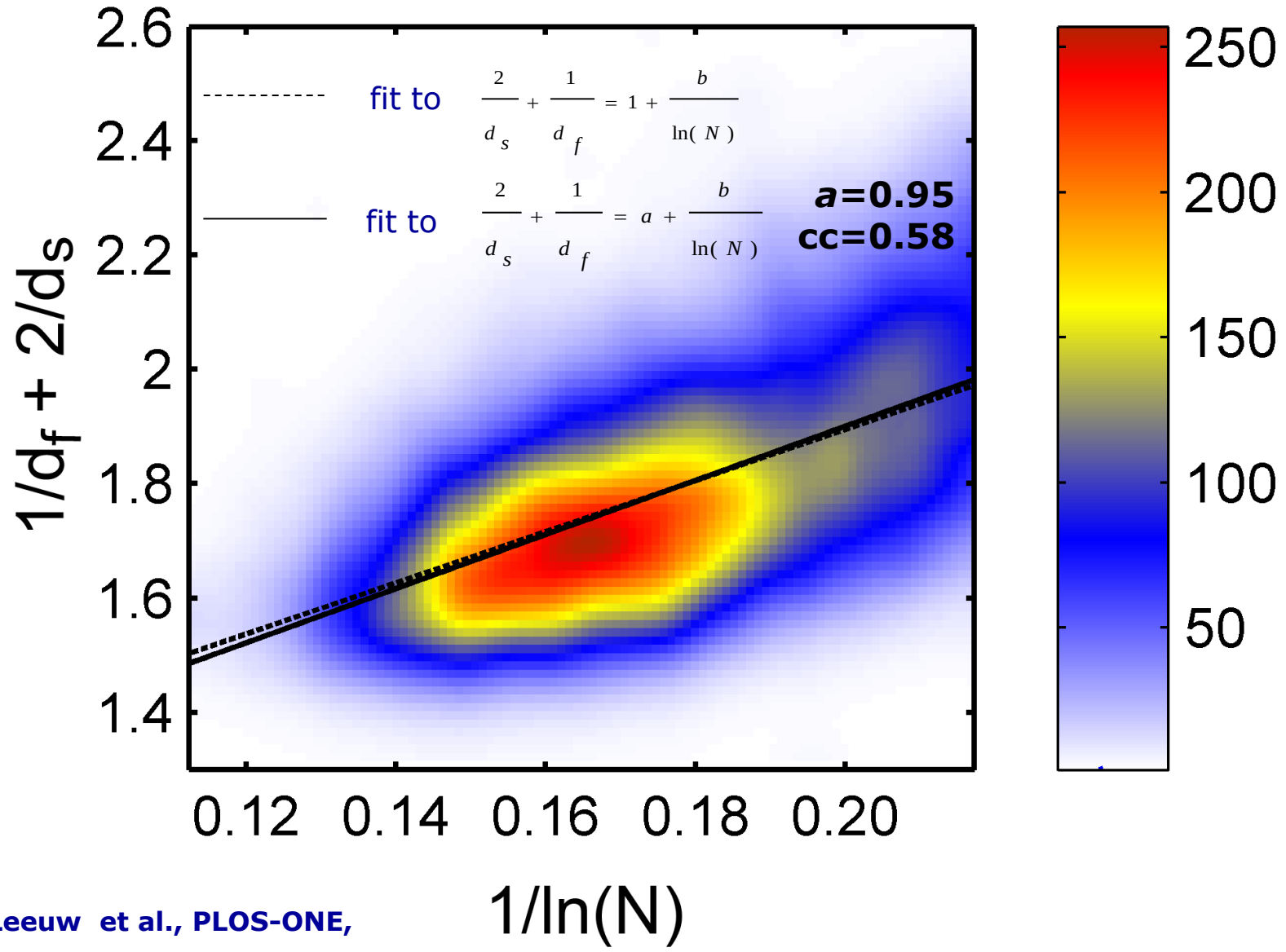
$$R^2 > 0.99$$

GNM cutoff length R_C=6 A

Structures to analyze : 4249

4249 proteins

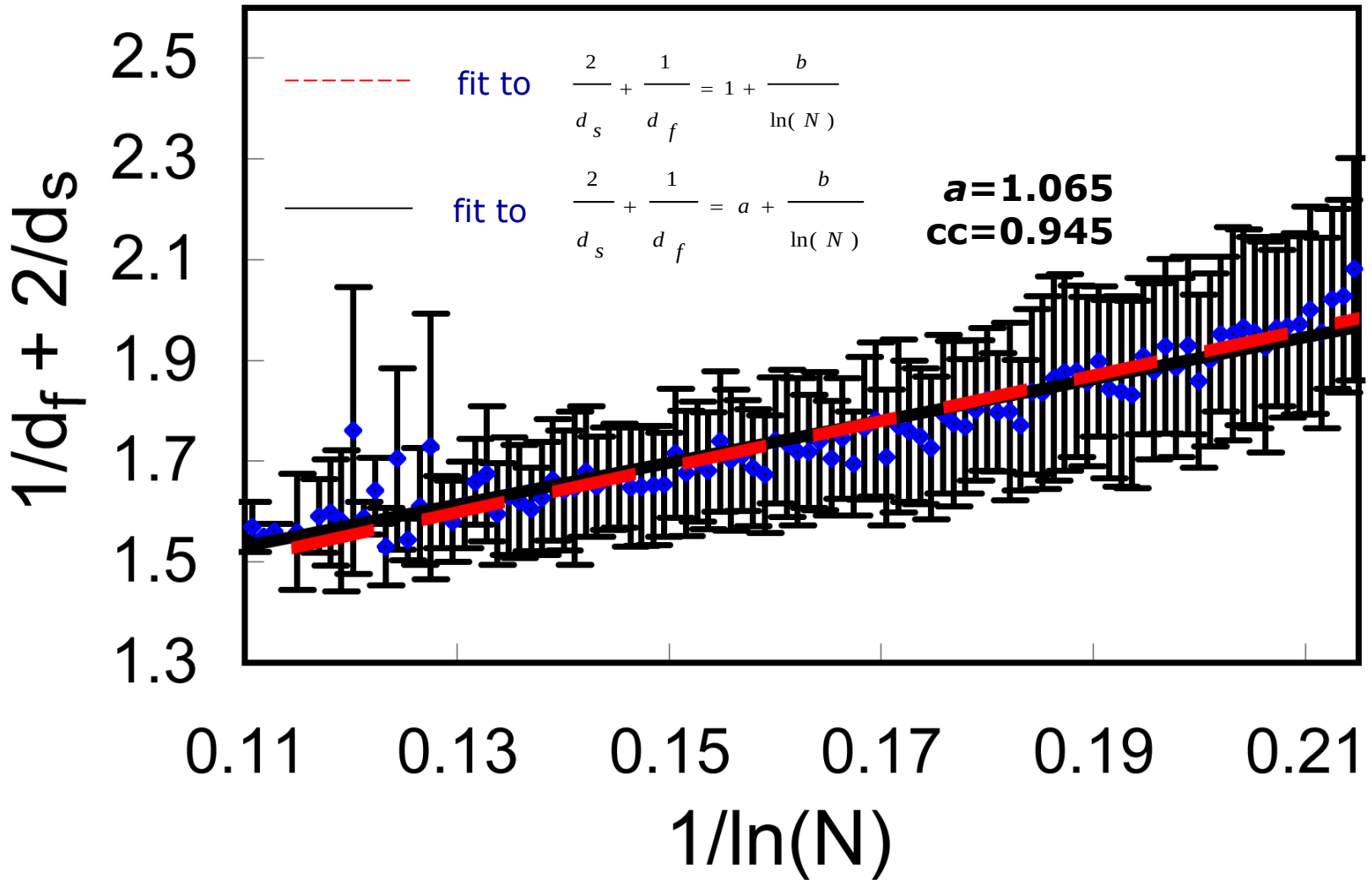
Colored histogram (100X100 bins)



M. de Leeuw et al., PLOS-ONE, 2009

4249 proteins

X-axis separated into 100 bins



M. de Leeuw et al., PLOS-ONE, 2009

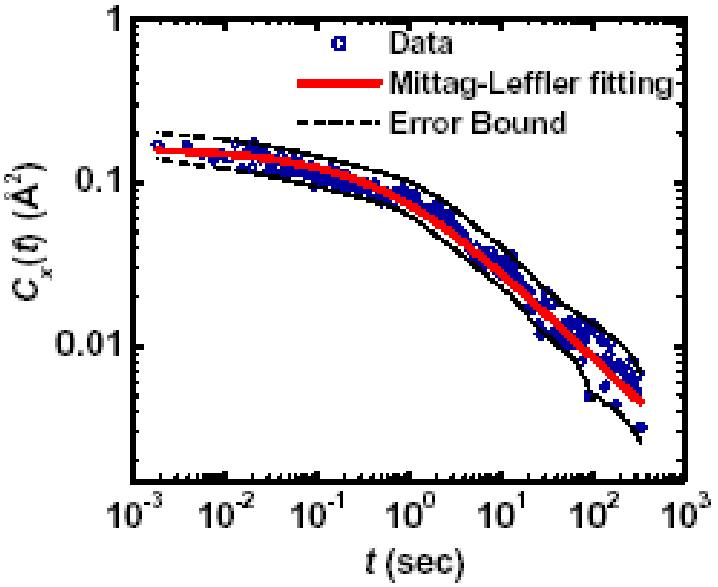
Dynamics

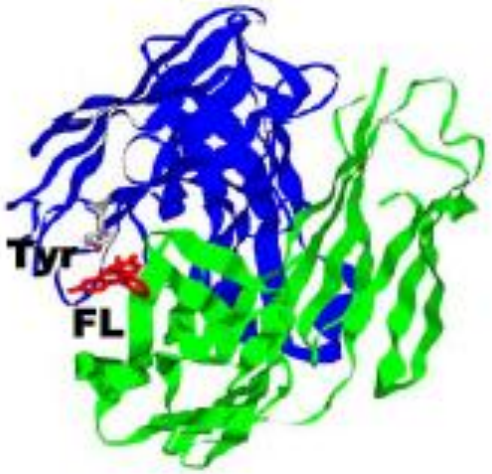
I) Time-autocorrelation function of the distance between two alpha-carbons

Motivation: Single molecule experiments in proteins (Xie and coworkers)

$$x(t) = X(t) - X_{eq}$$

Autocorrelation function $$C_x(t) = \langle x(t)\,x(0) \rangle$$





Displacement difference time-autocorrelation function

Two point masses (alpha-carbons), $\vec{l}$ and $\vec{l}\,'$

Positions in space $\vec{R}(\vec{l},t)$ and $\vec{R}(\vec{l}\,',t)$

Separation vector $\vec{X}(t) = \vec{R}(\vec{l},t) - \vec{R}(\vec{l}\,',t)$

Equilibrium spacing $\vec{X}_{\mathbf{eq}}$

Displacement difference vector $\vec{x}(t) = \vec{X}(t) - \vec{X}_{\mathbf{eq}}$

$$\vec{x}(t) = \vec{u}(\vec{l},t) - \vec{u}(\vec{l}\,',t)$$

Expansion in normal modes $\vec{u}(\vec{l},t) = \sum_{\alpha} \vec{u}_{\alpha}(t)\Psi_{\alpha}(\vec{l})$

