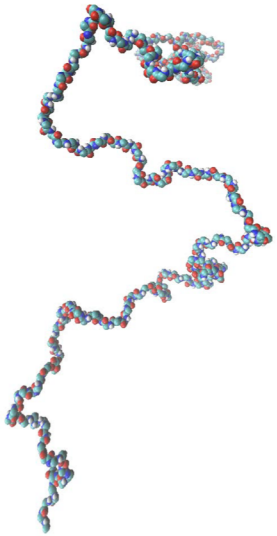
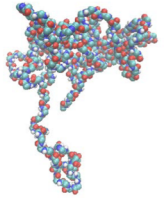
Advanced topics 4: Flexible systems and intrinsically disordered proteins





Tobin R. Sosnick

Dept. of Biochemistry & Mol. Biology, Chair
Institute for Biophysical Dynamics
University of Chicago

Graduate Program in Biophysical Sciences

BioSAXS 5
Nov. 6, 2019 Argonne Nat. Lab

Advanced topics 4: Flexible systems and intrinsically disordered proteins

References:

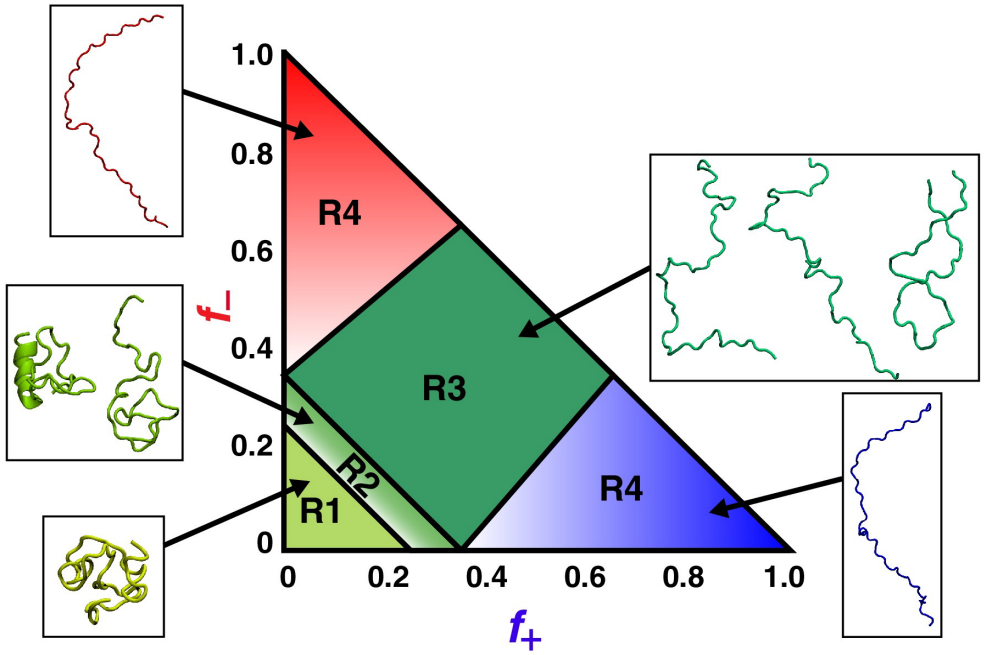
Small-angle Scattering

1. Riback, J.A., et al., *Innovative scattering analysis shows that hydrophobic disordered proteins are expanded in water.* Science, 2017. **358**(6360): p. 238-241.
2. Rambo, R.P. and J.A. Tainer, *Characterizing flexible and intrinsically unstructured biological macromolecules by SAS using the Porod-Debye law.* Biopolymers, 2011. **95**(8): p. 559-71.
3. Bernado, P. and D.I. Svergun, *Analysis of intrinsically disordered proteins by small-angle X-ray scattering.* Methods Mol Biol, 2012. **896**: p. 107-22.
4. Receveur-Brechot, V. and D. Durand, *How random are intrinsically disordered proteins? A small angle scattering perspective.* Curr Protein Pept Sci, 2012. **13**(1): p. 55-75.

Disordered proteins

1. Das, R.K., K.M. Ruff, and R.V. Pappu, *Relating sequence encoded information to form and function of intrinsically disordered proteins.* Curr Opin Struct Biol, 2015. **32**: p. 102-12.
2. van der Lee, R., et al., *Classification of intrinsically disordered regions and proteins.* Chem Rev, 2014. **114**(13): p. 6589-631.

Unfolded and disordered proteins



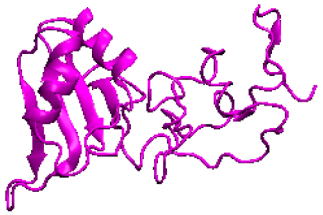
Das, Ruff, & Pappu
(2015) Curr Op Str Biol

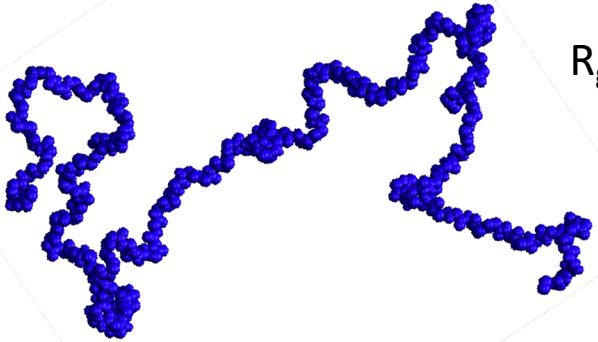
Relevance:

- Does protein folding typically initiate with a rapid hydrophobic collapse?
- ~1/3 of the proteome is intrinsically disordered: **IDP**(roteins), **IDR**(egions)
- Folding upon binding
- phase separation & disease
- Is water a good or poor solvent for proteins?

Scattering: compact versus unfolded polypeptides

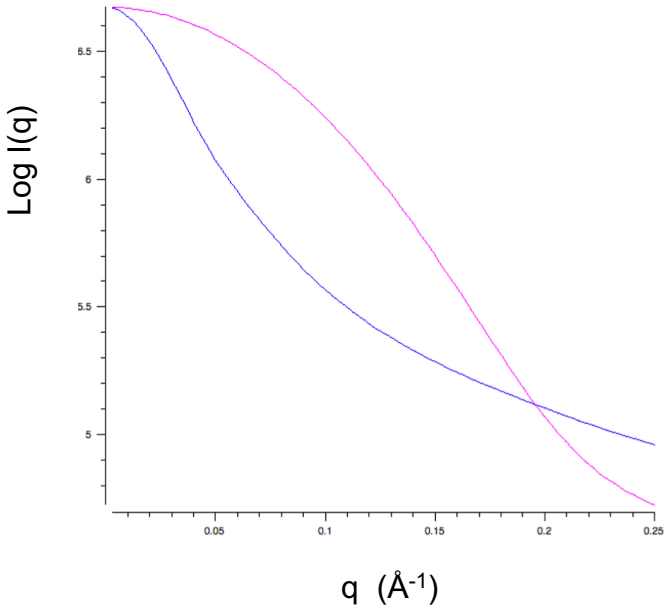
R_g=19 Å



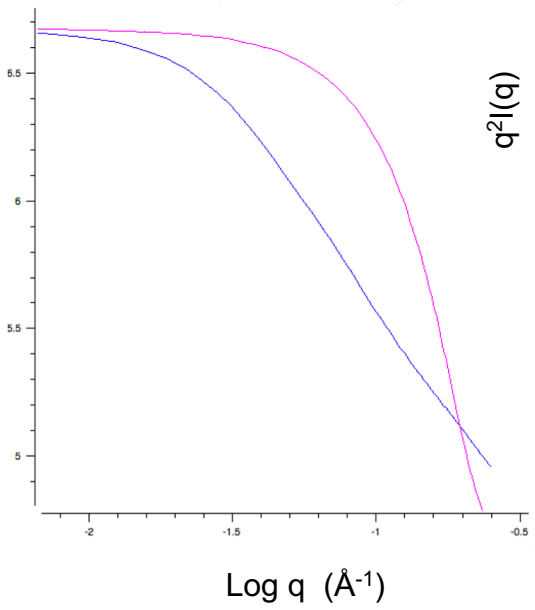


$R_g^{ensemble}$=49 Å

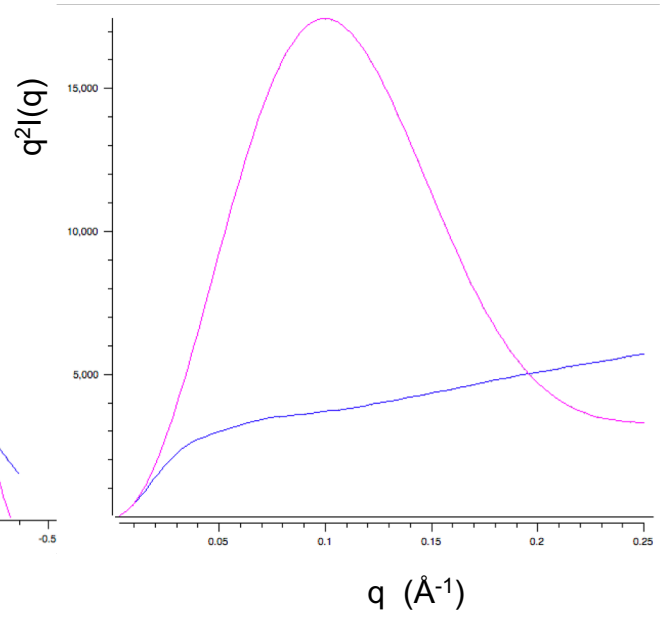
log-linear



log-log plot

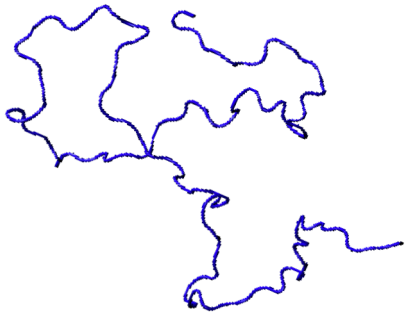


Kratky plot



Rescaling to remove size but keep shape information: Dimensionless Kratky Plot

Rg=100 Å



Rg=40 Å

Rg=10 Å

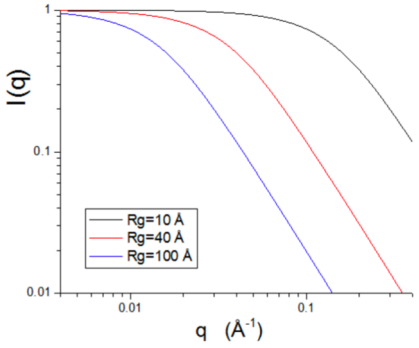
Debye formula for random walk:

