Atomistic simulations of intrinsically disordered proteins

Robert Best

Laboratory of Chemical Physics

NIDDK, National Institutes of Health

Acknowledgements

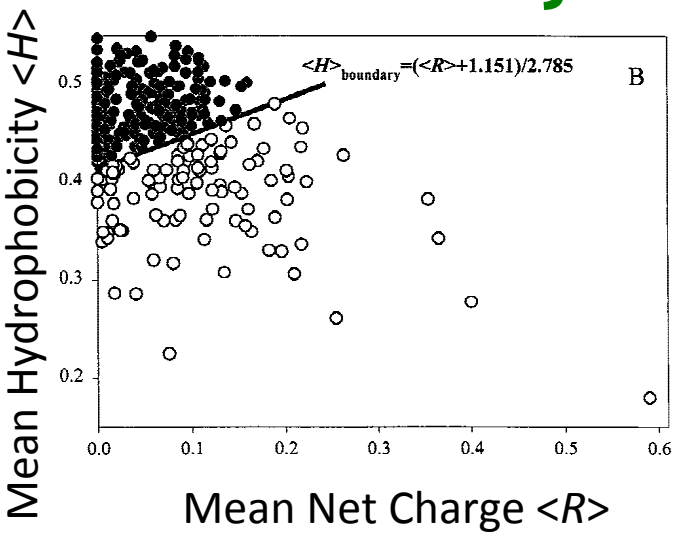
Group:

- **Wenwei Zheng** (NIH)

Collaborators:

- Gül Zerze, Jeetain Mittal (Lehigh University)
- **Alessandro Borgia**, Madeleine Borgia, **Ben Schuler** (University of Zürich) – smFRET, 2f-FCS
- **Alex Grishaev** (NIST) – SAXS
- **Klaus Gast** (University of Potsdam) - DLS
- Gerhard Hummer (NIH; now Max Planck for Biophysics)
- Magnus Kjaergaard, Birthe Kragelund (University of Copenhagen) – SAXS

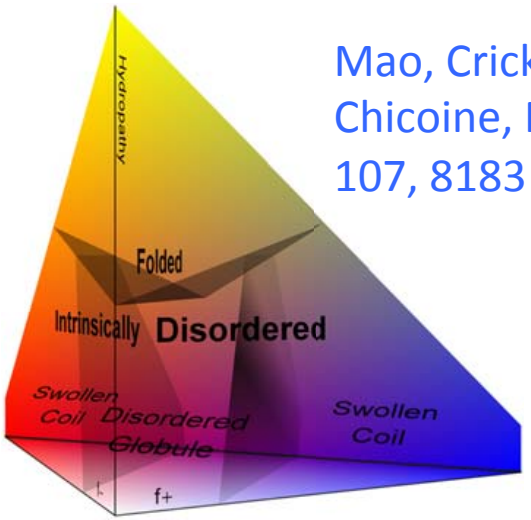
Intrinsically Disordered Proteins



Uversky, Protein Science, **11**, 739 (2002)

Factoids:

- Not folded (usually!)
- Low sequence complexity
- ~1/3 of eukaryotic proteome
- Often involved in signalling, e.g. Transcription factors



Mao, Crick, Vitalis, Chicoine, Pappu, PNAS, 107, 8183 (2011)

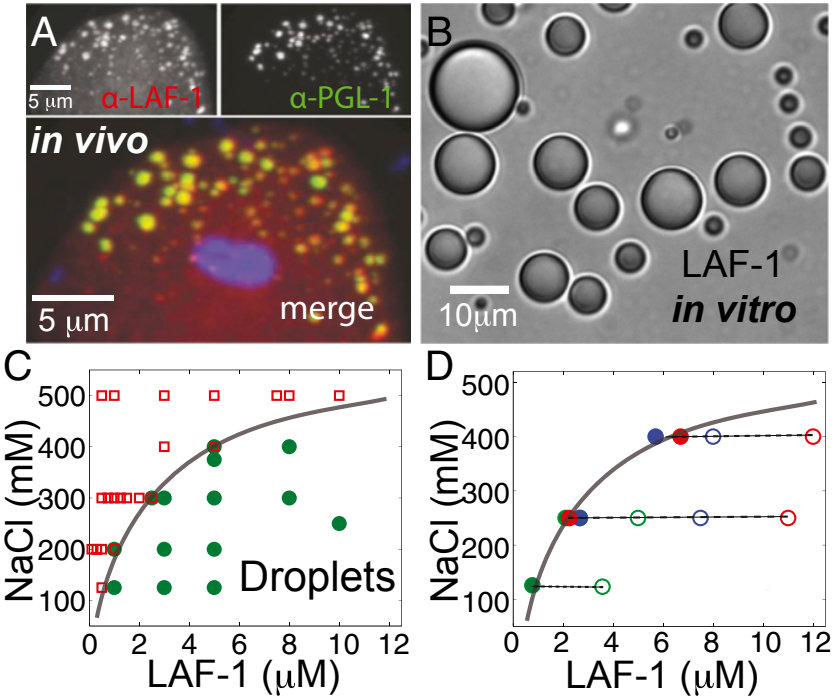
Challenges

- Relation between sequence properties and function?
- Challenging for conventional structural biology techniques
- Can molecular simulations help??

Intrinsically Disordered Proteins

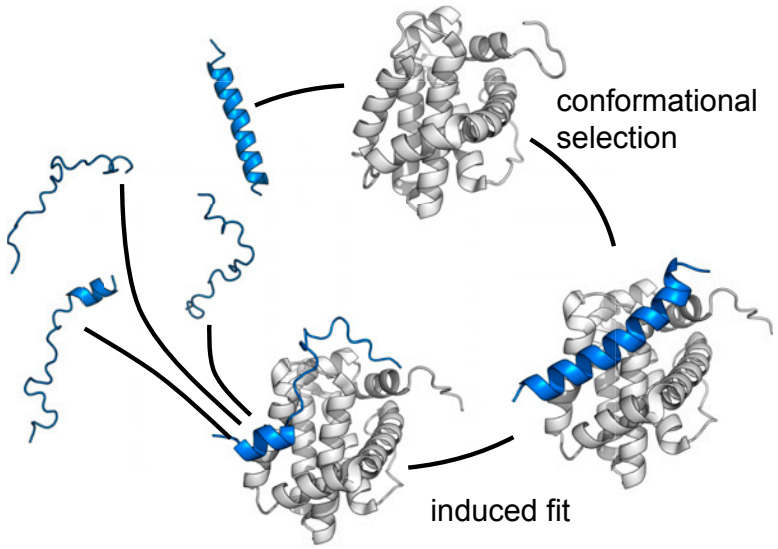
Problems where molecular simulation/ theory may help

"Granule" formation



Elbaum-Garfinkle et al, PNAS, **112**, 7189 (2015)

Coupled folding-binding



Rogers, PNAS, **111**, 15420 (2014)

Outline

1. An experimental controversy: protein collapse viewed via SAXS or FRET
2. Interpretation of SAXS and FRET by all-atom simulation
3. Interpretation of experiments in terms of molecular ensembles

1. An experimental controversy: protein collapse viewed via SAXS or FRET

1. Unfolded state collapse controversy

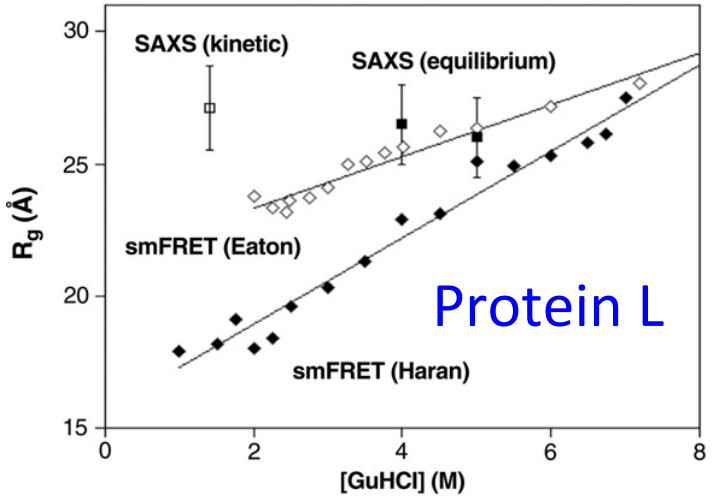
Exemplifies challenges of obtaining structural information on IDPs or unfolded proteins

Contents lists available at www.sciencedirect.com
Journal of Molecular Biology
journal homepage: http://ees.elsevier.com.jmb

Small-Angle X-ray Scattering and Single-Molecule FRET Spectroscopy Produce Highly Divergent Views of the Low-Denaturant Unfolded State

Tae Yeon Yoo [1]†, Steve P. Meisburger [2]†, James Hinshaw [3]†, Lois Pollack [2]*, Gilad Haran [4]*, Tobin R. Sosnick [1,5,6]* and Kevin Plaxco [7,8]*

JMB, **418**, 226 (2012)



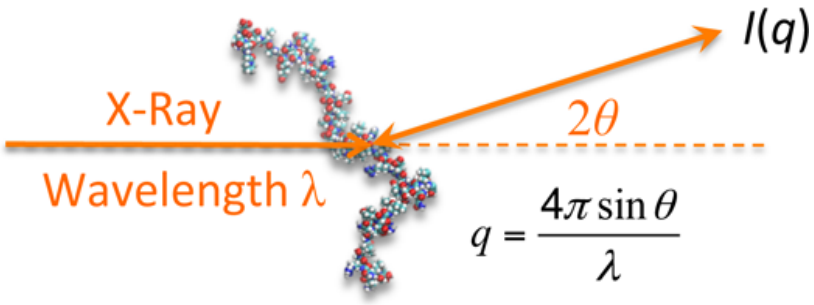
- FRET (and DLS): collapse at low [GdmCl]
- SAXS: no collapse!