+**disorder averaging**

$$\left\langle \vec{x}(t)\cdot\vec{x}(0)\right\rangle = \frac{2}{N}\sum_{\alpha}\left(1-\overline{\Psi}_{\alpha}(|\vec{l}-\vec{l}\,'|)\right)\left\langle \vec{u}_{\alpha}(t)\cdot\vec{u}_{\alpha}(0)\right\rangle$$

Granek & Klafter, PRL (2005)

Static fluctuations (t=0)

$$\left\langle \vec{x}^{\,2} \right\rangle \approx \frac{k_B T}{m\,\omega_o^{\,2}} \left(\frac{r}{b} \right)^{d_f\,(2/d_s - 1)} \sim r^{\,d_f\,(2/d_s - 1)}$$

Diverging with the real space separation between the two points ***r***

Longer separation distance ➡ ***Larger fluctuations***

Strongly overdamped fractons

$$\left\langle \vec{u}_\alpha(t)\cdot\vec{u}_\alpha(0)\right\rangle = \left\langle \vec{u}_\alpha^{\,2}\right\rangle_T \mathrm{e}^{-\frac{\omega_\alpha^{\,2}}{\gamma}t}$$

Where $m\gamma$ is the local friction (Rouse-like model).

Propagation length is $\xi(t) \sim t^{d_s/2d_f}$

$$\left\langle \vec{x}(t)\cdot\vec{x}(0)\right\rangle \sim \begin{cases} 1-\text{const.}\quad t^{1-d_s/2} & \text{short time} \quad s:\xi(t) << r \\ t^{-(d_s/d_l+d_s/2-1)} & \text{long times} \quad :\xi(t) >> r \end{cases}$$

d_l - Topological space dimension

r - Real space separation between the two points

l - Topological space separation between the two points

$$M \sim l^{d_l} \sim r^{d_f}$$

Granek & Klafter, PRL (2005)

Hydrodynamic interaction (Zimm)

Oseen hydrodynamic interaction tensor $\overleftrightarrow{O}(\vec{l},\vec{l}\,') \sim \frac{1}{\eta\, r}$

$$\frac{d}{dt}\vec{u}(\vec{l},t) = m\,\omega_o^2 \sum_{\vec{l}\,'} \overleftrightarrow{O}(\vec{l},\vec{l}\,') \sum_{\vec{l}\,''\in\vec{l}\,'} \left(\vec{u}(\vec{l}\,'',t) - \vec{u}(\vec{l}\,',t)\right) + \varsigma(\vec{l},t)$$

noise

- propagation length

$$\xi(t) \sim t^{\frac{d_s}{d_f\left(2-d_s+d_s/d_f\right)}}$$

- $$\left\langle \vec{x}(t)\cdot\vec{x}(0)\right\rangle \sim \begin{cases} 1 - \text{const.}\quad t^{\frac{2-d_s}{2-d_s+d_s/d_f}} & \text{short time} \quad \text{s} \\ t^{-\frac{2\,d_s/d_l-2+d_s}{2-d_s+d_s/d_f}} & \text{long} \quad \text{times} \end{cases}$$

Granek, Phys. Rev. E, Rapid comm. (2011)

MSD of dihedral angles

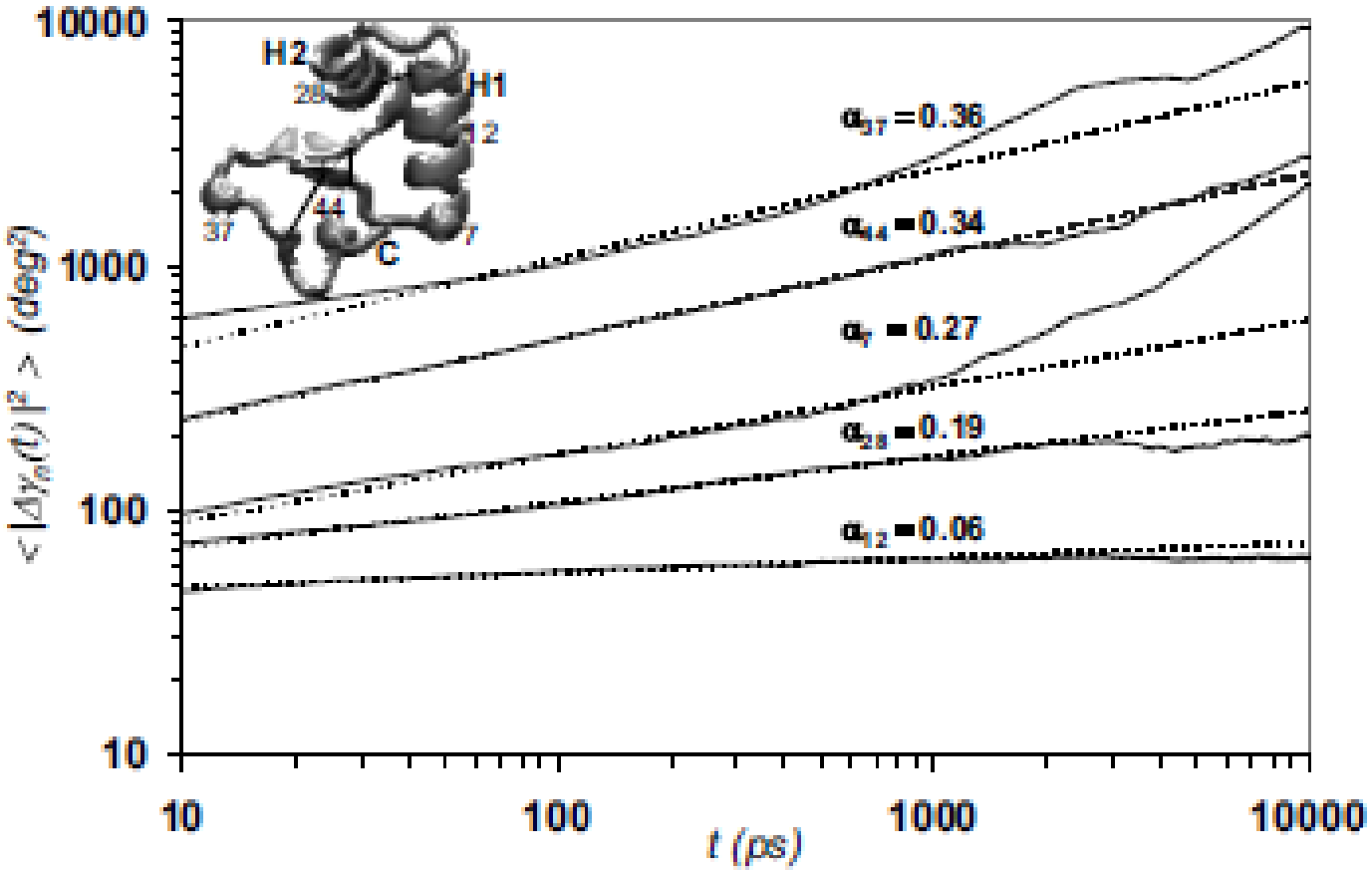
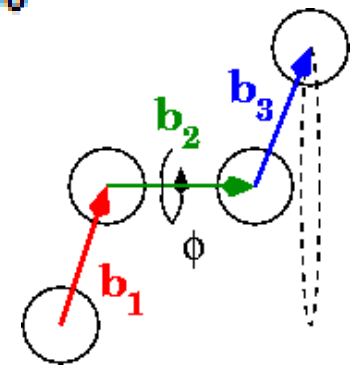


Fig. 1. Typical results for the mean-square displacement (MSD) of dihedral angles γ_n in crambin computed from MD on a double-log scale. For each dihedral angle γ_n, the MSD (solid lines) is compared with a power law (dotted lines) with an exponent α_n. *Inset* shows the main structural elements of crambin. The spheres represent the location of residues 7, 12, 28, 37, and 44, and the thick lines indicate the positions of the disulfide bridges.



Patrick Senet[a,b], Gia G. Maisuradze[a], Colette Foulie[b], Patrice Delarue[b], and Harold A. Scheraga[a,1]

II) Vibrational mean square displacement (MSD)

A specific alpha-carbon

$$\left\langle \Delta \vec{u}_\ell{}^2(t) \right\rangle_T \equiv \left\langle \left(\vec{u}_\ell(t) - \vec{u}_\ell(0) \right)^2 \right\rangle_T$$

