Take-Home Message

Cells often use multiple weak forces instead of one strong one

Indirect forces influence cellular behaviour

Reversibility is a sign of health; irreversibility a sign of disease

Forces and compartmentalisation

How to classify forces relevant in cells? Direct versus Indirect

Direct: Gravity, Coulomb, Dipole, VdW, H-bonds, Covalent and ionic bonds

- Operate between two molecules "independently" of others ⇒ pairwise additive
- Typically short-ranged (not gravity! or Coulomb force in a vacuum)

Indirect: Hydrophobic effect, Depletion, Capillary forces, Membrane-mediated, Fluctuation-induced forces

- Arise between many molecules simultaneously, strongly affected by other objects ⇒ NOT pairwise additive
- Can be long ranged (~ 1/R)
- Proportional to temperature (so pressure is *entropic*… PV = RT)

Direct forces

Gravitational force between two point masses:

$F = G M_1 M_2 / R^2$ $\quad G = 6.67\ 10^{-11}\ N.m^2/Kg^2$

Nearly negligible for cells, but they can sink under gravity

Coulomb force between charges in vacuum

$F = k Q_1 Q_2 / R^2$ $\quad k = 1 / 4\pi\varepsilon_0 = 9\ 10^9\ N.m^2 / C$

in a material ε_0 is replaced by $\varepsilon_0\,\varepsilon$, where ε is the relative permittivity of the material, and makes a big difference between water ($\varepsilon \sim 80$) and oil ($\varepsilon \sim 1$).

Bare force is usually screened out by ions under physiological conditions

Screened Coulomb force

$F = e^{-\kappa R} / R$ $\quad \kappa^{-1}$ = screening length ~ 1 nm in physiol. conditions

$F_G / F_C \sim G M_1 M_2 / k Q_1 Q_2 \sim 10^{-10}\ (10^{-27})^2 / 10^{10}\ (10^{-19})^2 \sim 10^{-36}$

so gravitational force between bare ions is utterly negligible.

Chemical and physical bonds

Bond Type	Strength (kJ/Mol)	Strength (k_BT/bond)	Length (nm)	Description
Covalent	500	200	0.154 (C-C)	Shared outer e^-
Ionic	~ 880	~ 355	0.276 (NaCl nn)	e^- donated/summed
Hydrogen	10 - 40	4 - 16	~ 0.176 (O…H)	small H - electroneg. atom
"Van der Waals"	~ 1	~ 0.4	$1/R^6$	fluctuating induced dipoles

NB 1 kJ/Mol ~ 0.4 kBT per particle

J. Israelachvili, Intermolecular and Surface Forces, Academic Press, 2nd ed. London 1992.

If Strength ~ k_BT, the force has no effect as thermal noise overwhelms it:

- Covalent/ionic bonds cannot be broken by (room) temperature (in vacuum)
- H-bonds may be broken by a large fluctuation
- VDW must occur as many bonds to have any effect

Indirect forces

Indirect forces originate in Coulomb's law (electrostatics) but depend on the temperature, and usually involve many weak interactions supporting each other rather than one dominant strong interaction.

This makes them hard to calculate, and almost impossible to guess their functional form or even their sign.

Importantly: indirect forces can be "free", i.e., they arise from the thermodynamics of the environment at room temperature and not on the atomic structure of the interacting molecules (cp. polymer not depending on monomer identity).

Entropic forces Today

hydrophobic effect of oily chains in water

depletion (molecular crowding)

fluctuation-induced forces

Membrane-mediated forces

curvature

composition

thickness mismatch

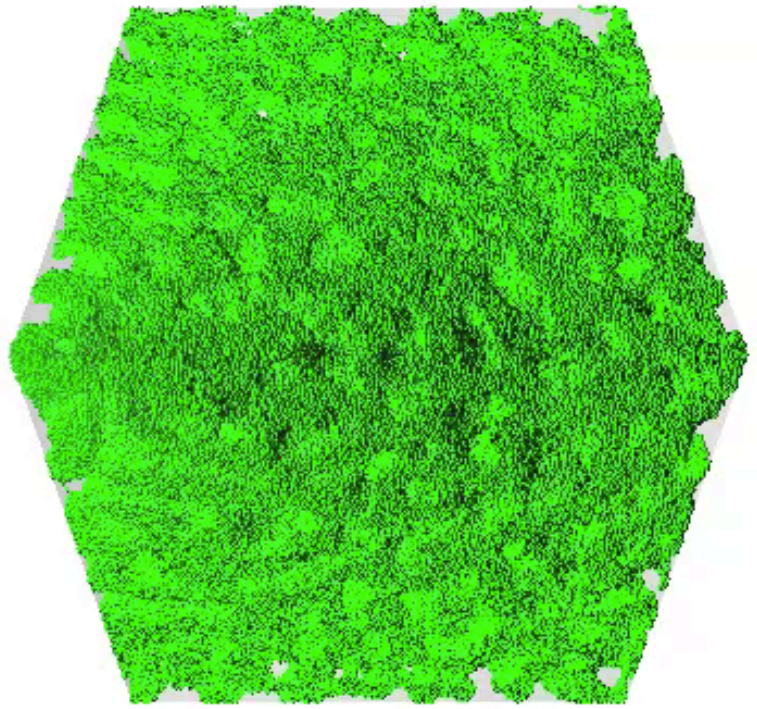
Hydrophobic effect

Water - an homogenous liquid with no bulk structure beyond ~ 1nm, that possesses the unique property of a hydrogen bonded network between the H_2O molecules.

This network has a huge entropy as the molecules continually make and break H-bonds.

The hydrophobic effect results from the prevention of H-bonding between water by the presence of non-polar molecules that drives them to aggregate in one place.

C. Tanford, The Hydrophobic Effect, Wiley, New York 1980



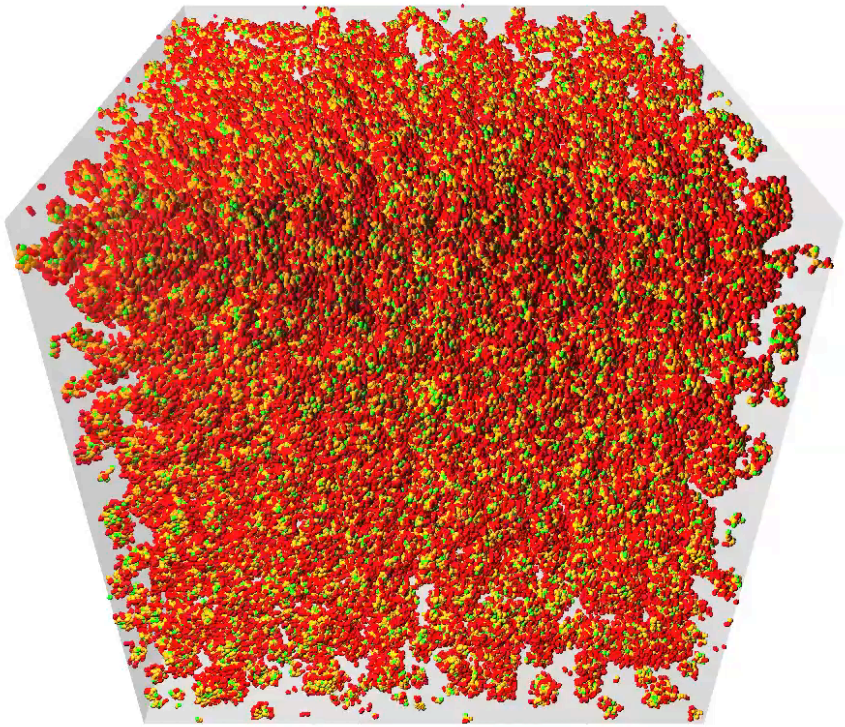
NB. Movies produced from DPD simulations and visualized with Povray and Quicktime. Water in the simulation box is invisible for clarity.

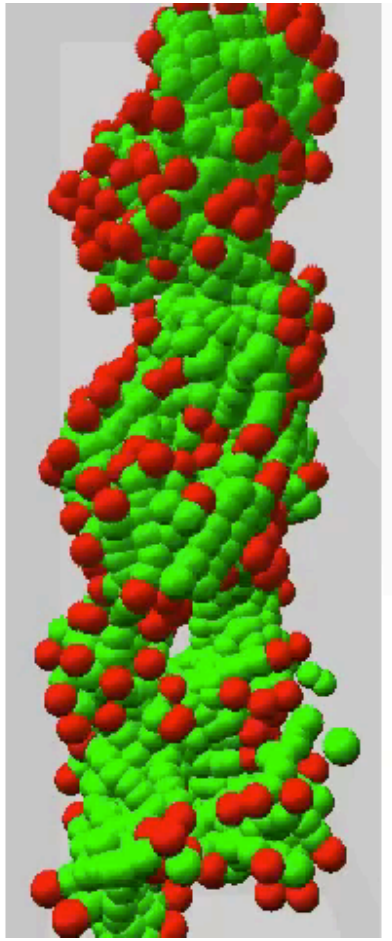
Mix oil and water ⇒ droplets appear by phase separation and create a (thin) bounding surface between the phases.

Amphiphiles in water form complex aggregates

Mix lipids and water ⇒ interfaces and *compartments* appear with interfacial width ~ lipid length

Compartments support gradients, and gradients can be used to do work, e.g., ion concentrations can differ across the neuronal plasma membrane.





Aggregate type is encoded in the molecular shape: no external control is needed: we expect that simulations will be useful in predicting lipid phases if we can capture their amphiphilic nature.

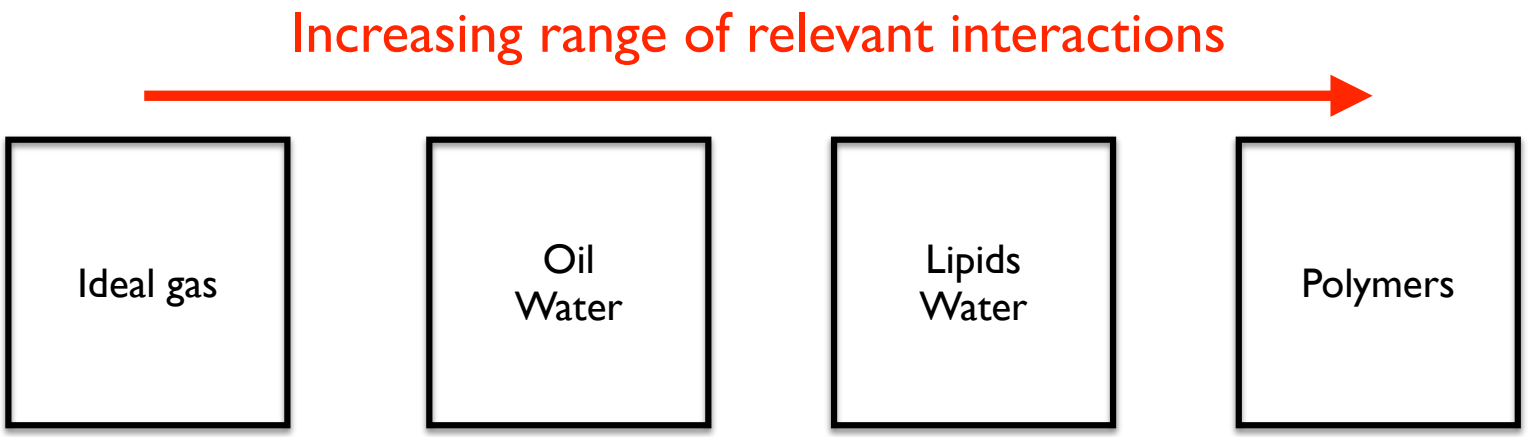
Proteins self-assemble into complex aggregates

- Oil in water can only form spheres with a smooth interface ~ size of oil molecule
- Lipids form aggregates with one dimension comparable to their size ~ 2 nm (micelles, planar bilayers, vesicles, cubic structures of bilayers, etc.)

What structures could emerge from the self-assembly of *large* molecules?

Folded proteins are one type of self-assembled aggregate; are there others?

Ask yourself: what is the range of the interactions driving the self-assembly?



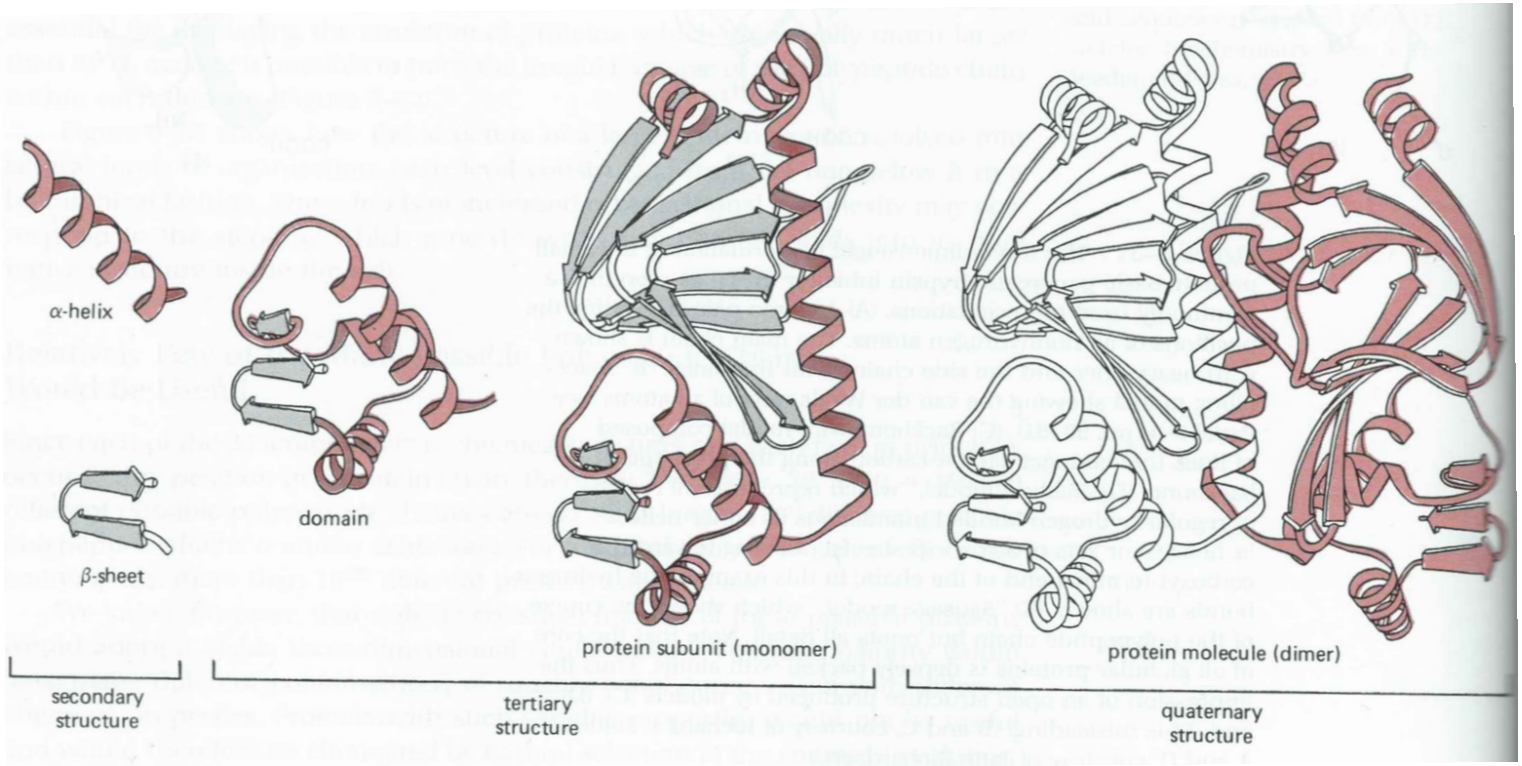
3D structure of proteins

Primary structure = sequence of amino acids

Secondary structure = H-bonding of contiguous aa's into α helices and β sheets

Tertiary structure = domains of globular units

Protein complexes = assembly of several proteins by non-covalent bonds

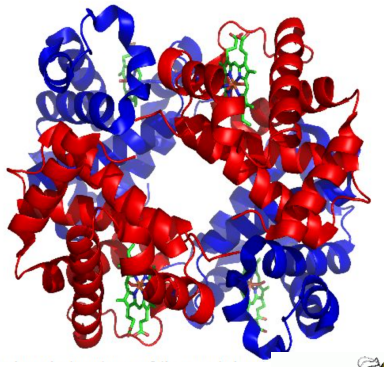


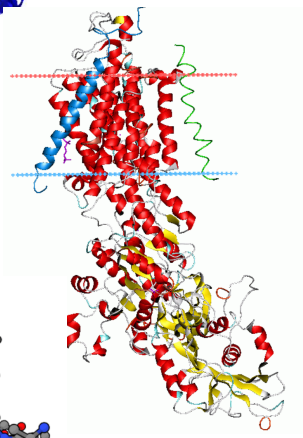
Alberts and Bray et al.

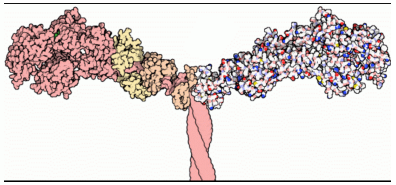
Folded proteins are molecular machines

What are some typical proteins?