Problem: hard to study unfolded proteins at low [GdmCl]

Implications

- Uncertainty over "correct" result because experimental outcomes differ. Problem for studying IDPs?
- Denaturation mechanism: standard model of "binding" of denaturant to protein appears to contradict SAXS outcome qualitatively. How do denaturants work?

Small-Angle X-ray Scattering



Scattering intensity

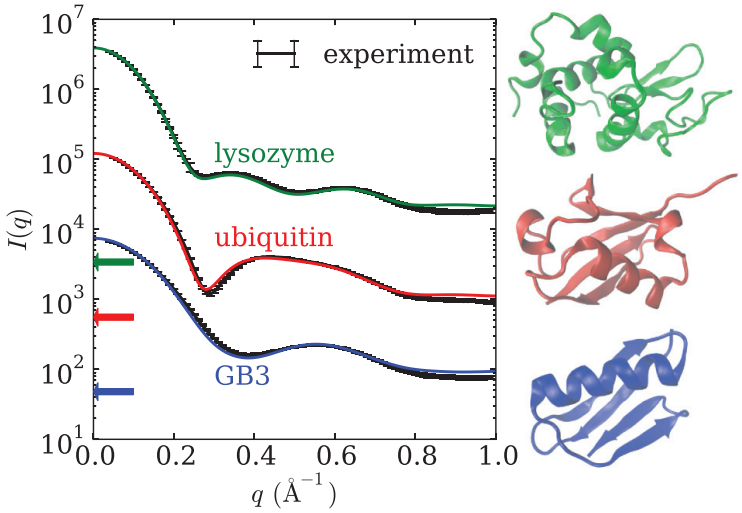
$$I(q) = 4\pi \int_0^\infty P(r) \frac{\sin(qr)}{qr} dr$$

Taylor expand, truncate, … Guinier Approximation

$$I(q) \approx I(0)\left(1 - \frac{q^2 r_g^2}{3}\right) + \ldots$$

$$\ln\left[I(q)\right] \approx \ln\left[I(0)\right] - \frac{qr_g^2}{3}$$

$$r_g^2 = \frac{1}{2}\left\langle r^2 \right\rangle$$



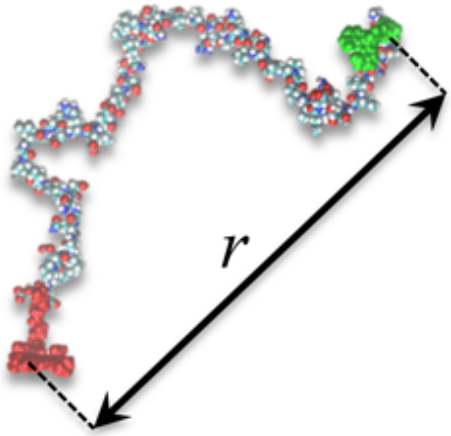
Must have:

1. Dilute Solution
2. Background Subtraction
3. qr_g very small

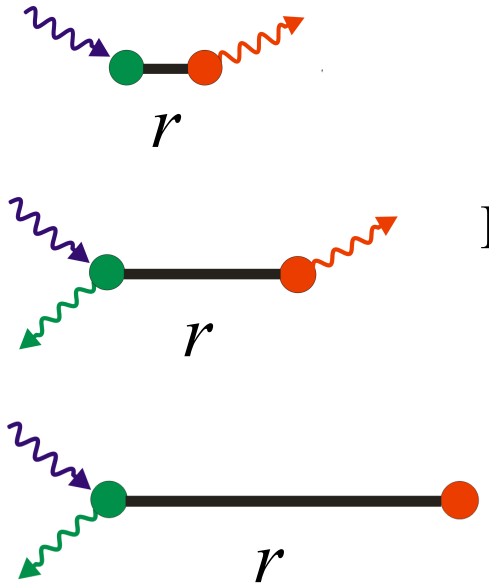
Förster Resonance Energy Transfer

FRET: sensitive to distance between *donor* and *acceptor* chromophores

Single-Molecule FRET
PNAS, **92**, 6264 (1996)



$$E(r) = \frac{1}{1 + \left(r / R_0\right)^6}$$



$$\mathrm{E} = \frac{n_A}{n_A + n_D}$$

What could be cause?

- Sticky chromophores in FRET?
- Preferential denaturant partitioning affecting SAXS?

  Zheng et al., JACS (in press)

- Details of experimental interpretation?

  Borgia et al., JACS (in press)

2. Interpretation of SAXS and FRET by all-atom simulation

Strategy

- Study 76-residue intrinsically disordered protein (ACTR), can cover complete denaturant range
- Compute SAXS, FRET, compare with experiment (A. Borgia, B. Schuler, A. Grishaev)
- Investigate molecular origins of observed signals – do they fit with experimental interpretation?

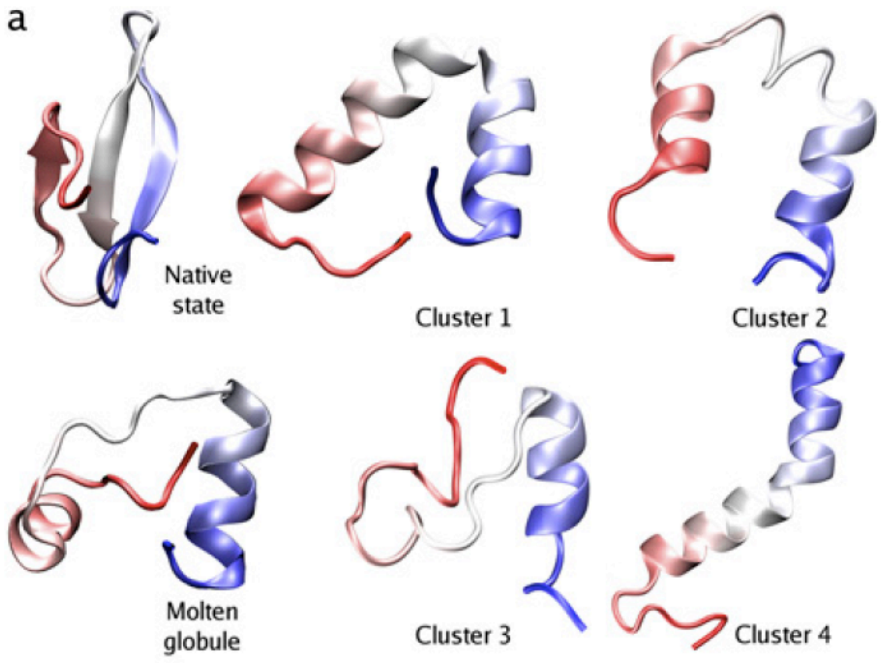
Key requirement is to have accurate energy function (force field)

All-atom Force fields

1. Secondary structure bias

10 μs simulation of fast-folding pin WW domain mutant with CHARMM 27 protein force-field and explicit water.

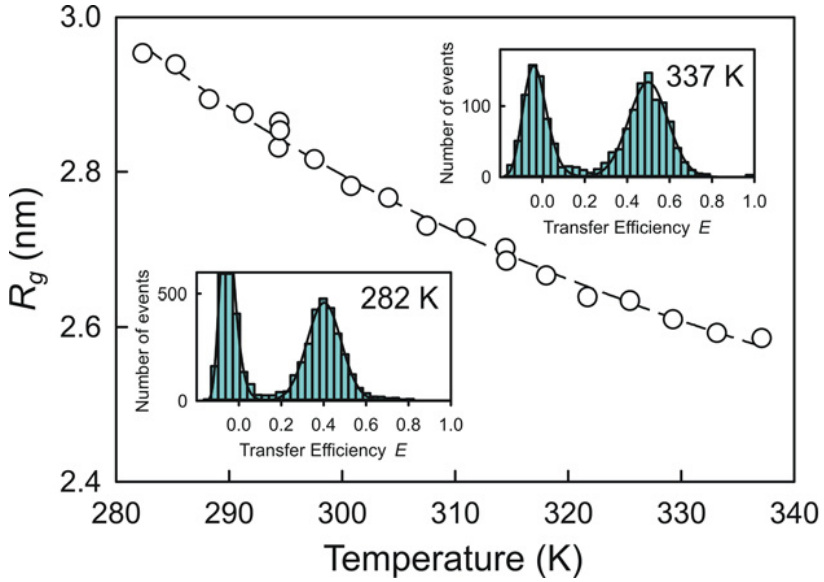
Experimental folding time ~13.3 μs.



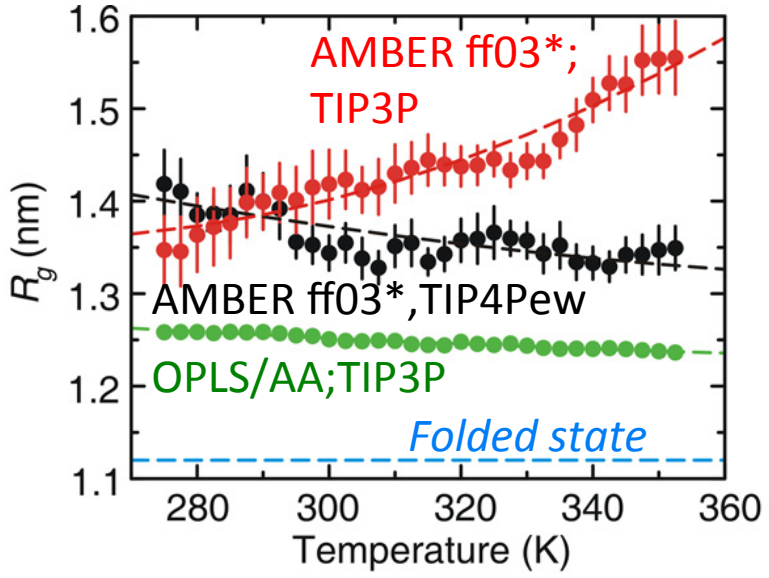
Freddolino, Liu, Gruebele & Schulten, Biophys. J. **94**, L75 (2008)

All-atom force fields
2. Collapse of unfolded state

Experiment



Simulation



Temperature-induced collapse of cold-shock protein

Nettels *et al. PNAS*, **106**, 20740 (2009)