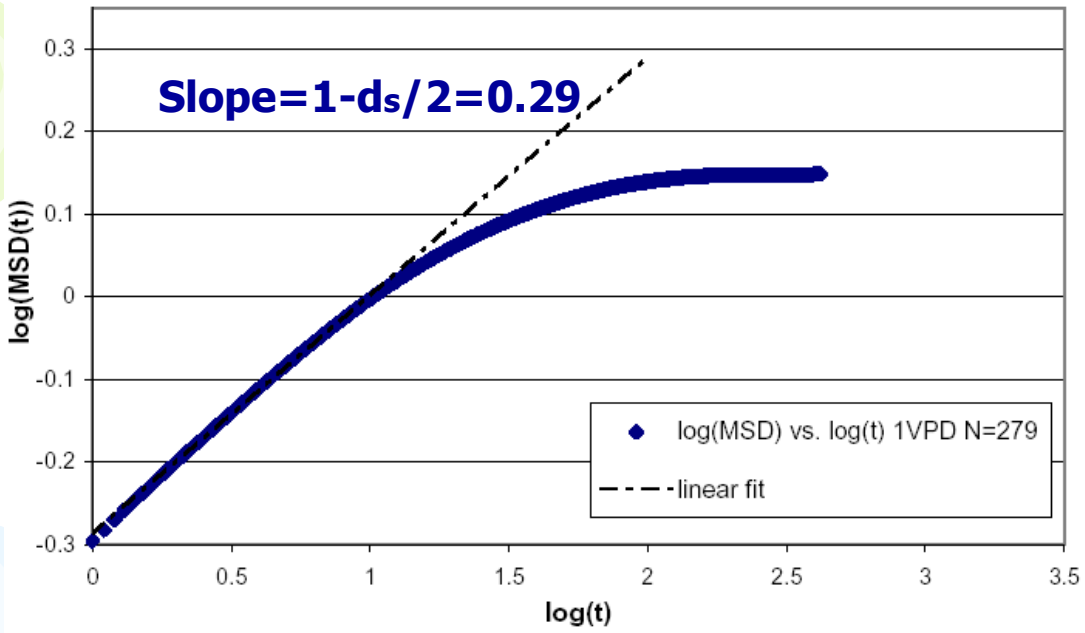
Average over all alpha-carbons

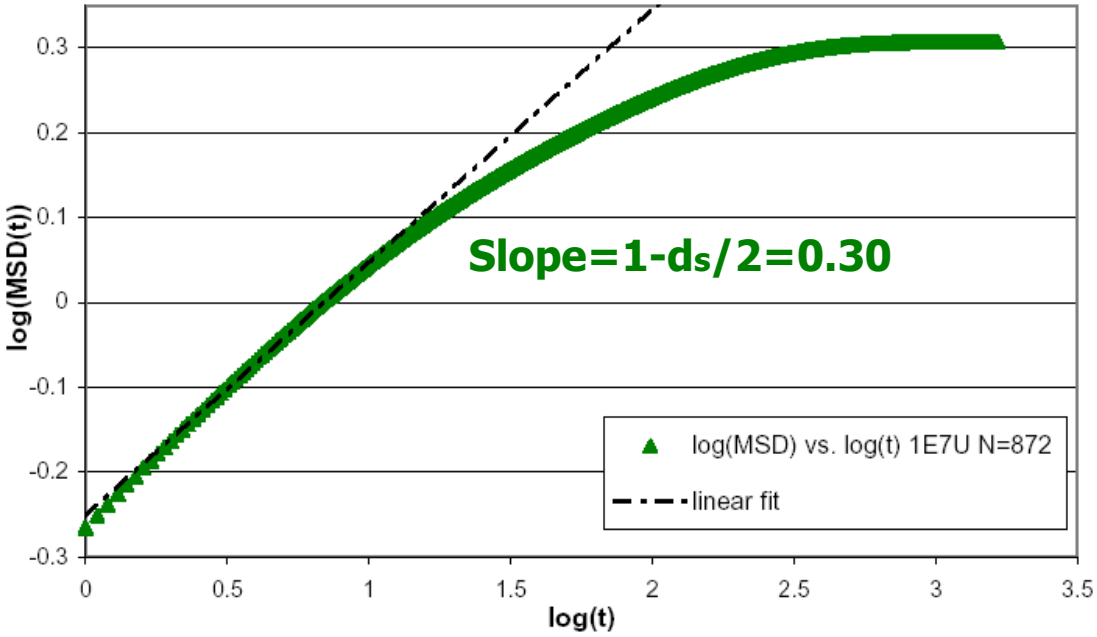
$$\left\langle \left\langle \Delta \vec{u}_\ell{}^2(t) \right\rangle_T \right\rangle_\ell = \frac{1}{N} \sum_\ell \left\langle \Delta \vec{u}_\ell{}^2(t) \right\rangle_T$$

In the fractal model:

$$\left\langle \left\langle \Delta \vec{u}_\ell{}^2(t) \right\rangle_T \right\rangle_\ell \sim t^{1-d_s/2}$$

Vibrational MSD





Dynamics

III) Random Walk on the GNM network

On Fractals

Probability of return to the origin $P_0(t)$ $\quad P_0(t) \sim t^{-d_s/2}$

Mean square displacement (MSD) $\left\langle r^2(t) \right\rangle$ $\quad \left\langle r^2(t) \right\rangle \sim t^{2/d_w}$ $\quad d_w = 2d_f / d_s$

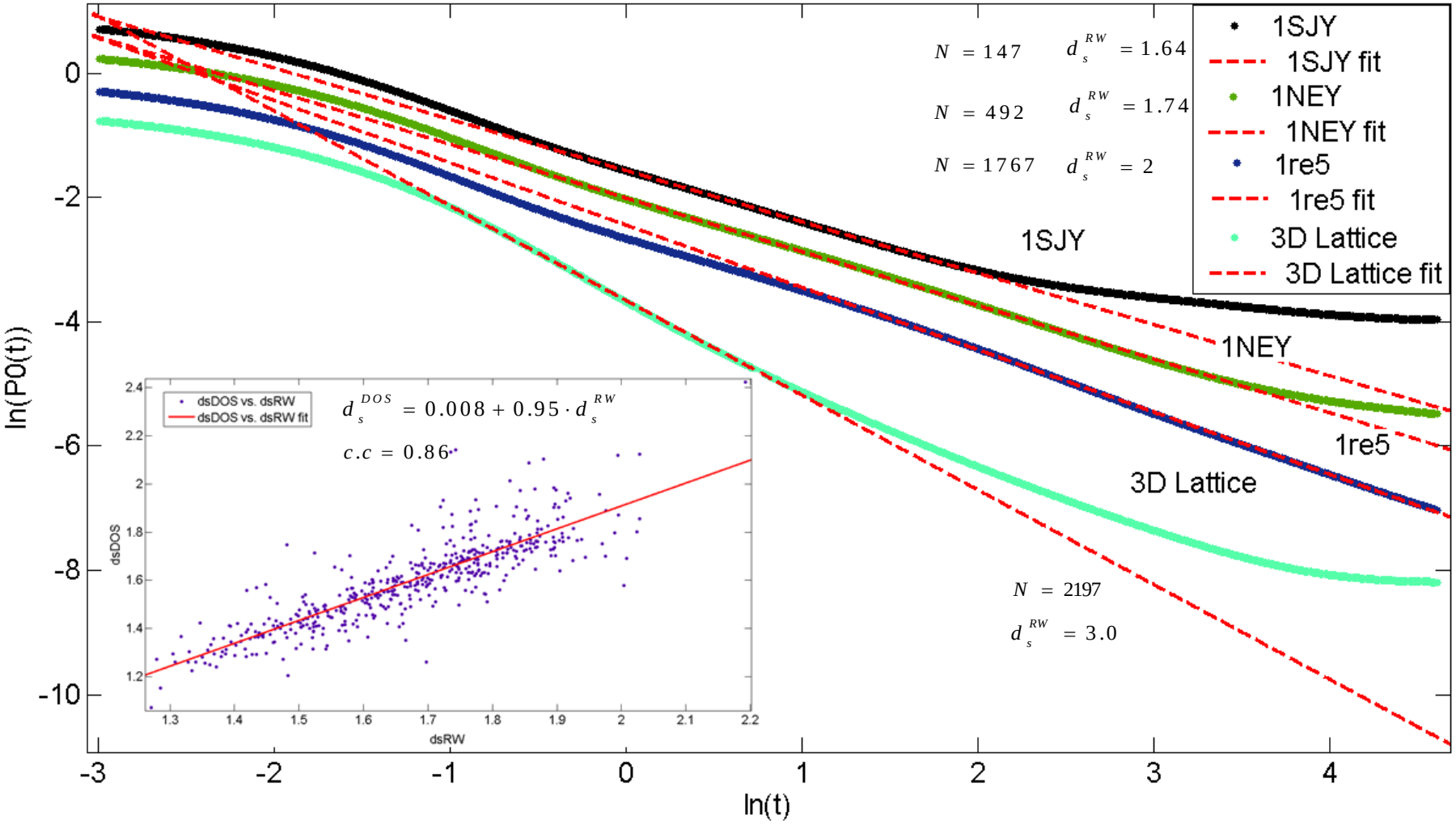
Mean first passage time (MFPT) $\quad \mathrm{MFPT} \sim r^{d_w - d_f}$

Note equivalence to vibrations:

$$\left\langle r^2(t) \right\rangle \longleftrightarrow \xi^2(t) \sim t^{d_s/d_f}$$

$$P_0(t) \longleftrightarrow \frac{d}{dt}\left\langle \Delta \vec{u}^{\,2}(t) \right\rangle_T \qquad \mathrm{MFPT} \longleftrightarrow \left\langle \vec{x}^{\,2} \right\rangle \sim r^{d_f(2/d_s - 1)}$$

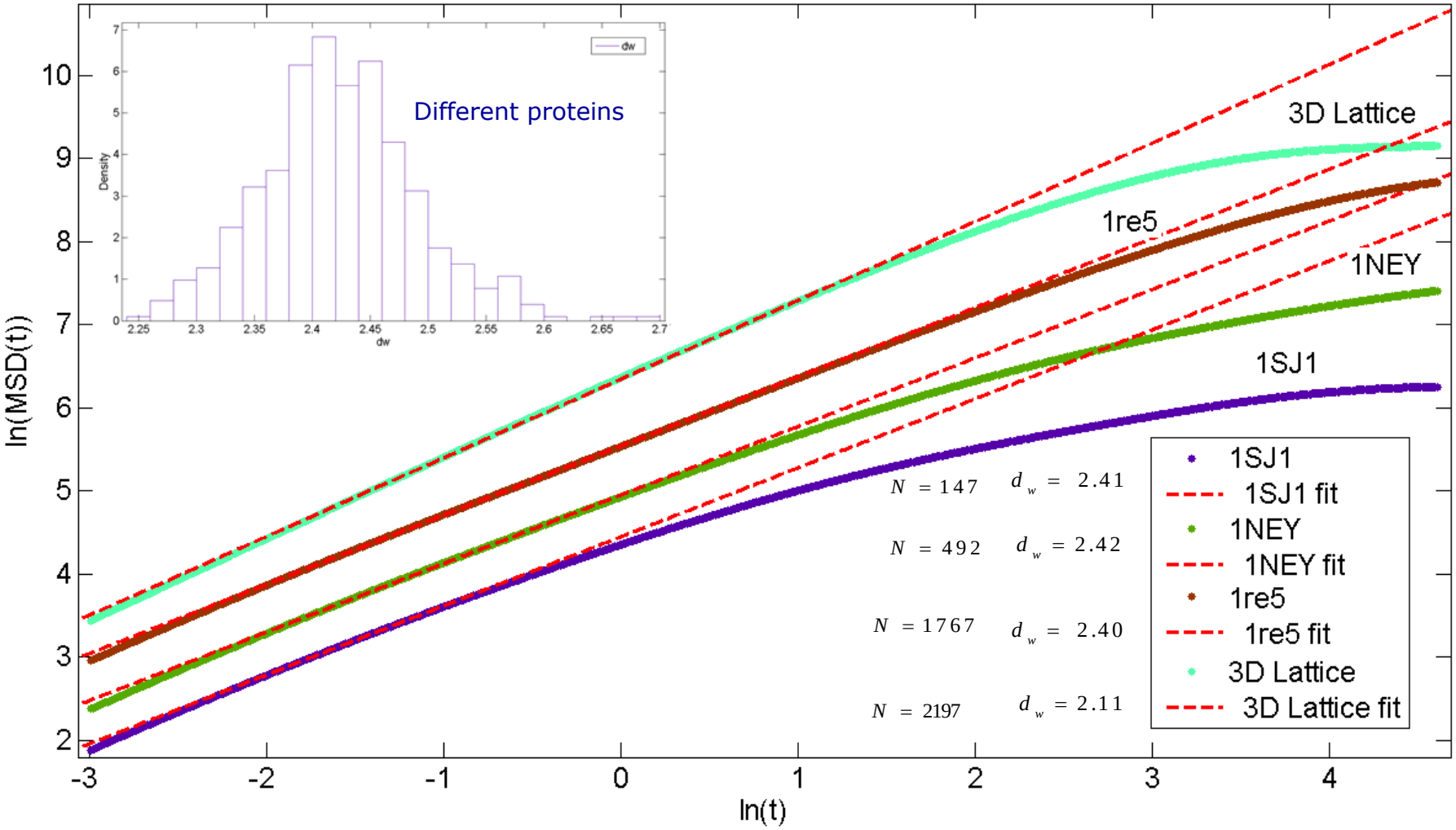
Probability of return to the origin $P_0(t) \sim t^{-d_s/2}$