- Haemoglobin - binds oxygen
- Actin - forms filaments, cytoskeleton
- Myosin motor - pulls cargo along a filament
- Na-K ATPase - pump ions against gradient
- Ion channels - allow ions to flow down gradient
- Respiratory chain proteins - makes ATP using energy of e- transport

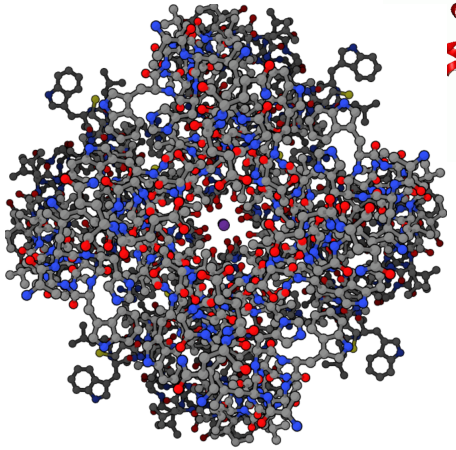






wikipedia.org

Machines because they are precisely arranged (mutate one aa and it can destroy the protein) and they operate relatively independently of other proteins



Proteins fold into precise, unique, lowest-energy states that are known from crystal structures: but is this always true of all proteins?

Are all proteins folded?

Do all proteins fold into lowest-energy states? ~25 years ago answer would have been yes.

Common picture of proteins was precisely-folded molecular *machines* that made lock-and-key binding with ligands or other proteins. Specific sequences of amino acids determined who bonded to whom, and that was it.

Very mechanical - very Newtonian - not thermodynamically accurate

There are other types of protein - *intrinsically-disordered proteins (IDP)*- that have no stable structure because of their amino acid sequence, but do exhibit biological activity, and play a fundamental role in organising cellular dynamics.

Examples

- Casein in milk, disordered structure makes digestion easier
- Calcineurin activates immune T cells on Ca binding, needs accessible binding sites, and surrounds its binding partner
- PSD-95, scaffold protein in neuronal spines, binds membrane receptors

Opposite extremes of protein models

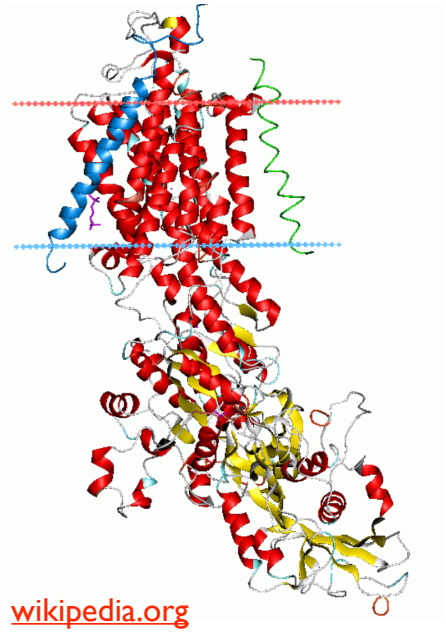
EPFL

Folded protein

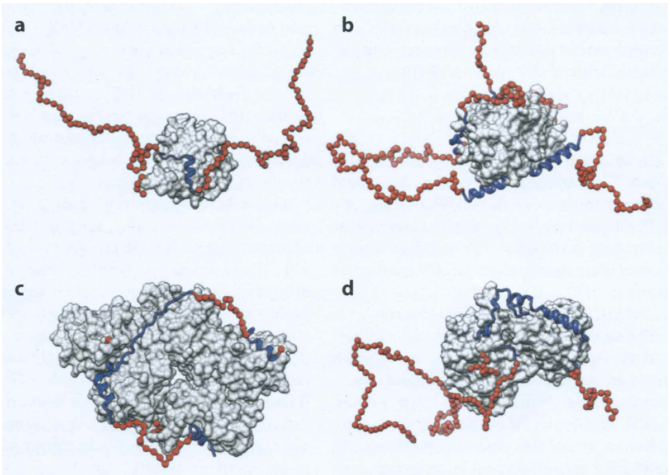
- Unique folded state
- Lowest energy (energy dominated)
- Precise shape
- Precise functions
- Disrupted by single aa mutation
- No model, need the actual protein

Intrinsically Disordered protein

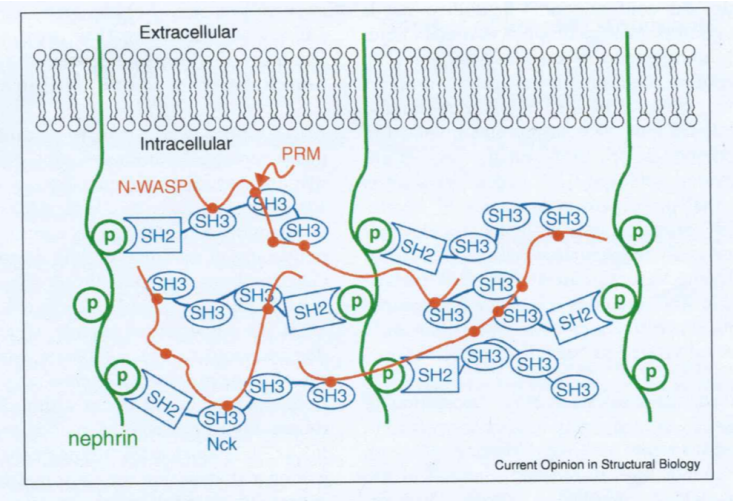
- No unique folded state
- Many conformations of similar energy (entropy dominated)
- Generic binding via multiple, weak sites
- Sequence not conserved but properties are
- Can fold/unfold on binding
- Model it as a phantom chain?



wikipedia.org



Oldfield and Dunker
Ann. Rev. Biochem.l. 83:553 (2014)



Chong and Forman-Kay,
Curr. Op. Struct. Biol. 41:180 (2016)

Intrinsically-Disordered Proteins

IDPs are common and behave very differently to folded proteins - human proteome has
30 - 50% disordered residues (mainly E, K, P, Q, S) *

- IDPs are long, flexible chains of amino acids (just like proteins that fold)
- No unique folded state/no average shape/no mean atom coordinates
- Contain regions or domains of low complexity (repeated amino acids)
- Mechanically floppy, fluctuate due to random thermal motion
- Multiple weak, non-specific binding sites
- Soluble at low concentration (contain few hydrophobic residues)
- Coalesce above a critical concentration into spherical, fluid droplets known as:

Biomolecular Condensates

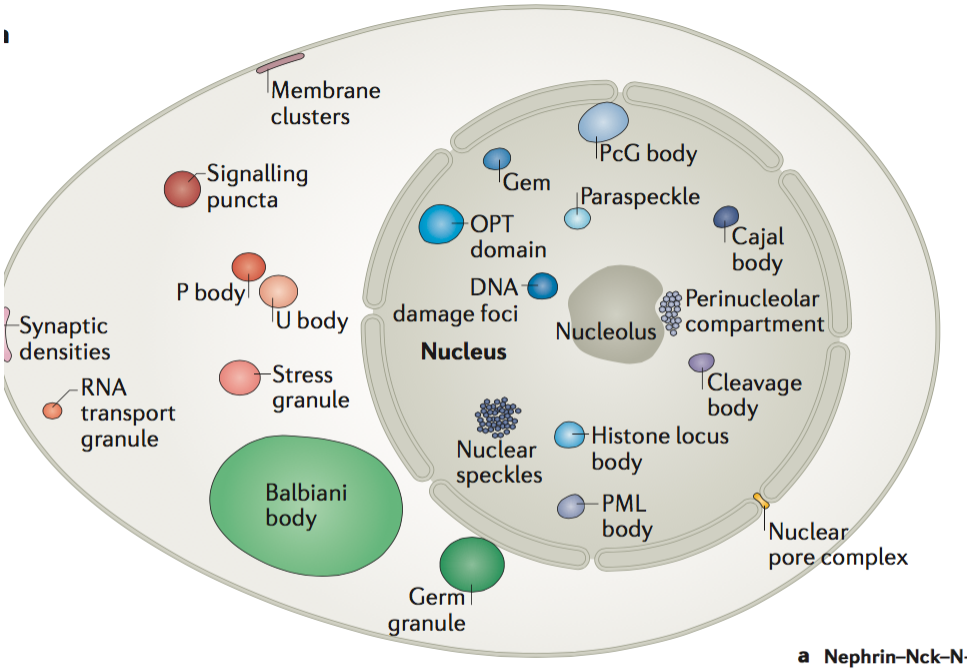
Intracellular Condensates

Membraneless Organelles

Protein droplets

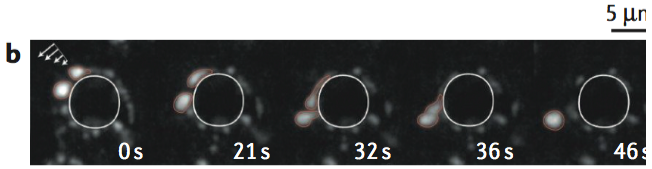
* C. J. Oldfield and A. K. Dunker, Ann. Rev. Biochem. 83:553 (2014)

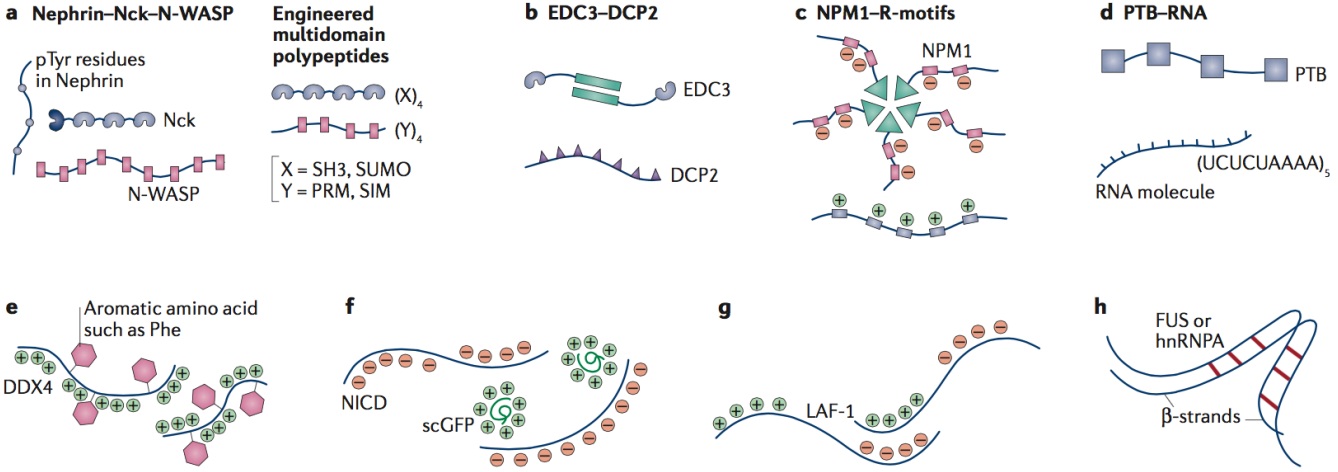
Biomolecular condensates: a new phase of cellular matter?



There's a lot!

All formed from long, flexible proteins (Intrinsically-Disordered Proteins)



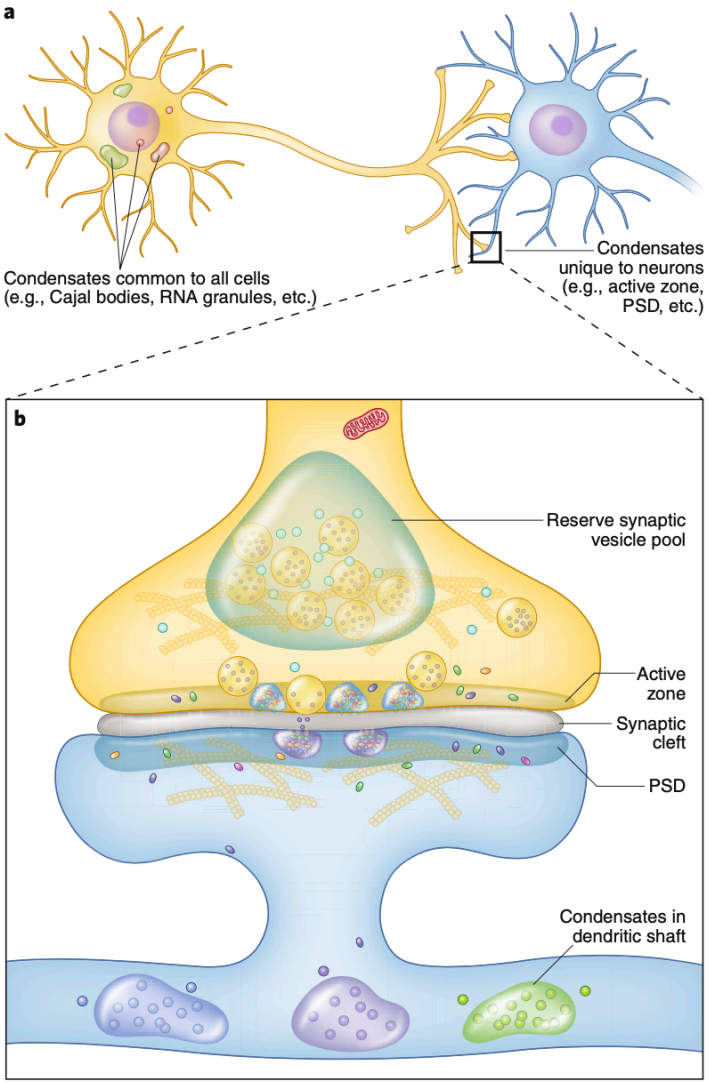


S. F. Banani et al. Biomolecular condensates: organizers of cellular biochemistry, Nat Rev. Mol. Cell Biol. 18:285 (2017)

… prevalent in neurons

Condensates in neurons:

- in the nucleus and cytoplasm[1]
- in axons where they clusters synaptic vesicles[2]
- in dendrites where receptors aggregate and bind neurotransmitter[3]



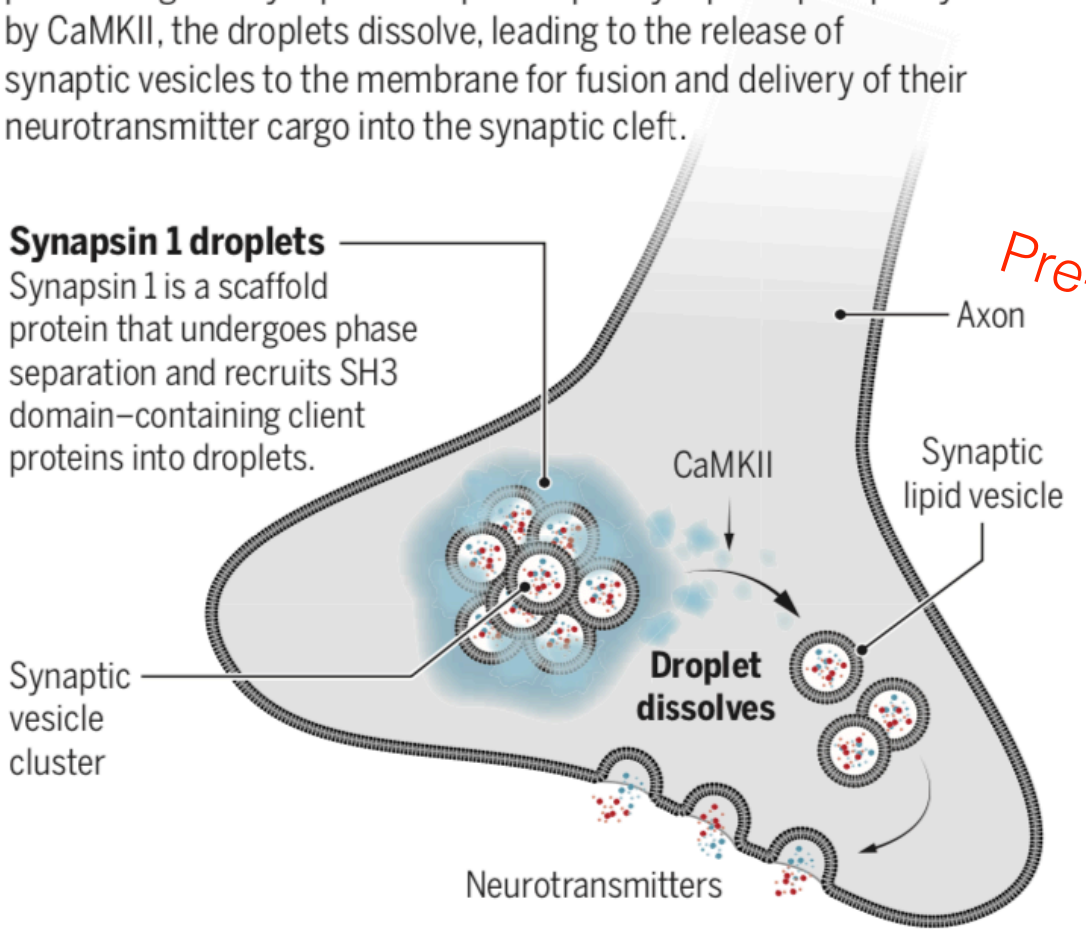
1 Alberti and Hyman, Bioessays 38:959 (2016)