Piana *et al.*, Curr. Opin. Struct. Biol. (2014)

Force field fixes

Secondary Structure bias:

- Corrected by adjusting torsion angle parameters against experimental data on peptides in water

  Garcia, Sanbonmatsu, PNAS, **99**, 2782 (2002)
  Best, Hummer. J. Phys. Chem. B, 113, 9004 (2009).
  Lindorff-Larsen *et al*, Science, **334**, 517 (2011)
  Best et al, JCTC, **8**, 3257 (2012)

Protein Collapse:

- Empirically adjusting protein-water interactions against experimental data

  Ashbaugh et al, J. Chem. Phys, **132**, 124504 (2010)
  Nerenberg, Jo, So, Tripathy, Head-Gordon, JPCB, **116**, 4524 (2011)
  Best, Zheng, Mittal, JCTC, 10, 5113-5124 (2014)

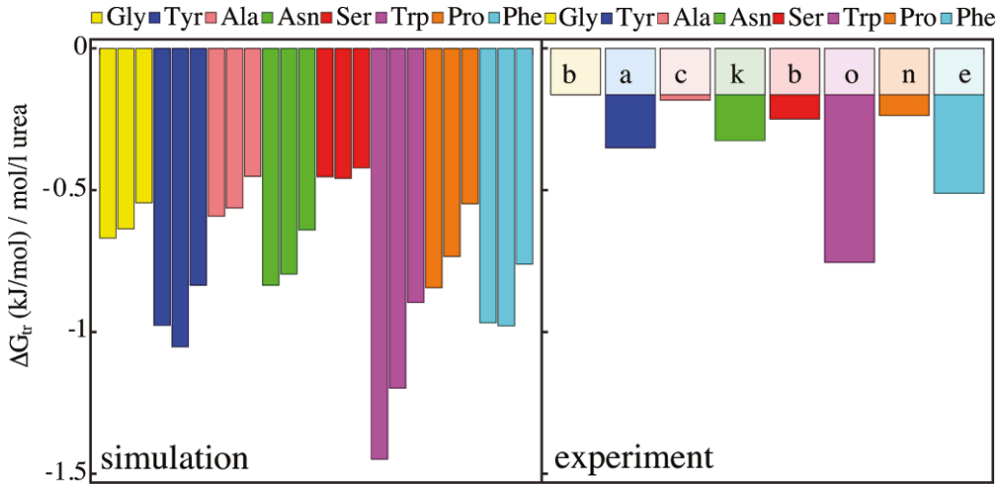
- Modifying water model

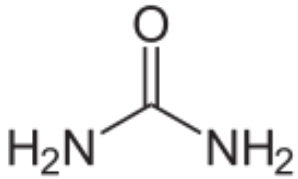
  Piana *et al.* JPCB, **119**, 5113 (2015)

- Adjusting Amide Lennard-Jones parameters (??)

  Yoo *et al.* JPC Lett. 2016, 7, 3812−3818 (2016)

Denaturant force field





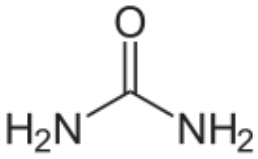
GROMOS protein/KBFF urea/SPC water

Horinek and Netz, JPC A, 115, 6125 (2011)

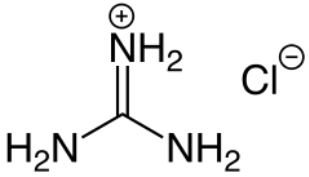
Urea-protein binding is too tight in most force fields

How to parametrize urea / GdmCl?

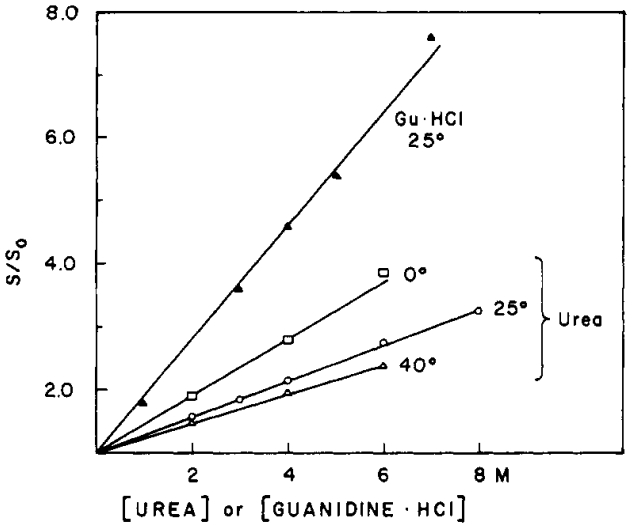
- Protein-water interactions good (ff03ws)
- KBFF model for urea accurate in TIP4P/2005
- Need to test/optimize protein-denaturant interaction



Urea



Guanidinium Chloride (GdmCl)



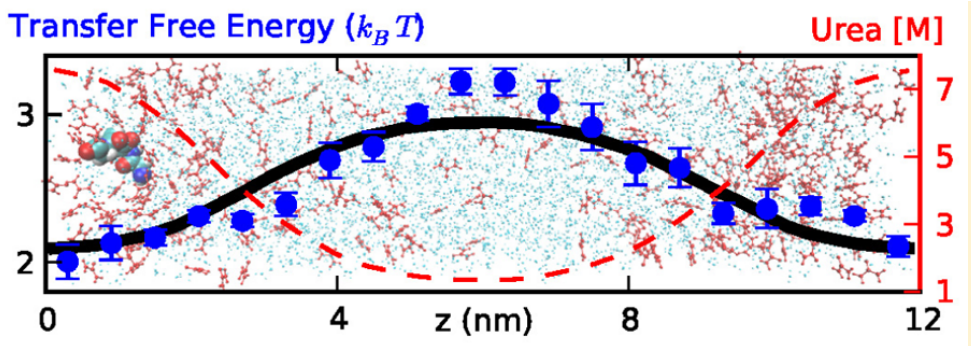
Target data: solubility of capped tetraglycine in denaturant

Transfer free energy from relative solubility:

$$\beta \Delta F_{\mathrm{tr}} = -\ln(S/S_0)$$

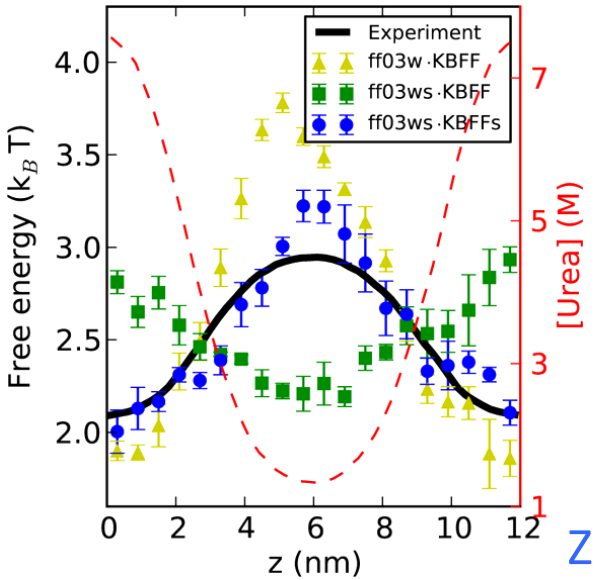
KBFF: Weerasinghe and Smith, JPCB, **107**, 3891 (2003)
Solubility: Robinson and Jencks, JACS, **87**, 2462 (1965)

Calculating transfer free energies



Urea restraint:

$$U_{\mathrm{d}}(z) = k\left[\cos\frac{2\pi(z - z_{\mathrm{c}})}{z_{\max}} + 1\right]$$



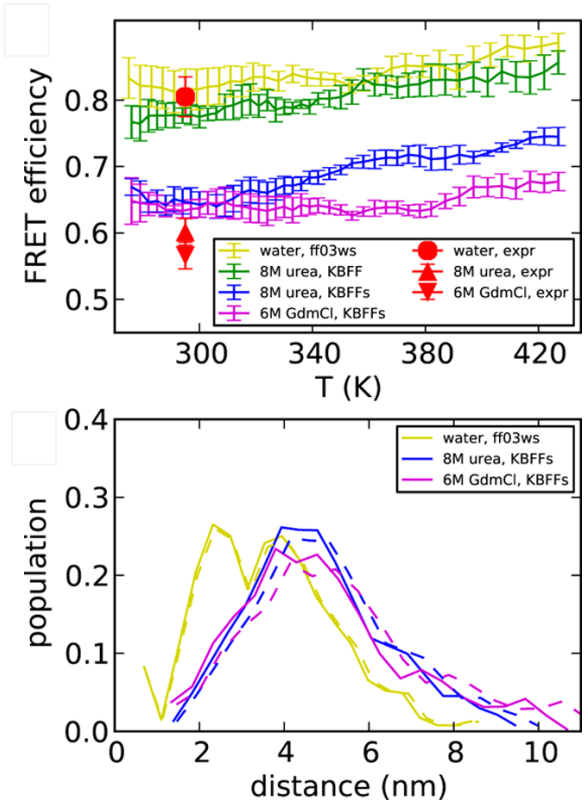
ATGEE transfer free energies

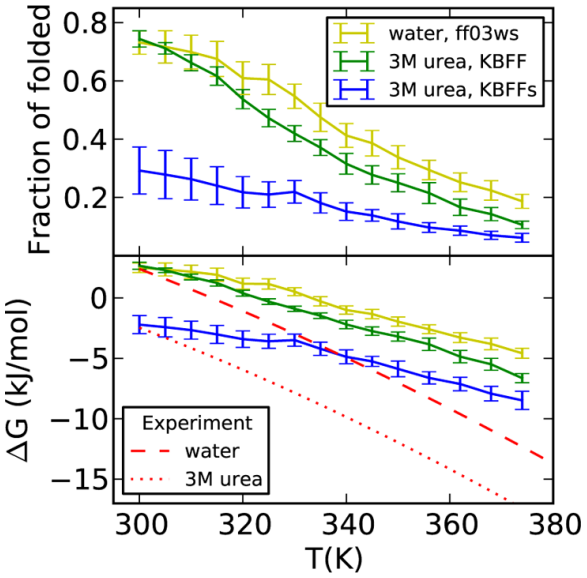
denaturant model	force field	transfer free energy ($k_{\mathrm{B}}T$)
Urea		
KBFF	ff03w·TIP3P	−1.99 (0.11)
$\mathrm{Amber}_{\mathrm{D}}$	ff03w	−3.03 (0.27)
KBFF	ff03w	−1.83 (0.08)
KBFF	ff03ws	0.43 (0.10)
KBFFs	ff03ws	−0.81 (0.12)
experiment		−0.77[26]