$$I(q) = \frac{2I_o\left(e^{-(qRg)^2} - 1 + (qRg)^2\right)}{(qRg)^4}$$

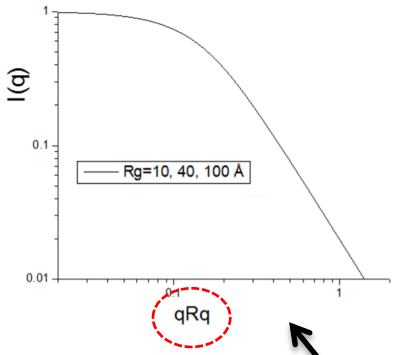
$$x = qRg$$

$$I(q) = \frac{2I_o\left(e^{-x^2} - 1 + x^2\right)}{x^4}$$

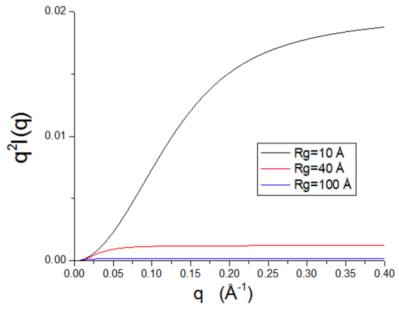
log-log Plot



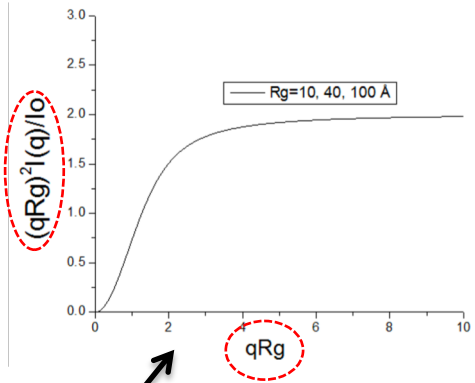
Dimensionless log-log Plot



Kratky plot



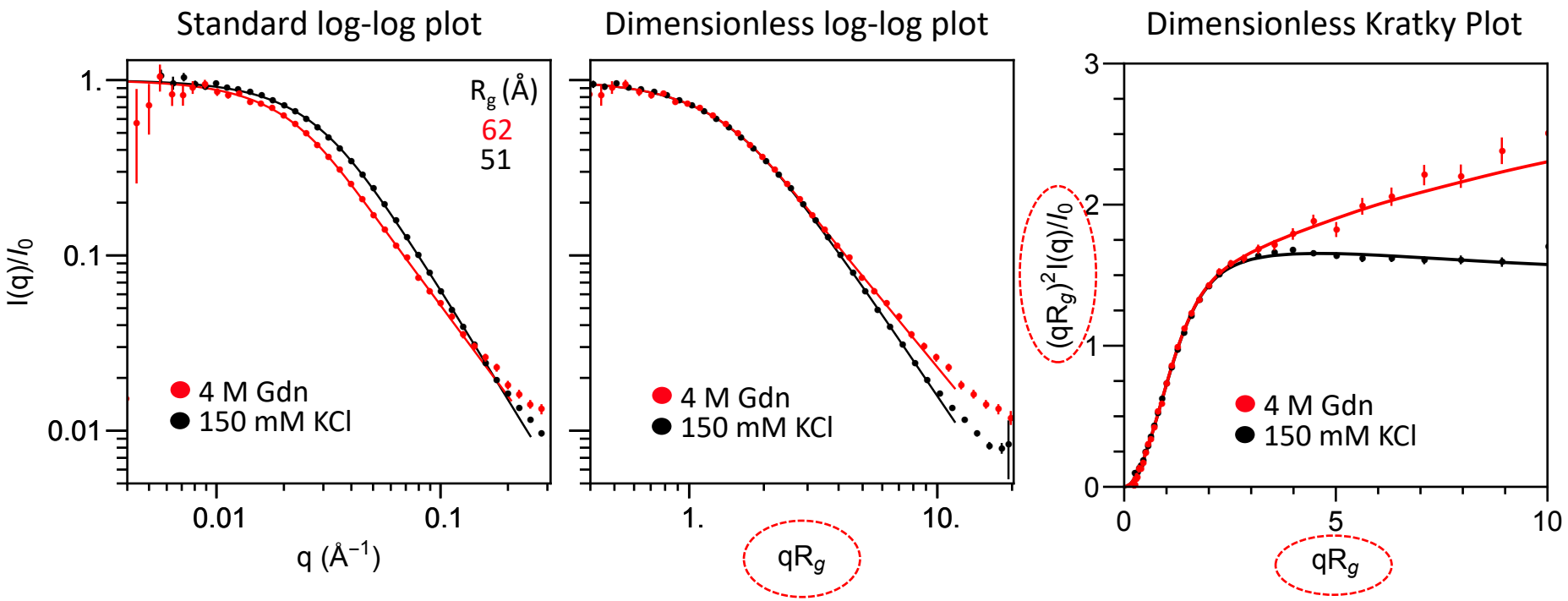
Dimensionless Kratky plot



Report only on shape

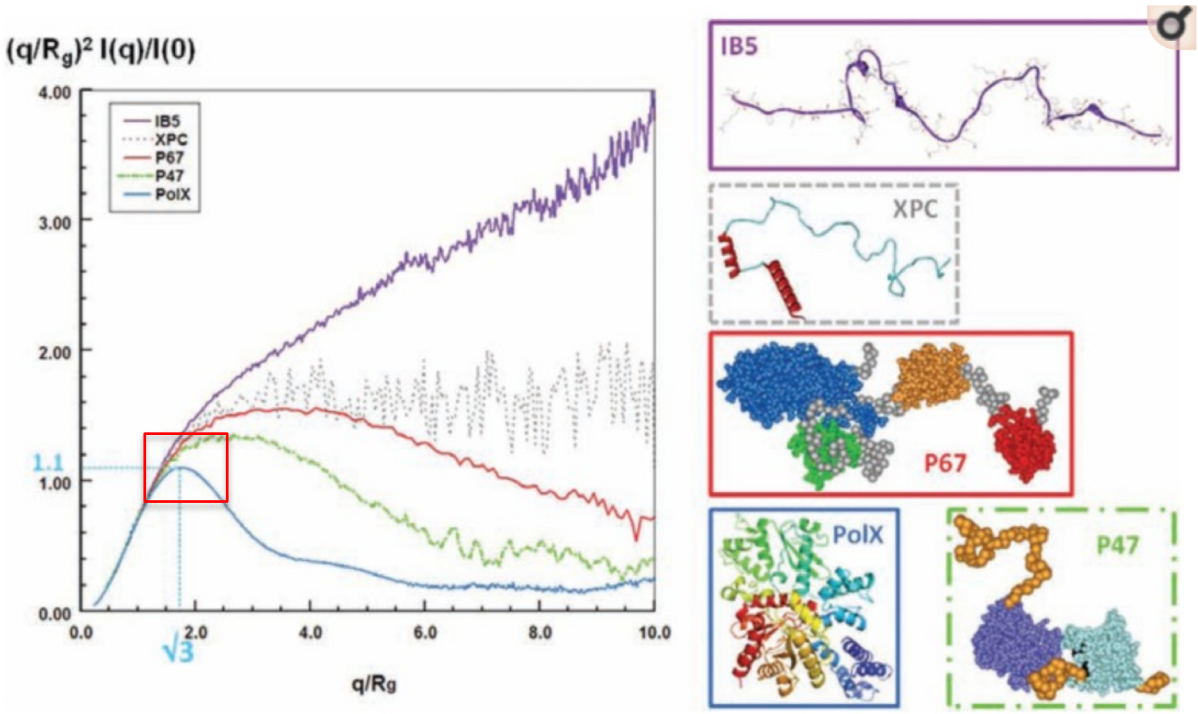
More on scaling and shape

IDP in water and high denaturant concentration



Shapes are different – but how?