2 Milovanovic et al. Science 361: 604 (2018)
Wu et al. Molecular Cell 73:971 (2019)

3 Chen et al. Nature Neuroscience 23:301 (2020)

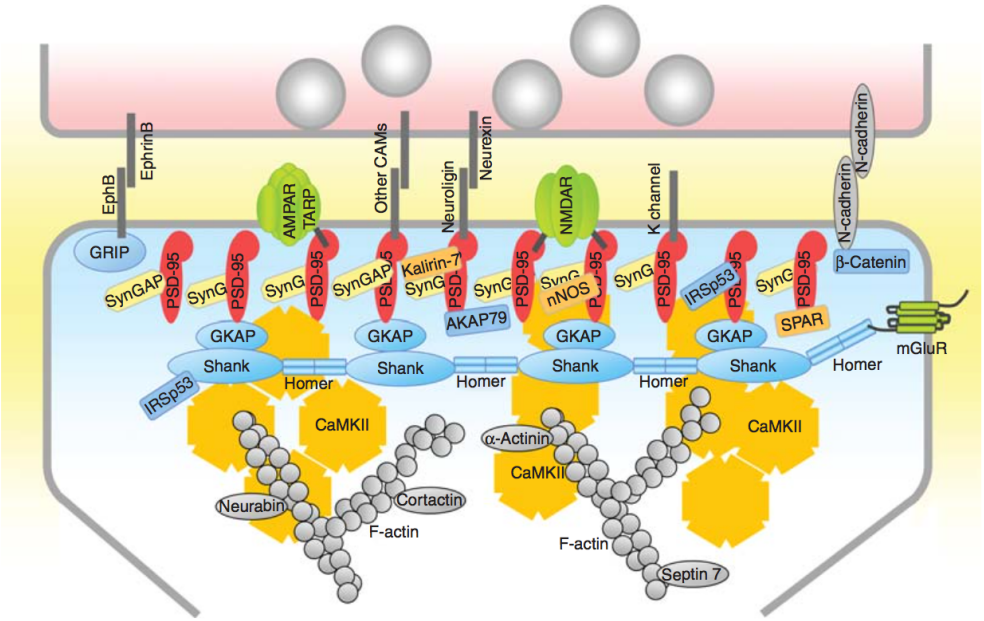
Clustering of synaptic vesicles

Synaptic lipid vesicles are proposed to be clustered together by partitioning into synapsin 1 droplets. Upon synapsin 1 phosphorylation by CaMKII, the droplets dissolve, leading to the release of synaptic vesicles to the membrane for fusion and delivery of their neurotransmitter cargo into the synaptic cleft.

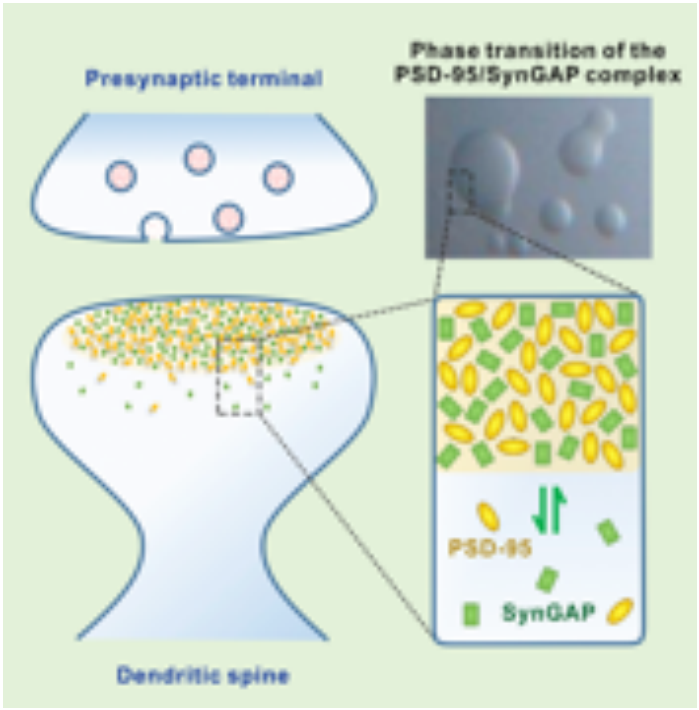


Pre-synaptic phase separation

Boczek and Alberti, Science 361:548 2018



Post-synaptic phase separation



The Postsynaptic Organization of Synapses,
M. Sheng and E. Kim,
Cold Spring Harbour Persp. Biol. 2011;3:a005678

SynGap and PSD-95 phase separate into droplets above a threshold of 100 µM in vitro
(protein concentration inside ~ 100 x bulk)

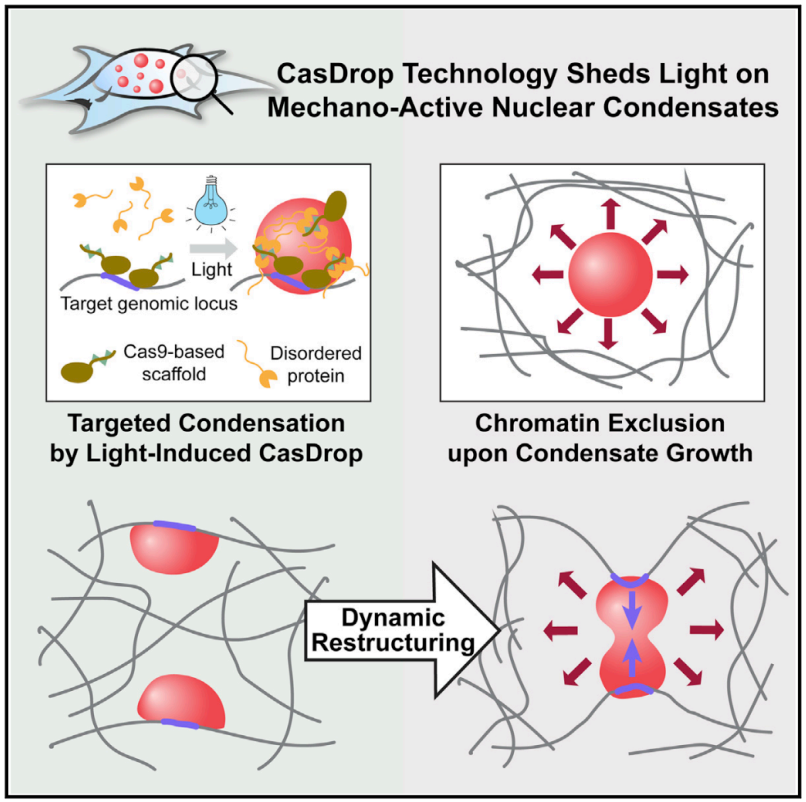
PSD is ~100s nm wide x 30 - 50 nm thick, associated with the post-synaptic membrane at synapses; it has ~ 30-60 protein species and no bounding membrane; biochemical changes underlie learning.

Zeng et al., Cell 166:1163-1175 (2016)

… with unusual mechanical behaviour

Cell

Liquid Nuclear Condensates Mechanically Sense and Restructure the Genome



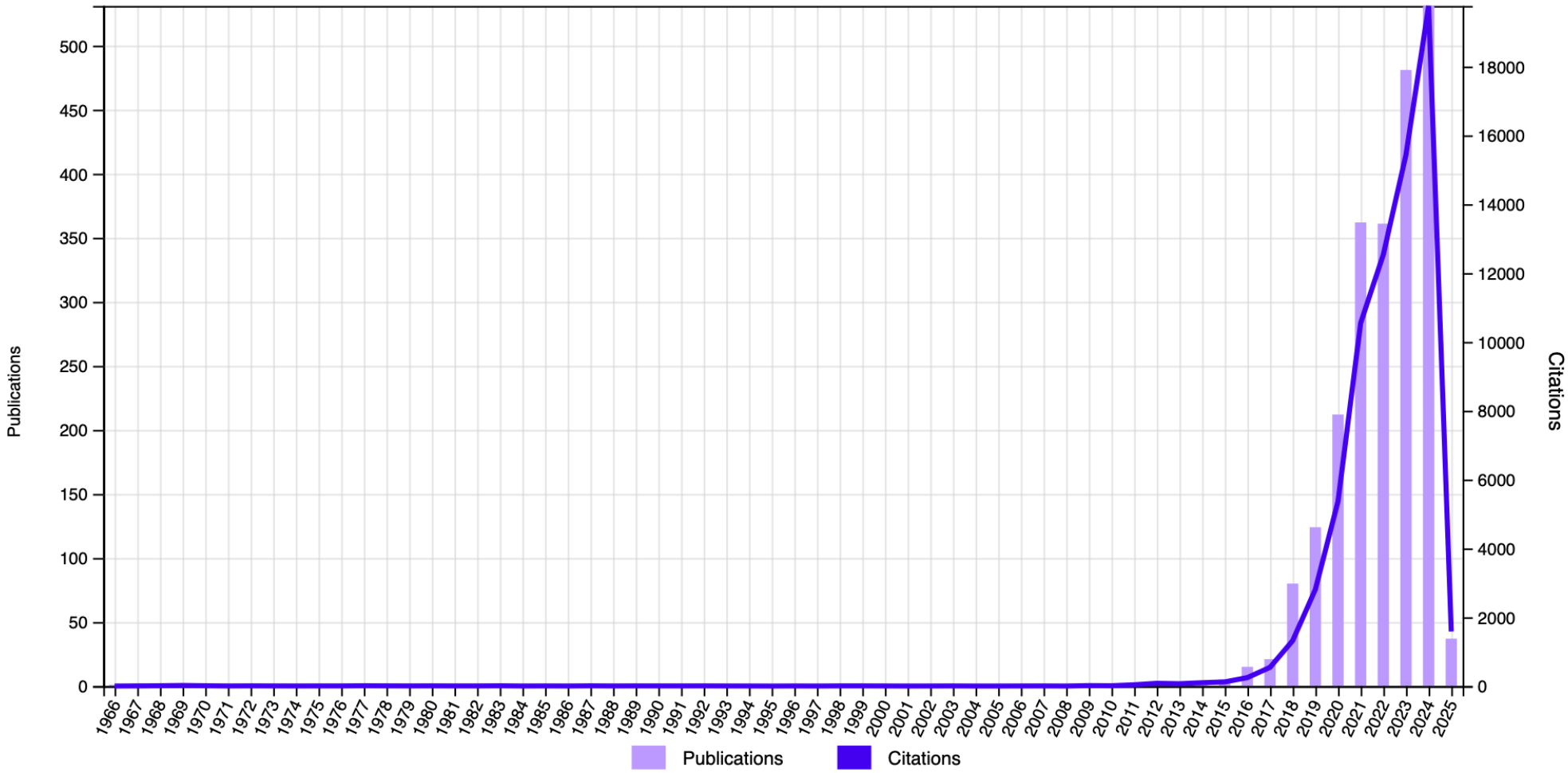
Highlights

- The CasDrop system enables controlled liquid condensation at specific genomic loci
- The IDR-driven condensates grow preferentially in regions of low chromatin density
- Condensate formation leads to mechanical exclusion of non-targeted chromatin
- Condensates pull in targeted genomic loci, serving as mechanical chromatin filters

Shin et al. Cell 175:1481 (2018)

Quote of the Day

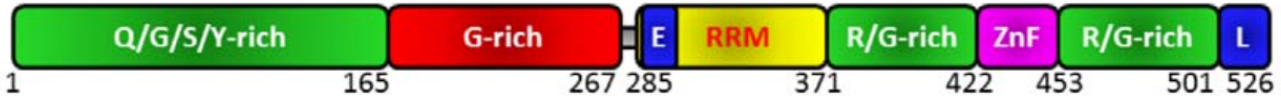
disorder can do things structure can't (K Dunker)



Web of Science search: biomolecular condensate OR membranless organelle

Observations from the literature

- Eukaryotes have more IDRs than prokaryotes; 30-50% of human proteome are disordered regions [1]
- Membraneless organelles compartmentalise biochemistry [5]
- Disease-associated proteins are rich in disorder [1] and loss of fluidity accompanies disease [2]
- FUS, 526 residues: N-term. LC seq (QGSY) - RNA recog, motif; C-term LCD; phase separated FUS is disordered [3]; RRM interacts with membranes [6]

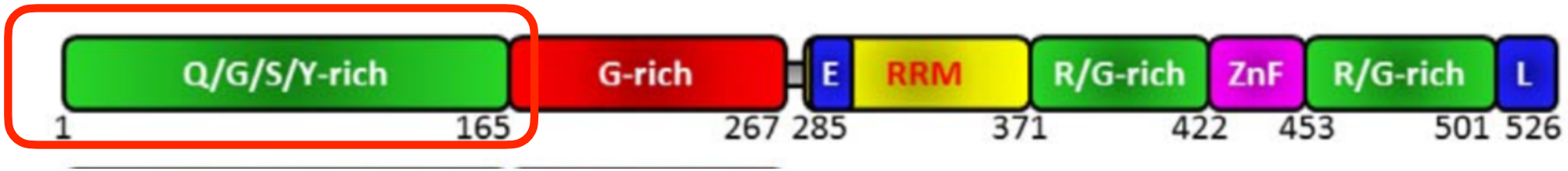


- p53 is disordered at the N, C termini while the central DNA-binding region is ordered; it has tens of binding sites at each terminus (> 40 at N term.) [1]
- BRCA1 has ~40 binding sites in its central IDR, and ~ 400 binding partners [1]
- IDRs are rich in Ser, Arg, Tyr; PTM sites are often located in IDRs [1]
- Replication Protein A has an IDR linker with a *variable* sequence but *conserved rigidity and disorder* [1]
- PSD-95 and SynGap spontaneously phase separate into fluid droplets in vitro [4]

IDR = intrinsically-disordered region
LC = low complexity, repeated amino acids
PTM = post-translational modification

[1] Oldfield and Dunker, Ann. Rev. Biochem. **83**:553 (2014)
[2] Alberti and Hyman, Bioessays **38**:959 (2016)
[3] Burke et al. Molecular Cell **60**:231 (2015)
[4] Zeng et al. Cell **166**:1163 (2016)
[5] Holehouse and Pappu, Biochemistry DOI: 10.1021/acs.biochem.7b01136 (2018)
[6] Lu et al. Bioarxiv **doi:** https://doi.org/10.1101/122671

Fused in Sarcoma (FUS)



low-complexity domain

ASNDYTQQA TQSYGAYPTQ PGQGYSQQSS QPYGQQSYSG YSQSTDTSGY

GQSSYSSYGQ SQNTGYGTQS TPQGYGSTGG YGSSQSSQSS YGQQSSYPGY

GQQPAPSSTS GSYGSSSQSS SYGQPQSGSY SQQPSYGGQQQS YGQQQSYN

PPQGYGQQNQ YNSSSGGGGG

FUS LCD has 37 Q, 43 S, 10 T, **24 Y,** 27 G, 11 P
(only 2 charged D)

Uniform, regular spacing of aromatic residues (e.g., Y) is a conserved feature of IDRs

RNA-binding protein:

- DNA repair
- RNA transcription, splicing, transport
- multiple PTMs at N terminus
- involved in the disease ALS
- phase separates in vitro
  = good model system

Burke et al. Molecular Cell 60:231 (2015)
Murthy et al. Nature Struct. Mol. Biol., 26:637 (2019)
Esteban-Hofer et al. Biophys. J. 123:538 (2024)