Zheng *et al.*, *J. Chem. Theor. Comput.*, **11**, 5543 (2015)

Testing urea force field

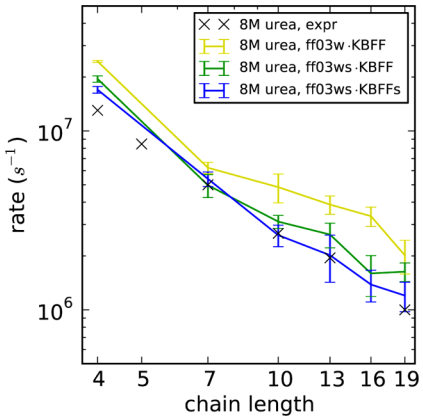
Dimensions of Csp M34





Stability of Trp Cage Miniprotein

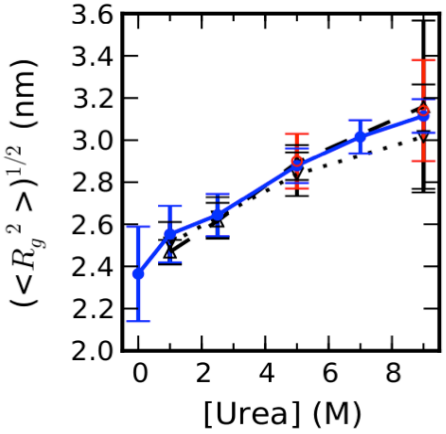
Triplet quenching of $C(AGQ)_nW$

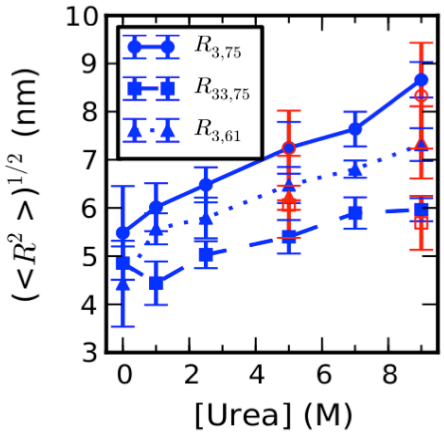


Zheng *et al.*, *J. Chem. Theor. Comput.*, **11**, 5543 (2015)

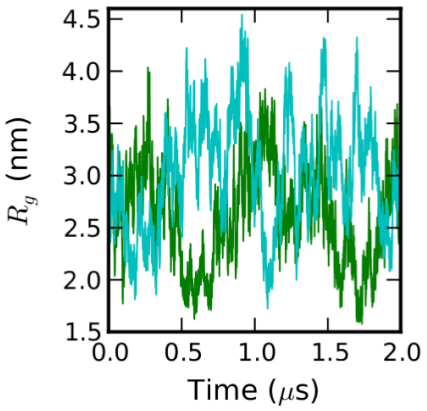
All-atom simulations of ACTR in urea and Gdmcl

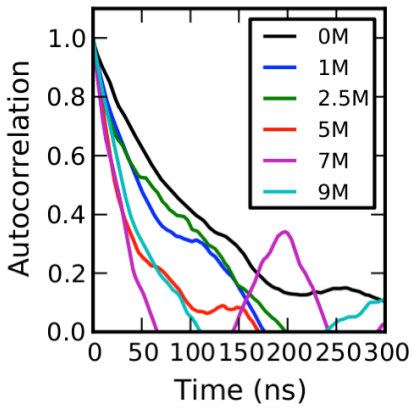
Brute force MD simulations. No free energy barrier, no advantage to enhanced sampling*

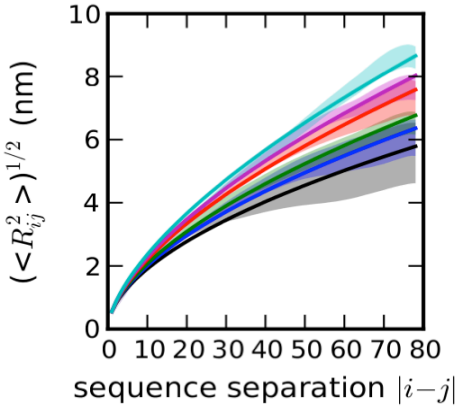




ACTR expands in urea







Increase of scaling exponent with denaturant concentration:
~0.5 in water
~0.6 at high [urea]

Denaturation mechanism

How to characterize weak binding?

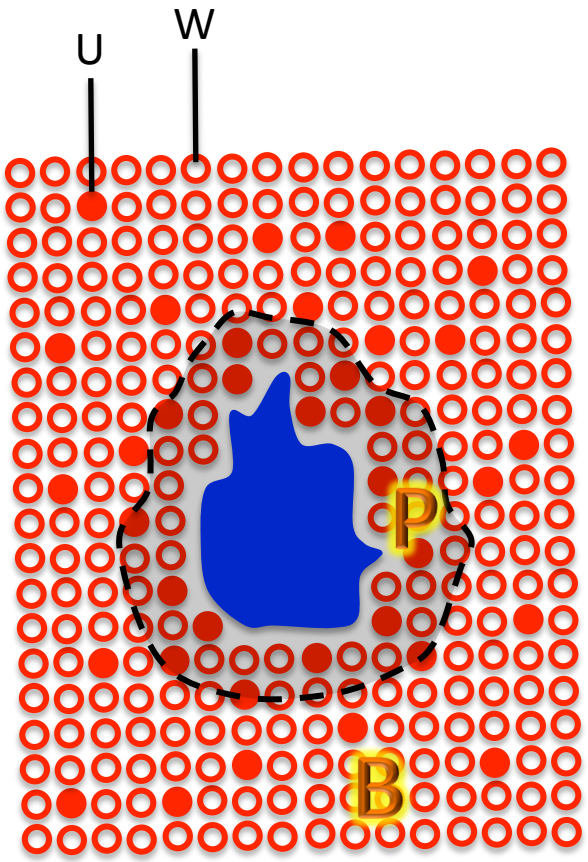
Preferential interaction coefficients:

$$\Gamma_{UP} = (\partial m_U / \partial m_P)_{\mu_U}$$

From simulation:

$$\Gamma_{UP} = \langle n_U^P - n_W^P \left(\frac{n_U^B}{n_W^B}\right) \rangle$$

Can decompose into "group" contributions using Voronoi analysis



Denaturation mechanism

How to characterize weak binding?

Preferential interaction coefficients

$$\Gamma_{UP} = (\partial m_U / \partial m_P)_{\mu_U}$$

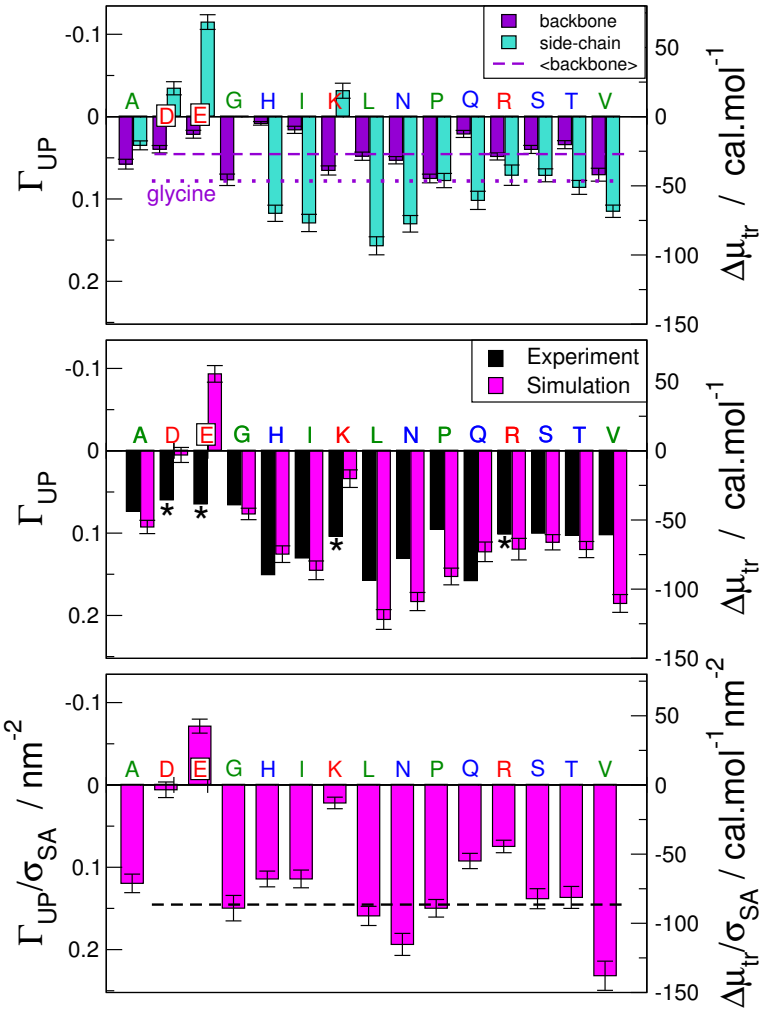
From simulation:

$$\Gamma_{UP} = \langle n_U^P - n_W^P \left(\frac{n_U^B}{n_W^B} \right) \rangle$$