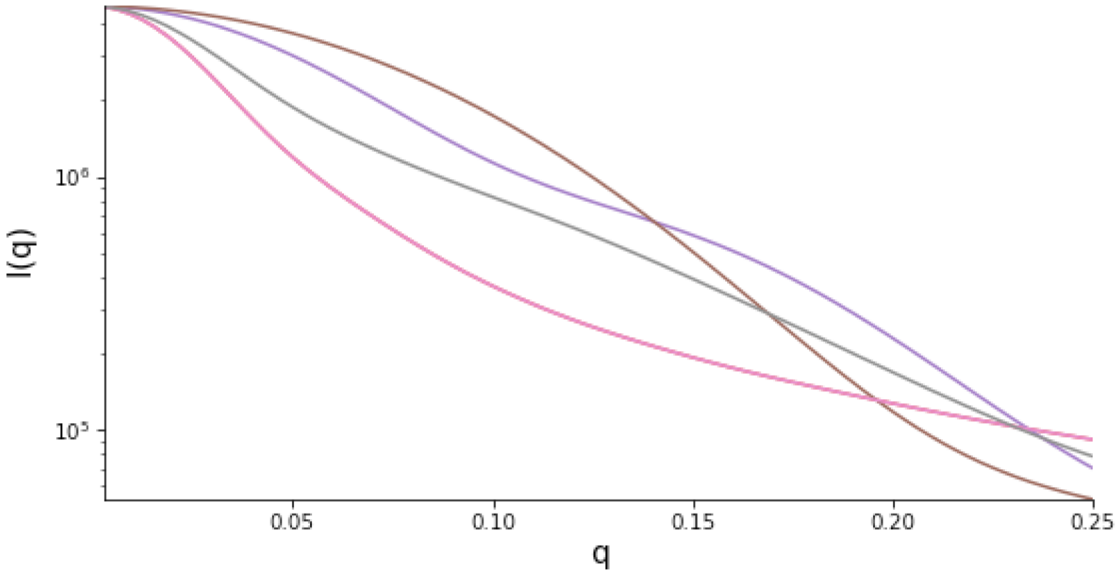
Useful examples from globular to unfolded

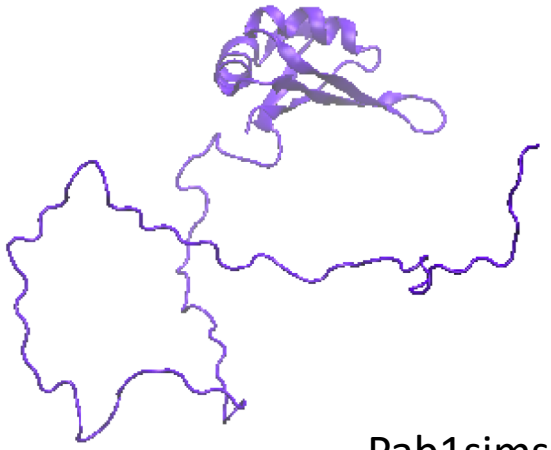


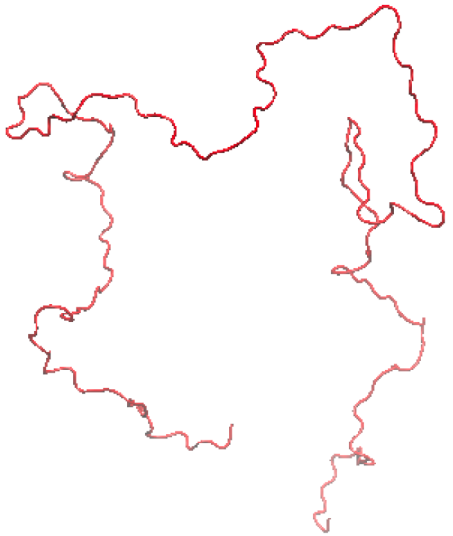
Normalized Kratky plots. The scattering pattern of globular proteins in a normalized Kratky plot exhibits a bell-shaped profile with a clear maximum value of 1.104 for $qR_g=\sqrt{3}$, regardless of the size of the protein, and are all nearly superimposable in the q range $0<qR_g<3$. Conversely, for a random chain, the curve rises with increasing angle, to nearly reach a plateau between 1.5 and 2 and may further increase at q>0.2-0.3 $Å^{-1}$, depending on the persistence length and the internal structure of the protein. Bell-shaped profile of a globular protein (PolX, blue line); curve of a protein consisting of several domains tethered by linkers with rather compact conformations ($p47^{phox}$, dotted green line) or extended conformations ($p67^{phox}$, continue red line); curve of a fully disordered protein with very short elements of secondary structure (XPC dotted grey line); and curve of a fully disordered and extended protein with short segments of polyproline repeats (salivary protein IB5, continue purple line).

Receveur-Brechot, V. and D. Durand, *How random are intrinsically disordered proteins? A small angle scattering perspective.* Curr Protein Pept Sci, 2012. **13**(1): p. 55-75.

Examples of mixed systems



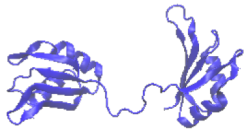


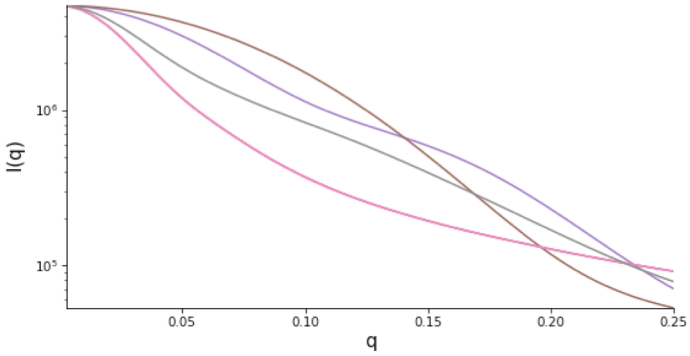


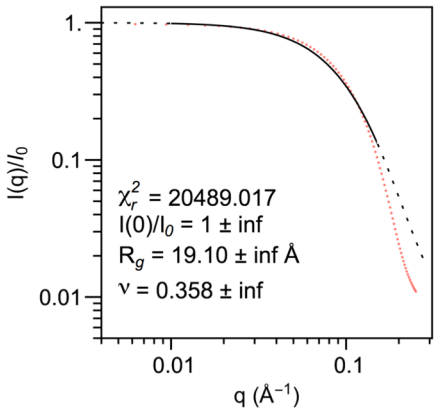
Pab1sims.vmd

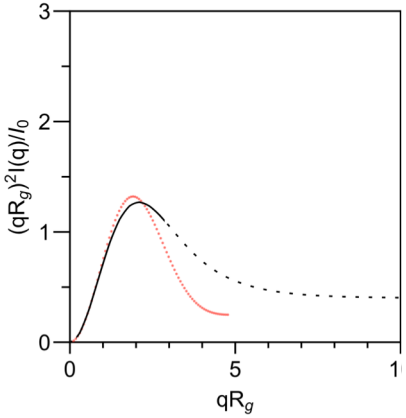
Examples of mixed systems

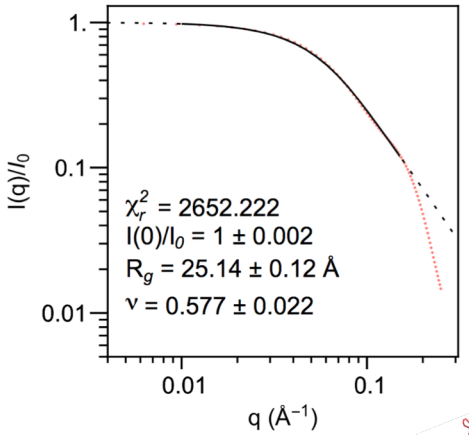
(shapes, Fit)