Reuveni et al., PNAS (2010)

RW Mean square displacement (MSD) $\left\langle r^2(t) \right\rangle \sim t^{2/d_w}$

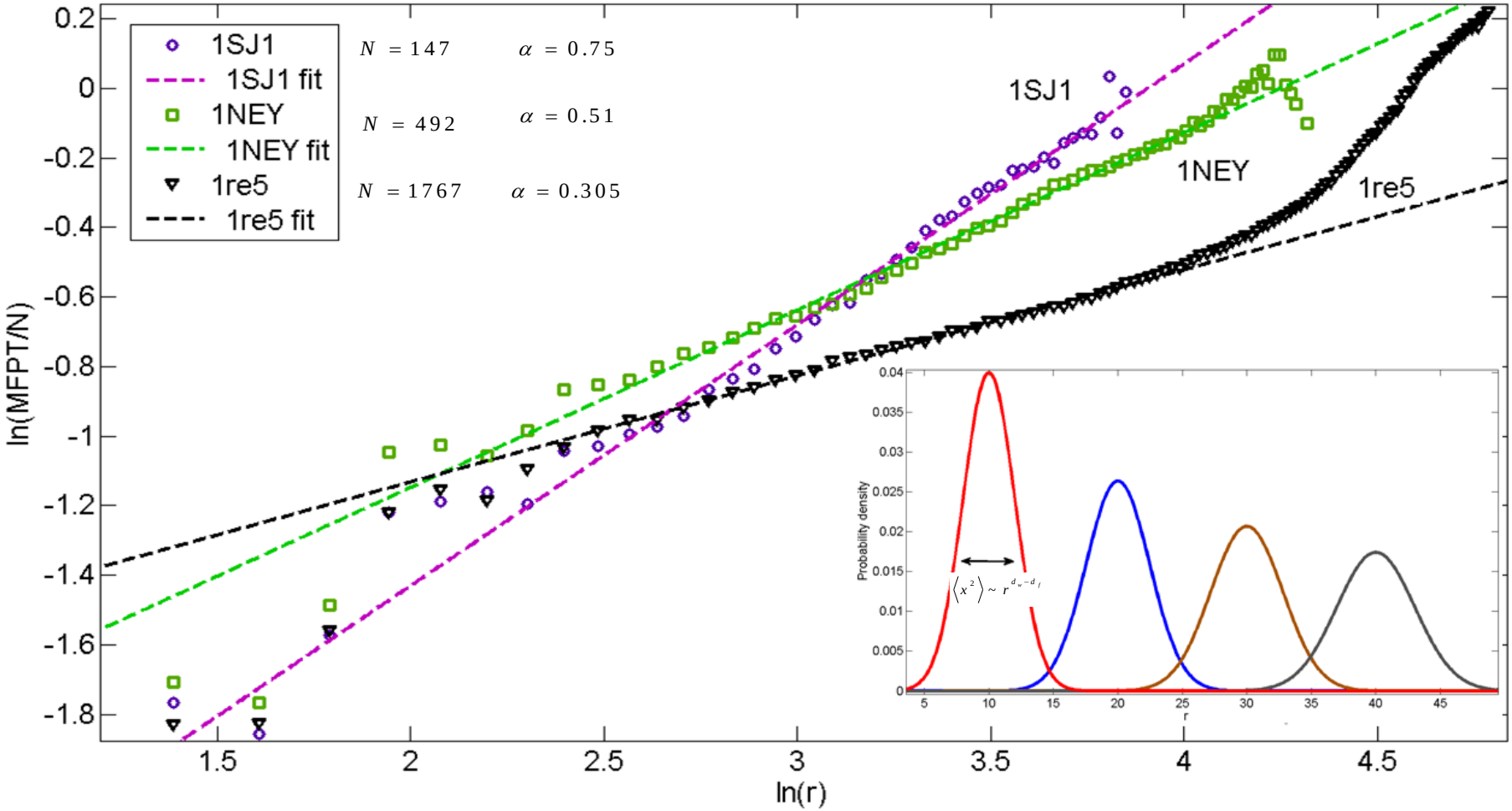
Propagation length (vibrations) $\xi^2(t) \sim t^{2/d_w}$



Reuveni et al., PNAS (2010)

Mean first passage time $\mathrm{MFPT} \sim r^{\alpha} \qquad \alpha = d_w - d_f$

Variance of distance fluctuations $\left\langle x^2 \right\rangle \sim \frac{MFPT}{N} \sim r^{\alpha}$

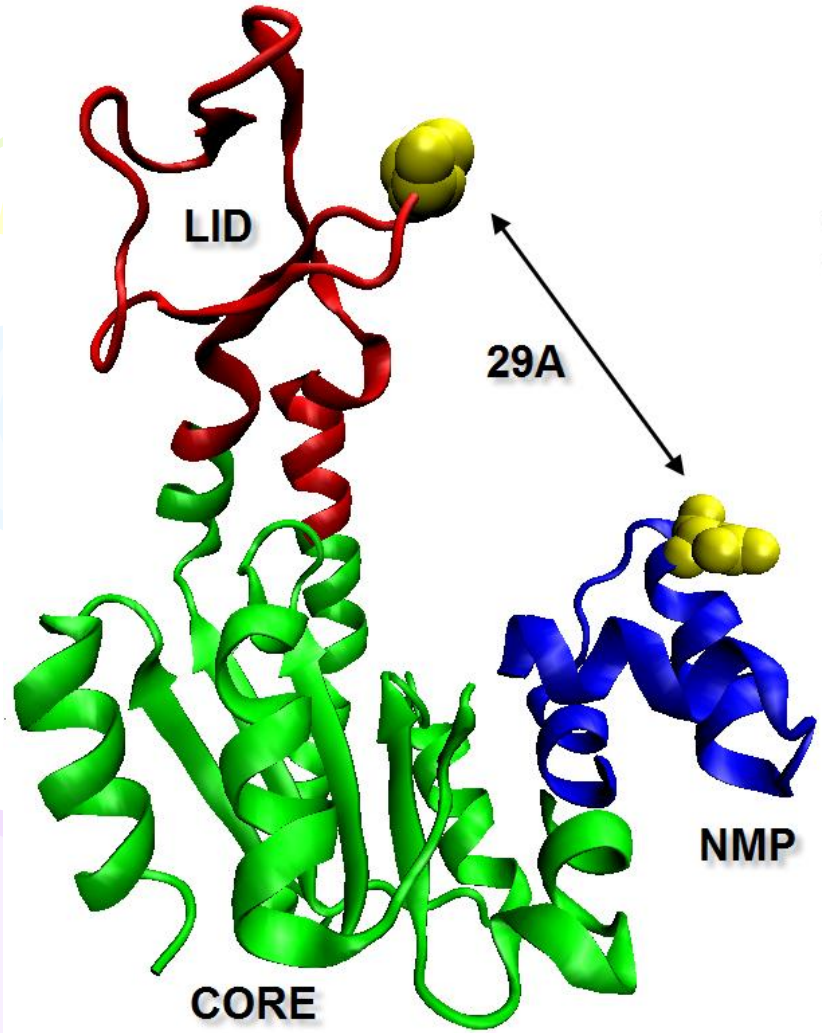


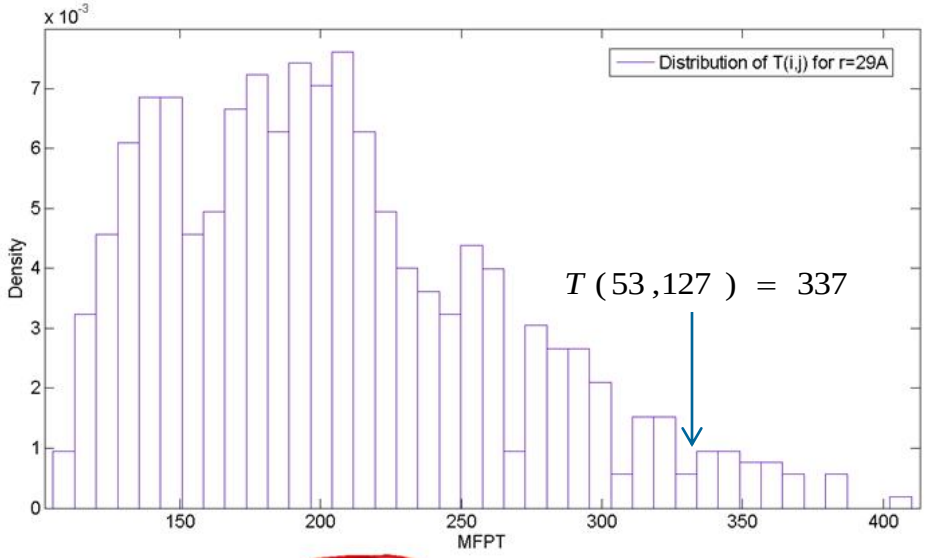
Reuveni et al., PNAS (2010)

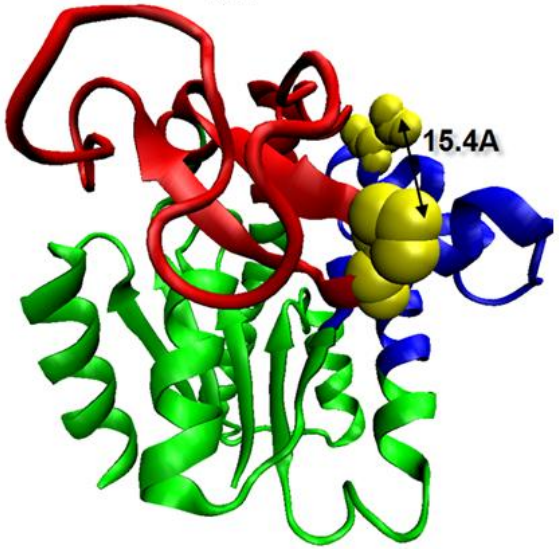
What about specificity?