Can decompose into “group” contributions using Voronoi analysis

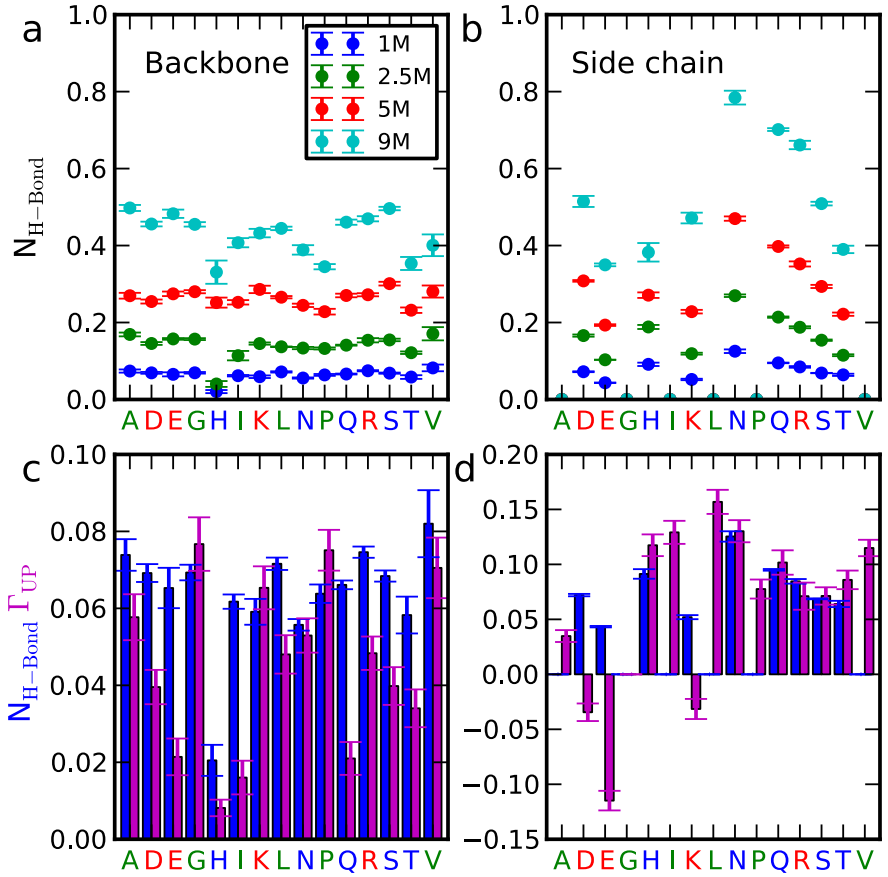
Relation to transfer free energies:

$$\Delta\mu_{tr} \approx -RT\Gamma_{UP}$$



Denaturation mechanism

Hydrogen bonding plays dominant role, especially for backbone



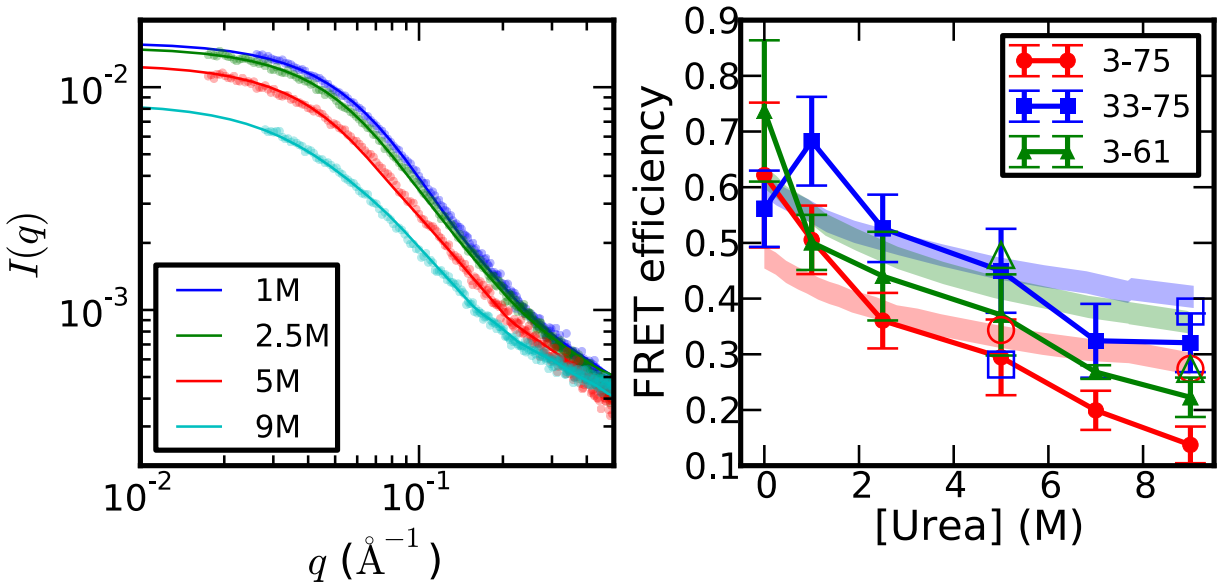
Comparison with experiment

SAXS: all-atom calculation of scattering intensity from pair-distance distribution functions

Köfinger, Hummer, PRE, **87**, 052712 (2013)

FRET: calculated (initially) from CA-CA distance of labelled residues using Förster theory, accounting for dye linker length

T. Förster, Ann. Phys, **6**, 55 (1948)



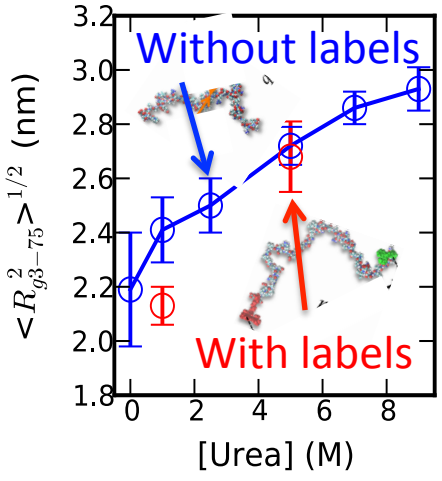
$$E = \left\langle \frac{1}{1+(\frac{R}{R_0})^6} \right\rangle$$

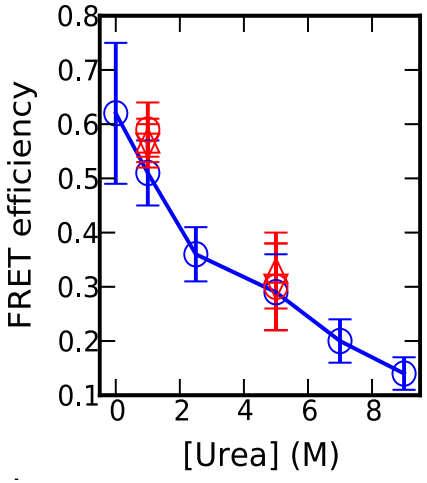
Comparison with experiment

Simulation	FRET χ^2	Experiment *0 M*	*1.0 M*	*2.5 M*	*5.0 M*	*7.0 M*	*9.0 M*
	0.0 M	**0.92**	1.44	2.81	5.02	6.51	7.95
	1.0 M	1.89	1.67	3.25	6.67	9.18	11.7
	2.5 M	3.75	**1.16**	**0.27**	1.07	2.16	3.45
	5.0 M	7.26	3.65	1.29	**0.28**	**0.26**	0.52
	5.0 M (L)	37.9	24.4	13.6	6.19	3.55	2.08
	7.0 M	95.5	63.2	36.8	17.8	10.6	6.06
	9.0 M	66.3	46.2	29.7	17.2	12.1	8.43
	9.0 M (L)	8.38	5.16	2.69	1.08	0.56	**0.30**
	SAXS χ^2	*1.0 M*	*2.5 M*	*5.0 M*	*9.0 M*		
	1.0 M	1.59	3.99	5.98	1.90		
	2.5 M	**1.53**	**3.26**	4.79	1.54		
	5.0 M	1.96	4.25	3.94	1.26		
	5.0 M (L)	2.43	5.73	**3.77**	1.17		
	9.0 M	2.71	6.67	3.78	**1.16**		
	9.0 M (L)	3.35	9.97	4.27	1.19		

Experiment agrees best with simulations at similar urea concentration

Is collapse induced by FRET labels?





- R_g same with/without labels

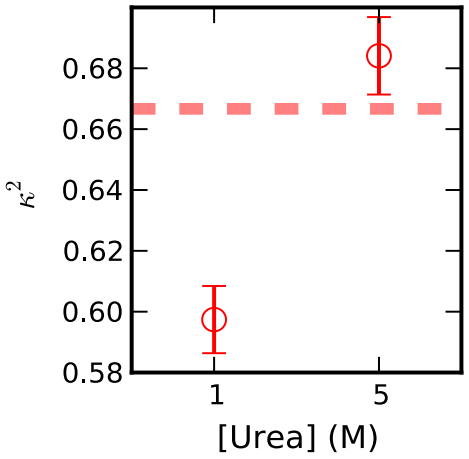
- Same FRET efficiency with explicit labels, and label dynamics