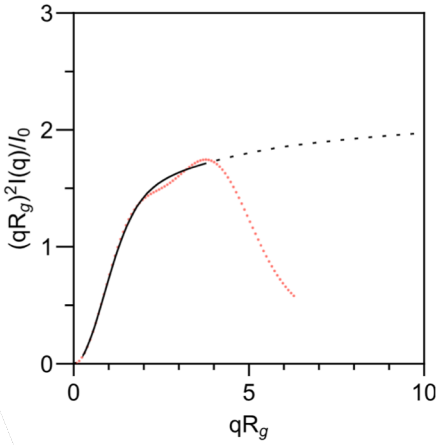


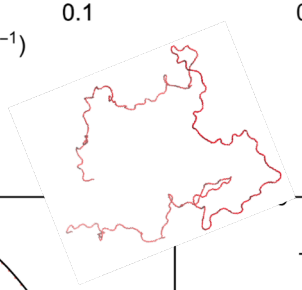


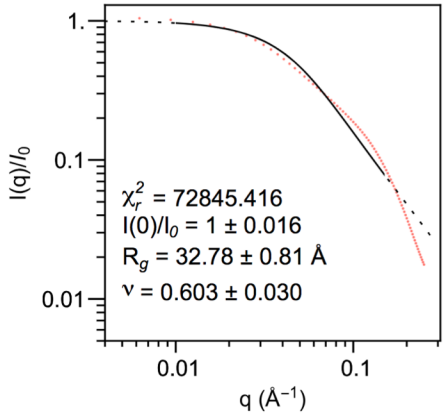


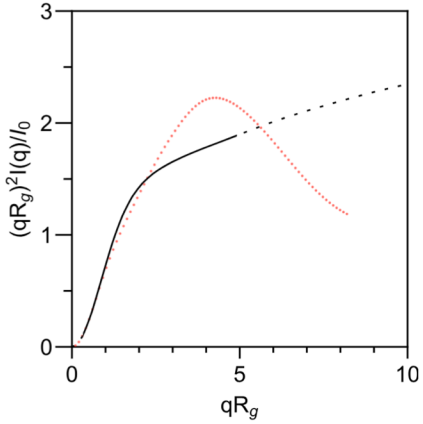


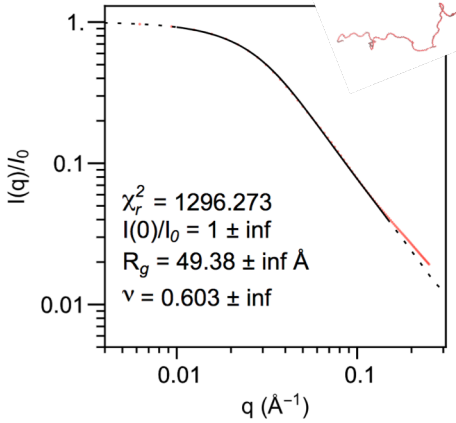


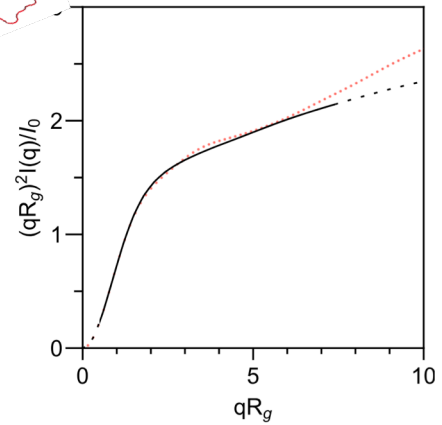




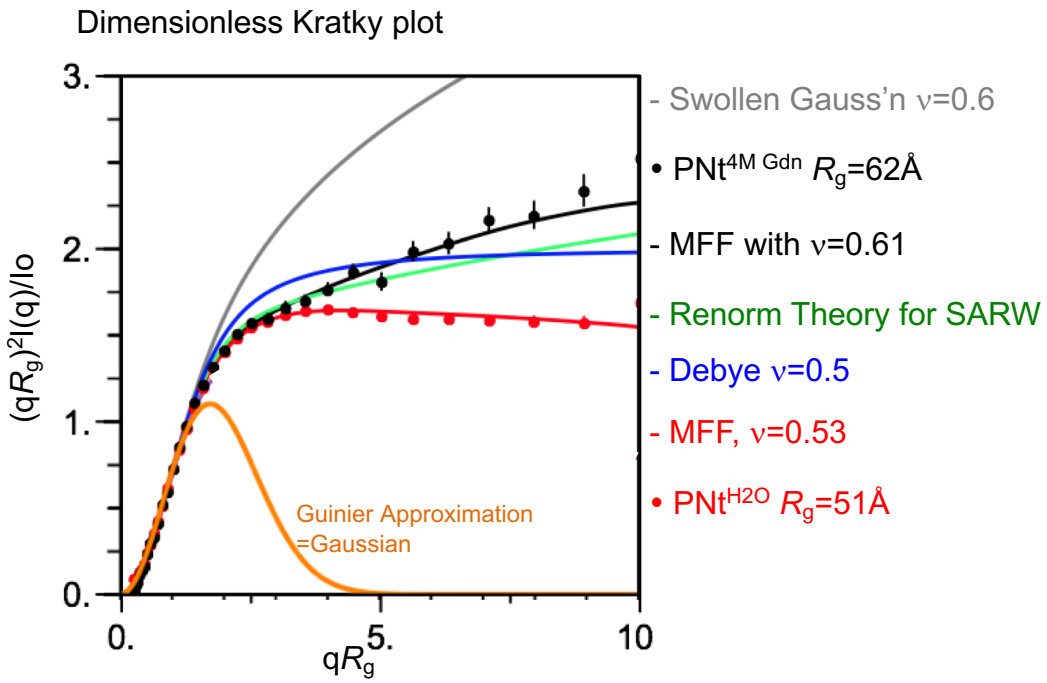








Scattering from Random Walks (allowed to cross) and self-avoiding RW (not allowed to cross)



Debye formula for random walk:

$$I_{Debye}(q) = \frac{2I_o\left(e^{-(qRg)^2} - 1 + (qRg)^2\right)}{(qRg)^4}$$

Swollen Gaussian coil model

$$P_{SwollenGausCoil}(R_{ij}) \propto e^{-\left(3R_{ij}^2 \,/\, 3\langle R_{ij}^2\rangle\right)}$$

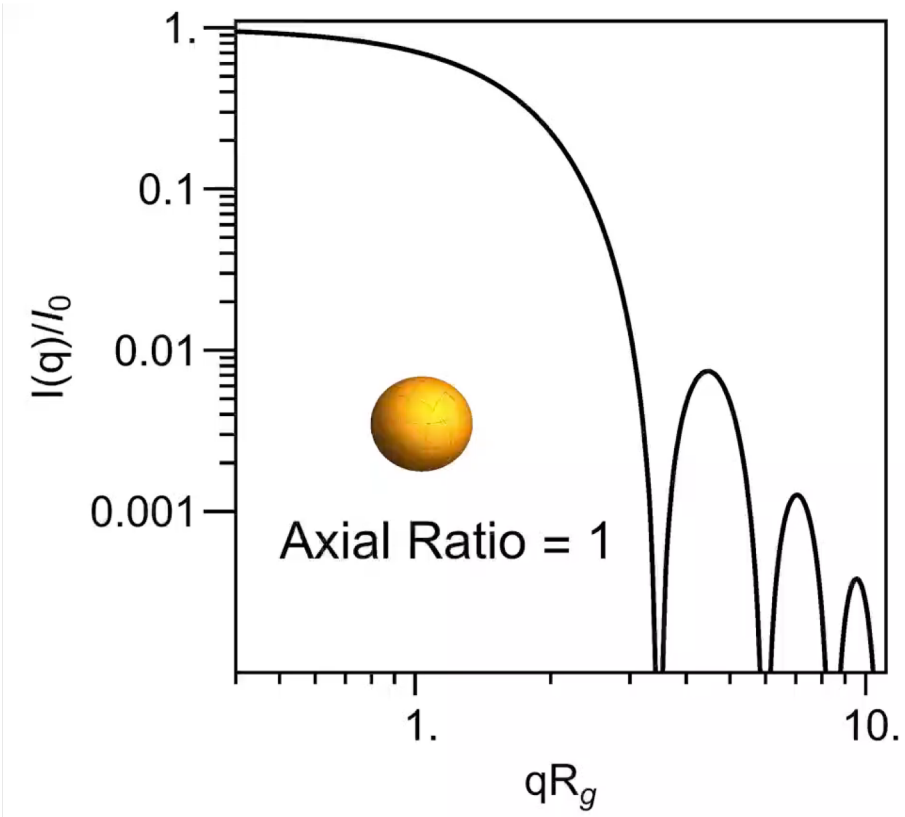
$$\left\langle R_{ij}^2 \right\rangle = a^2 \left|i-j\right|^{2\nu}$$

a=segment length

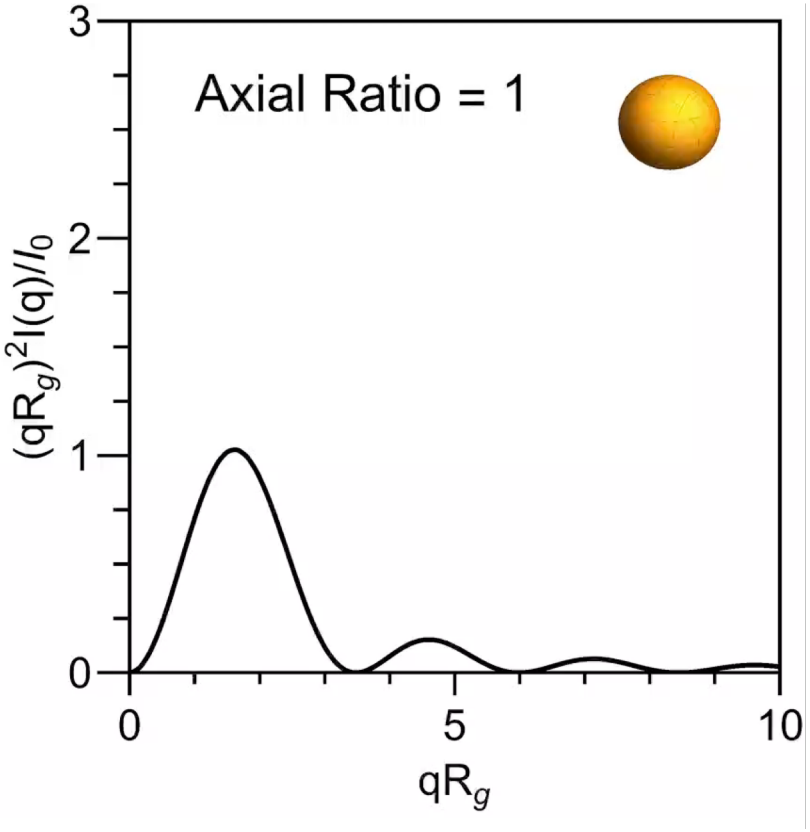
Molecular Form Factors (MFF) describe shape of objects