“Sticky” Tyrosines are highly conserved in FUS LCD

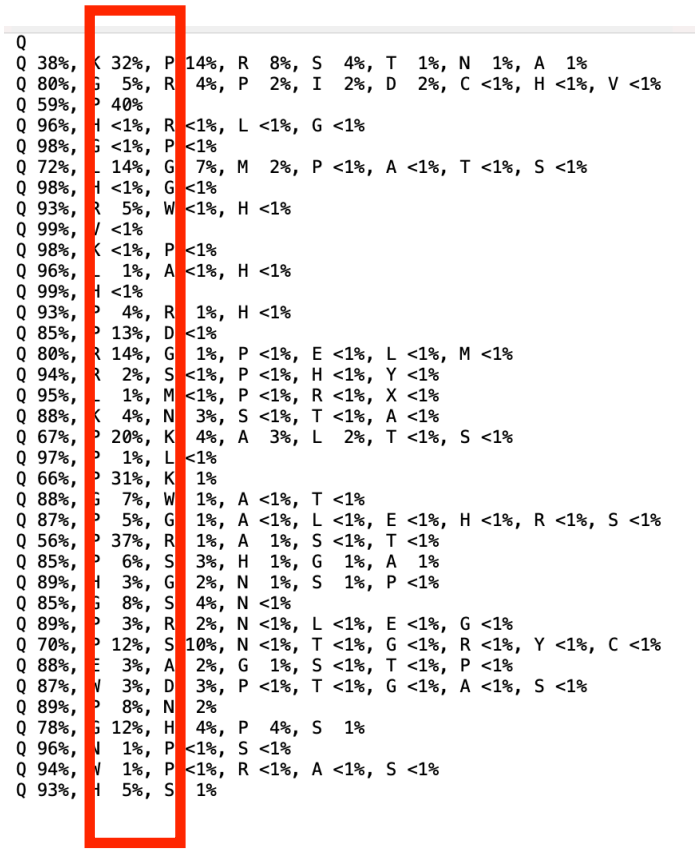

```
shillcoc$ awk 'BEGIN{n=0} {if($2=="Y") {n++; print}} END{print n}' residues.txt
   5       Y     -0.656     7    -0.968, -0.396  8,6       e                99/126       Y 95%, D  2%, H  1%, F  1%
  13       Y     -1.466     9    -1.643, -1.413  9,9       e       f       121/126       Y
  16       Y     -1.155     8    -1.413, -0.968  9,8       e       f       122/126       Y 98%, C <1%, L <1%
  24       Y     -0.817     7    -1.106, -0.549  8,7       e               123/126       Y 96%, F  1%, C <1%, I <1%
  32       Y     -0.968     8    -1.250, -0.693  9,7       e       f       122/126       Y 97%, C <1%, H <1%, A <1%
  37       Y     -1.305     9    -1.643, -1.106  9,8       e       f       121/126       Y 98%, E <1%, X <1%
  40       Y     -1.143     8    -1.413, -0.968  9,8       e       f       122/126       Y 97%, D  1%, A <1%
  49       Y     -0.788     7    -1.106, -0.549  8,7       e               124/126       Y 96%, S <1%, D <1%, R <1%, F <1%
  54       Y     -1.321     9    -1.643, -1.106  9,8       e       f       124/126       Y 99%, L <1%
  57       Y     -1.334     9    -1.643, -1.106  9,8       e       f       122/126       Y 99%, C <1%
  65       Y     -0.952     8    -1.250, -0.693  9,7       e       f       122/126       Y 97%, C <1%, S <1%, V <1%
  74       Y     -0.399     6    -0.832, -0.045  7,5       e               125/126       Y 95%, C <1%, M <1%, W <1%, L <1%, G <1%, I <1%
  80       Y     -0.608     7    -0.968, -0.396  8,6       e               122/126       Y 89%, C  7%, M  1%, F <1%, S <1%
  90       Y     -1.188     8    -1.413, -0.968  9,8       e       f       126/126       Y 98%, F <1%, C <1%
  96       Y     -0.984     8    -1.250, -0.693  9,7       e       f       126/126       Y 97%, A <1%, S <1%, N <1%
  99       Y     -0.590     7    -0.968, -0.230  8,6       e               125/126       Y 95%, C  2%, L <1%, S <1%, R <1%
 112       Y      0.261     5    -0.230,  0.751  6,4       e               119/126       Y 91%, G  1%, H  1%, A <1%, S <1%, W <1%, F <1%, Q <1%, P <1%
 121       Y     -0.595     7    -0.968, -0.230  8,6       e               117/126       Y 94%, C  1%, F  1%, Q <1%, S <1%
 129       Y      0.129     5    -0.396,  0.423  6,4       e               116/126       Y 93%, C  1%, H  1%, A <1%, V <1%, R <1%, G <1%
 135       Y      0.935     3     0.423,  1.232  4,3       e               114/126       Y 85%, C  6%, L  1%, S  1%, G <1%, D <1%, T <1%, A <1%, M <1%, Q <1%
 142       Y     -1.242     9    -1.643, -1.106  9,8       e       f       100/126       Y 99%, F <1%
 148       Y     -0.726     7    -1.106, -0.396  8,6       e               111/126       Y 96%, M <1%, H <1%, L <1%, F <1%
 154       Y     -0.033     5    -0.549,  0.423  7,4       e                97/126       Y 92%, S  3%, C  2%, Q  1%, H  1%
 160       Y     -0.164     5    -0.549,  0.167  7,5       e                97/126       Y 92%, C  2%, S  2%, Q  1%, H  1%, G  1%
```

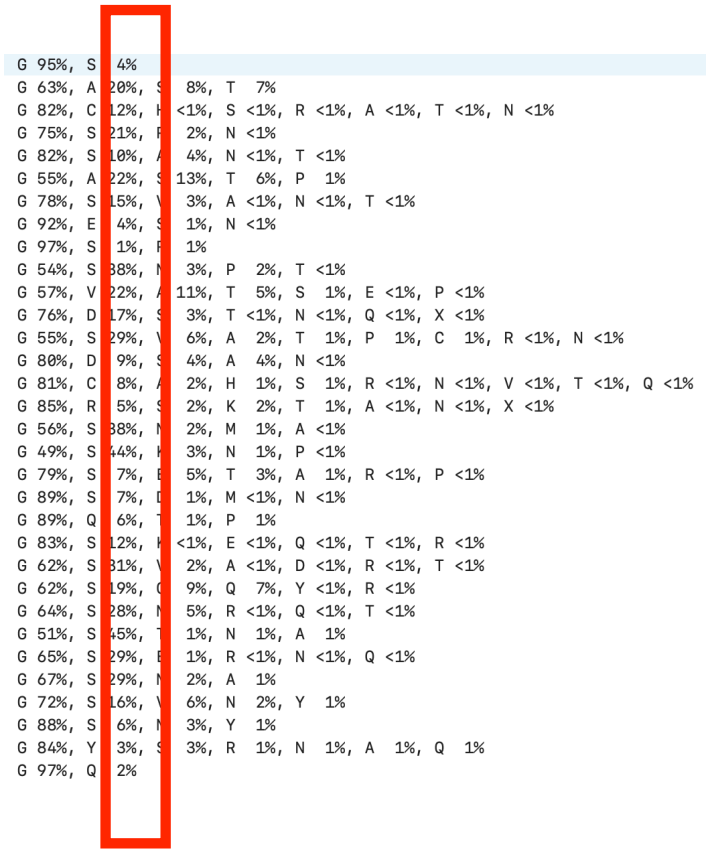
residue #

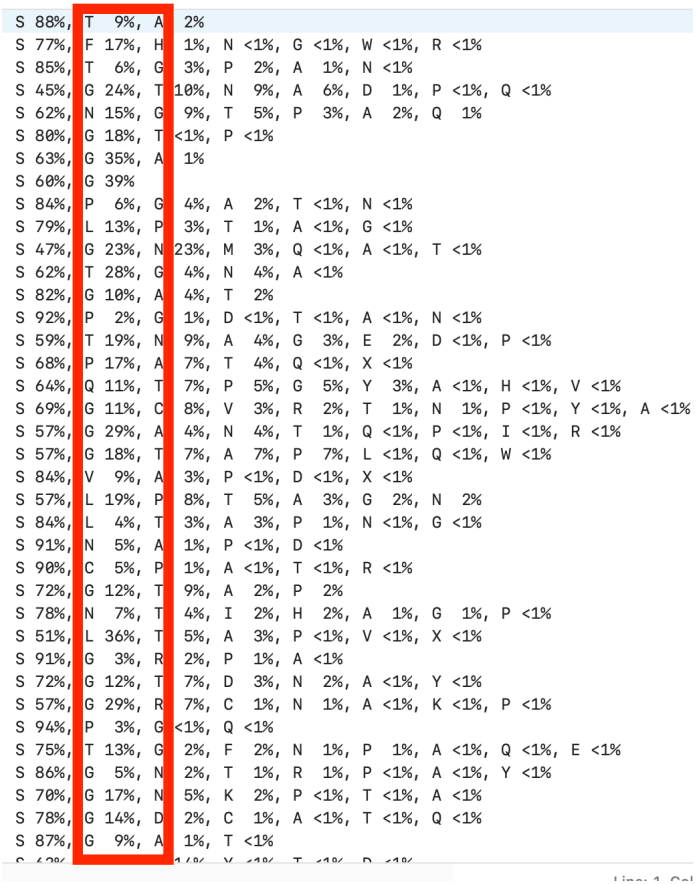
species

conservation score

QGS are much less conserved







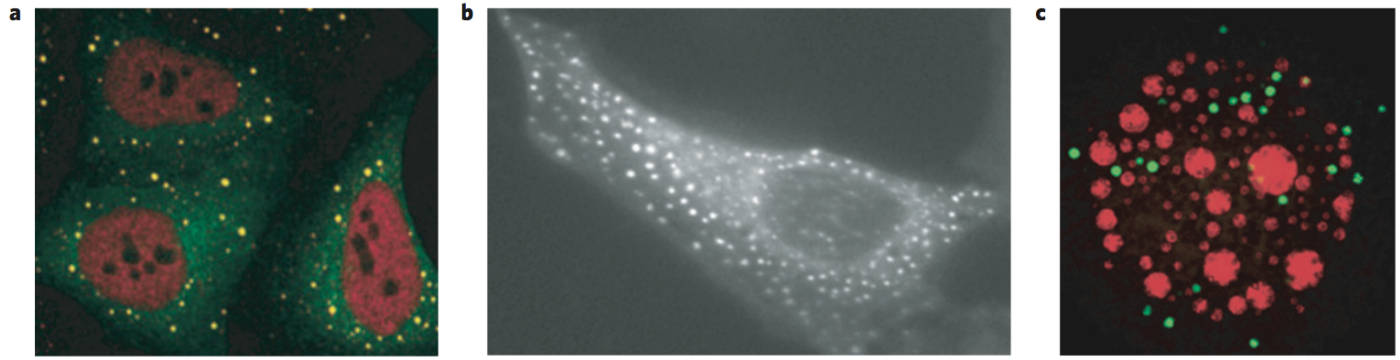
… and QGS are rarely (<1%) mutated into Y in FUS LCD (save S76Y, G168Y, 3%)

Molecular properties of FUS

- Tyrosines are uniformly distributed along the LCD of FUS
- They provide sticky sites that transiently bind molecules together
- Other residues also contribute to weak, non-specific binding of IDPs

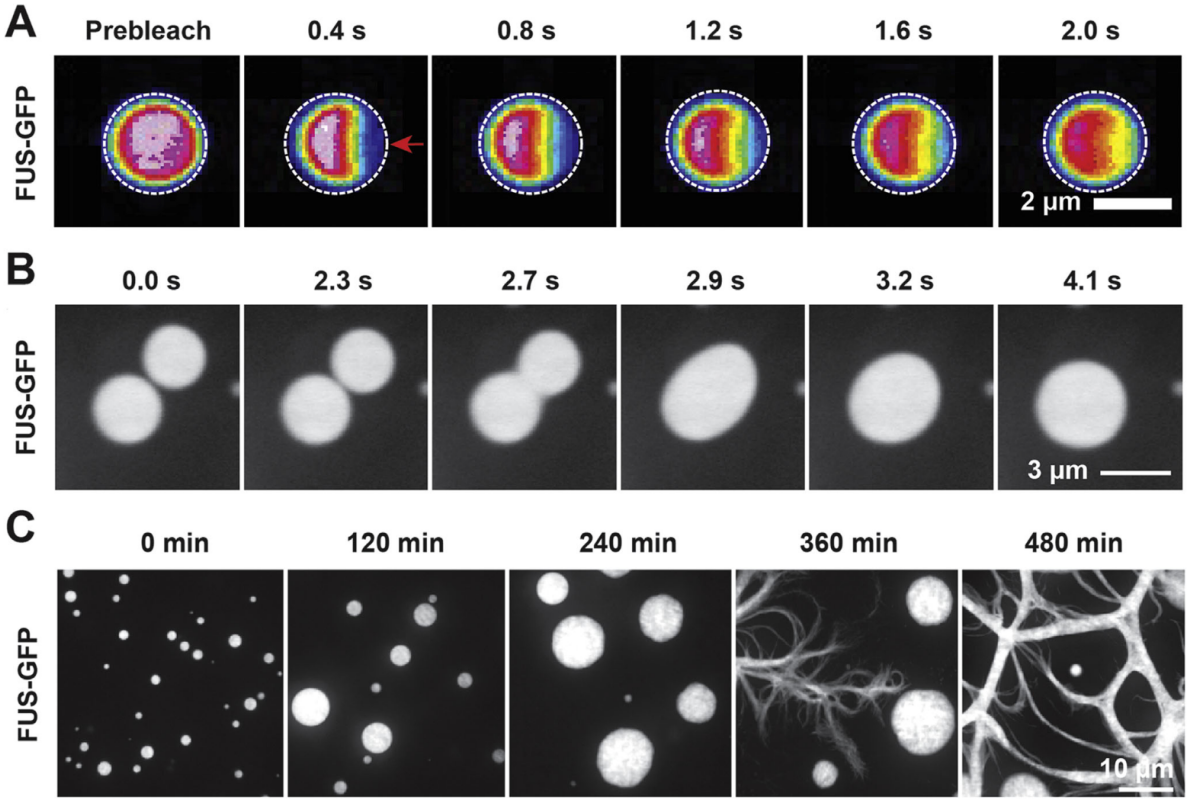
Maraldo et al. Trends in Biochemical Sciences 49:633 (2024)

Physical properties of condensates



Brangwynne et al.
Nature Phys. 11:899 (2015)

Figure 1 | Examples of membrane-less bodies in cells. a, P bodies (yellow) in tissue culture cells (adapted from ref. 63, NPG). **b**, Purinosomes (adapted from ref. 3, AAAS). **c**, Nucleoli (red) and histone locus bodies (green) in the nucleus of a large *X. laevis* oocyte (adapted from ref. 14, NPG).



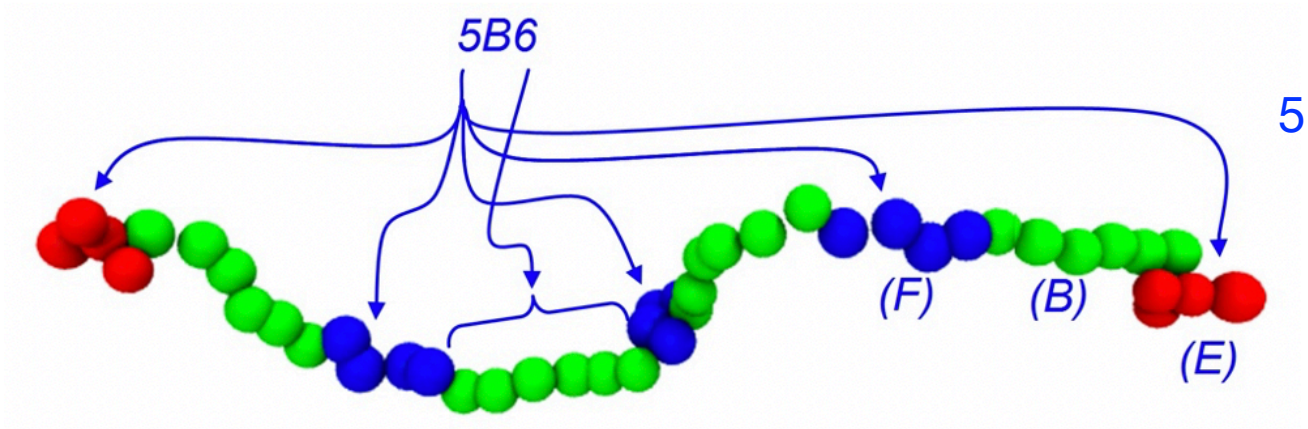
A) FRAP of FUS-GFP after bleaching at red arrow (in vitro)

B) Merging of FUS-GFP droplets under flow (in vitro)

C) Aging of droplets into fibrous aggregates (in vitro)

Alberti and Hyman,
Bioessays 38:959 (2016)

Reduce biological complexity to a model



5B6 = 5 sticky sites separated by 6 backbone beads

IDP ~ uncharged, semi-flexible polymer in good solvent conditions, with punctate sticky sites whose affinity ε^* can be tuned ($\varepsilon = 0$ = no attraction, $\varepsilon > 0$ = increasing attraction)

Sticky sites represent attractive residues, e.g., Y.

Mutations: we can add/remove sticky sites to represent mutating Y.