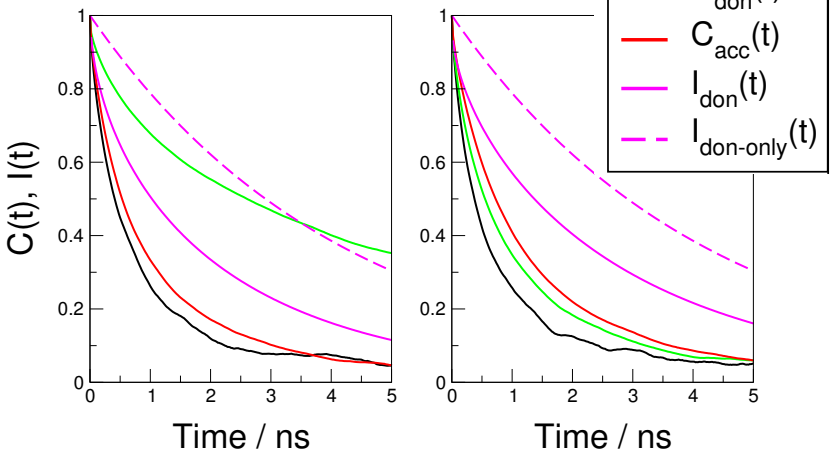
$$k_{ET}(x) = \frac{3}{2} k_D R_0^6 \frac{\kappa^2(x)}{R^6(x)}$$

$$I(t) = \left\langle \exp\left[- \int_0^t (k_D + k_{ET}(t_0 + \tau)) d\tau \right] \right\rangle_{t_0}$$

$$\langle E \rangle = 1 - k_D \int_0^{t_{max}} I(t) dt$$

Best et al, Biophys J 2015

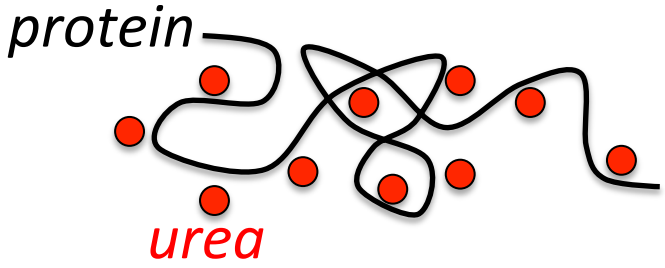




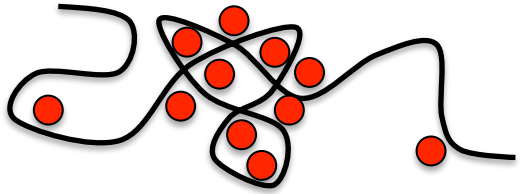
Rapid reorientation of chromophores

$\langle \kappa^2 \rangle \approx 2/3$

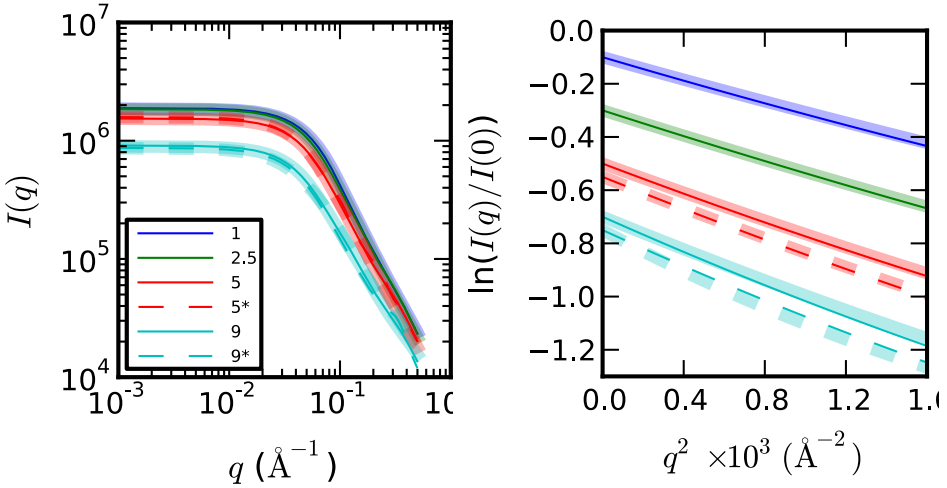
Does denaturant association affect apparent R_g?



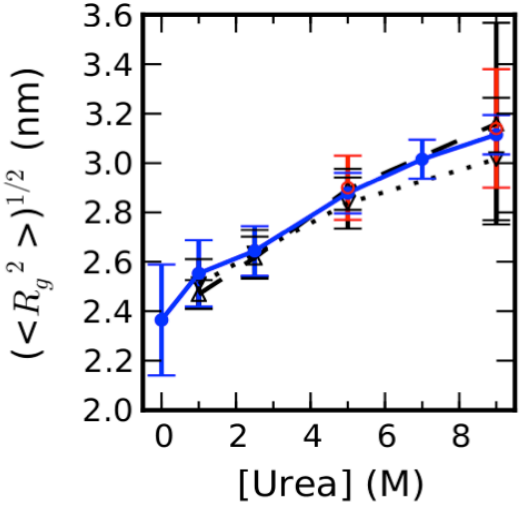
(a) Uniform association



(b) Cooperative binding



R_g from Guinier fit = true R_g



Little difference between
- Explicit solvent model (thick lines)
- Implicit (CRYSOL) solvent model (thin lines)

Part 2 – Conclusions

- An IDP, ACTR, expands as denaturant concentration increases
- This is consistent with both SAXS and FRET data
- FRET chromophores do not appear to *cause* collapse in water
- SAXS Guinier analysis should faithfully report R_g

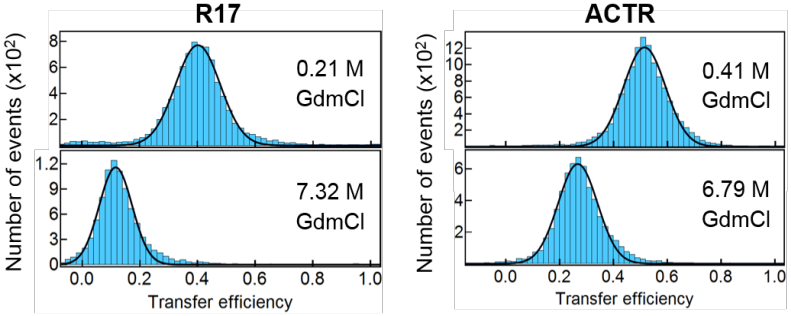
Is there an issue of experimental interpretation? Inverse problem…

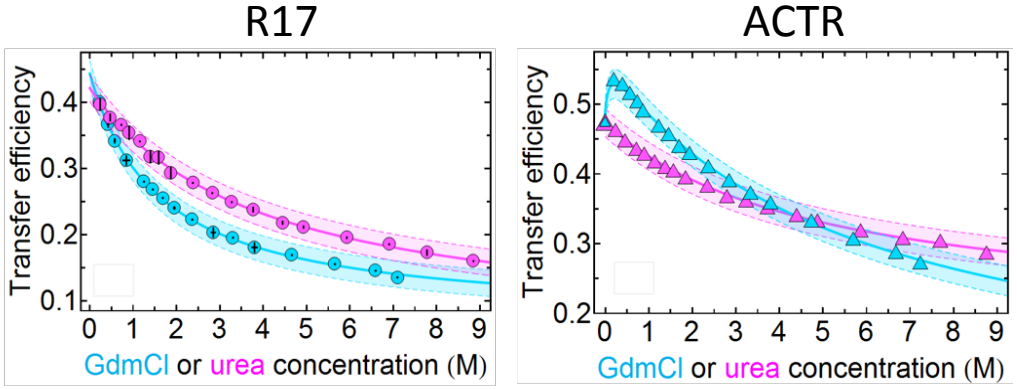
3. Interpretation of experiments in terms of molecular ensembles

Strategy

- Study IDP ACTR, as well as Spectrin R17 mutant by FRET and SAXS using both Urea and GdmCl *under identical experimental conditions*
- Both proteins can be studied at very low denaturant without populating folded state
- Do simplest analysis of each type of experiment to extract R_g
- Fit both FRET and SAXS data jointly to a molecular ensemble

FRET





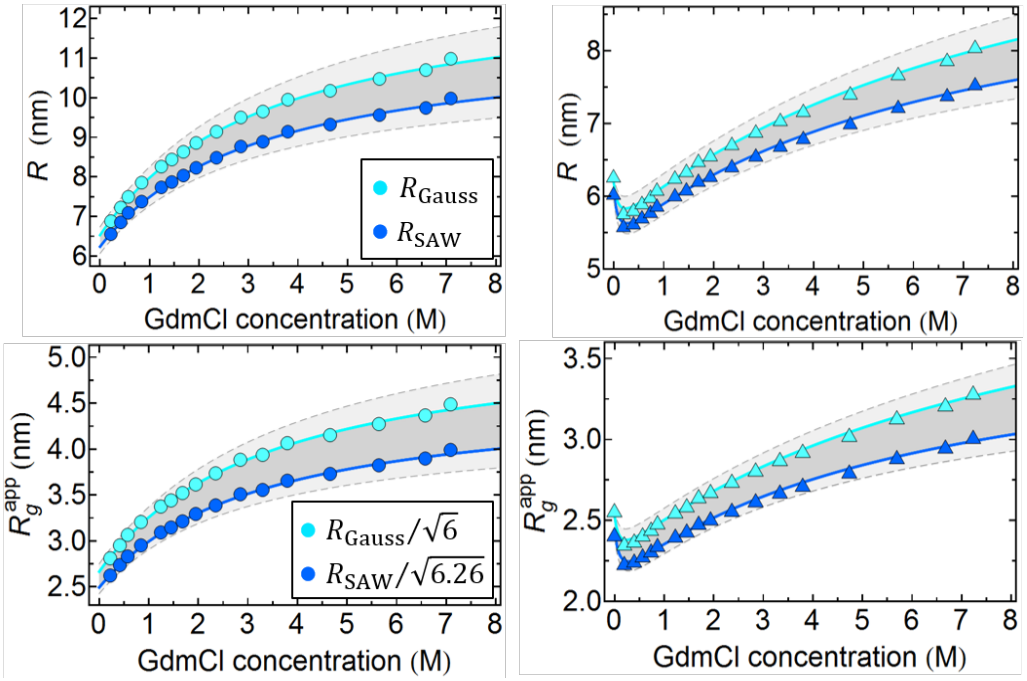
FRET efficiency decreases with [denaturant], suggesting expansion

Inferred R and R_g increase with [denaturant]