E.g., ellipsoid, axial ratio

Standard Log-log plot

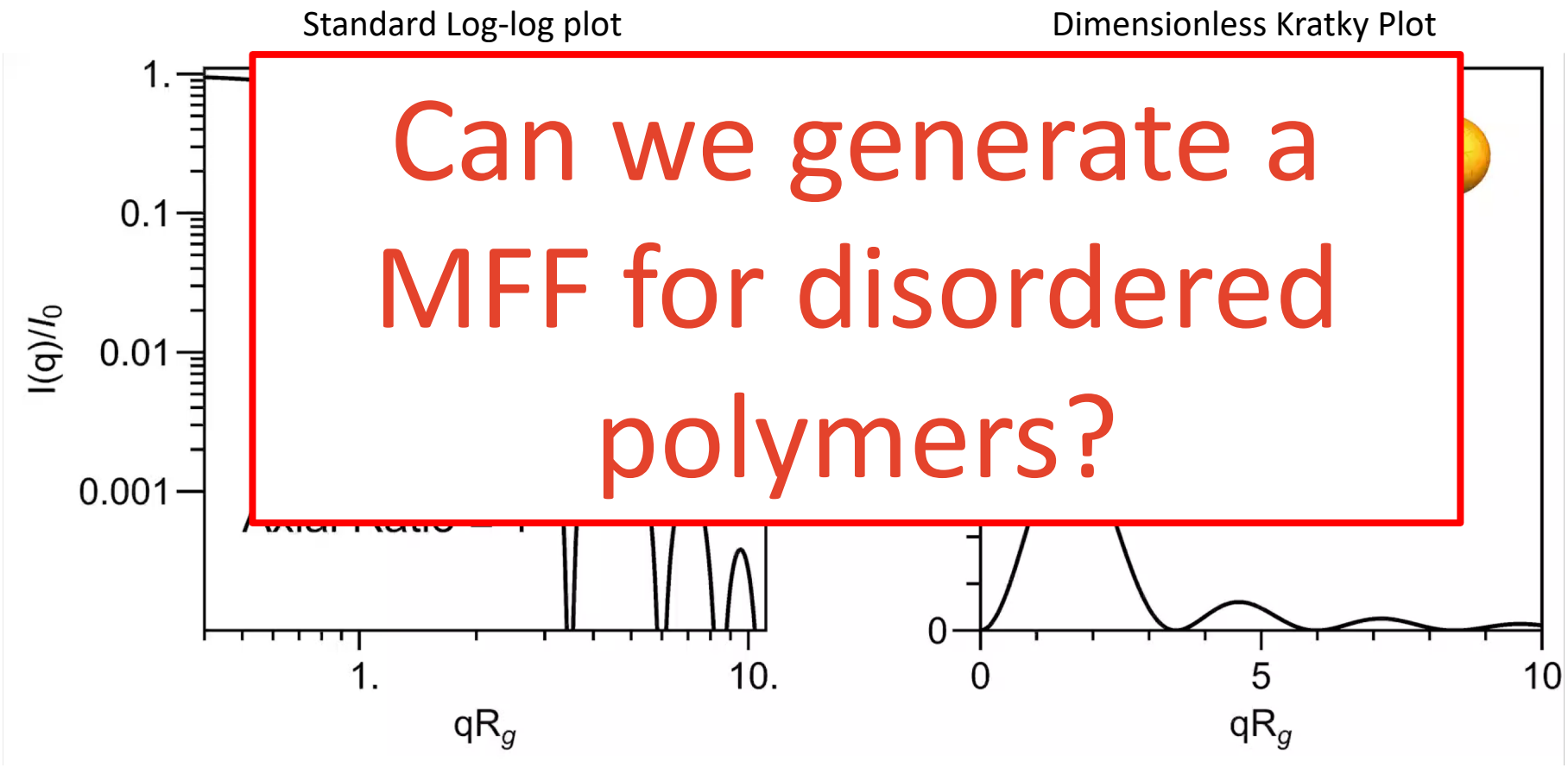


Dimensionless Kratky Plot



Molecular Form Factors (MFF) describe shape of objects

E.g., ellipsoid, axial ratio



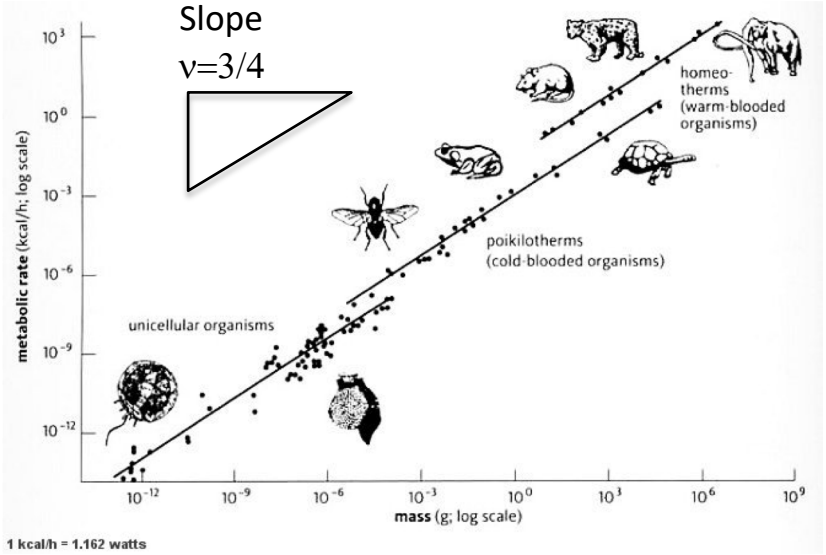
Can we generate a MFF for disordered polymers?

Allometry and scaling laws in Biology

General form for a biological variable for animals

Variable $\propto$ (Size)$^{\nu}$

Kleiber's Law: Metabolic Rate $\propto$ (Mass)$^{3/4}$



http://universe-review.ca/I10-83-metabolic.jpg

Spheres

Mass $\propto R^3$

or

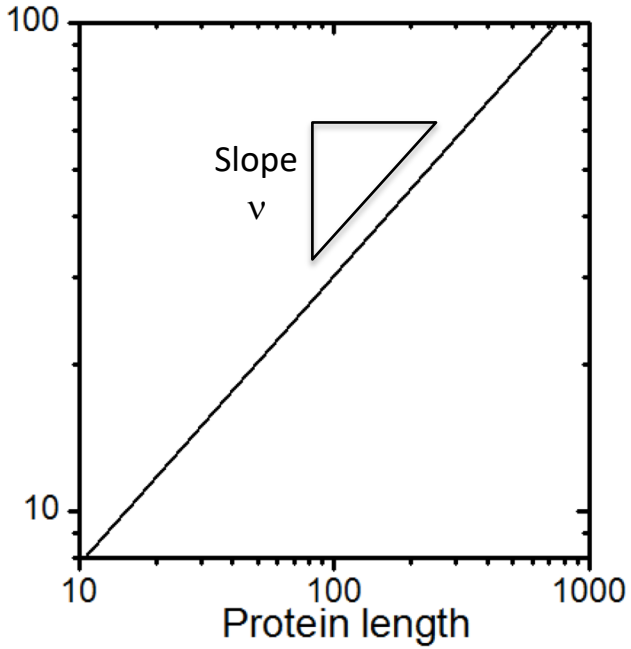
$R \propto M^{1/3}$

$\nu=1/3$

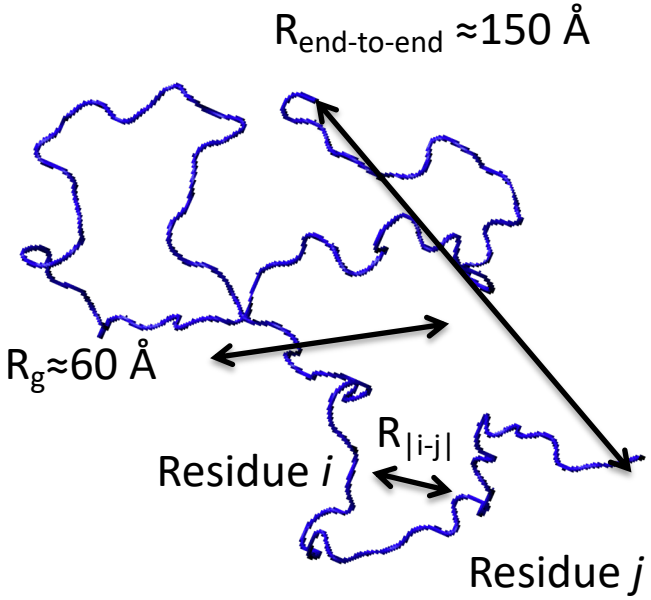
Polymer Physics: scaling laws and Flory exponent ν
Good and poor solvents

Solvent quality: Flory exponent ν

$R_g \propto N^{\nu}$
or
$R_{|i-j|} \propto |i-j|^{\nu}$



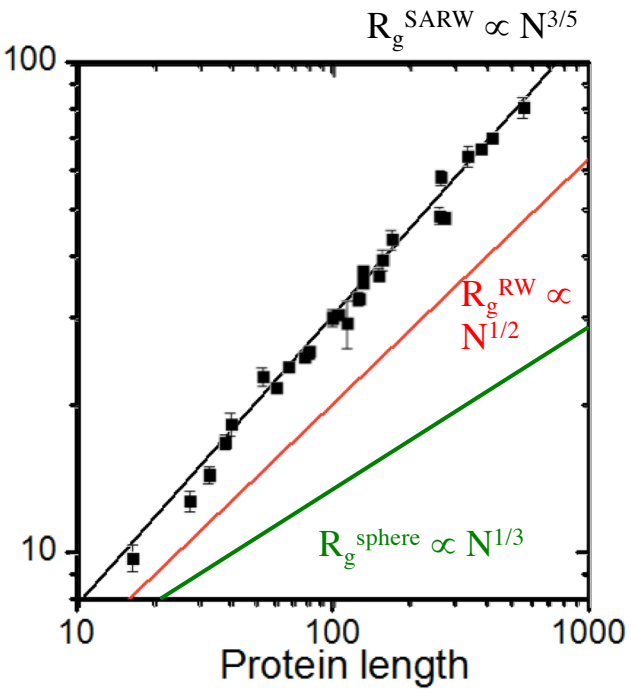
Or intra-residue spacing |i-j|



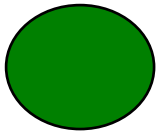
Polymer Physics: scaling laws and Flory exponent ν
Good and poor solvents

Solvent quality: ν

$R_g \propto N^{\nu}$



ν		
3/5	Self Avoiding random walk (SARW) $E_{prot\text{-}water} > E_{prot\text{-}prot}$	**Self-avoiding** Random Walk
Good solvent (contracted) $0.5<\nu\leq0.6$		
1/2	Random Walk behavior: Θ solvent $E_{prot\text{-}prot} \sim E_{prot\text{-}water}$	Random Walk (non-physical)
Poor (collapsed) $1/3<\nu<1/2$		
1/3	Globule $E_{prot\text{-}prot} > E_{prot\text{-}water}$	