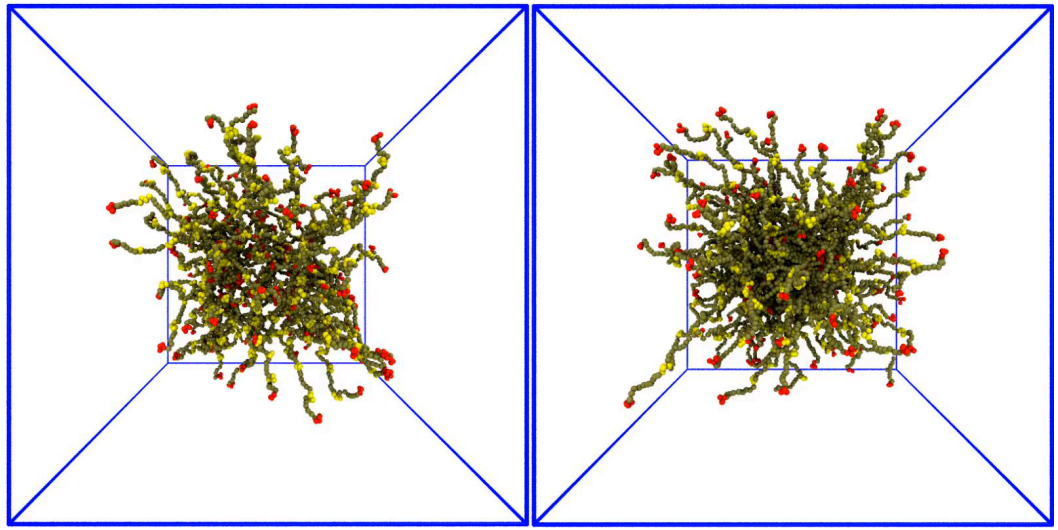
Water ~ spherical bead (always invisible)

$\varepsilon^* = (a_{EW} - a_{EE})/a_{EW} \geq 0$ in terms of the DPD conservative force parameters between sticky sites and solvent

Model IDPs self-assemble into fluid droplets

Spontaneous phase separation without hydrophobic interactions,

(i.e. in a good solvent)

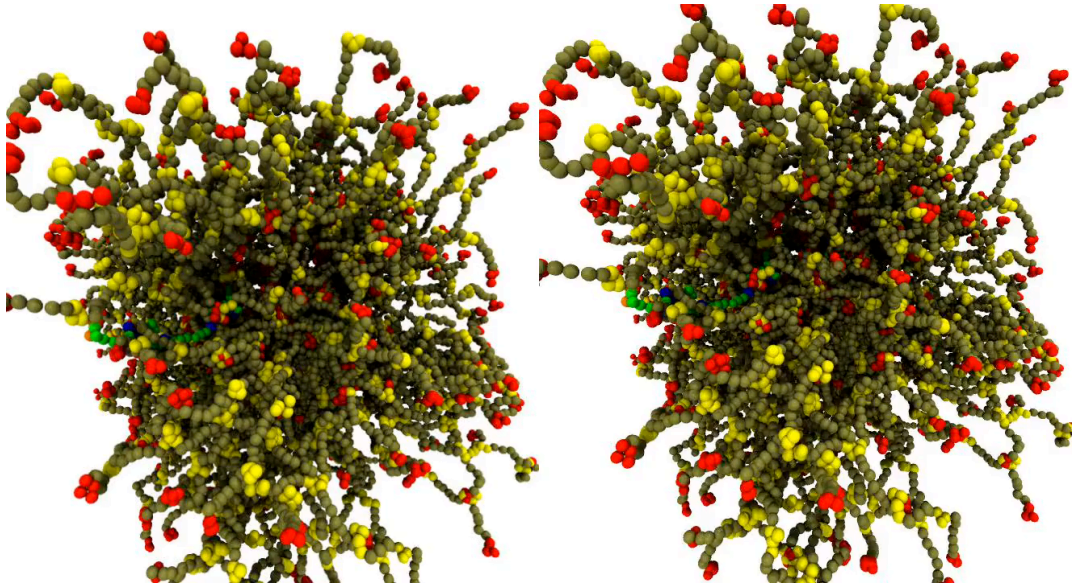


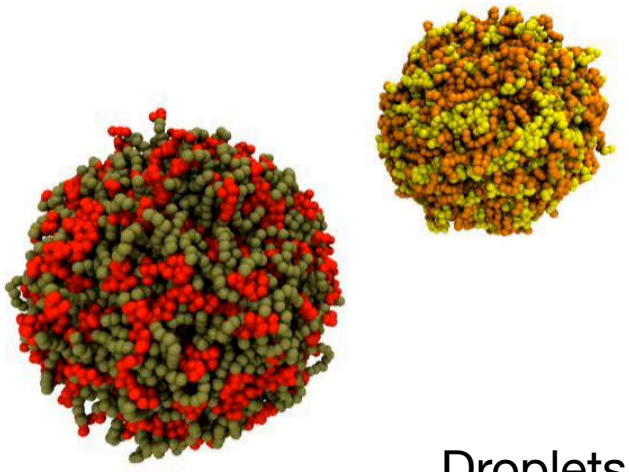
6B6 6B10

Conform. entropy increases with increasing spacing

Dense phase is a fluid,

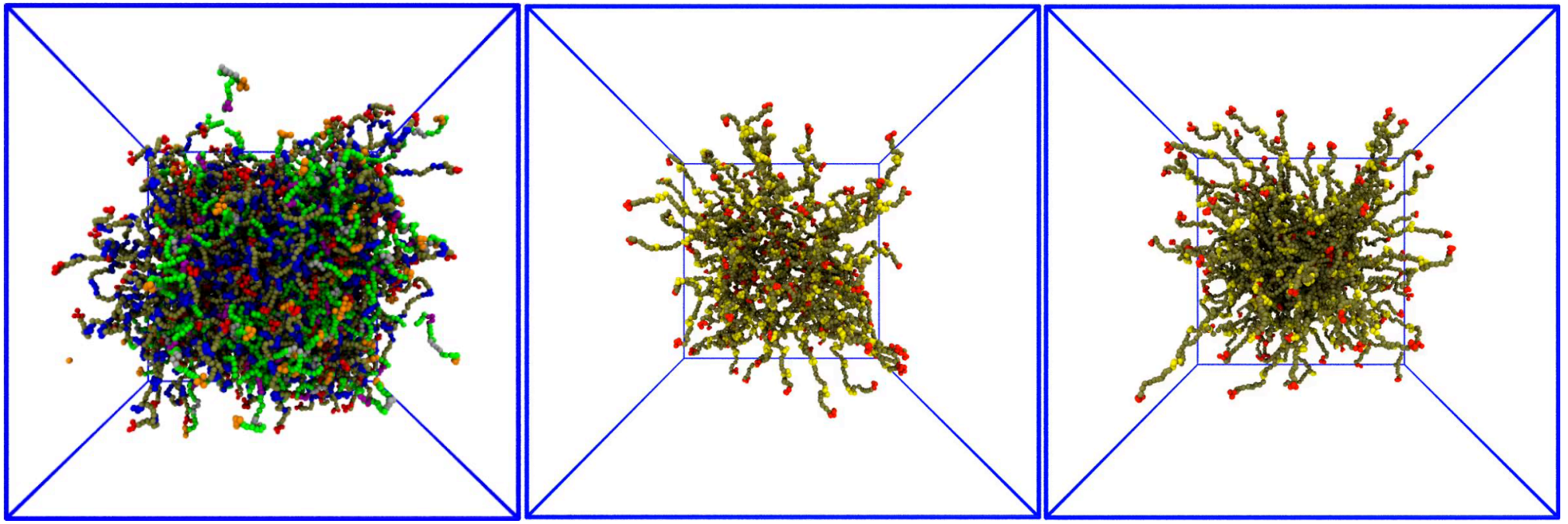
polymers fluctuate and diffuse





Droplets fuse

Affinity and separation of sticky sites are correlated



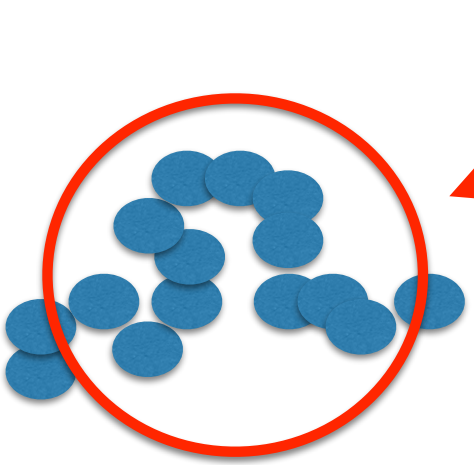
6B6 + 6B6 no affinity

6B6 with weak affinity

6B10 with same weak affinity

NB All cases are in explicit good solvent, which is invisible

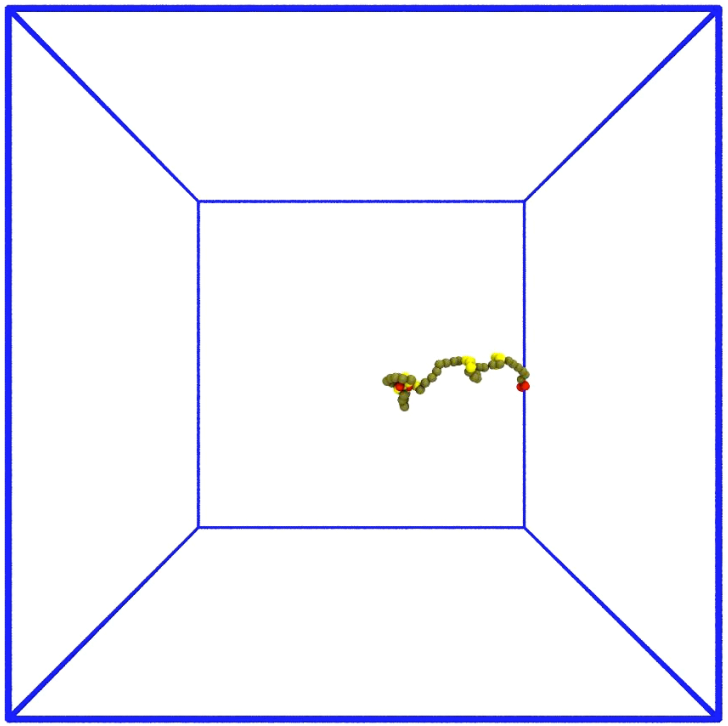
What is the *size* of a fluctuating polymer?



$$< R_{ee}^2 > = N.a^2 + < \sum r_i \, . \, r_j >$$

$$<R_g^2> = 1/N \sum (R_i - R_{cm})^2 = 1/2N^2 \sum R_{ij}^2$$

$$\left\langle \frac{1}{R_H} \right\rangle = \frac{1}{N^2} \sum_{i \neq j} \left\langle \frac{1}{r_{ij}} \right\rangle$$



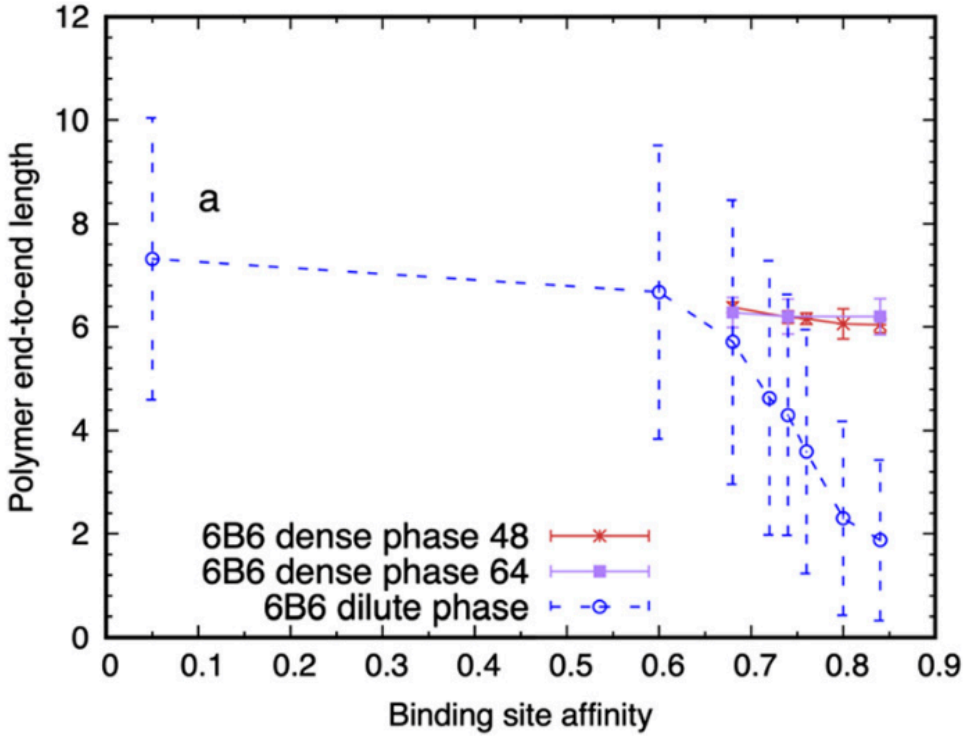
A simple way to assess structure in a disordered protein is to measure its hydrodynamic radius (R_h). The R_h is the radius of an idealized sphere that would diffuse at the same rate as the molecule of interest, and is based on the Stokes-Einstein relation in Eq. 1, where k_B is the Boltzmann constant, T is the temperature, η is the viscosity, and D is the translational diffusion coefficient. Thus, although the R_h is not a true measure of the radius of a nonglobular protein, as its diffusion is related to its nonspherical shape, it is very useful as a simple measure of compaction in disordered proteins.

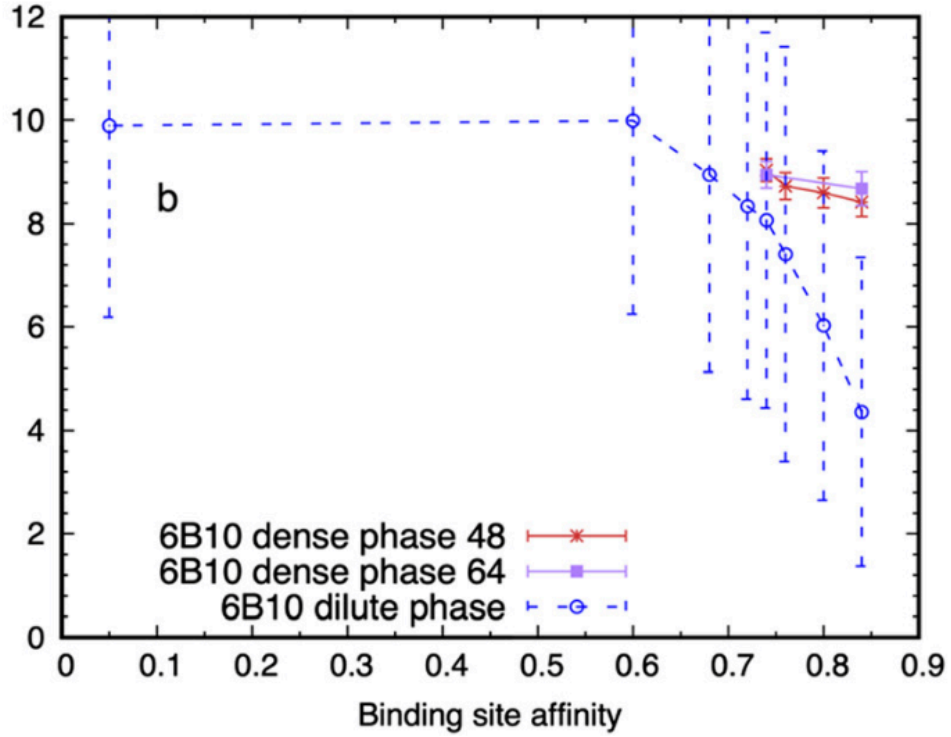
$$R_h = \frac{k_B T}{6\pi\eta D}. \qquad (1)$$

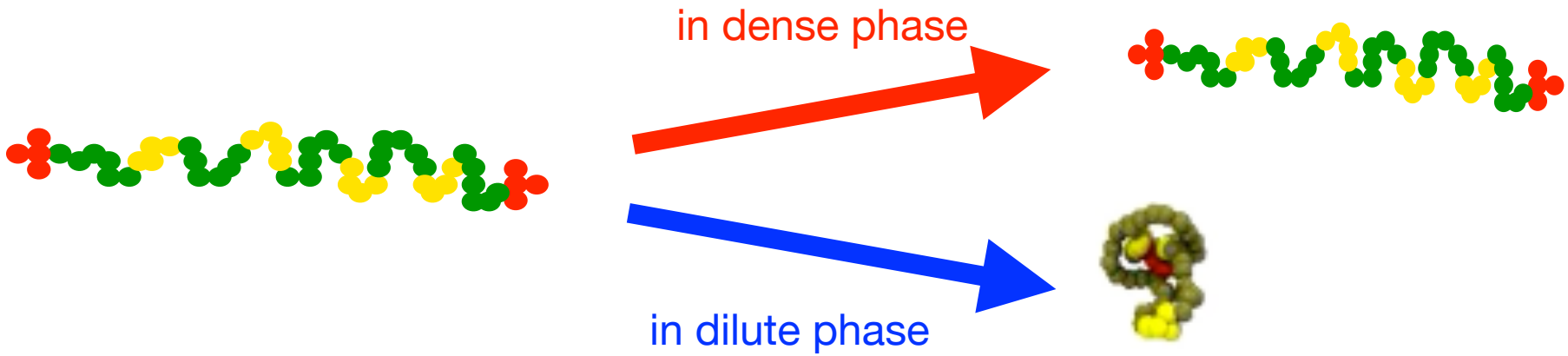
Warning this is not trivial

Marsh and Forman-Kay, Biophys. J 98:2383 (2010)