Adenylate kinase, N=214

$MFPT\ (r = 29\ \text{A}) = 205\ \text{(a.u.)}$, but







Reuveni et al., PNAS (2010)

Structure with bound ATP and AMP (substrate)

Dynamic Structure Factor

Neutron Spin-Echo Studies of Hemoglobin and Myoglobin: Multiscale Internal Dynamics

J. Mol. Biol. (2010) 397, 423–435

Jyotsana Lal[1], Peter Fouquet[2], Marco Maccarini[2] and Lee Makowski[1]*

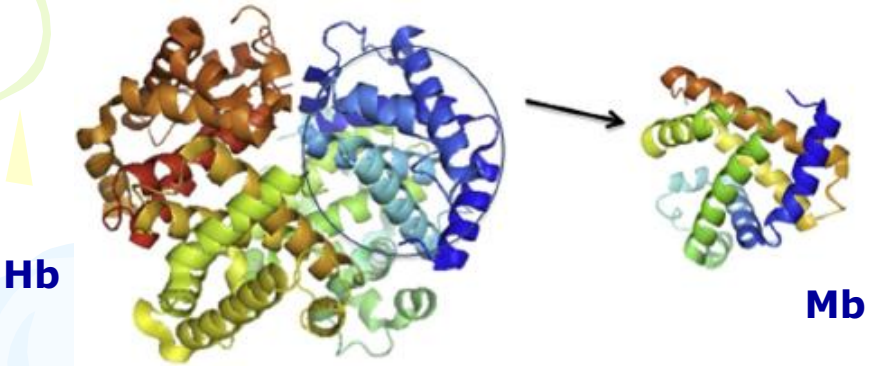
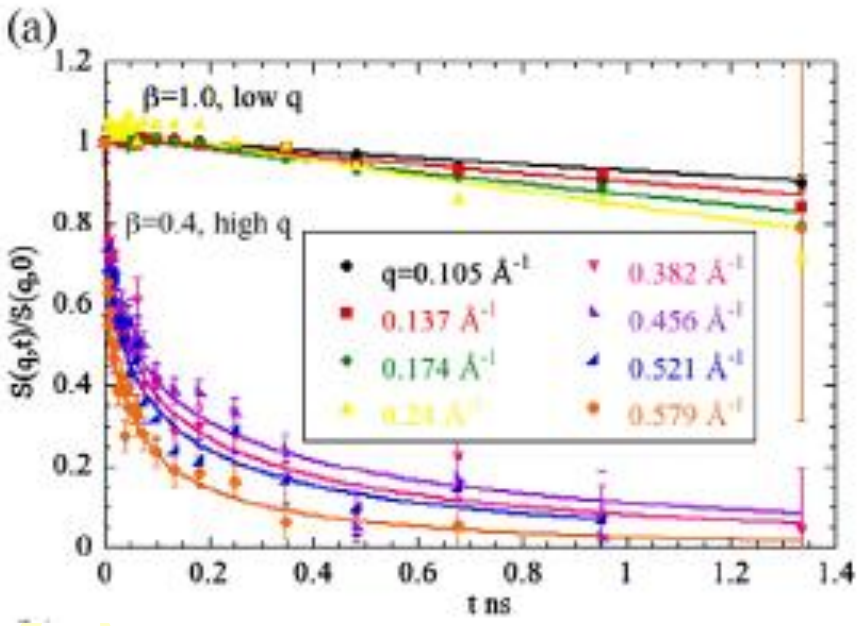
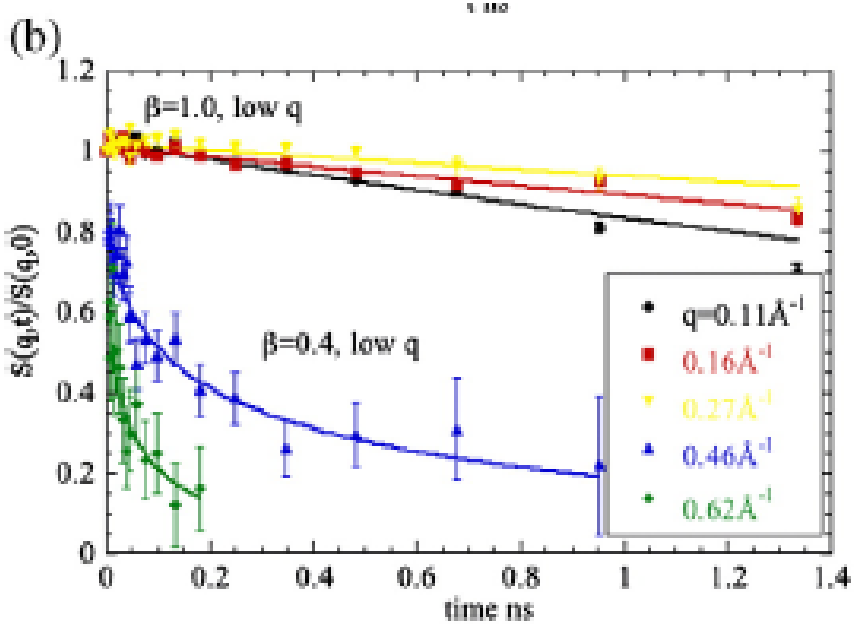
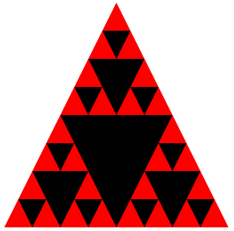


Fig. 5. Normalized intermediate scattering functions $S(q,t)/S(q,0)$ of (a) Mb and (b) Hb in D_2O measured with the NSE spectrometer IN11c (Institut Laue-Langevin) at several wavevectors q for a concentration of 150 mg/ml and a temperature of 288 K. Continuous lines indicate fits of the data to a single-exponential decay function (low q regime) and a stretched-exponential decay function (high q regime).





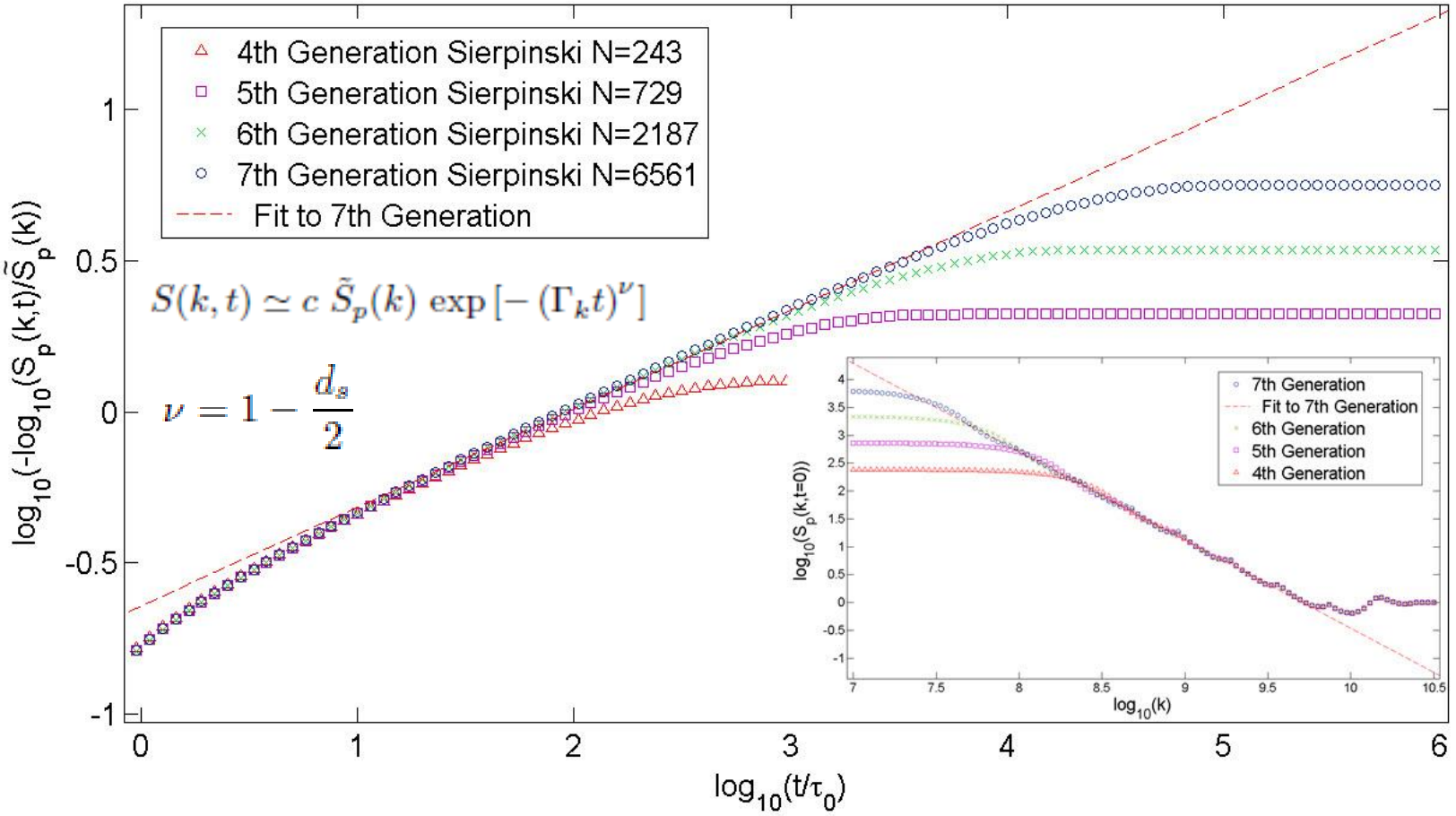
Sierpinski gasket – Dynamic Structure Factor