$V=\frac{4}{3}\pi R^3$

Chemically denatured Proteins:
R_g scales as a self-avoiding random walk

Kohn et al., (2005) *PNAS*

General scaling behavior of proteins

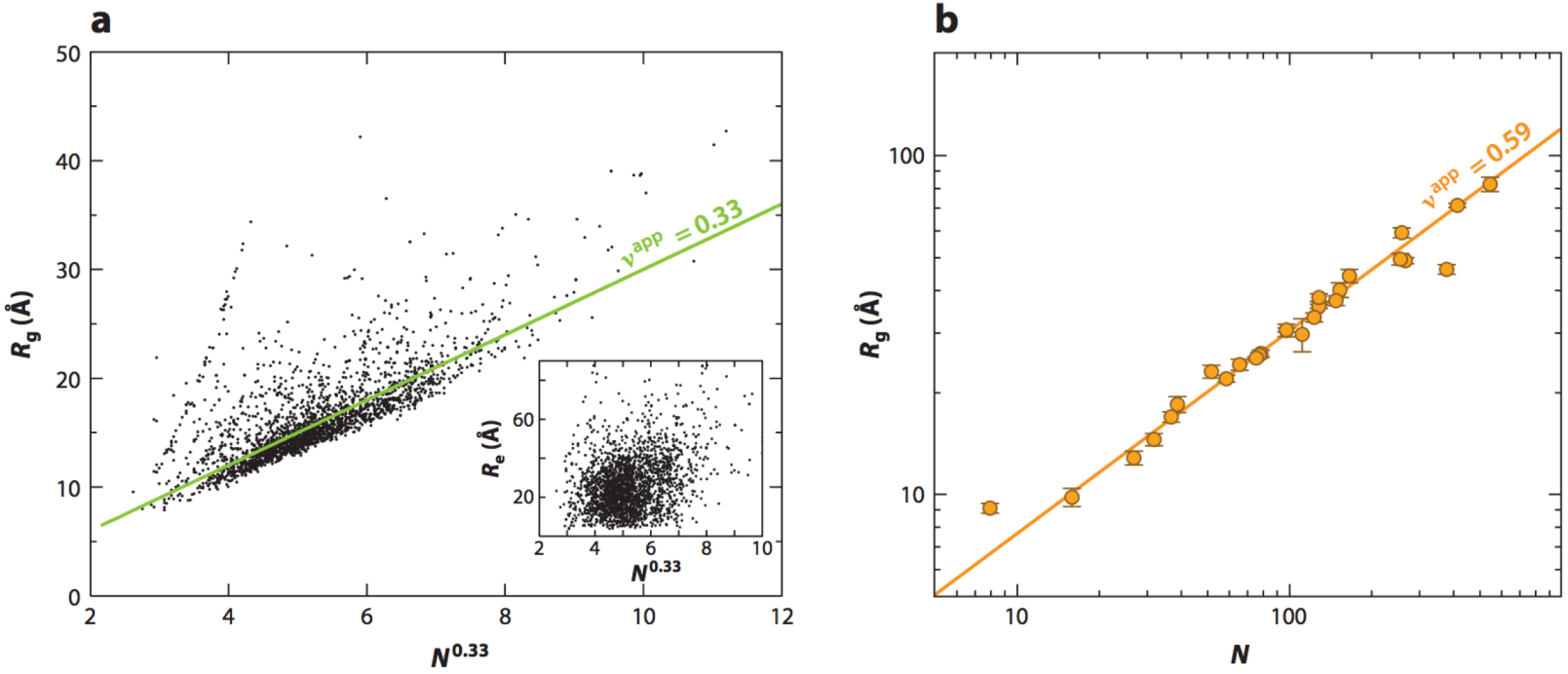


Figure 3

(*a*) Scaling of radius of gyration (R_g) with chain length N for folded proteins based on ~2,400 nonredundant structures taken from PDBSELECT25 (50). While the radius of gyration shows reasonable agreement with $\nu_{app} \approx 0.33$, the end-to-end distance shows a poor correlation (*inset*). (*b*) Scaling behavior for chemically denatured proteins based on data from Reference 74. The unfolded state under strongly denaturing conditions is well described by a self-avoiding random chain ($\nu_{app} \approx 0.59$).

Holehouse & Pappu (2018)
Annual Review Biophysics

Obtaining a MFF for disordered polymers
Polymer Physics: scaling laws and Flory exponent ν
Good and poor solvents

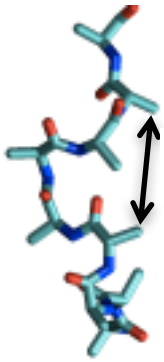
Debye formula for random walk

$$I_{Debye}(q) = \frac{2I_o\left(e^{-(qRg)^2} - 1 + (qRg)^2\right)}{(qRg)^4}$$

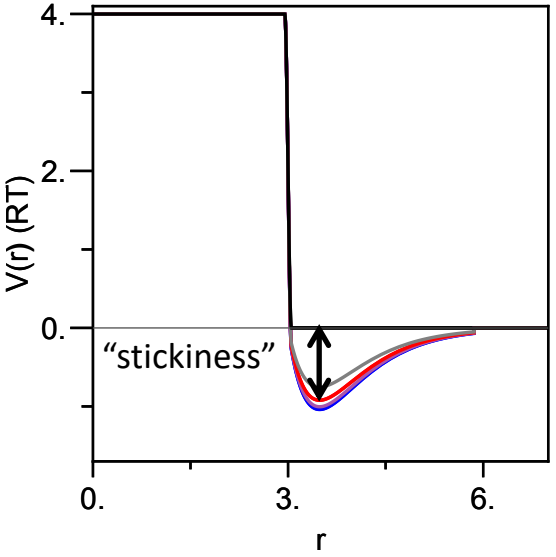
non-physical inter-penetration of chain

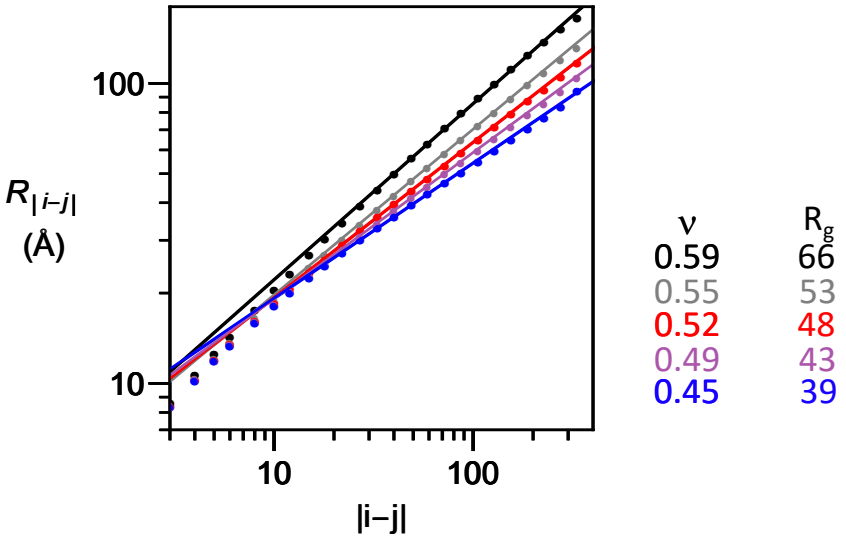
Can we generate a general MFF for realistic polymers?

Simulations generate ensembles used to make an MFF



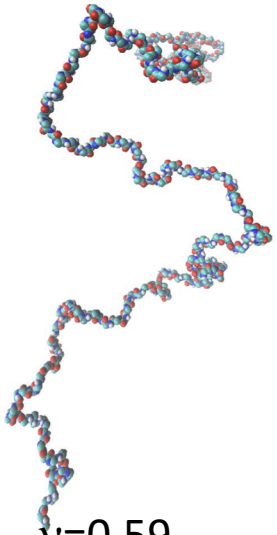
Vary Cβ–Cβ interaction





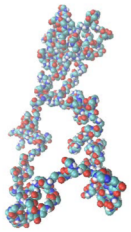
ν	R_g
0.59	66
0.55	53
0.52	48
0.49	43
0.45	39

Hard sphere



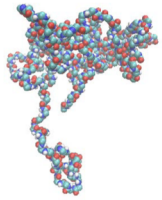
ν=0.59

A little sticky



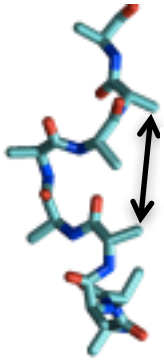
ν=0.52

A little more sticky



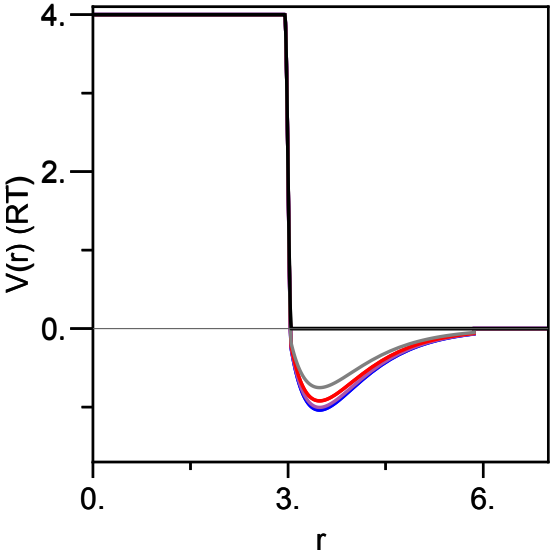
ν=0.45

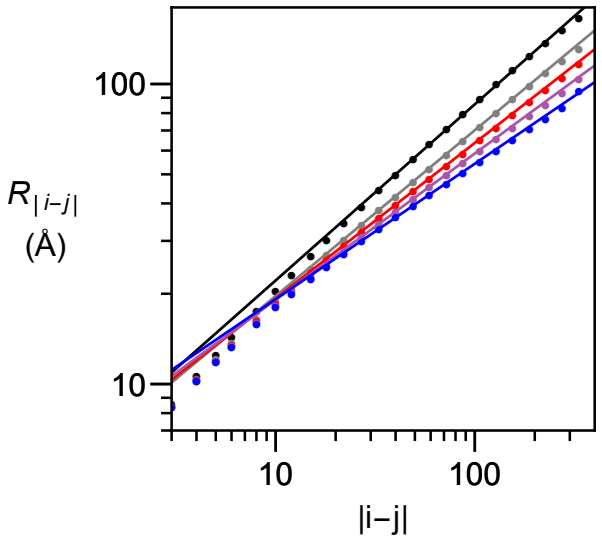
Simulations generate ensembles used to make an MFF



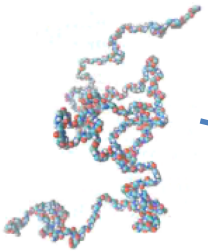
V(r)