Model IDPs fluctuate as random coils in dense phase





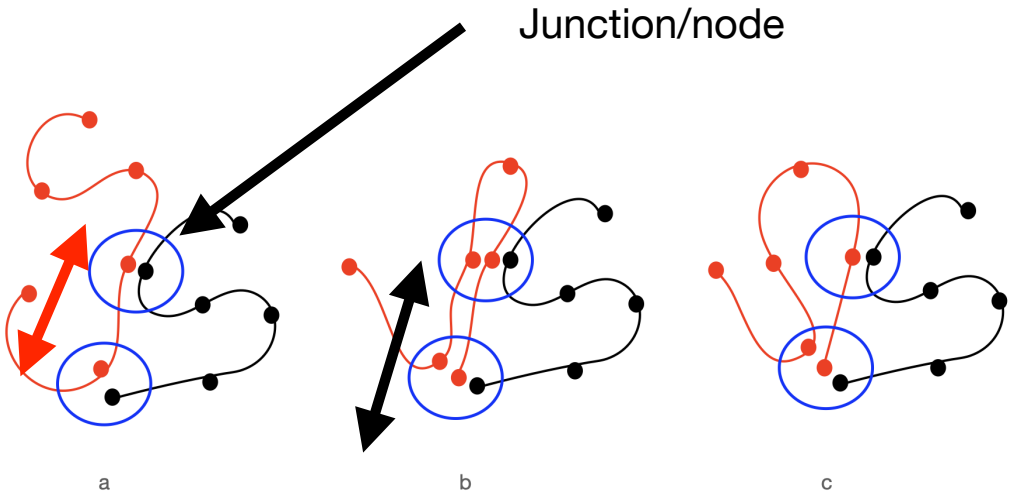


Dense phase has a selected length scale

— a selected length scale appears between junctions/nodes where sticky sites transiently bind to each other (red/yellow)

— heterogeneous spatial mass distribution

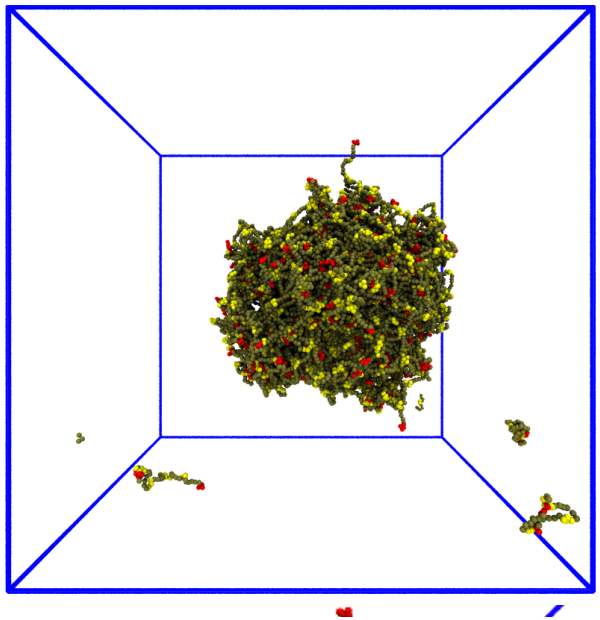
— multiple timescales for binding / unbinding / diffusion of polymers in dense phase

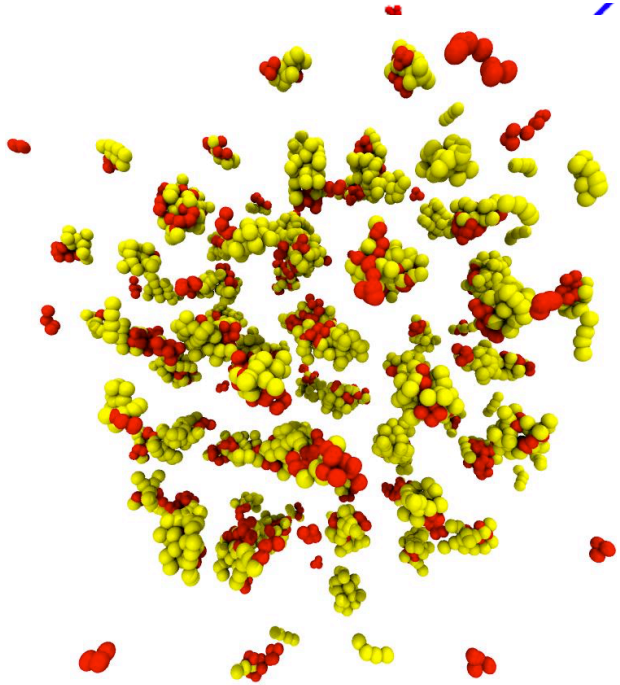


Neighbour node separation

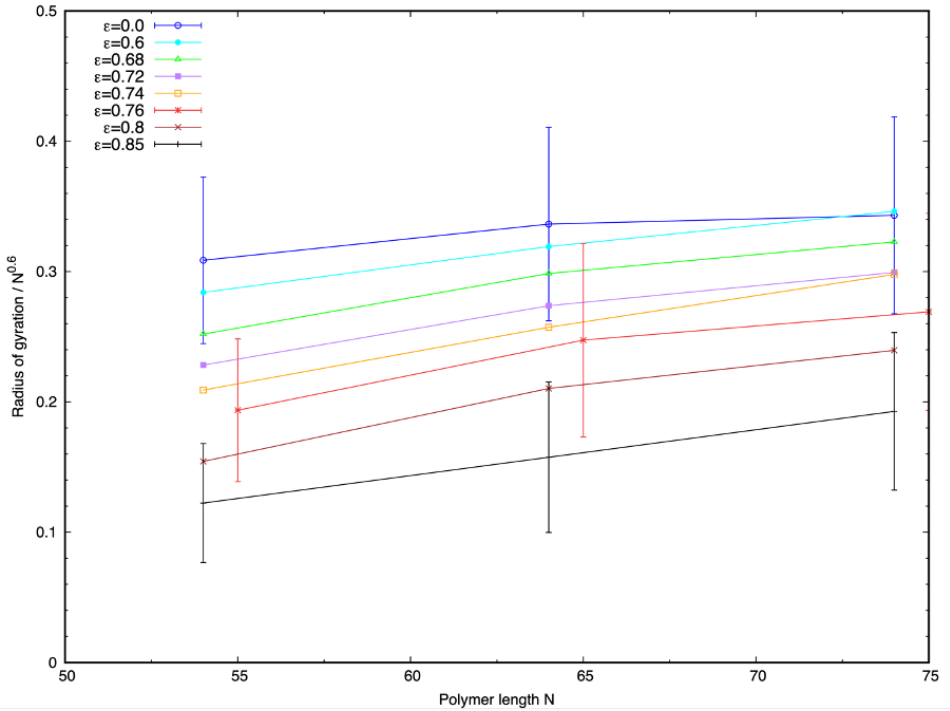
Node separation

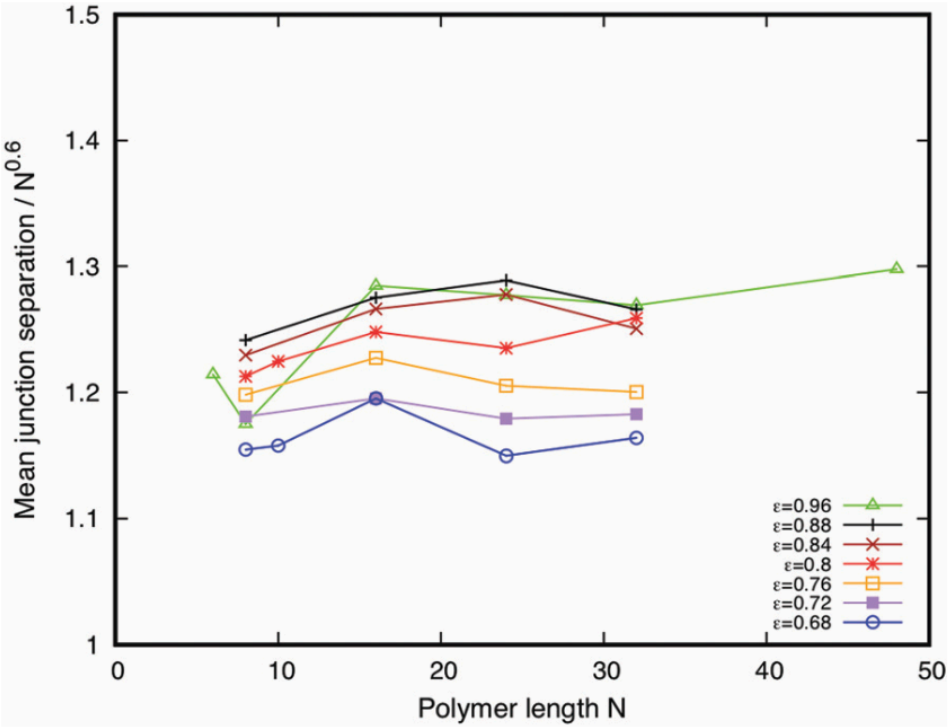
Node mass = 2





Conformational fluctuations of model IDPs





Scaling of the radius of gyration of a single polymer in the dilute phase with binding site separation length (6Bm, m = 6, 8, 10) for a range of binding site affinities.

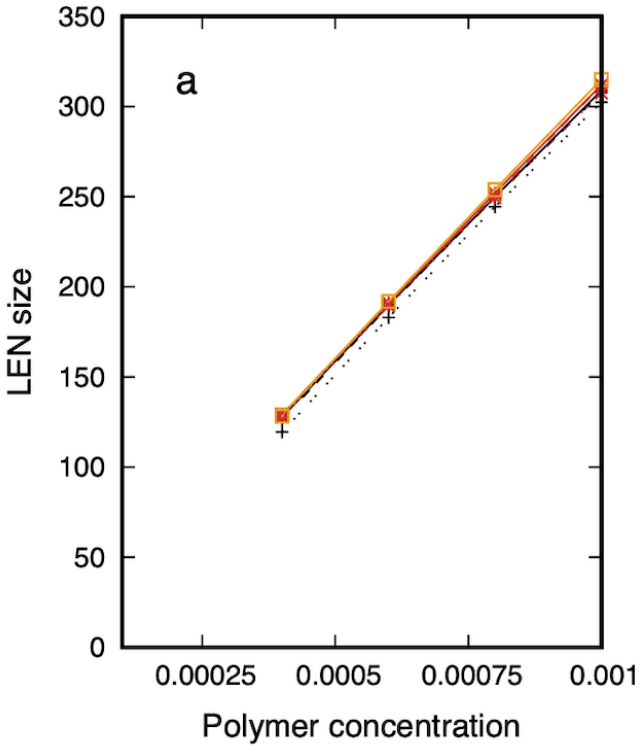
Values are divided by the expected value for a self-avoiding random walk - $R_g \sim N^{0.6}$

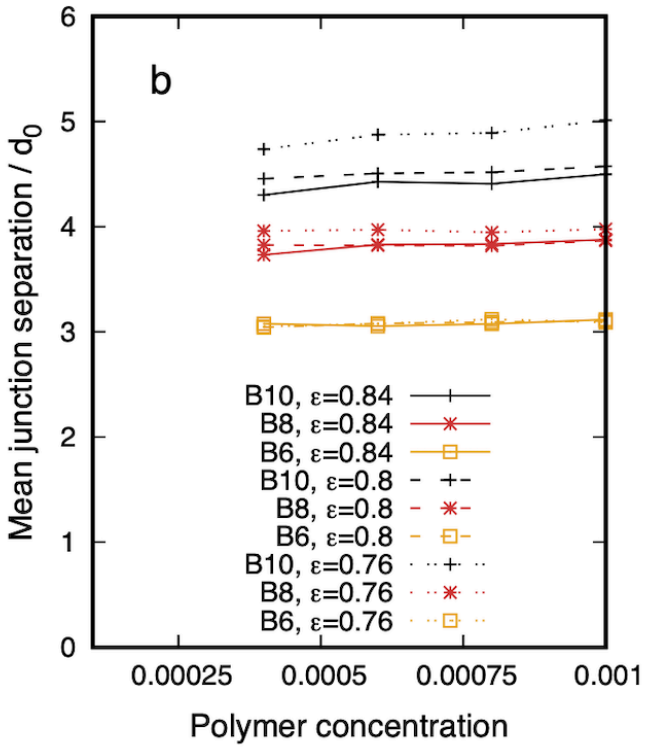
Shillcock et al., Soft Matter 18:6674 (2022)

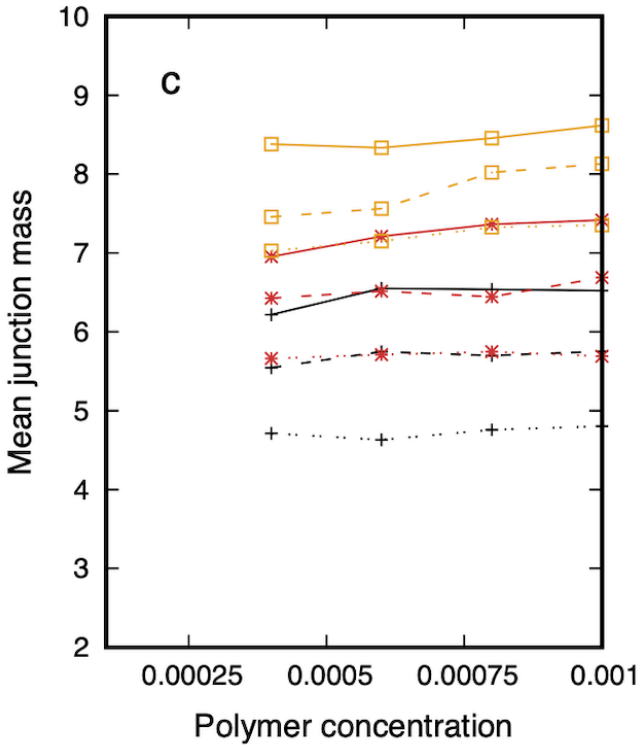
Scaling of the mean node separation in the dense phase (2BN, telechelics) with polymer length for a range of binding site affinities.

Values are divided by $N^{0.6}$

Shillcock et al., Soft Matter 16:6413 (2020)







Dense phase spatial structure depends on sticky site distribution along the polymer more than their affinity

Junction mass distribution depends on spacing and affinity of sticky sites

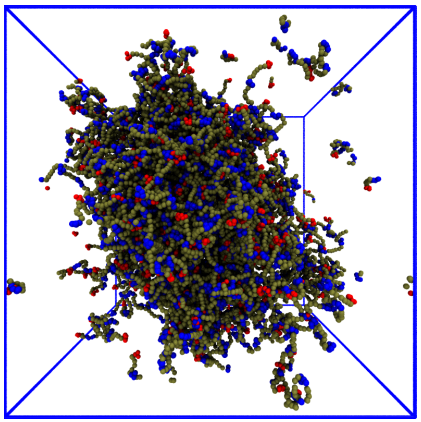
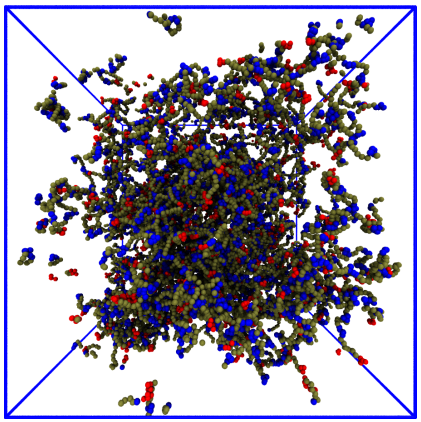
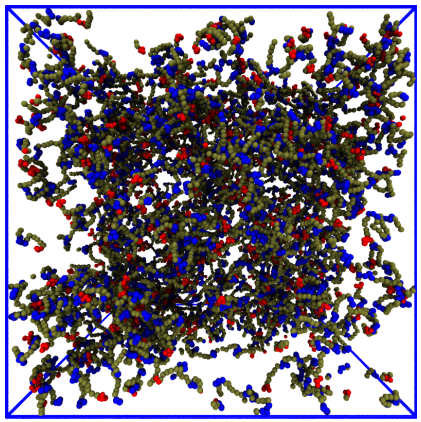
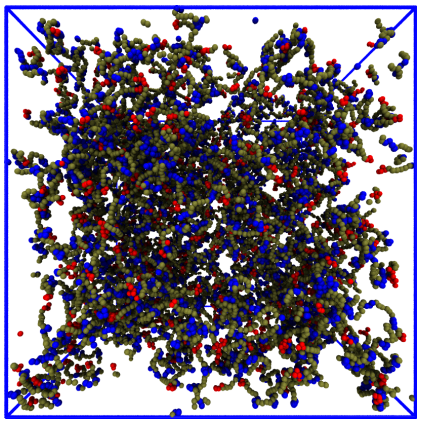
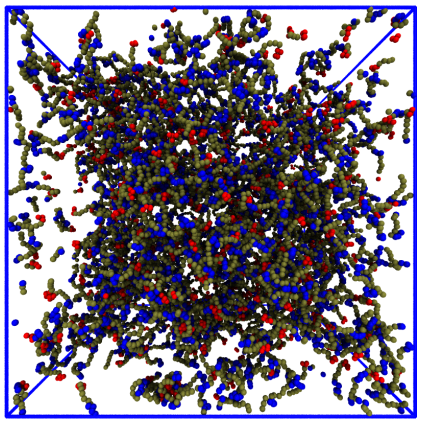
All measurements are made on the **LEN**

= **largest equilibrium network**

= largest cluster of connected polymers at each timestep

Shillcock et al. Soft Matter 18:6674 (2022)