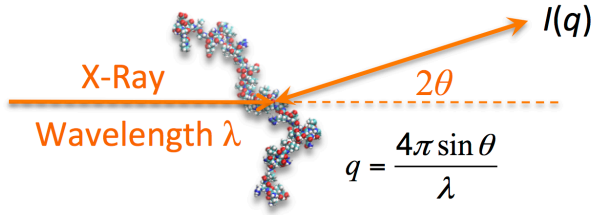
$$\langle E \rangle = \int_0^\infty E(r)P(r)\,dr$$

$$E(r) = 1/(1 + r^6/R_0^6)$$

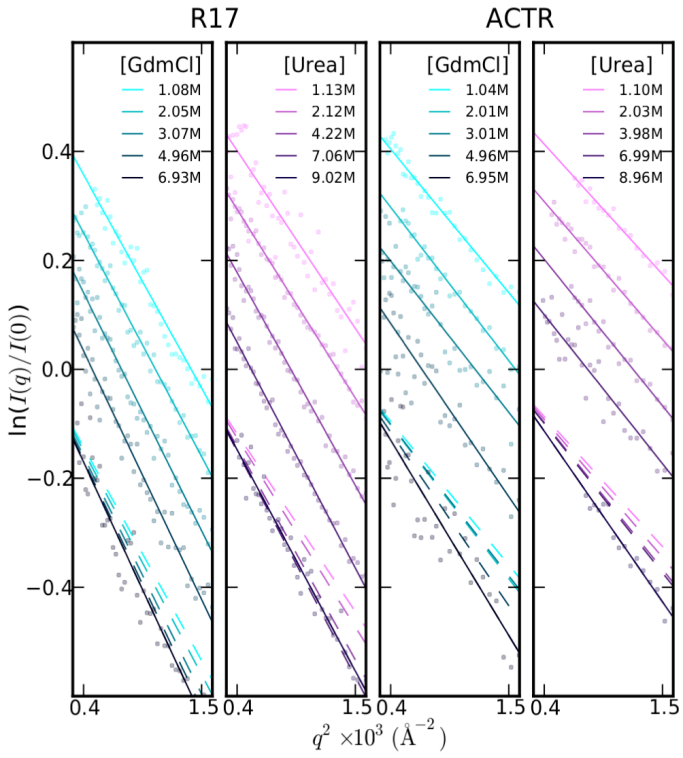
$$R = \langle r^2 \rangle^{1/2}$$



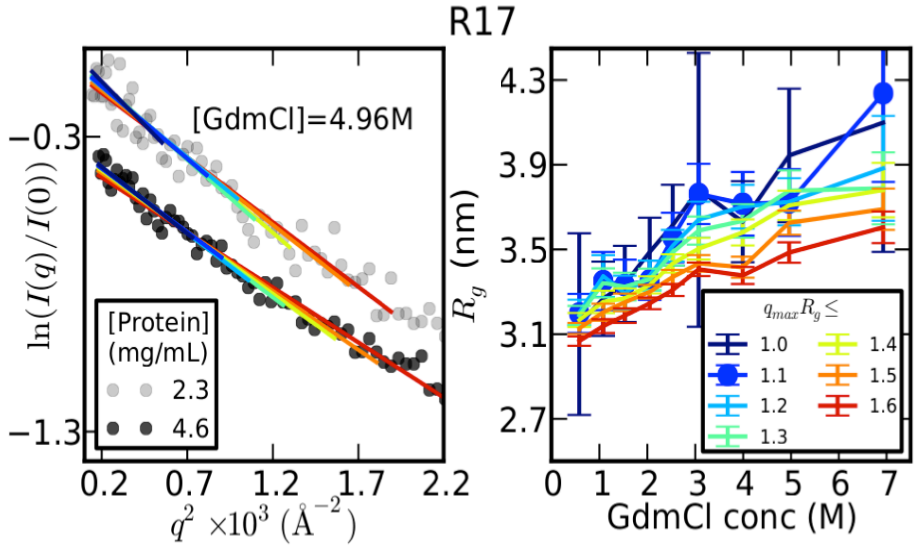
SAXS



$$I(q) \propto \exp\left[-\frac{q^2 r_g^2}{3}\right]$$ Guinier analysis



Guinier fits show expansion



Strong sensitivity of Guinier fit to fitted range

- For folded proteins $q_{max}R_g < 1.3$ is safe
- For IDPs, q_{max} needs to be much less, no more than 1.1

Ensemble fitting

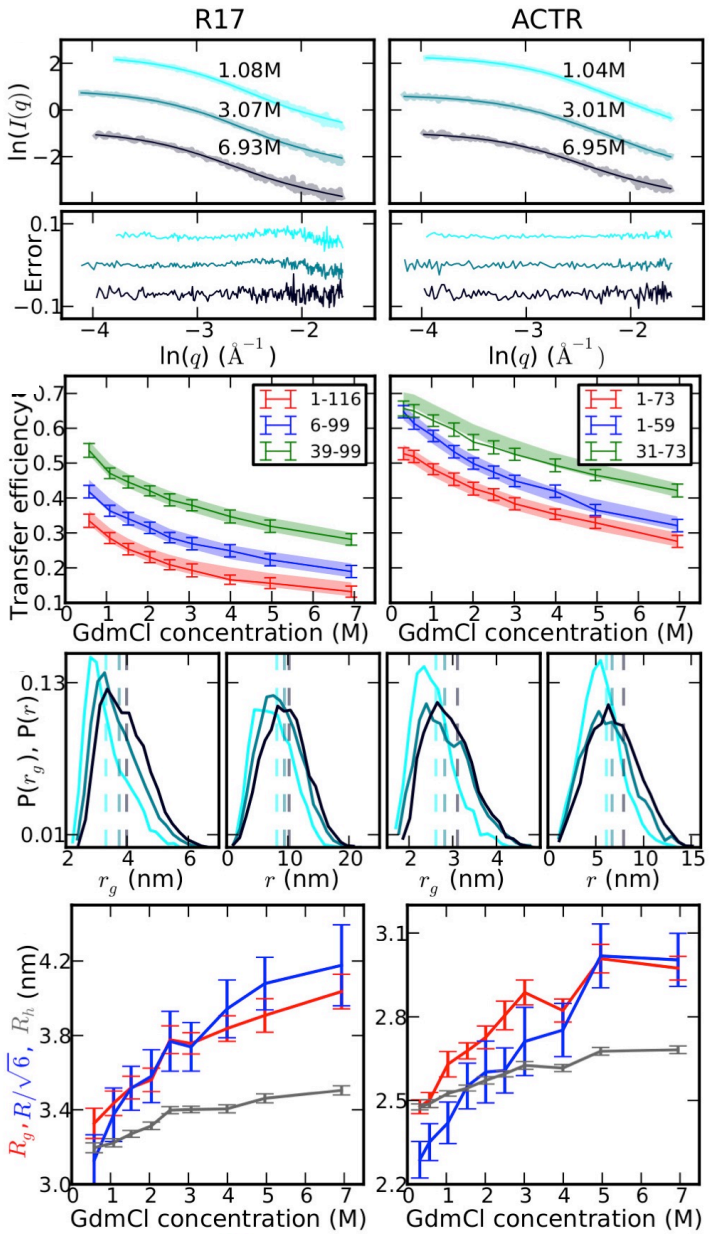
- Rather than assuming a polymer model (FRET) or Guinier approximation (SAXS) to get R_g, use an **explicit molecular ensemble** (e.g. generated by simulation)
- Here we use ABSINTH implicit solvent model generate initial ensemble $\{x_1, x_2, x_3, \ldots, x_N\}$ using REMC to sample. Vitalis, Pappu, J. Comput. Chem. **30**, 673 (2008)
- Reweight the ensemble to fit data (minimize χ^2) by assigning weights $\{w_1, w_2, w_3, \ldots, w_N\}$
- Avoid overfitting by "regularization"

$$G(\{w_i\}) = 0.5\chi^2 - T_{\text{fit}} S(\{w_i\})$$

$$S(w_i) = -\sum_i w_i \ln w_i$$

Köfinger and Hummer, JCP, **143**, 243150 (2015)

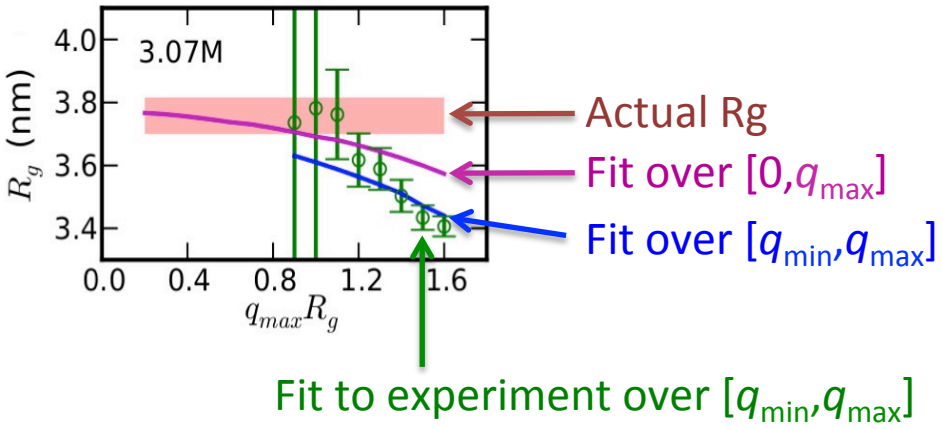
Other groups working on "Ensemble" methods: Vendruscolo, Weare, Cavalli, Head-Gordon, Boomsma