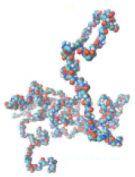
Change Cβ–Cβ interaction

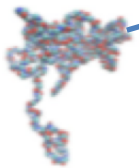




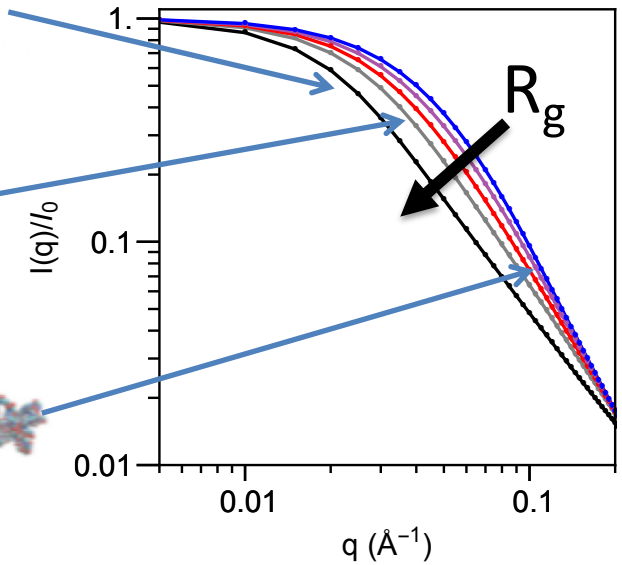
ν	R_g
0.59	66
0.55	53
0.52	48
0.49	43
0.45	39



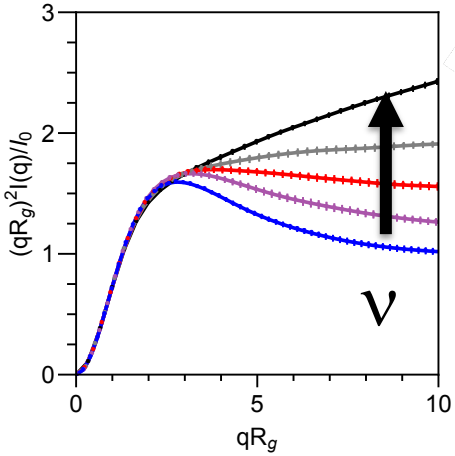


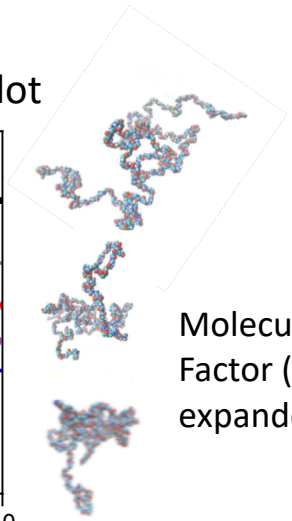


Standard Log-Log Plot



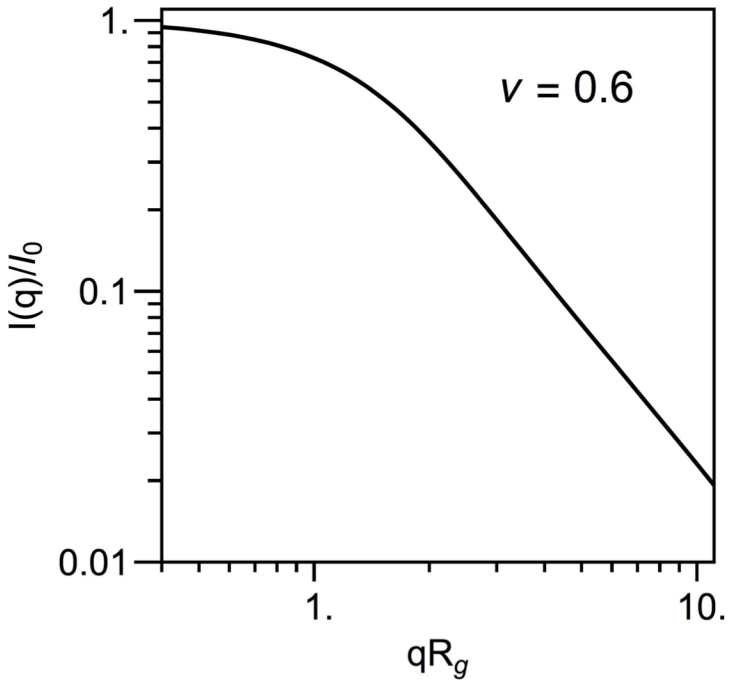
Dimensionless Kratky Plot

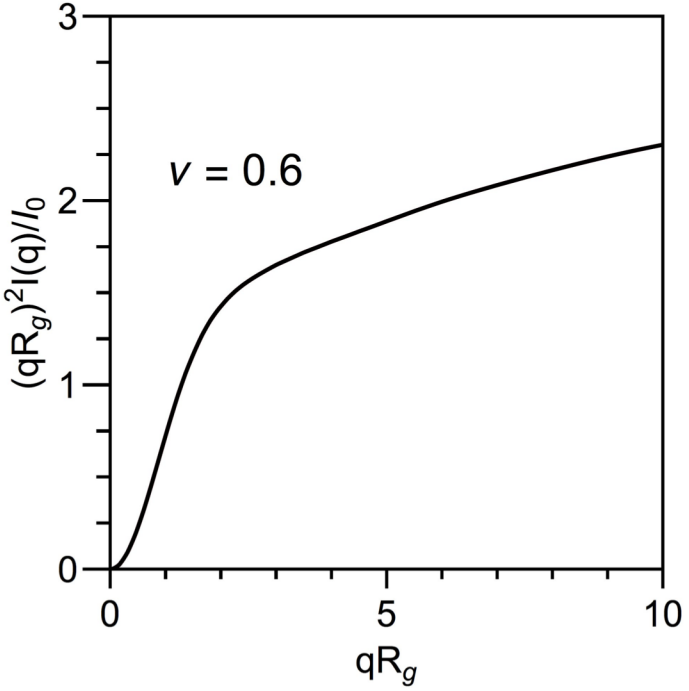




Molecular Form Factor (MFF) for expanded polymers

MFF for disorder systems: Obtain R_g and ν from a single SAS measurement



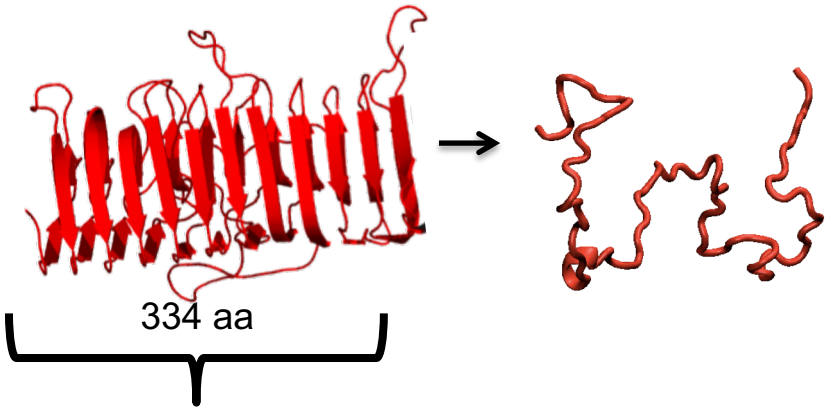


How well does it work?

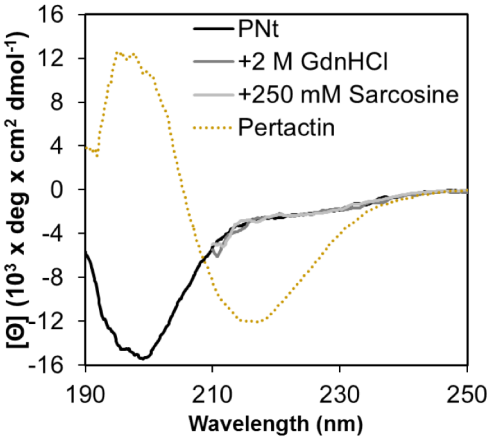
Riback et al. Science 2017

How well does our MFF work?
PNt, 334 residue, low charge hydrophobic IDP
P. Clark (U Notre Dame)