- **No translational and rotational diffusion**

$\tilde{S}(k)$ - "Forzen network" static structure factor

$S(k,t)$ - Dynamic structure factor

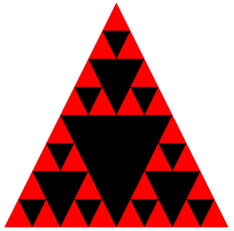


- **Stretched exponential!**

Reuveni, Klafter, Granek, to be published

$\tau_0 = 10.53$ ps

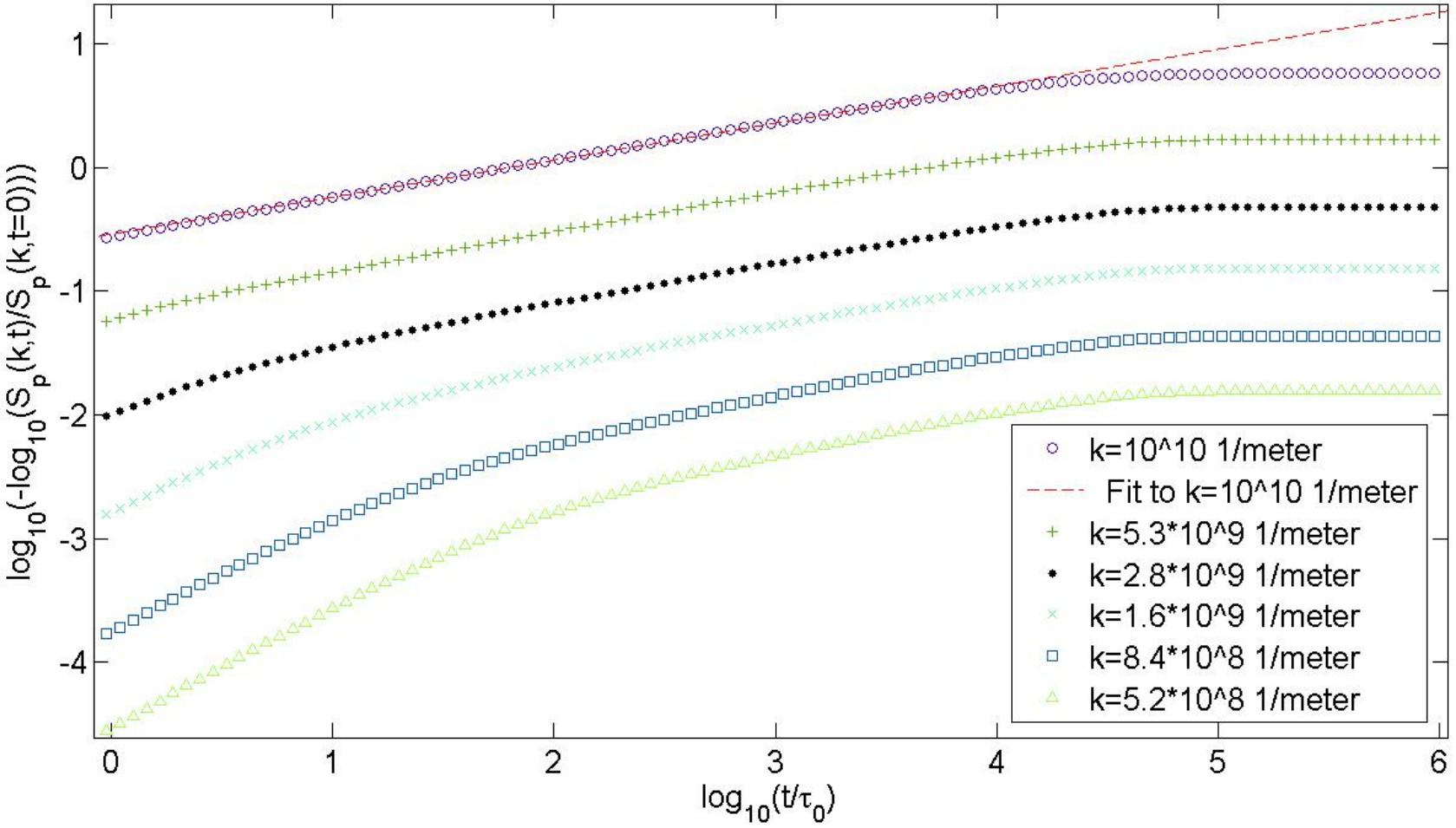
Sierpinski gasket – Dynamic Structure Factor



- **No translational and rotational diffusion**

$S(k,0) \equiv S(k)$ - Static structure factor

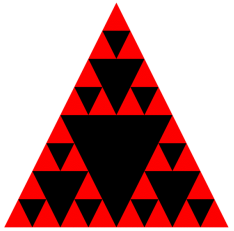
$S(k,t)$ - Dynamic structure factor



Reuveni, Klafter, Granek, to be published

$\tau_0 = 10.53$ ps

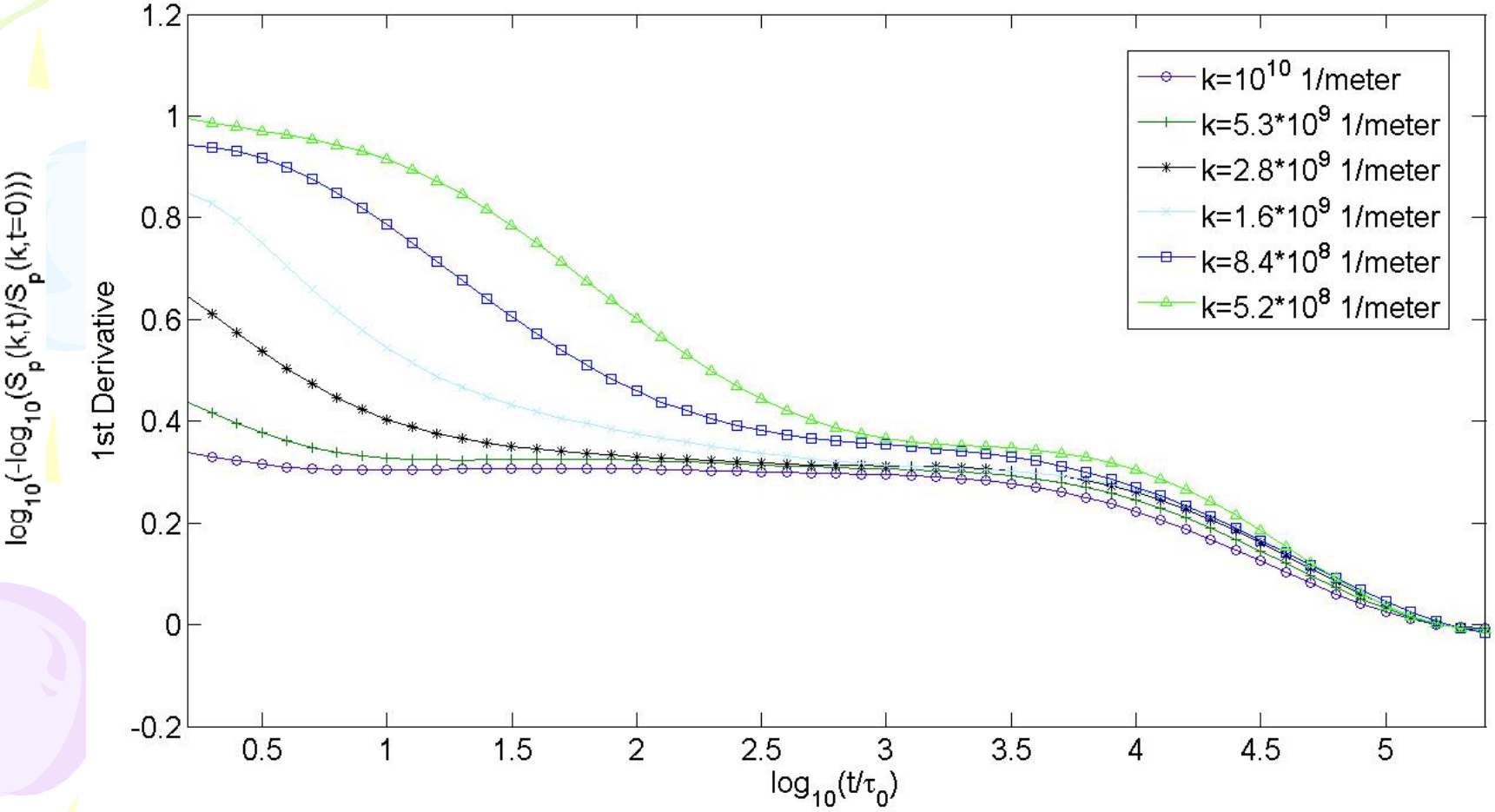
Sierpinski gasket – Dynamic Structure Factor