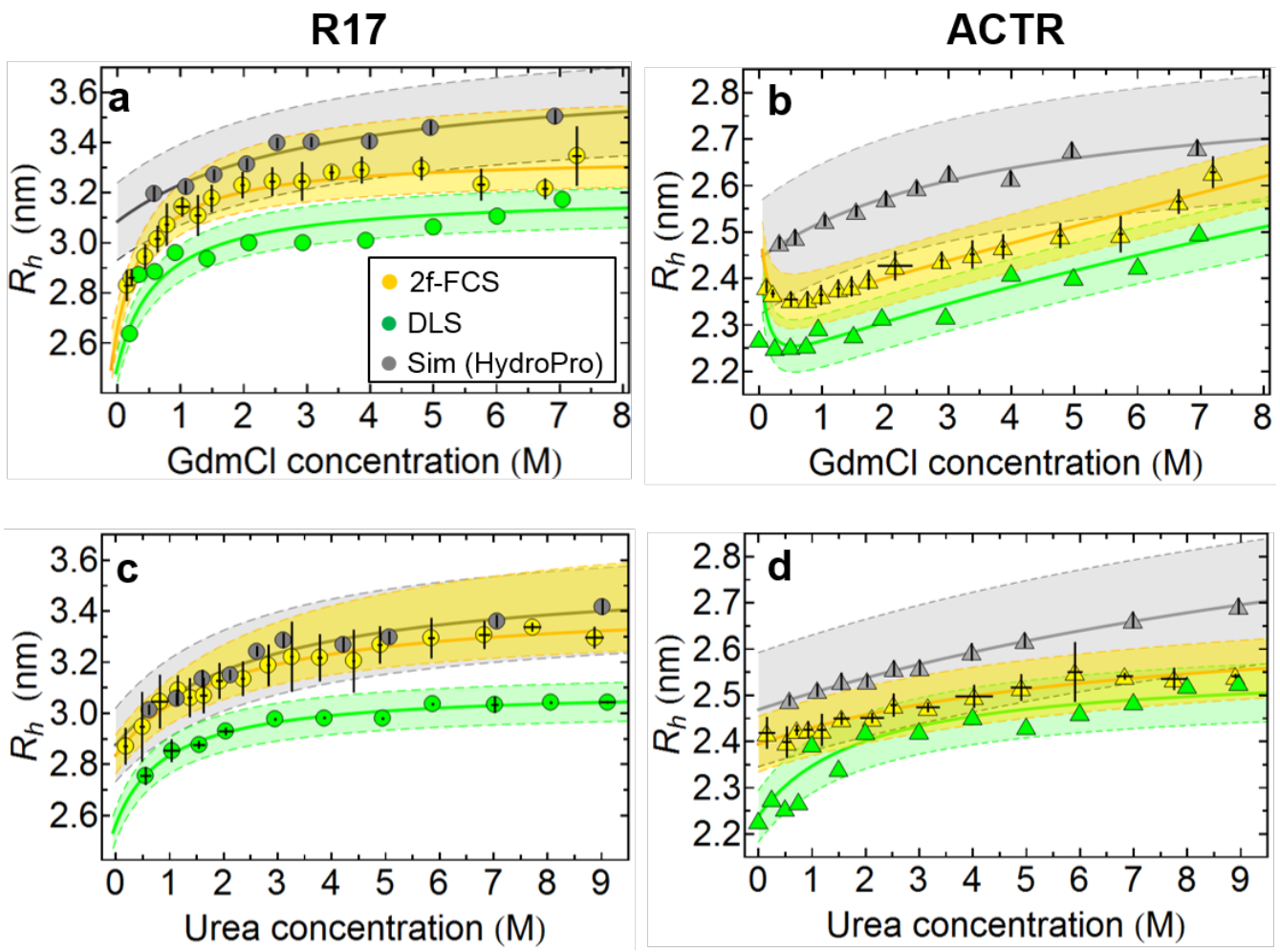
Ensemble fits

Ensemble fits also suggest expansion

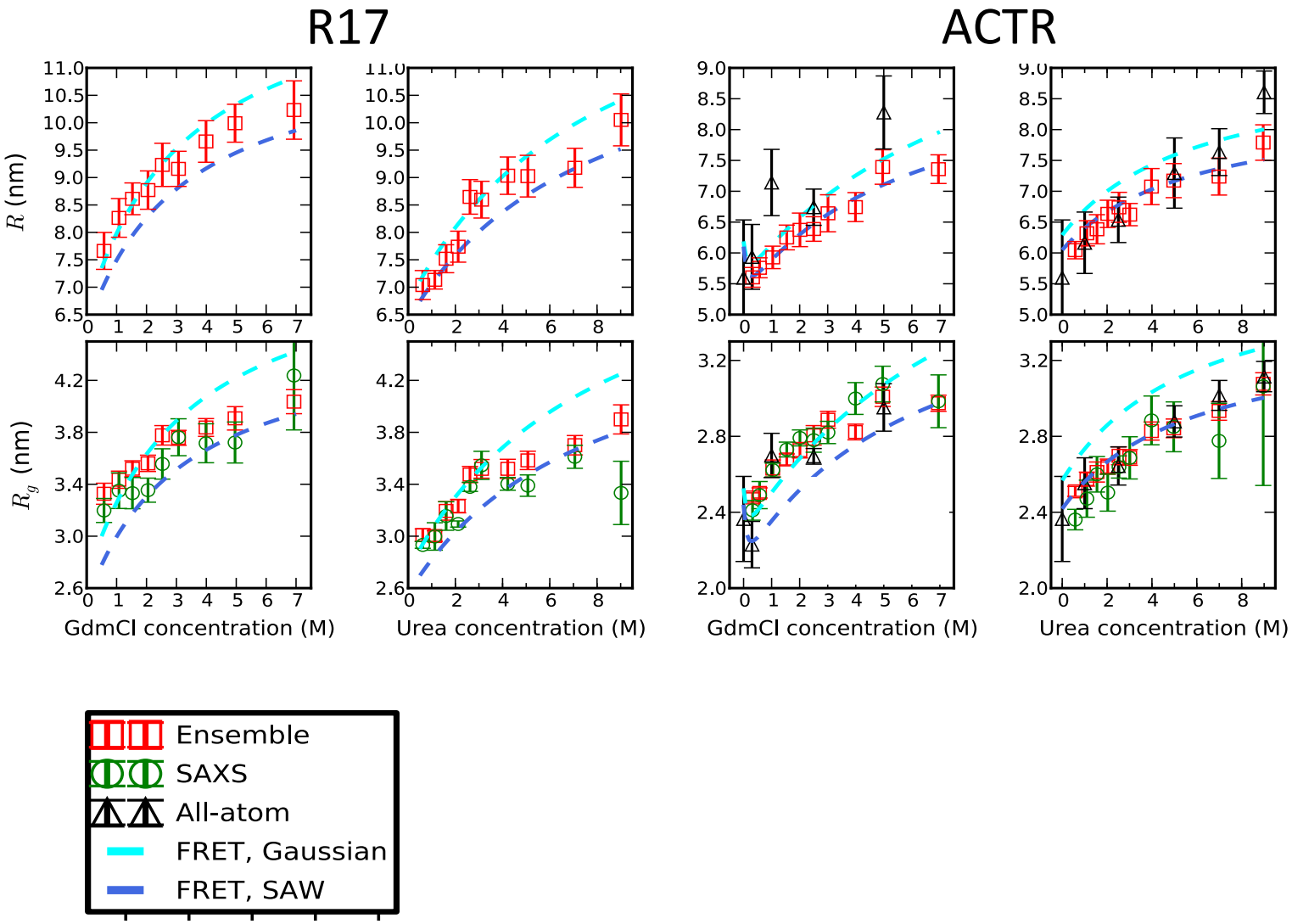
We can use the molecular ensembles to test Guinier analysis (no noise)



Validation – Hydrodynamic Radii



Summary


R17
ACTR
R (nm)
R_g (nm)
GdmCl concentration (M)
Urea concentration (M)
GdmCl concentration (M)
Urea concentration (M)
Ensemble
SAXS
All-atom
FRET, Gaussian
FRET, SAW

Statistical significance of observed expansion (from Bayesian Information Criterion)

Fit to a constant versus a straight line

|ΔBIC| of 10 or more is highly significant

Protein/Denaturant	*FRET SAW	*FRET Gaussian	*SAXS	*Ensemble	#Ensemble FRET	#Ensemble SAXS
[Denaturant]≥0M						
R17d/GdmCl	-171.0	-174.5	-19.5	-62.7	-38.9	-38.8
R17d/Urea	-173.9	-176.1	-145.5	-163.3	-58.8	-69.6
ACTR/GdmCl	-156.4	-173.9	-69.1	-206.5	-55.8	-113.0
ACTR/Urea	-97.5	-100.5	-23.4	-140.3	-21.0	-85.7
[Denaturant]≥3M						
R17d/GdmCl	-2.3	-2.5	**1.1**	-5.6	**1.1**	**-1.4**
R17d/Urea	-11.8	-12.2	**0.8**	-9.0	-7.0	-3.5
ACTR/GdmCl	-9.2	-11.1	-2.3	-3.2	-4.7	**0.7**
ACTR/Urea	-4.9	-5.2	**1.1**	-29.8	-2.8	-22.9

Boldface: Statistically Insignificant ΔBIC

Part 3 Conclusions

- Both FRET and SAXS indicate expansion for R17 and ACTR
- Likely origins of earlier discrepancy:

  - Standard FRET analysis using Gaussian chain may overestimate change of R_g

  - Guinier fits to obtain R_g for IDPs are very challenging and not recommended

  - Most of the expansion occurs at high [denaturant] and would therefore be difficult to observe for stable foldable proteins in SAXS

  - Ensemble Fits may be a more “robust” analysis

Part 3 Conclusions

- Both FRET and SAXS indicate expansion for R17 and ACTR
- Likely origins of earlier discrepancy:

  - Standard FRET analysis using Gaussian chain may overestimate change of R_g

  - Guinier fits to obtain R_g for IDPs are very challenging and not recommended

  - Most of the expansion occurs at high [denaturant] and would therefore be difficult to observe for stable foldable proteins in SAXS

  - Ensemble Fits may be a more “robust” analysis

Acknowledgements

Group:

- Wenwei Zheng (NIH)

Collaborators:

- Gül Zerze, Jeetain Mittal (Lehigh University)
- Alessandro Borgia, Madeleine Borgia, Ben Schuler (University of Zürich) – smFRET
- Alex Grishaev (NIST) – SAXS
- Gerhard Hummer (NIH; now Max Planck for Biophysics)
- Magnus Kjaergaard, Birthe Kragelund (University of Copenhagen) – SAXS