Dense phase is grossly independent of IDP affinity

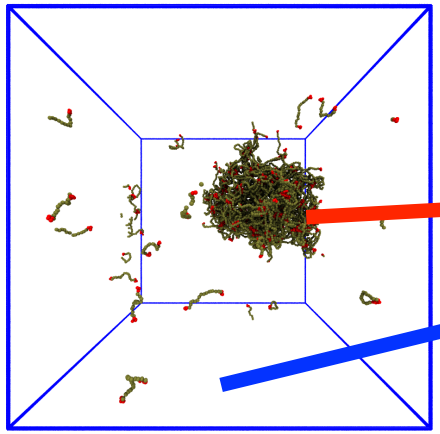
ε = 0.68 ε = 64 ε = 0.6 ε = 0.56 ε = 0.52

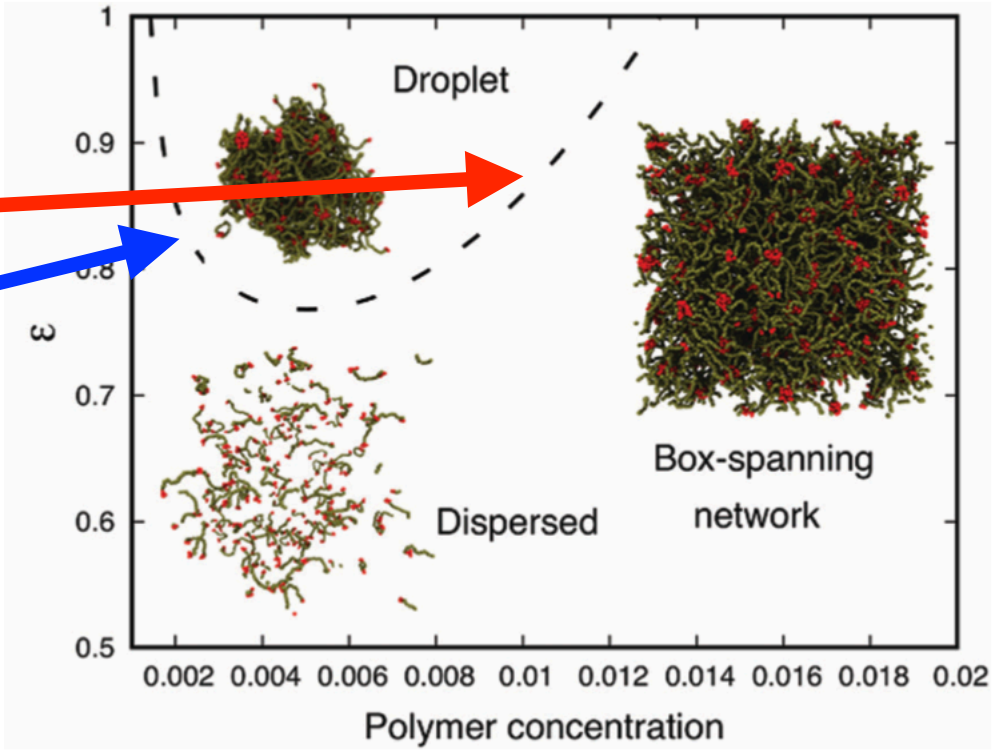
Observable	ε = 0.72	0.68	0.64	0.6	0.56	0.52	0.48
# polymers in dense/dilute phase	457 / 3	436 / 25	347 / 114	47 / 414	21 / 440	14 / 446	No network
<Node sep>	6.17 ± 2.20	6.00 ± 2.19	5.85 ± 2.20	5.74 ± 2.26	5.82 ± 2.19	5.66 ± 2.14	
<Neigh sep>	3.47 ± 0.78	3.59 ± 1.01	3.90 ± 1.41	4.42± 1.87	4.66 ± 1.89	4.62 ± 1.88	
<Mass>	6.30 ± 3.48	5.20 ± 2.58	4.20 ± 1.62	3.50 ± 0.80	3.37 ± 0.59	3.30 ± 0.54	

Not phase separated below a threshold affinity

RNA4a/r0, RNA3/r3

How do we know it's a phase?





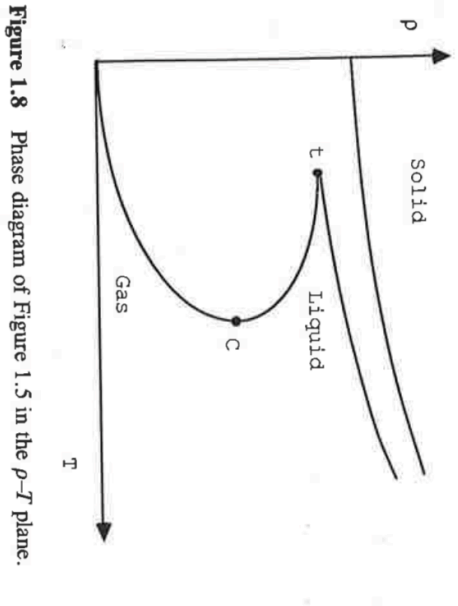


Figure 1.8 Phase diagram of Figure 1.5 in the ρ–T plane.

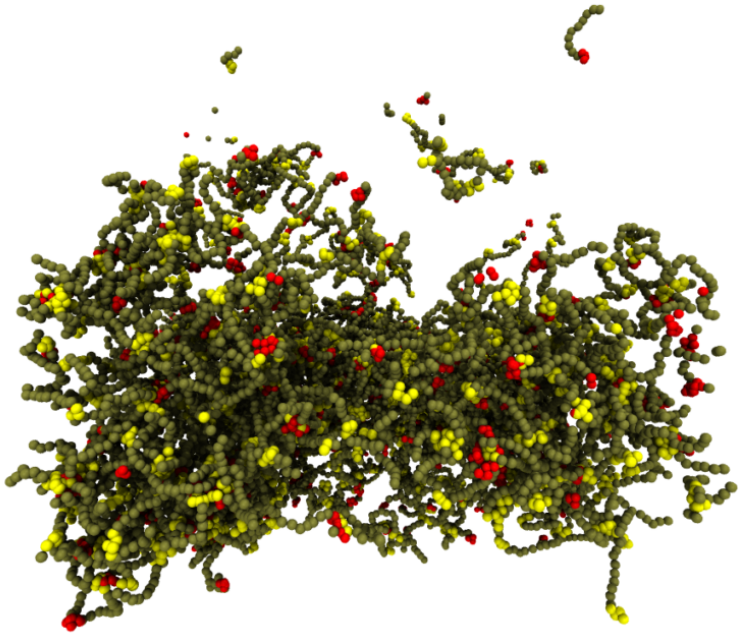
Higher T ~ lower ε

Thermodynamics:

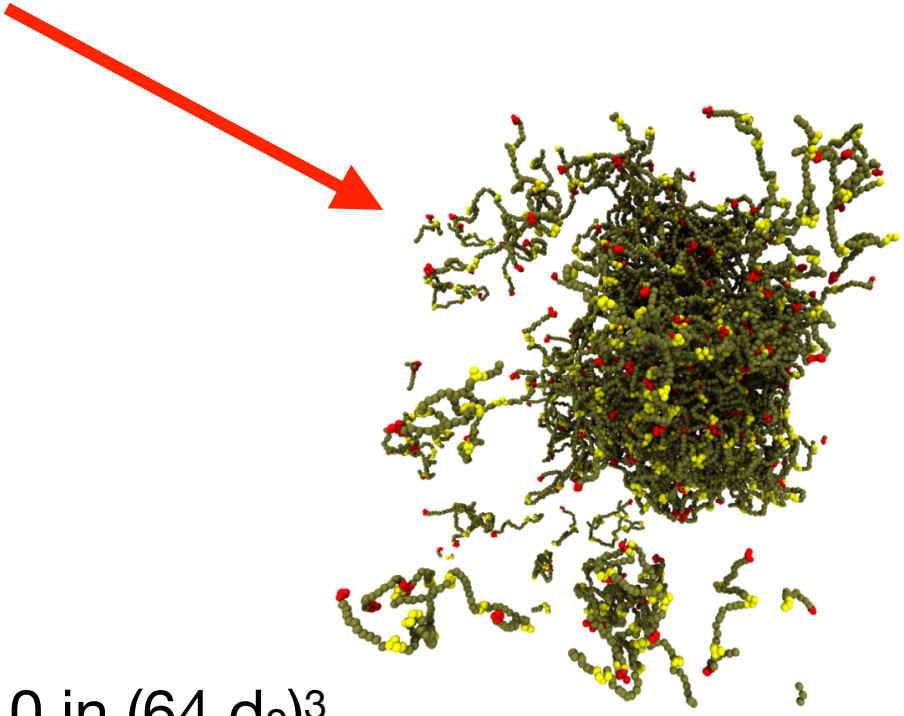
" A ***phase*** is an homogeneous region of matter bounded by surfaces across which its properties change discontinuously, e.g., solid, liquid or gas."

Shillcock et al., Soft Matter 16:6413 (2020)
Shillcock et al., Soft Matter 18:6674 (2022)

The Lever rule



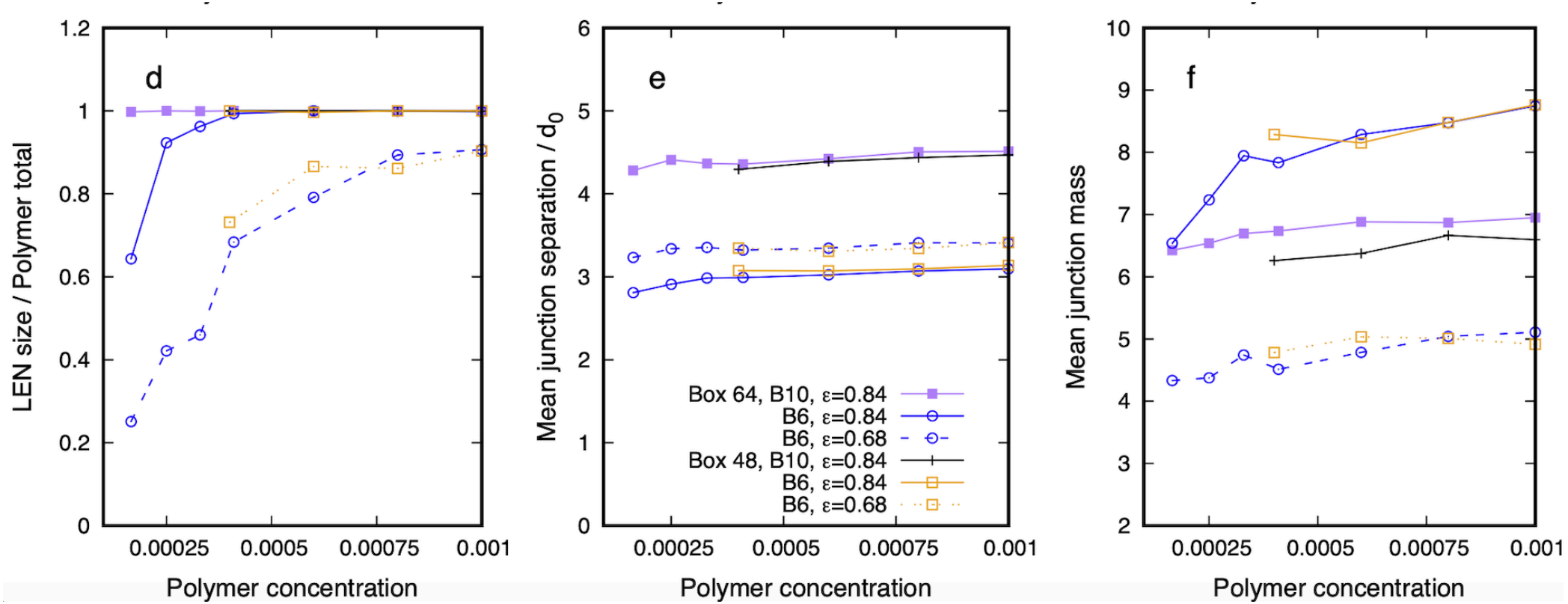
309 6B10 in (48 d_0)^3



313 6B10 in (64 d_0)^3

Dense phase shrinks in a larger box to keep dilute phase conc. constant.

Structure is independent of system size



Dense phase properties are independent of the simulation box size, and polymer concentration.

This is a key check of any simulation to ensure the box size is not influencing the results.

What do experiments find?

4 observations on FUS:

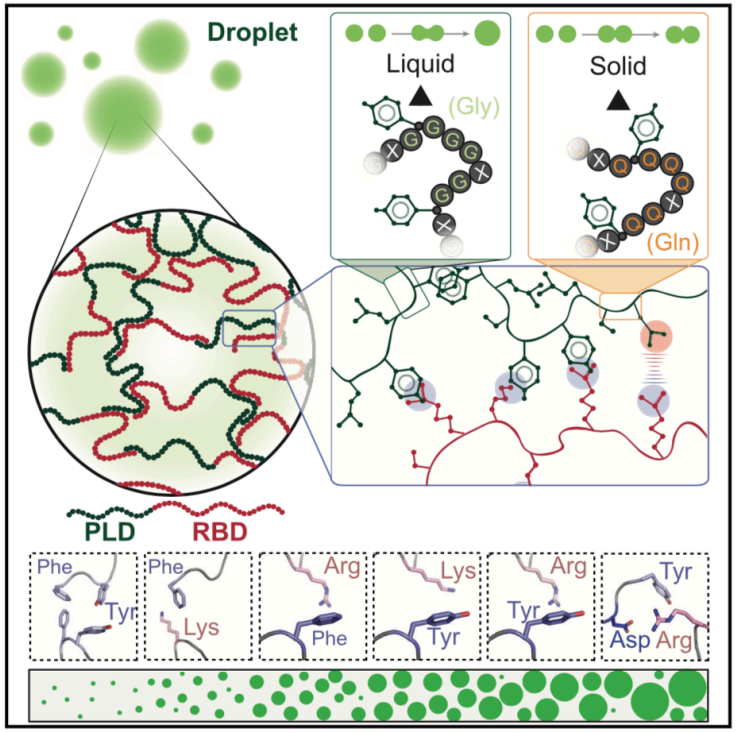
- Dense phase is mainly solvent (unlike oil/water)
- IDPs fluctuate as much in the dense phase as dilute phase
- Mass distribution (= polymers binding to each other) is heterogeneous
- PTMs at the termini of IDPs are involved in regulation

In vitro experiments on FUS LCD

Little/no secondary structure in condensed phase,
ES interactions are unimportant,
FUS LCD interactions are transient, weak, *multivalent*
(Burke et al. Mol. Cell. 60:231 2015)

Phase separated droplets of FUS LCD are ~ 65% solvent by volume, with protein concentration 2 - 30 mM
(Murthy et al. Nature. Struc. Mol. Biol. 26:637 (2019)

Disordered NTD and CTD are required for full FUS LLPS, but NTD alone is sufficient
(Wang et al. Cell 174:1 (2018)



J. Wang et al., Cell, 174:688 (2018)

Translational diffusion is slower (x10-100) than in bulk (Burke, 2015)

Ensemble structure of the N-terminal domain (1–267) of FUS in a biomolecular condensate

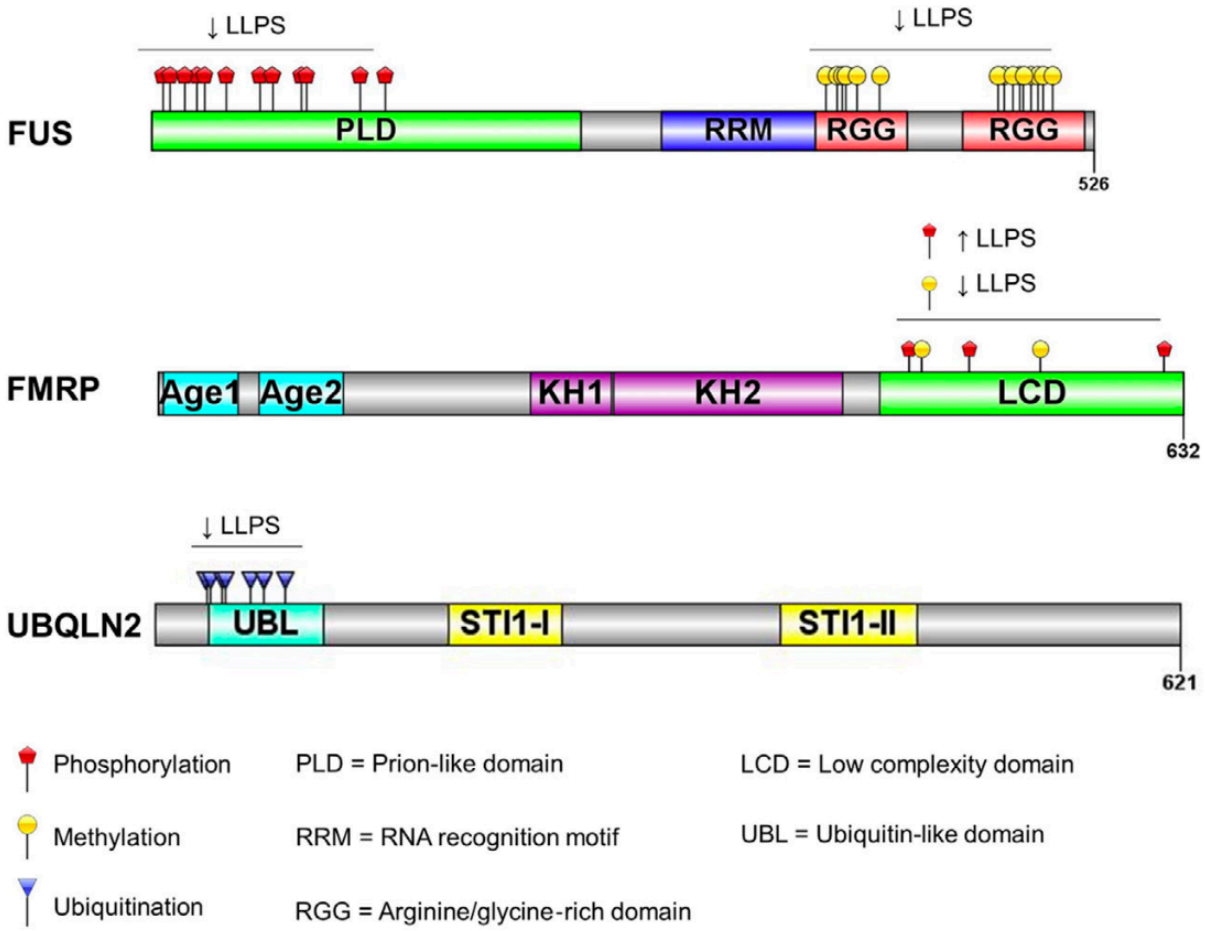
- Specific intramolecular interactions are *negligible* for FUS NTD in the condensed state
- QGSY-rich domain (1 - 165) is *more extended* and the RGG1 domain (166 - 267) more compact than average.
- Some intramolecular interactions in FUS NTD are *replaced* by intermolecular interactions during LLPS.