Pertactin: 539 aa β-helix

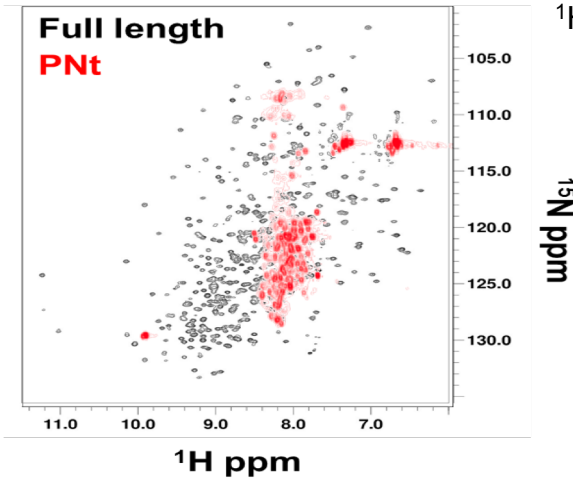


CD: unfolded

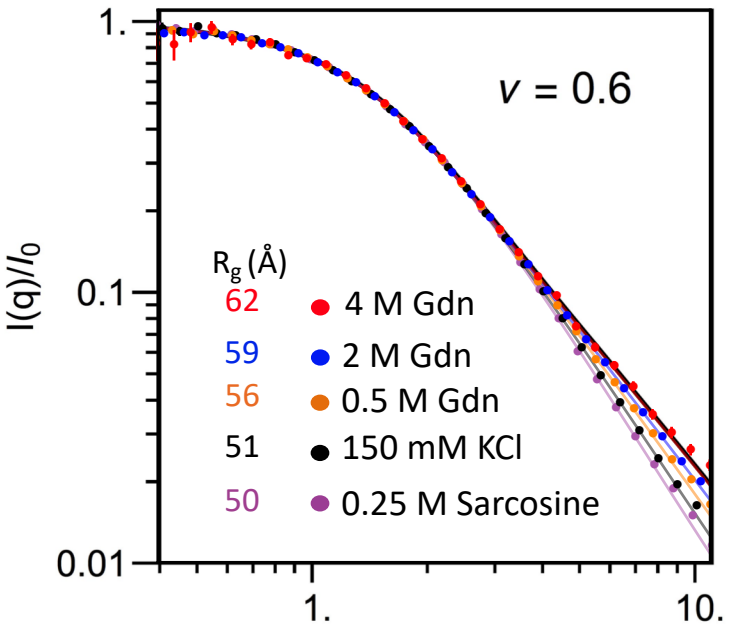


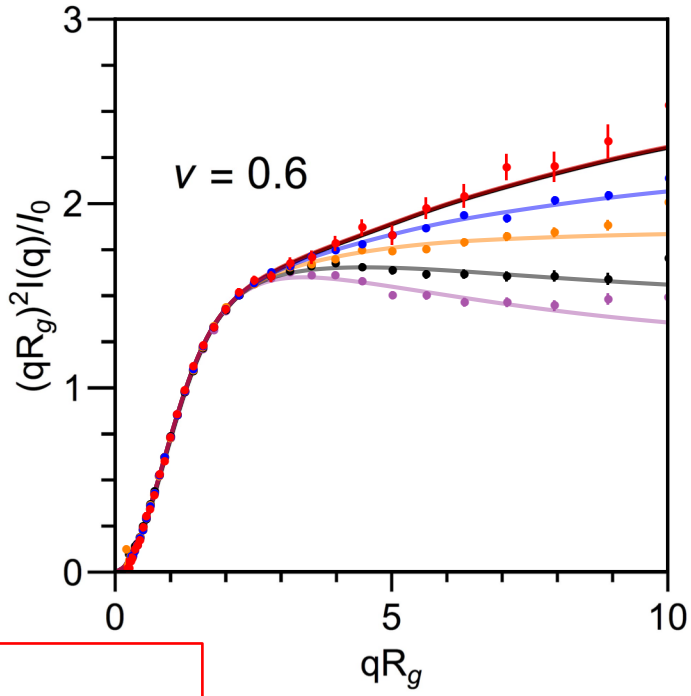
Chemical shifts typical of IDP

^{1}H ^{15}N HSQC:



Measuring R_g and ν for an IDP using the MFF





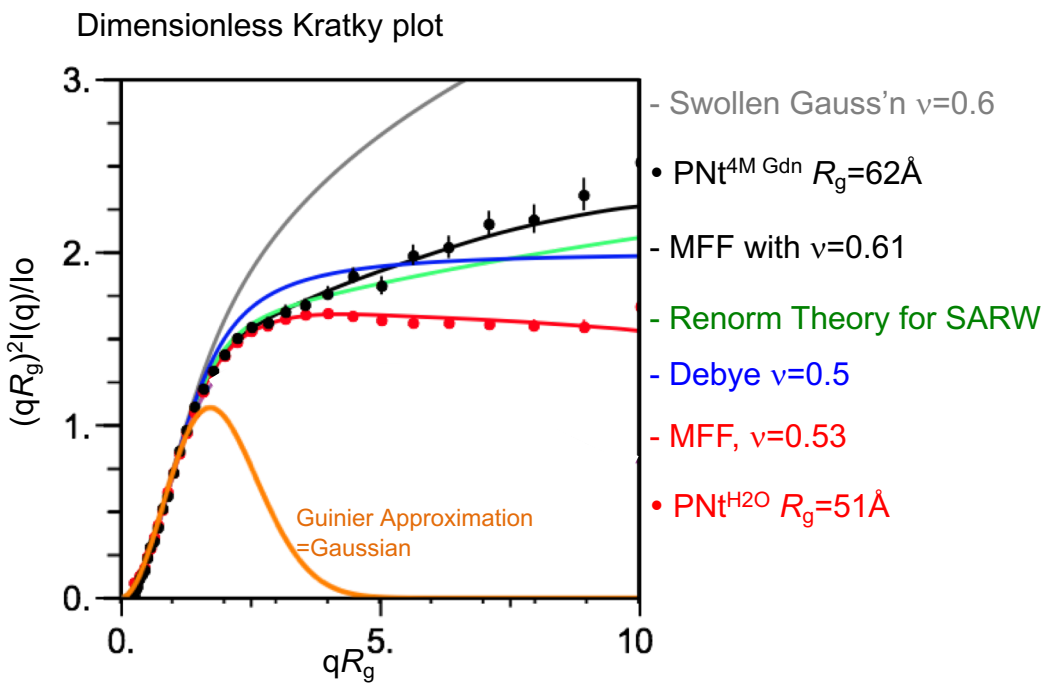
IT WORKS!

LET'S USE IT!

Fit out to qR_g~5 is generally sufficient
Hydration may be an issue at q>0.15Å^{-1}

Riback et al. Science 2017

Scattering from Random Walks (allowed to cross) and self-avoiding RW (not allowed)



Debye formula for random walk:

$$I_{Debye}(q) = \frac{2I_o\left(e^{-(qRg)^2} - 1 + (qRg)^2\right)}{(qRg)^4}$$

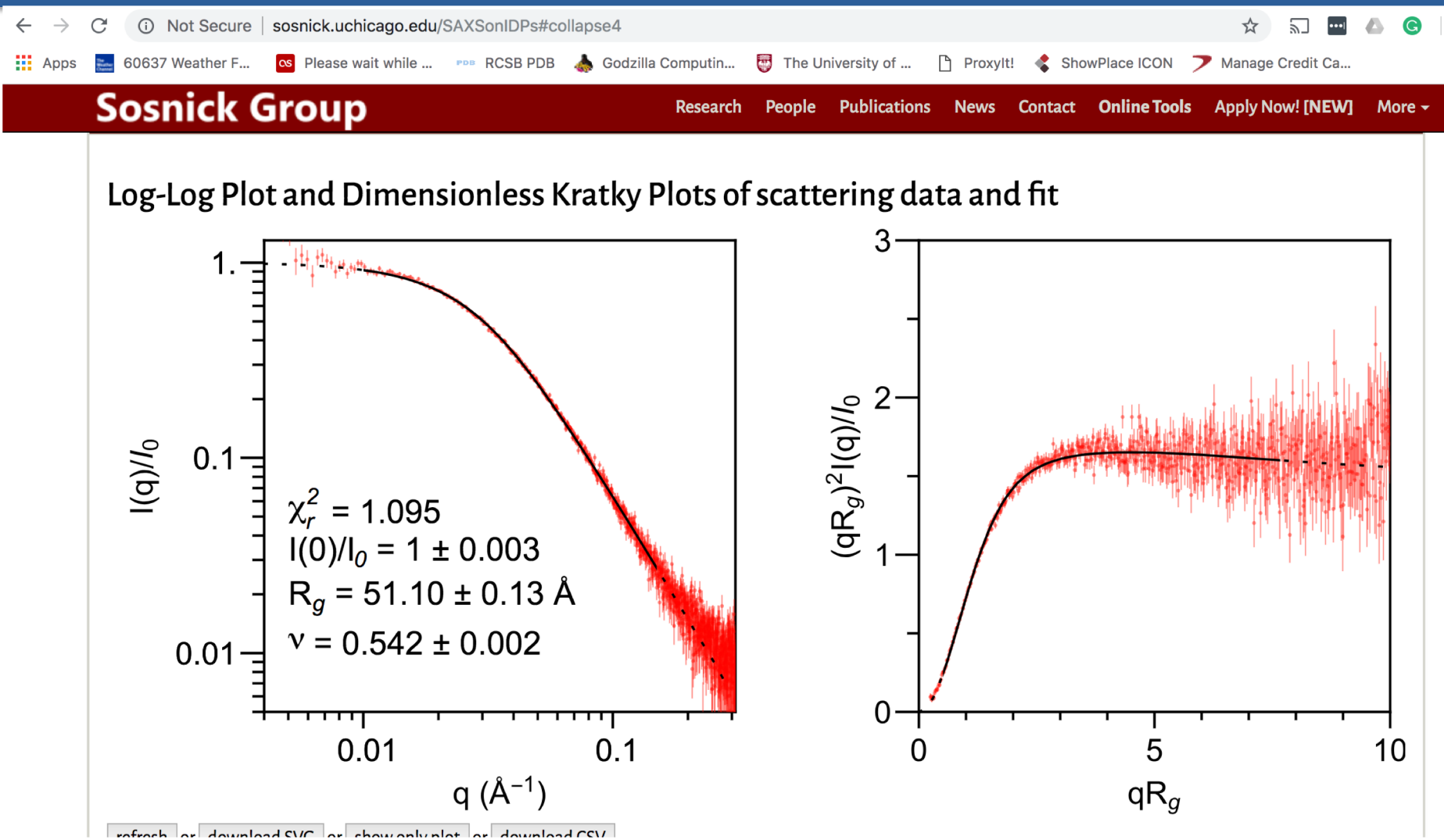
Swollen Gaussian coil model

$$P_{SwollenGausCoil}(R_{ij}) \propto e^{-\left(3R_{ij}^2 \,/\, 3\langle R_{ij}^2\rangle\right)}$$

$$\left\langle R_{ij}^2 \right\rangle = a^2 \left|i-j\right|^{2\upsilon}$$

a=segment length

Website for fitting disordered polymers
http://sosnick.uchicago.edu/SAXSonIDPs



Acknowledgements

Collaborators:
Notre Dame: Prof. Patricia Clark, Dr. Micayla Bowman

UChicago: Prof. Karl Freed,

Drs. Joshua Riback, John Jumper
Adam Zmyslowski

NIGMS

NSF GRP

Srinivas Chakravarthy BioCAT Beamline
APS: DOE, NIH