- **No translational and rotational diffusion**

$S(k,0) \equiv S(k)$ - Static structure factor

$S(k,t)$ - Dynamic structure factor



Reuveni, Klafter, Granek, to be published

$\tau_0 = 10.53$ ps

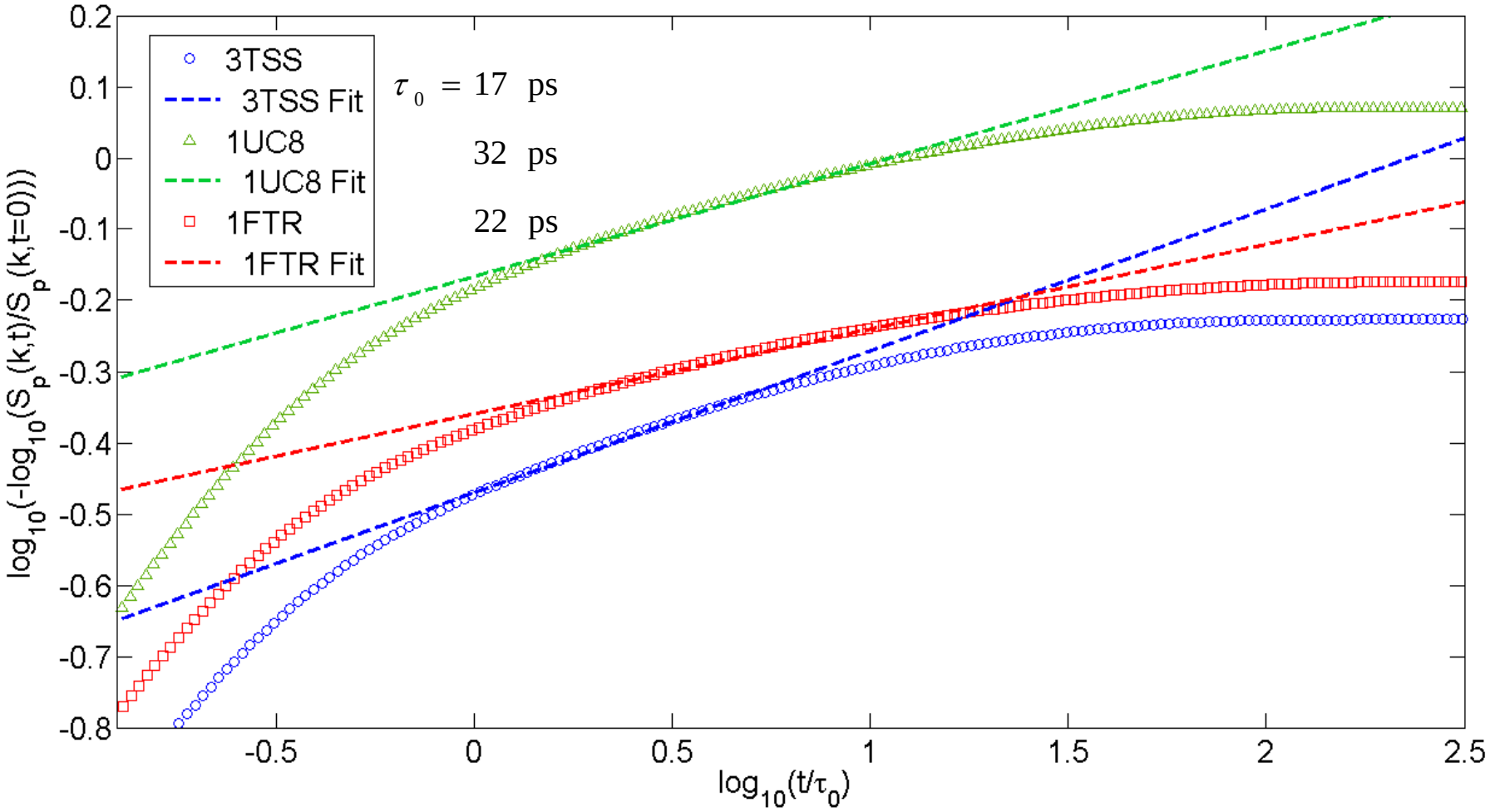
Proteins – Dynamic Structure Factor

- **No translational and rotational diffusion**

$S(k,0) \equiv S(k)$ - Static structure factor

$S(k,t)$ - Dynamic structure factor

$kb = 5$, $b = 5$ Å

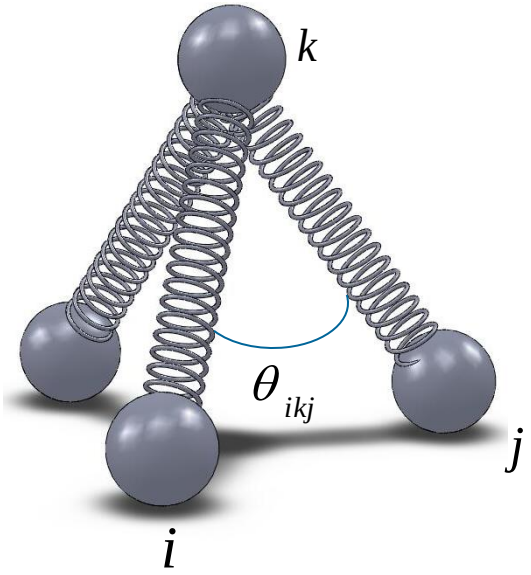


Reuveni, Klafter, Granek, to be published

Tensorial Elasticity Model

$$H_{TEM} = \frac{1}{2} m \omega_o^2 \sum_{i,j}^{N} \Gamma_{ij} \left[\left(\vec{u}_i - \vec{u}_j \right) \cdot \hat{r}_{ij} \right]^2 + \frac{1}{2} B \sum_{i,j,k}^{N} \Gamma_{ik} \Gamma_{jk} \left(\delta\theta_{ikj} \right)^2$$

Anisotropic network model (ANM) Bond-bending



Advantage:

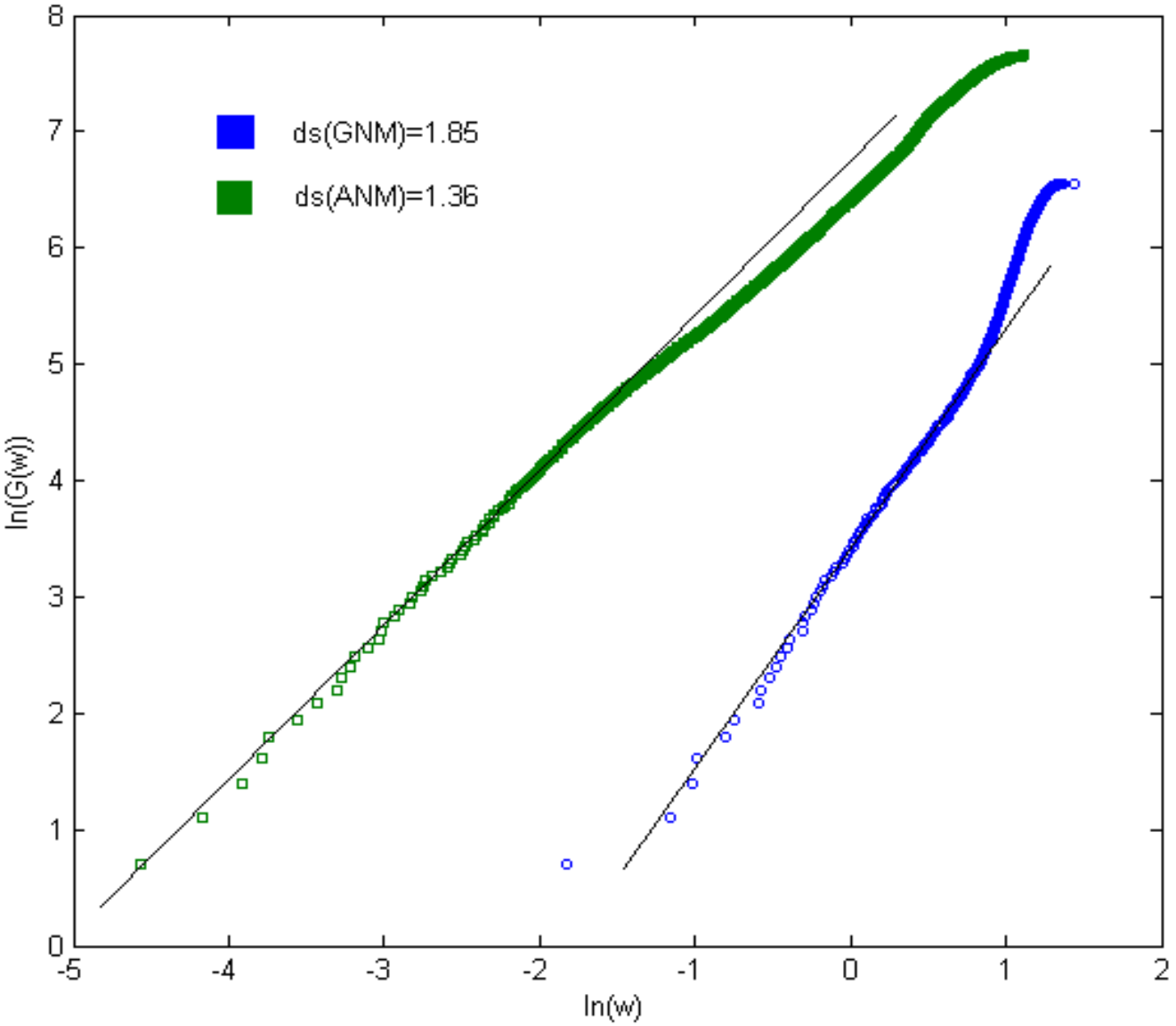
- **Allows to compute anisotropic displacement fluctuations.**

Disadvantage:

- **Complicated.**
- **Involves an additional free parameters (B).**

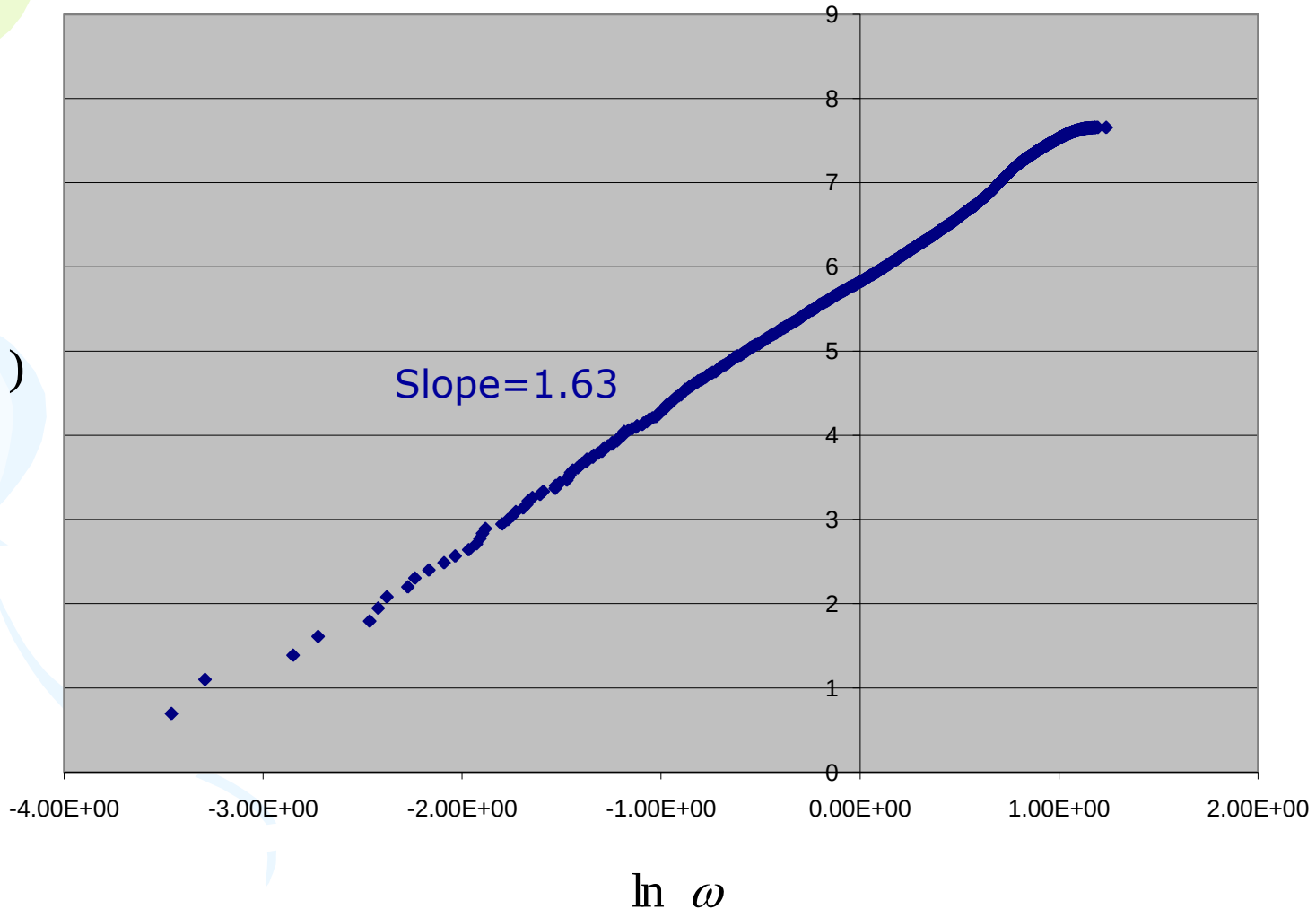
Ben-Halevi et al., to be published

Anisotropic Network Model (ANM) for Methyltransferase (2NQ5).



Ben-Halevi et al., to be published

Tensorial Model for Methyltransferase (2NQ5).