modeling and three such restraints for model validation, we have found that FUS NTD behaves as a random-coil polymer under good-solvent conditions in both the dispersed and condensed state. Conformation distribution in the biomolecular condensate is virtually indistinguishable from the one in an unrestrained ensemble, with the latter one being based on only residue-specific

Prediction: Shillcock et al., Soft Matter 18:6674 (2022)

Experiment: Esteban-Hofer et al., Biophys. J. 123:538 (2024)

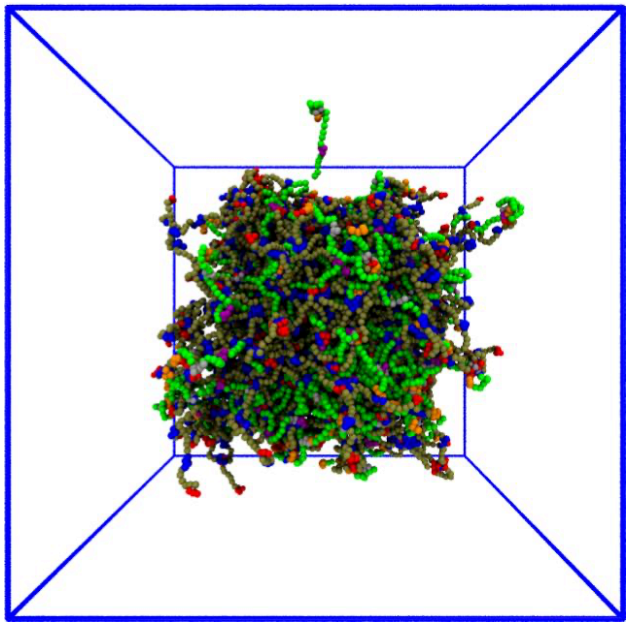
PTMs frequently cluster at the termini of IDPs



Why?

- accessibility for regulation
- has largest effect on conformational fluctuations

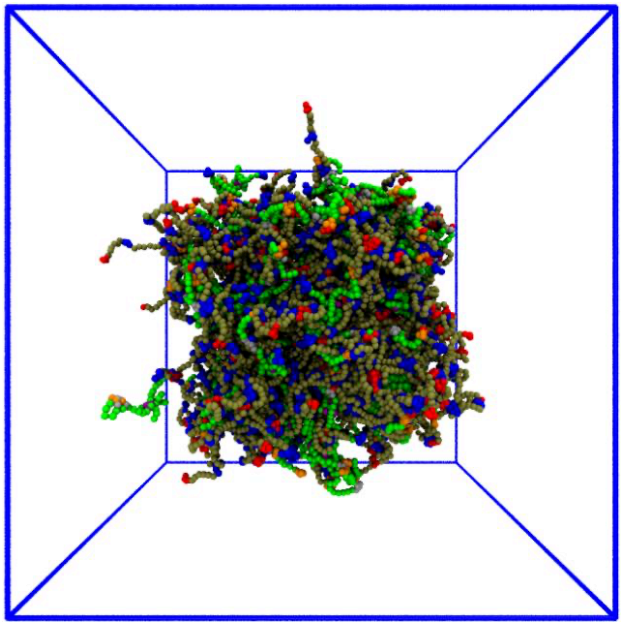
Why might PTMs cluster at termini? Regulation?



5555

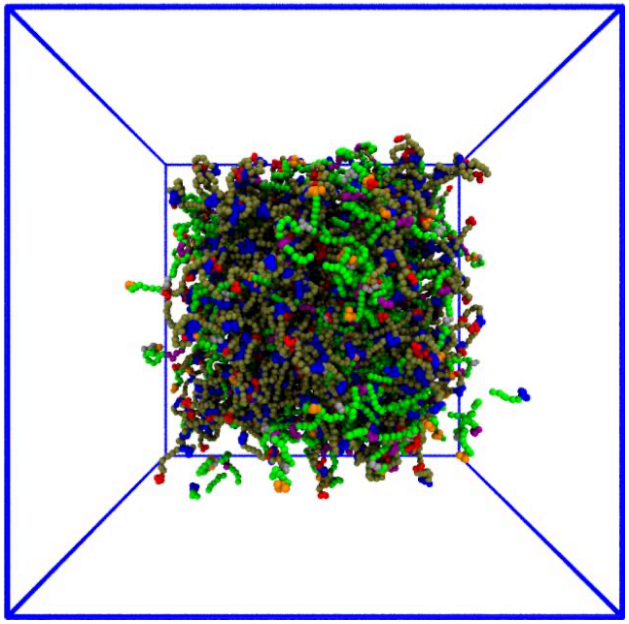
All affinities equal

(both species identical)



55825

One endcap inert in minor species



552525

Both endcaps inert in minor species

Story so far …

- Single-component IDP with multiple sticky sites phase separates (above a threshold affinity) into a dense phase
- “Dense” phase is ~70% solvent
- Spatial length-scale is selected by distribution of sticky sites
- IDPs show random coil fluctuations in dense / dilute phases
- Dense phase structure is independent of IDP affinity (above threshold)
- Sticky endcaps are required for dense phase formation:
  - Enabling / disabling one endcap destabilises the dense phase
  - Disabling both endcaps eliminates phase separation

Caveat

Commentary

TRANSPARENT PROCESS OPEN ACCESS THE EMBO JOURNAL

On the role of phase separation in the biogenesis of membraneless compartments

Andrea Musacchio*

erroneous. Nonetheless, the arguments presented here provide compelling evidence that the majority of PS claims have been based on highly incomplete tests that ignored more plausible drivers of macromolecular concentration. This criticism applies to the vast majority of PS claims in the literature, as can be easily verified retro-

CellPress **Molecular Cell**

Commentary

Intrinsically disordered regions (IDRs): A vague and confusing concept for protein function

Kevin Struhl[1,*]
[1]Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA
*Correspondence: kevin@hms.harvard.edu
https://doi.org/10.1016/j.molcel.2024.02.023

The term "intrinsically disordered region" (IDR) in proteins has been used in numerous publications. However, most proteins contain IDRs, the term refers to very different types of structures and functions, and many IDRs become structured upon interaction with other biomolecules. Thus, IDR is an unnecessary, vague, and ultimately confusing concept.

Musacchio, EMBO J. 41: e109952 (2022)
Hedtfeld et al. Mol. Cell 84:1 (2024)

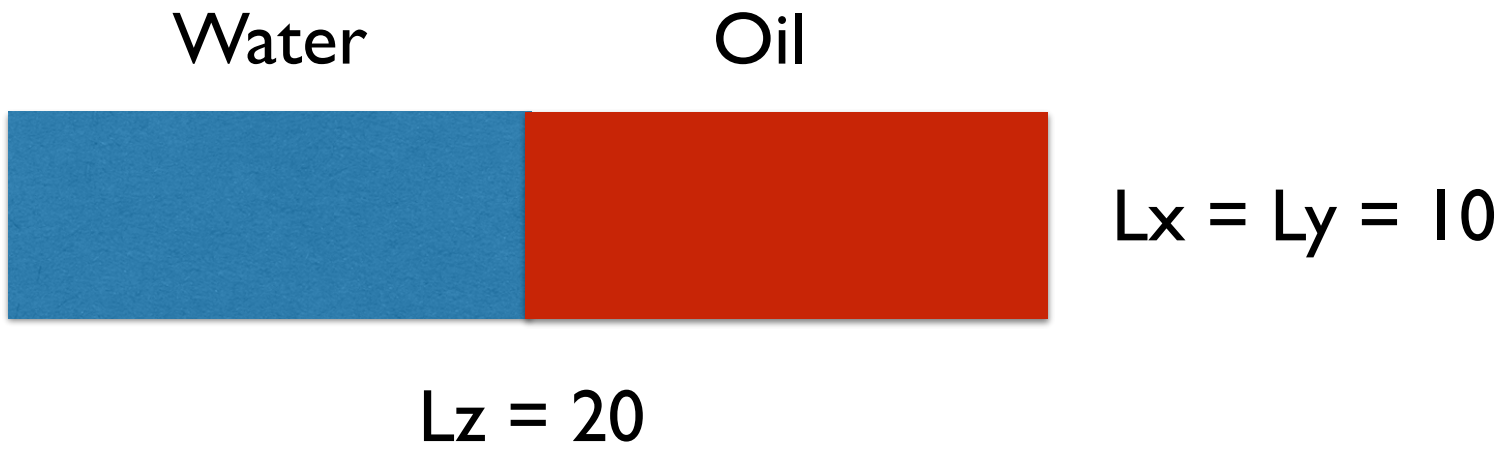
Break / Exercise

Measuring interfacial tension with an "Analysis" object

The **dmpcas** file (analysis state file) contains general measurements of temperature, pressure, bond lengths, polymer end-to-end lengths, etc.

Analysis Objects are used to measure detailed observables for specific sysetms, e.g., micelle, vesicle planar membrane, planar interface, etc

The **interface** analysis object measures the surface tension of a planar interface between two fluids, e.g.



Measuring interfacial tension with an "Analysis" object (dmpci.if1)

```
Bead     W
         0.5
         25
         4.5

Bead     O
         0.5
         75    25
         4.5    4.5

Bond     O  O    128     0.5


Polymer Water   0.5    " (W) "
Polymer Oil     0.5    " (O O O O) "

Box      10   10   20        1 1 1
```

Create two bead types to make the oil and water molecules, and make their mutual repulsion large (75 here)

```
Analysis
Type          interface
Times         10000   20000
Polymers      Oil     Water
Thresholds    0.1   0.9
Normal        0 0 1
```

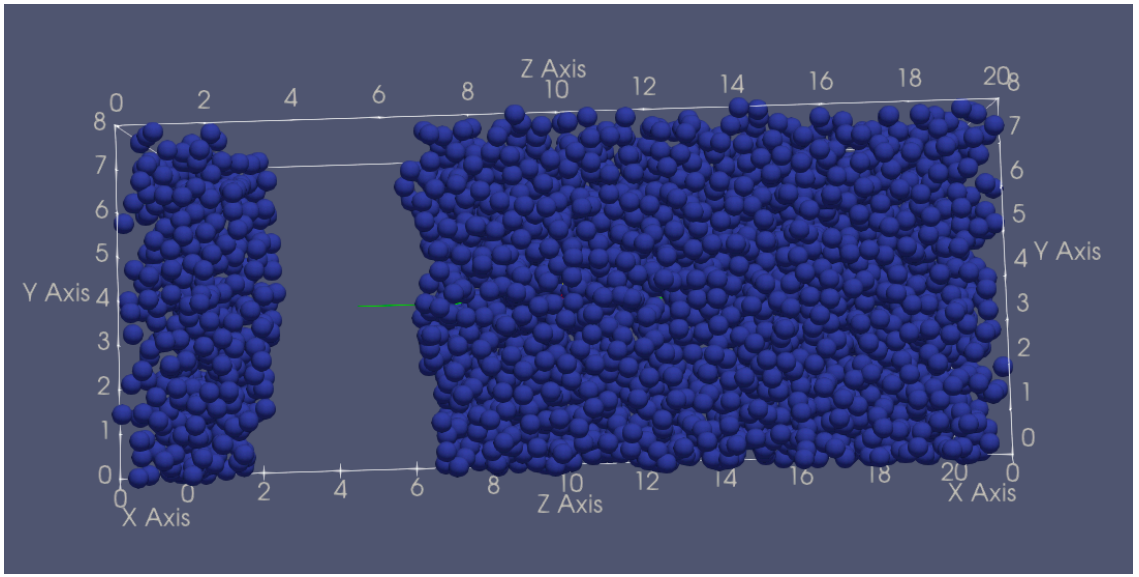
The interface analysis object requires arguments that define: duration of analysis, the two species names, min/max densities to define bulk phases, and the interface normal

A long thin box ensure rapid equilibration and a single interface

Time averaged data are in the dmpcas file

```
Oil Density Profile
3.1401562       0.087261254
3.1107813       0.083767926
1.0951563       0.077216436
0.00296875      0.0075567514
0               0
0.000625        0.0037759519
0.7215625       0.097862407
3.0770313       0.12016539
3.1026562       0.097183355
3.126875        0.089820785
3.1201562       0.086509319
3.1096875       0.095171002
3.0765625       0.11034171
2.9228125       0.11775529
2.9664062       0.11597024
3.025625        0.13017641
3.0596875       0.12186258
3.1145313       0.096497627
3.1089062       0.097112989
3.1178125       0.097838455
```

```
Water Density Profile
0               0
0.010625        0.014967413
1.3253125       0.08658506
2.8479688       0.082999641
2.8473437       0.090243134
2.8395312       0.084746146
1.695           0.088539583
0.02796875      0.029997965
0.000625        0.0030618622
0.00078125      0.0034053898
0.0025          0.0057282196
0.00359375      0.010792567
0.03859375      0.047979437
0.14015625      0.066748383
0.10140625      0.069719548
0.06609375      0.069028378
0.043125        0.051307133
0.0090625       0.021100115
0.0003125       0.0021875
0               0
```



Density profiles along Z axis

```
Summed High Density
3.2779687       0.047511049

Summed High Location
10.6            5.9497899

Summed Low Density
2.3703125       0.064100598

Summed Low Location
3.88            1.9963968
```

Max / min densities and location

Stress profile,
and surface tension

```
Interfacial Tension
6.2609146       5.4967365

Total Stress Profile
0.31304573      0.27483682

Oil-Oil Stress Profile
-0.34455214     0.19631646

Oil-Water Stress Profile
0.86872295      0.064329826

Water-Water Stress Profile
-0.27504417     0.075758405
```

Time series data are in the dmpcaa file

dmpcaa.interface1.if1 — file name includes the analysis object's type, an integer representing how many have been created, and the runid

```
Analysis
    Type         interface
    Times        10000  20000
    Polymers     Oil  Water
    Thresholds   0.1  0.9
    Normal       0  0  1
10100       0           6           3.11049     0.028125    2           2.85417     2.29688     6           3.25        15          8.37211
10200       0           6           3.07187     0.0270833   2           2.86458     2.39062     2           3.23438     8           8.02143
10300       0           6           3.11298     0.0270833   2           2.85938     2.45312     6           3.25        18          10.5926
10400       0           6           3.10379     0.0270833   2           2.79688     2.42188     6           3.23438     11          8.72927
10500       0           6           3.125       0.028125    2           2.85938     2.39062     2           3.25        10          8.25609
10600       0.00520833  6           3.15104     0.0333333   1           2.76042     2.48438     2           3.375       18          10.1377
10700       0           6           3.14062     0.0291667   2           2.84896     2.42188     6           3.35938     12          5.67956
10800       0           1           3.14773     0.0302083   1           2.80208     2.34375     6           3.3125      0           1.50118
```

```
pISimBox->GetCurrentTime(), "Time");
m_MajorLowMean,   "MajorLow");
m_MajorMidPoint,  "MajorMid");
m_MajorHighMean,  "MajorHigh");
m_MinorLowMean,   "MinorLow");
m_MinorMidPoint,  "MinorMid");
m_MinorHighMean,  "MinorHigh");
m_SumMinDensity,  "SumMin");
m_SumMinId,       "SumMinLoc");
m_SumMaxDensity,  "SumMax");
m_SumMaxId,       "SumMaxLoc");
m_SurfaceTension,"SurfaceTension");
```

These are the observables

The two species are (arbitrarily) labelled the major and minor species.

The value of/ location of their highest and lowest densities are given.

Exercise

1) Repeat the simulations using longer oil polymers:

   (8 O), (16 O)

How does the surface tension change?

2) Return to (4 O), and reduce the aOW repulsion parameter to 50, 40, 30.

How does the surface tension change?