$\ln G(\omega)$

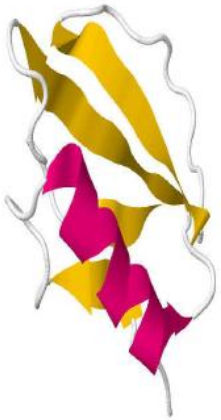
$\ln \omega$

$m\omega_o^2 / B = 1\ \mathrm{A}^{-2}$

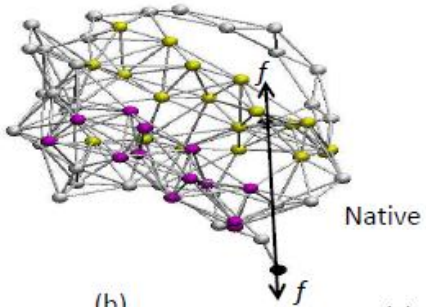
Ben-Halevi et al., to be published

Unfolding by Pulling Force

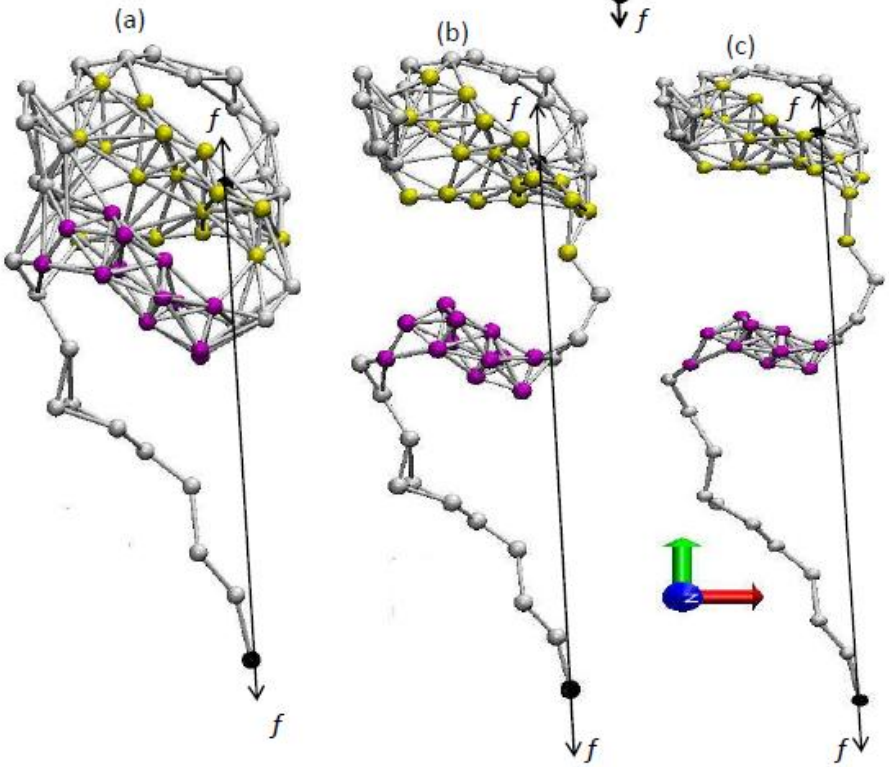
Cymotrypsin Inhibitor (CI2)

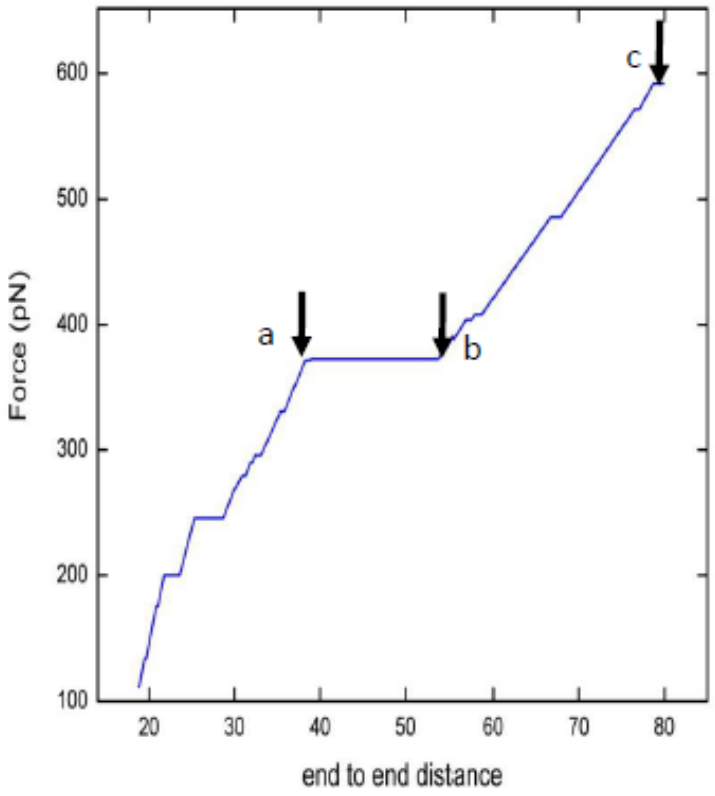


GNM of CI2



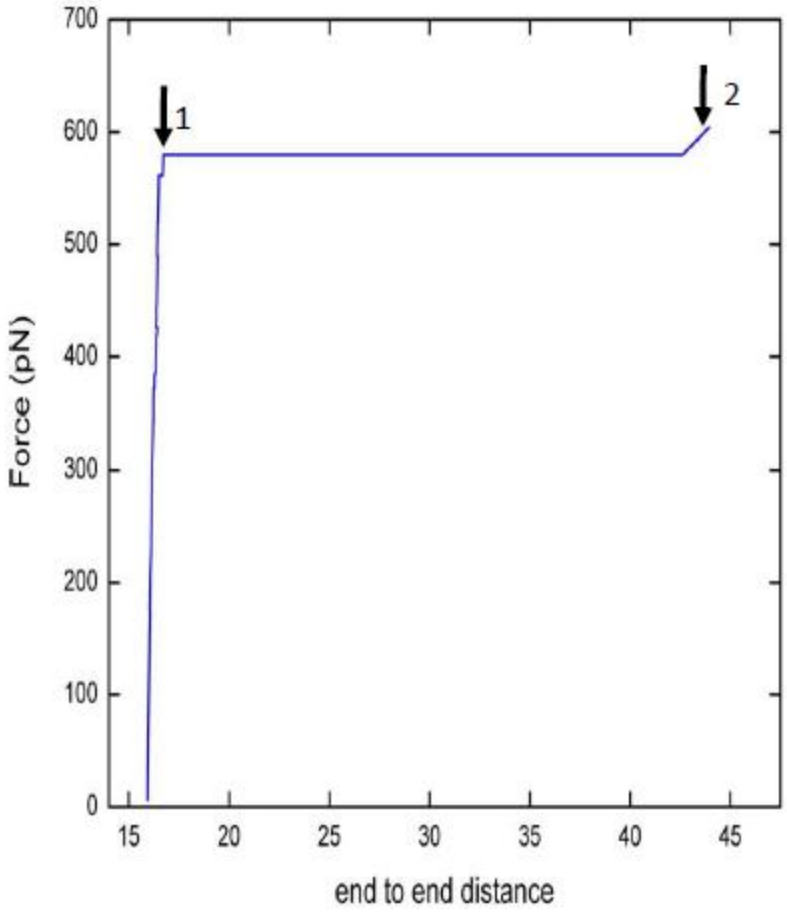
force at the end pair

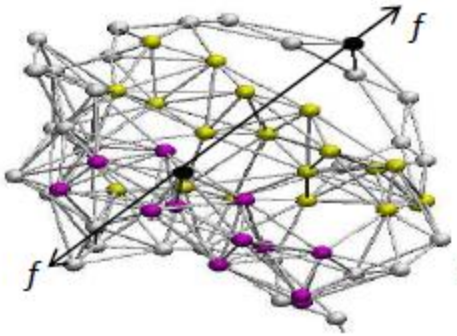




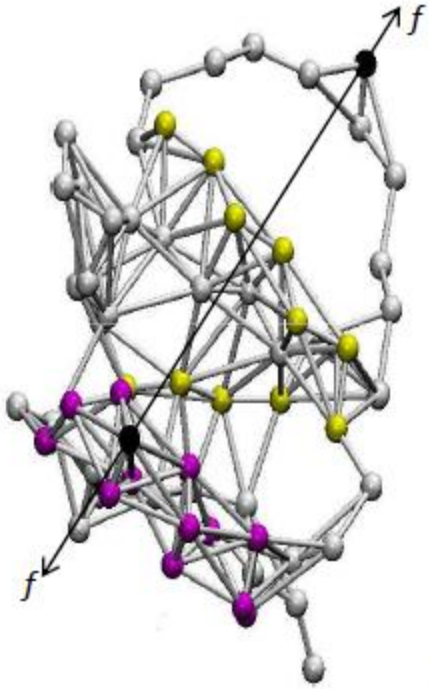
Srivastava & Granek, to be published

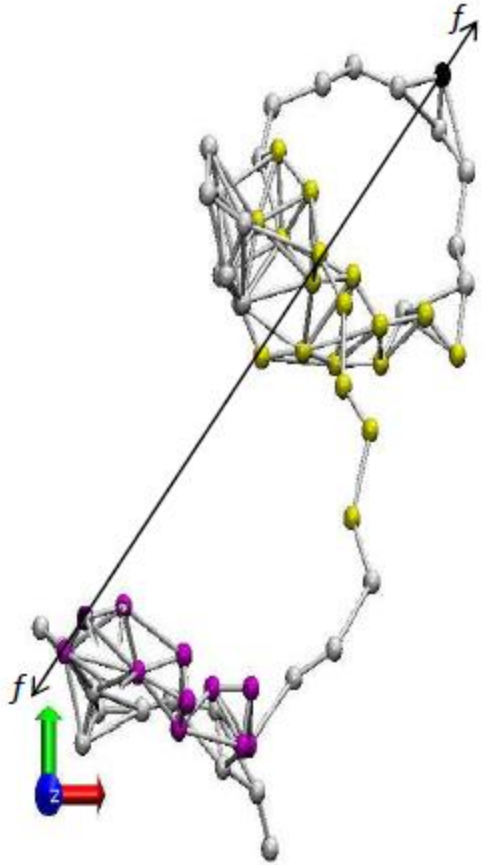
Residue pair15 and 40





Native





Conclusions:

- Novel approach for **vibrations** in **folded proteins** based on their **fractal** nature → Provides a description on a universal level.
- Folded proteins are **marginally stable**: they exist in a thermodynamic state close to unfolding, which allows for large scale motion without unfolding.
- The above criterion leads to a universal "**equation of state**", verified for about 5,000 proteins.
- The fractal-like properties of proteins lead to **Anomalous dynamics/ strange kinetics:** autocorrelation of separation; vibrational MSD; random walk MSD, return probability & mean first passage time, dynamic structure factor.

Questions to address:

- Relation to enzymatic/biological activity.
- Role of water.
- More realistic potentials, e.g.: Tensorial elasticity, distinguishing between backbone bonds, H-bonds, electrostatic bonds, etc.
- Response and unfolding under force.

Coworkers

Tel-Aviv University

Shlomi Reuveni

Joseph Klafter

Ben-Gurion University

Marina de Leeuw

Roee Ben-Halevi

Amit Srivastava

RG

Funding: ISF, DIP

Thank You!